American Economic Journal: Applied Economics 2022, 14(3): 329–366
https://doi.org/10.1257/app.20180385

Breastfeeding and Child Development†

By Emla Fitzsimons and Marcos Vera-Hernández*

We show that children who are born at or just before the weekend are less likely to be breastfed, owing to poorer breastfeeding support services in hospitals at weekends. We use this variation to estimate the effect of breastfeeding on children's development in the first seven years of life, for a sample of births of low-educated mothers. We find large effects of breastfeeding on children's cognitive development but no effects on health or noncognitive development during the period of childhood we consider. Regarding mechanisms, we study how breastfeeding affects parental investments and the quality of the mother-child relationship. (*JEL* I12, I14, I18, J13, J16, J24)

There is little doubt that conditions in early childhood can have long-lasting effects on human capital, reinforcing the intergenerational transmission of wealth as well as human capital (See Almond, Currie, and Duque 2018; Almond and Currie 2011; Black and Devereux 2011; Currie and Almond 2011; Cunha, Heckman, and Schennach 2010; Cunha and Heckman 2007; Case, Lubotsky, and Paxson 2002). However, much less is known about the key contributors to the intergenerational gap. Breastfeeding has the potential to play a key role both because of claims regarding its beneficial effects on child development and its stark socioeconomic gradient—in 2018, 72 percent of children whose mothers were college graduates were being breastfed at 6 months, compared to 45 percent of children whose mothers had less than

* Fitzsimons: UCL Institute of Education and Institute for Fiscal Studies, London (email: e.fitzsimons@ucl.ac.uk); Vera-Hernández: University College London and Institute for Fiscal Studies, London (email: m.vera@ucl.ac.uk). Seema Jayachandran was coeditor for this article. We thank Douglas Almond, Michael Anderson, Manuel Arellano, Orazio Attanasio, Martha Bailey, Michael Baker, Jo Blanden, Richard Blundell, Stéphane Bonhomme, Olympia Bover, Jesús Carro, Ken Chay, Gabriella Conti, Janet Currie, Emilia Del Bono, Maria Iacovou, Toru Kitagawa, Victor Lavy, Valerie Lechene, Sandra McNally, Alan Manning, Ellen Meara, Costas Meghir, Guy Michaels, Adam Rosen, Uta Schönberg, Rachel Soloveichik, John van Reenen, and Jeffrey Wooldridge for very helpful discussions, as well as participants at seminars in Alicante, UC Berkeley, Bologna, Bristol, CEMFI. Duke, City University of London, Erasmus University, London School of Economics, the IFS, the NBER Spring Meeting, Barcelona Summer Forum, and UPenn. We are extremely grateful to Inês Black, Nancy Daza Báez, Soledad Giardili, Jeremy McCauley, and Carys Roberts for outstanding research assistance and to Rebecca Hamlyn, Jon Johnson and Rachel Rosenberg for assistance with data. We are grateful to the Centre for Longitudinal Studies, UCL Institute of Education for the use of the Millennium Cohort Study data and to the UK Data Service for making them available, as well as The Picker Institute Europe for making the 2007 Maternity Users Survey available. We also thank NHS Digital for access to the Hospital Episode Statistics under data sharing agreement CON-205762-B8S7B. This paper has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 695300-HKADeC-ERC-2015-AdG), and from the Economic and Social Research Council (grant reference ES/H021221/1). Our warmest and most tender thanks to Corinne Vera, who was born on a Saturday and inspired this paper! All errors are the responsibility of the authors.

† Go to https://doi.org/10.1257/app.20180385 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

high school education (CDC 2020). However, with the exception of one randomized controlled trial (Kramer, Aboud, et al. 2008; Kramer, Fombonne, et al. 2008; Kramer et al. 2001) that randomized 31 hospitals in Belarus into a health care worker assistance program (designed by UNICEF) to increase breastfeeding or a control, most of the claims about breastfeeding's beneficial effects on child development are based on observational studies. The challenge is to define an empirical strategy that provides credible causal evidence, thus helping to understand its role in child development.

This paper estimates the causal effects of breastfeeding on child development at various ages up to age seven. It exploits the authors' novel observation that in the United Kingdom the timing of birth affects breastfeeding for low-educated mothers. In particular, amongst this group of mothers, breastfeeding rates are lower for those who give birth just before or early into the weekend compared to those who give birth at any other time during the week. We argue that this is because the provision of infant feeding support in UK hospitals is lower during weekends than during the week. Without early hands-on support at the hospital, it is much more difficult for successful breastfeeding to be established. At the same time, we provide extensive evidence that maternal and birth-related characteristics do not vary by timing of birth, nor do a range of other hospital maternity services vary by timing of birth. Timing of delivery therefore provides a credible source of exogenous variation that we use as an instrumental variable (IV) for breastfeeding.

Our estimates, based on the UK Millennium Cohort Study (University of London, Institute of Education, Centre for Longitudinal Studies 2022), show that breastfeeding has large positive effects on the cognitive development of children whose mothers have relatively low levels of education, of around 0.5 of a standard deviation, though the confidence interval is wide. We detect no evidence of any benefits for children's health, though we note that most health outcomes are self-reported. These stark findings hold after a battery of robustness tests, including alternative sample selections and the inclusion or exclusion of hospital fixed effects.

Though there are some caveats to the findings, and we do not claim to provide the definitive answer on the subject, we believe that our paper breaks ground in providing important evidence that breastfeeding matters for children's cognitive development. We also note that while the effects on cognition are large, they are around half the size of estimates from the well-known randomized controlled trial of Kramer, Aboud, et al. (2008) in Belarus and the 10-year follow-up of a randomized controlled trial of specially supplemented formula milk (Isaacs et al. 2011). Also consistent with our results, Kramer et al. (2001) find only weak effects on health in childhood and null effects on noncognitive skills (Kramer, Fombonne, et al. 2008).

Several features unique to the UK health system contribute to the validity of our empirical strategy because they limit the ability of women to choose when they deliver. This is a context in which 98 percent of births are in public hospitals (and births are fully covered by the public insurance), which conform to guidelines of the National Institute of Clinical Excellence (NICE). C-sections are only allowed for medical reasons, and indeed the rate of C-sections was 17 percent in 2000 (very close to the World Health Organization recommendation of 15 percent).[1] Also, labor inductions are only undertaken if there are medically indicated reasons for them.[2] In addition, and also in contrast to the United States, expectant women do not have a preassigned midwife or obstetrician who is present at delivery, alleviating concerns that health care professionals schedule the delivery at convenient times (nonrandomly).

We focus on low risk, vaginal deliveries, thereby excluding C-sections and children who were placed in intensive care. This is both because breastfeeding skills and required support is different for the excluded group, and is also in order to focus on a sample for which health care is relatively uncomplicated. Apart from providing evidence that emergency C-sections do not vary by day of the week, we also provide evidence that maternal and birth-related characteristics do not vary by timing of birth. Alongside this, we provide several pieces of evidence that other hospital maternity services do not vary by timing of birth. First, we use the administrative hospital records of all births in public hospitals in England during the sample period, covering approximately half a million births, to show that the hospital readmission rate in the 30 days after birth is virtually the same for weekend and weekday born babies. Second, we use our main data source to show that a comprehensive set of hospital services relating to labor and delivery do not differ by timing of birth. Third, the fact that we find that breastfeeding affects cognition but not health reinforces the claim that hospital services do not differ by timing of birth.

There is a vast literature on the importance of the early years for later outcomes (Gertler et al. 2014; Heckman and Mosso 2014; Currie and Almond 2011; Walker et al. 2011; Cunha, Heckman, and Schennach 2010; Cunha and Heckman 2007; Heckman 2007; Heckman and Masterov 2007). Our paper makes an important contribution to at least four strands of this literature. The first relates to the importance of hospitals and maternity care for later outcomes. Two studies consider the effects of medical treatments at birth for very low birth weight newborns, finding lower one-year mortality rates (Almond et al. 2010) and higher school test scores and grades (Bharadwaj, Løken, and Neilson 2013). Other studies consider the length of hospital stay postpartum, finding no impacts on health (Almond and Doyle 2011), and the effects of improved hospital postneonatal mortality rates and access to hospitals for Blacks in the 1960s/1970s, finding improvements in their academic and cognitive skills as teenagers (Chay, Guryan, and Mazumder 2009). In contrast, we focus not on medical care but on maternal care in the form of breastfeeding. Moreover, our results are applicable to healthy newborns and not just to those with particular health risks.

A second contribution is to the literature on the optimal timing of interventions in the early years. We show that though breastfeeding is not a form of medical care, hospital policy–specifically, breastfeeding support–can influence it significantly.

[1] See http://www.nationmaster.com/country-info/stats/Health/Births-by-caesarean-section.

[2] The 2001 NICE Clinical Guidelines on Induction of Labour specify that women should be offered a labor induction in the following situations: prolonged pregnancy (41 weeks or more), pregnancy complicated by diabetes, and pre-labor rupture of the membranes. In uncomplicated pregnancies, induction of labor prior to 41 weeks gestation should be considered if (1) resources allow (2) the woman has a favorable cervix and (3) there are compelling psychological or social reasons.

Given the evidence we provide on its importance for cognitive development, this raises the question as to how and when policy to increase breastfeeding rates should be targeted. Rather than focusing solely on the provision of infant feeding support in maternity wards, a more integrated approach to providing information on breastfeeding to expectant women would, in underpinning subsequent hospital support, be likely to be more effective. In this respect, our paper supports the view that pre-natal interventions are important (Carneiro, Løken, and Salvanes 2015; Almond and Currie 2011; Currie and Almond 2011).

Third, our findings contribute to the literature that explores the pathways to improved long-term outcomes. Milligan and Stabile (2011) find that early cash transfers increase children's test scores, without improving health. This is consistent with Field, Robles, and Torero (2009) who find that iodine supplementation in pregnancy increases schooling by a year and a half despite not improving health. This evidence suggests that improving health is not a prerequisite to improving cognition in the early years. Our paper reinforces this by showing that cognitive development can increase without commensurate improvements in health.

Finally, our paper contributes to understanding the importance of nutrition for later outcomes. While links between nutrition and development have been documented, much of the literature focuses on developing countries and/or on extreme shocks, such as famines, making it difficult to extrapolate to everyday circumstances in developed countries.[3] The few studies in developed countries that consider margins more responsive to policy, point to a positive effect of nutrition on later outcomes. For instance, Dahl and Lochner (2012) and Milligan and Stabile (2011) find that increased economic resources *in utero* improve children's later cognition, most likely due to better early nutrition. Hoynes, Schanzenbach, and Almond (2016) find improvements of expanded nutritional resources *in utero* and in early childhood on adult health.

The rest of the paper is organized as follows. Section I summarizes how breastfeeding can improve child development, alongside previous literature on the topic; in Section II we discuss the data, and in Section III the identification strategy, including evidence from hospital records. Section IV reports the results of the First Stage, concluding that the probit model provides a much better fit than a linear First Stage. Section V presents the main results of the paper and compares them with findings from the only existing randomized trial in this area. Section VI provides robustness tests, including a falsification exercise, and conclusions are in Section VII. Note, throughout, we make extensive use of appendices to provide more in-depth analyses of particular issues.

[3] For studies in developing countries see Barham, Macours, and Maluccio (2013); Barham (2012); Martorell et al. (2010); Field, Robles, and Torero (2009); Maccini and Yang (2009); Maluccio et al. (2009); Behrman and Rosenzweig (2004); Glewwe and King (2001). For studies on effects of exposure to extreme conditions such as famine on later outcomes such as test scores, employment and life expectancy see Scholte, van den Berg, and Lindeboom (2015); Ampaabeng and Tan (2013); Kelly (2011); Van den Berg et al. (2010); Almond (2006); van den Berg, Lindeboom, and Portrait (2006) find lower test scores for students exposed to Ramadan in early pregnancy. Almond and Mazumder (2011) find that observance of fasting during Ramadan has long-term health effects.

I. Background

In this section, we discuss the potential channels through which breastfeeding might improve child development, as well as provide an overview of some of the related literature.

A. *Mechanisms*

The literature has emphasized two main mechanisms with the potential to explain the effect of breastfeeding on child development: the first relates to the compositional superiority of breast milk over formula milk owing to the presence of particular fatty acids, and the second relates to mother-child interaction.

The compositional superiority of breast milk over formula milk is mainly due to the presence of two long-chain polyunsaturated fatty acids: Docosahexaenoic Acid (DHA) and Arachidonic Acid (AA). Around one half of the brain is made up of lipid, much of which is DHA and AA (Gerber Medical 2013; Grantham-McGregor, Fernald, and Sethuraman 1999). They are major parts of the neuron membranes, the core components of the nervous system, and their content affects membrane fluidity and the functioning of membrane-associated proteins such as transporters, enzymes and receptors (Fernstrom 1999). During the first year of life, infants require large quantities of DHA and AA for brain development (Clandinin et al. 1981). DHA and AA are naturally present in breast milk and are easily absorbed due to the triglyceride structure of breast milk.

These potential benefits of breast milk are exacerbated in our period by the fact that although DHA and AA were permitted in formula milk in the EU since 1996, it was not until August 2001 that one of the two big producers of formula introduced DHA and AA into its milk.[4] The majority of the children in our sample were therefore not exposed to this supplemented formula. Instead, the available formula milk required infants to produce DHA and AA from other components of the milk. This synthesis requires sufficient enzyme capacity, which young infants generally do not have (Koletzko et al. 2008; Uauy and De Andraca 1995), resulting in lower absorption of DHA and AA from formula than from breast milk.

The second mechanism through which breast milk may be more beneficial for children's development than formula milk is due to increased mother-child interaction. First, breastfeeding increases skin-to-skin contact, which might promote secure attachment (Britton et al. 2006). Second, breastfeeding triggers beneficial hormonal responses in mothers, potentially reducing stress and depression, which might improve quality of care (Reynolds 2001; Uauy and Peirano 1999). Third, breastfeeding involves direct physical contact and regular interaction with the mother every day, which may stimulate cognitive development.

[4] Authors' analysis of market reports and advertisements in midwifery journals shows that one of the two largest producers of infant formula milk in the United Kingdom started DHA and AA supplementation in August 2001, and the second largest producer started in 2002. Only 11 percent of children in our sample were born in August 2001 or later.

Two other possible, less-studied mechanisms relate to maternal labor supply, and the use of antidepressants. On the former, mothers might stop breastfeeding to return to work. Extending breastfeeding and postponing the return to work might affect children's cognitive development through more mother-child interaction, time in formal/informal childcare, as well as income effects (especially if the postponement has longer term career effects). On the latter, the use of antidepressants while breastfeeding might entail risks to the baby (Pinheiro et al. 2015). If breastfeeding mothers with postpartum depression choose not to take medication, it could affect the duration and severity of postpartum depression and maternal-child interactions.

B. *Related Literature on Breastfeeding*

Studies in economics considering the relationship between breastfeeding and children's outcomes are nonexperimental, using various methods to control for selection bias—propensity score matching (Rothstein 2013; Belfield and Kelly 2012; Borra, Iacovou, and Sevilla 2012; Quigley et al. 2012), maternal fixed effects (Der, Batty, and Deary 2006; Evenhouse and Reilly 2005), and IVs (Del Bono and Rabe 2012; Baker and Milligan 2008). Closest in nature to our empirical approach are the latter two using IVs. In the UK context, Del Bono and Rabe (2012) exploit the rollout of the Baby Friendly Hospital Initiative, BFI (WHO, UNICEF), a program implementing best practice in breastfeeding support at the hospital level through following 'Ten Steps to Successful Breastfeeding', using as an instrument the distance from the mother's home to the closest hospital that voluntarily implemented the BFI program. In Canada, Baker and Milligan (2008) exploit large increases in maternity leave entitlements as an IV, showing that it increased breastfeeding in the first year of life by more than one month.

The general consensus from this literature is that there is a small, positive association between breastfeeding and cognitive development, with often insignificant associations between breastfeeding and noncognitive development, and between breastfeeding and health. However, evidence cited in the American Academy of Pediatrics (2012), drawing mainly on Ip et al. (2007), and the medical literature, highlights the benefits of breastfeeding for a range of infant and maternal health outcomes, though it should be noted that the majority of the evidence is observational and does not have a convincing empirical strategy to deal with unobserved confounders. Notwithstanding the general consensus from the medical literature, some epidemiological studies that exploit data and contexts in which breastfeeding is not positively related to socioeconomic status tend to find that the effects of breastfeeding on cognitive development are more robust than on health (Brion et al. 2011; Daniels and Adair 2005).

There is just one study that uses experimental variation to identify the effects of breastfeeding on children's outcomes, that of Kramer et al. (2001). The intervention, the Promotion of Breastfeeding Intervention Trial (PROBIT), is based on the WHO Baby Friendly Hospital Initiative, which provided health care worker assistance for initiating and maintaining breastfeeding randomly across 31 hospitals in Belarus in the late 1990s. The effects on health—both in the first 12 months of life and the medium-term—are weak or nonexistent (Kramer et al. 2009, 2007, 2001).

On the other hand, there are very large effects, of one standard deviation or higher, on cognition at age 6.5 years (Kramer, Aboud, et al. 2008).[5]

II. Data

The main data source is the Millennium Cohort Study (MCS), a rich longitudinal birth cohort study covering the United Kingdom,which follows approximately 19,000 babies born at the beginning of the noughties.[6] We use data from each of the surveys conducted up to 7 years of age (9 months (2000/2001), 3 years (2004/2005), 5 years (2006), 7 years (2008)).

To provide supporting evidence on our identification strategy, we utilize two additional datasets: the Hospital Episode Statistics of 2000–2001 (NHS Digital) and the Maternity Users Survey of 2007 (Healthcare Commission, Picker Institute Europe 2009). The Hospital Episode Statistics contains all births in English public hospitals, and we use the sample of births corresponding to the same period of MCS births in England (September 2000–August 2001; around half a million births in total). This administrative dataset allows us to compute readmission rates, a widely used statistic measuring hospital quality, by day of the week of birth, providing evidence on the validity of our empirical strategy. We use the Maternity Users Survey of 2007, a postal survey of around 26,000 mothers three months after giving birth, to analyze how feeding support at the hospital varies by day of the week of birth, which is also key to our identification strategy.

In our sample selection, we drop multiple births, those who were not born in a hospital and those born in Northern Ireland. For reasons explained in Section III, we focus on a sample of vaginal deliveries, dropping those born through C-sections and those who were placed in intensive care after delivery. However, in Section VI we show that our results are robust to including them. Unless otherwise indicated, we drop high-educated mothers (for reasons explained in Section IIIA)[7], which leaves us with an analysis sample of 5,809 children.

In our main data source, the MCS, children took age-appropriate tests administered by trained interviewers—the Bracken School Readiness (age 3) and British Ability Scales (ages 3, 5, 7). These measures offer a distinct advantage over parental-reported measures (Fernald et al. 2009). Children's behavioral (noncognitive) development was measured via maternal-report using the Strengths and Difficulties Questionnaire (SDQ), a validated behavioral screening tool (ages 3, 5, 7). Children's health includes maternal-reported measures of morbidity and

[5] They only report intention-to-treat estimates. The effect of one standard deviation on cognition is based on the authors' own computations of the Wald estimator based on the data reported for three months of exclusive breastfeeding.

[6] Born between September 1, 2000 and August 31, 2001 in England and Wales and between November 22, 2000 and January 11, 2002 in Scotland and Northern Ireland.

[7] We define low educated = 1 if highest qualification is at or below National Vocational Qualification (NVQ) Level 2. This includes academic qualifications at or below the secondary school examination taken at age 16, or occupational/vocational qualifications at or below NVQ level 2 (there are 5 levels of NVQ ranging from Level 1, on basic work activities, to Level 5 for senior management). We also include as low educated those with unknown NVQ level but who left school before age 17; high educated = 1 otherwise.

chronic conditions (ages 9 months, 3, 5, 7 years). Details on the measures are in online Appendix A.

Within the above developmental domains—cognitive skills, noncognitive skills and health—we aggregate multiple measures within and across ages into a summary index, following Anderson (2008). In this way, our results provide a statistical test for whether breastfeeding has a "general effect" on development that is robust to concerns about multiple inference (Hoynes, Schanzenbach, and Almond 2016; Kling, Liebman, and Katz 2007; Liebman, Katz, and Kling 2004). To create summary indices for cognition, we combine cognitive scores at age 3 (expressive language and school readiness), age 5 (expressive language, pictorial reasoning, visuospatial) and age 7 (numerical, verbal and visuospatial) into a single cognitive index. The index is a weighted mean of the standardized scores of each test, with the weights calculated to maximize the amount of information captured in the index by giving less weight to outcomes that are highly correlated with each other. For noncognitive outcomes, we combine the standardized scores of the strength and difficulties questionnaire at ages 3, 5, and 7. For health, we combine seven health indicators measured at each wave (including asthma, hay fever, eczema, wheezing, ear infections (age 3 only), obesity, and long-standing health conditions).

Breastfeeding duration is measured using information on how old the child was when (s)he last had breast milk, so it relates to any breastfeeding, regardless of exclusivity.[8] Figure 1 shows spikes in the number of babies breastfed at discrete points in time–for (at least) 30 days, 60 days, and 90 days, displaying a relatively large spike at 90 days. Our measure of breastfeeding therefore takes the value one if the infant was breastfed for at least 90 days, and zero otherwise. Note the recommendation in the United Kingdom at the time was to breastfeed exclusively for at least 16 weeks, or 112 days. However, if we took the cutoff to be 112 days, we would allocate zero to those who were breastfed for 90 days, which is the more relevant empirical threshold (in any case, in Section VI we confirm our results using 60 and 120 days). We opted for a binary indicator of breastfeeding, rather than a continuous measure, for three reasons: (1) comparability with previous literature, (2) the distribution of the number of days that a child was breastfed has a large mass point at zero, and is very concentrated on focal numbers of days (e.g., 30, 60, 90), and (3) as our instrument is based on support received at hospital, it is unlikely to explain differences in breastfeeding duration in the upper part of the distribution (e.g., between 90 and 120 days), by which time correct attachment and positioning would likely have been acquired.

III. Identification Strategy

In this section we discuss the rationale underlying our choice of instrument, and provide evidence on the validity of the exclusion restriction. We also provide graphical evidence on the relationship between timing of delivery and breastfeeding, as

[8] Exclusive breastfeeding, defined as breast milk only (no water, no formula milk, no solids) cannot be accurately defined from the data, due to lack of information on water intake. However, if we relax the definition and consider it to be breast milk (no formula milk, no solids), we estimate that of those being breastfed at 90 days, around two thirds are being exclusively breastfed.

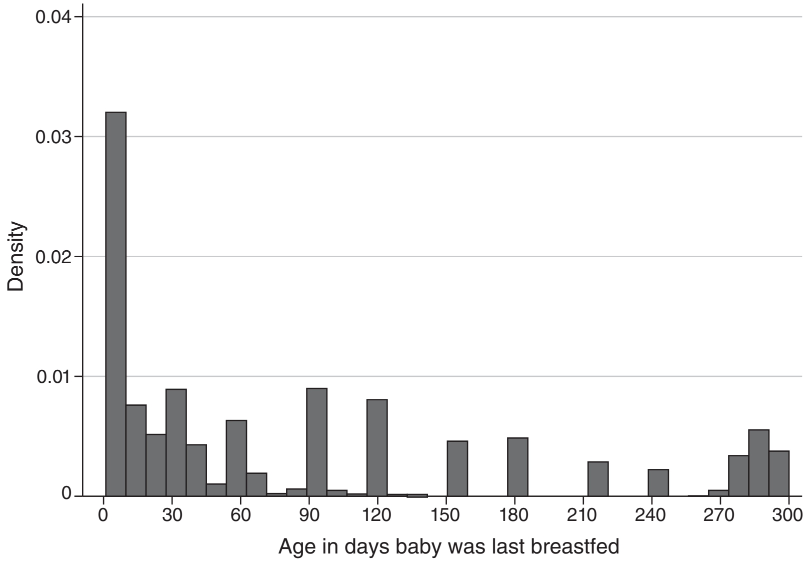

Figure 1. Breastfeeding Duration in Days

Notes: Mothers who never initiate breastfeeding were excluded: 45.7 percent. Sample comprises low-educated mothers, but excludes children born through caesarean sections (either emergency or planned) and children placed in intensive care.

Source: Millennium Cohort Study

well as between timing of delivery and child development, to precede the formal analysis.

A. *Instrument Rationale*

Breastfeeding is a skill that requires practice and learning early on. If not learned successfully in the very early postpartum period, serious damage to the nipples can quickly occur, resulting in pain/infection for the mother, and/or failure to thrive for the baby, and ultimately the premature cessation of breastfeeding. Many studies highlight the importance of hospital support and policies and procedures in the early postpartum period as key determinants of breastfeeding success–such as skin-to-skin contact straight after birth (Renfrew et al. 2009; Bolling et al. 2007), increased "Baby-Friendly" hospital practices, and other maternity-care practices (Del Bono and Rabe 2012; DiGirolamo, Grummer-Strawn, and Fein 2008; Merten, Dratva, and Ackermann-Liebrich 2005). Similarly, UNICEF asserts that " … putting resources into supporting women to breastfeed successfully would be hugely cost effective to the NHS, as well as preventing the distress and pain felt by a mother who has a bad experience of breastfeeding." (Renfrew et al. 2012).

In the United Kingdom, at the time our sample of children was born, infant feeding support was provided by midwives, nurses, and clinical support workers as part of their daily duties. As staff weekend working hours are more expensive, staff duties are limited to the core services of labor, delivery, and maternal and child health–at

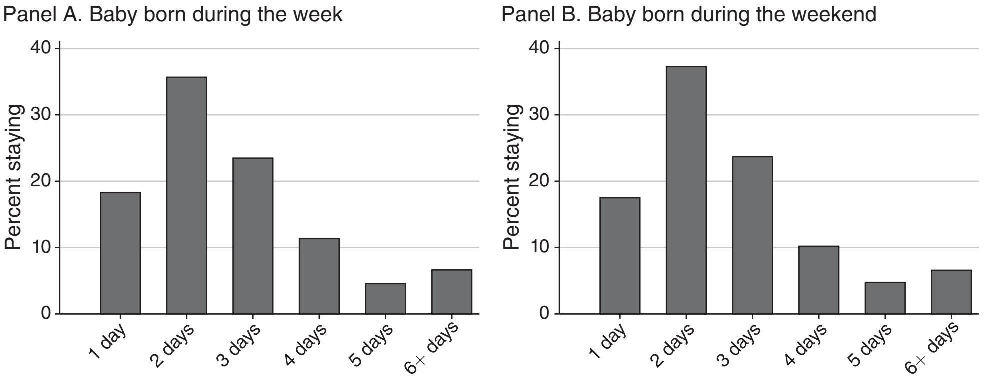

FIGURE 2. LENGTH OF STAY IN HOSPITAL

Notes: Sample comprises low-educated mothers, but excludes children born through caesarean sections (either emergency or planned) and children placed in intensive care.

Source: Millennium Cohort Study

the expense of infant feeding support. The average length of stay is virtually the same for weekday and weekend births (46.80 hours and 44.12 hours, respectively, see also Figure 2, panels A and B), so mothers most exposed to reduced feeding support are those who give birth on Fridays, followed by Saturdays, and, to a lesser extent, Sundays. More generally, exposure to weekend feeding support increases as the week progresses (online Appendix Figure F1).

We use the UK Maternity Users Survey (MUS 2007) to provide evidence to support our claim that breastfeeding support is lower at weekends. The survey asks mothers, among other things, "Thinking about feeding your baby, breast or bottle, did you feel that midwives and other carers gave you consistent advice/practical help/active support and encouragement?" Stark differences emerge when we split the sample by education status.[9] Columns 1-3 of Table 1 show that low-educated mothers of children born on Friday or Saturday are less satisfied with the infant feeding advice obtained in hospital compared to mothers of Monday-borns. This pattern is broadly mirrored in breastfeeding rates, as measured in the MCS, where column 6 reports lower breastfeeding rates for children born on Friday, Saturday, and Sunday (and similarly, but weaker, on mixed feeding in the first few days as reported by the MUS 2007 in column 4).[10] These significant differences are essential to our identification strategy.[11]

[9] In the MUS, as we do not observe highest qualification level, we define low educated = 1 if left full-time education at or before age 16; high educated = 1 if left full-time education after age 16. This might overestimate (underestimate) the true proportion of high (low) educated, as those who left full-time education before age 16 may, through subsequent occupational/vocational training, have an NVQ Level above 2 as their highest qualification level.

[10] Concerning breastfeeding, the MUS only asks if the child was ever put to the breast and how was the child fed in the first few days after birth.

[11] The difference on Sunday between columns 1–3 and column 6 may be due to the different time periods (columns 1–3 relate to 2007 (MUS); column 6 relates to 2000/01 (MCS)).

TABLE 1—BREASTFEEDING SUPPORT AND BREASTFEEDING RATES BY DAY OF BIRTH

	Low educated					
Data source →	MUS 2007					MCS 2000-2001
Day of birth ↓	Received consistent advice	Received practical help	Received active support	Mixed feeding in the first few days	Some breastfeeding in the first few days	Breastfed for at least 90 days
	(1)	(2)	(3)	(4)	(5)	(6)
Sun.	0.004	−0.014	−0.016	0.005	−0.006	−0.055
	(0.022)	(0.023)	(0.023)	(0.018)	(0.026)	(0.022)
Tue.	−0.022	−0.021	−0.024	0.020	0.011	−0.025
	(0.022)	(0.024)	(0.023)	(0.018)	(0.026)	(0.022)
Wed.	−0.007	−0.006	−0.018	−0.000	0.006	−0.015
	(0.022)	(0.023)	(0.023)	(0.018)	(0.026)	(0.022)
Thurs.	−0.007	−0.011	−0.021	0.003	−0.031	−0.028
	(0.022)	(0.024)	(0.024)	(0.018)	(0.026)	(0.022)
Fri.	−0.095	−0.083	−0.084	0.035	0.023	−0.058
	(0.022)	(0.023)	(0.023)	(0.018)	(0.026)	(0.021)
Sat.	−0.028	−0.066	−0.052	0.027	−0.003	−0.06
	(0.022)	(0.024)	(0.023)	(0.018)	(0.026)	(0.021)
Monday mean	0.814	0.784	0.796	0.125	0.568	0.265
p-value joint	0.000	0.001	0.006	0.263	0.548	0.018
p-value Fri.–Sun.	0.000	0.000	0.001	0.150	0.659	0.015
Observations	4,914	4,772	4,813	5,275	5,275	5,809

Notes: The table reports coefficients from an OLS regression over day of week dummies (Monday omitted). The dependent variable is listed at the top of the column. Columns 1–5 are from the Maternity Users Survey (MUS). Column 6 is from the Millennium Cohort Study (MCS). All columns exclude emergency and planned C-sections, and column 6 additionally excludes babies placed in intensive care. Standard errors in parentheses.

Sources: Maternity Users Survey and Millennium Cohort Study

Neither pattern—differences in breastfeeding support or rates by day of the week of birth—exists for high-educated women (see online Appendix Table F1). Possible reasons for this include: facing time constraints, midwives target the high educated; the high educated are more likely to seek out help from midwives; the high educated can benefit more from the same level of support as they have more information beforehand; and the high educated rely less on hospital support, due to easier access to private lactation consultants after discharge, peer community groups, and information pamphlets, for instance.[12]

B. *Study Sample*

Given the above evidence, from here on we focus on the sample of low-educated mothers for whom hospital feeding support matters significantly for breastfeeding.

[12] We can rule out that differences in reporting by education are due to selection effects (in particular that the more educated go to better hospitals). We can control for hospital fixed effects in the main analysis using the MCS data, and when we do, we find the same pattern between breastfeeding rates and timing of birth as when we omit them. This is not surprising: as women register at their nearest hospital at around 12 weeks gestation, hospital choice is not related to day of labor onset. The Choice and Book system introducing hospital choice to NHS patients began in 2005; its precursor, the London Patient Choice Project, only started in October 2002 (Dawson et al. 2004).

We also exclude planned caesarean sections. This is mainly because they do not take place on weekends.

For the main analysis, we also exclude emergency caesareans and babies who had been in intensive care units (ICU), thereby focusing on a sample of low-risk vaginal deliveries. This is for two reasons. First, because breastfeeding skills and support are different for both of these; and second, to focus on a sample for which health care is relatively uncomplicated.[13] Reassuringly, however, the distribution of emergency caesareans and ICUs does not vary by day of the week (online Appendix Table B1). Moreover, in Section VI, we show that our results prevail when we include them.

C. *Validity of Exclusion Restriction*

Before discussing the validity of the exclusion restriction, we define the exclusion restriction that we use in the analysis. First, we define $Hour_i$ as the number of hours between Sunday 00:01am and the hour of child i's birth (0 refers to the first hour of Sunday and 167 to the last hour of Saturday):

$$Hour_i = 24 \times DayBirth_i + TimeBirth_i, \tag{1}$$

where $DayBirth_i$ is day of the week of birth of child i (Sunday is 0 and Saturday is 6), and $TimeBirth_i$ is the hour of birth of child i (in 24 hour format). Second, we define $Exposure_i$ as the share of hours falling in a weekend, in the interval between the infant's birth and 45 hours later (the average length of stay in hospital).[14]

For the exclusion restriction to hold, it is necessary that any unobserved variables that affect the outcomes of interest are uncorrelated (conditional on covariates) with the excluded variable (*Exposure* or *Hour*). Potential unobservable variables might be (1) mothers' and children's characteristics, and (2) hospital maternity care practices. Although the assumption cannot be tested, it is informative to consider the correlation between observed variables and the exclusion restriction. If such correlations were important, it would be difficult to maintain the assumption of absence of correlation between the exclusion restriction and the unobserved variables that may affect the outcomes of interest. Hence, in what follows we examine the correlation between (1) mothers' and children's characteristics and (2) hospital maternity care practices and Exposure (and with Hour in online Appendix B).

Maternal and Child Characteristics by Timing of Birth.— A potential concern is that mothers who are more exposed to the weekend are somehow different from those who are not. To shed light on this, Table 2 shows the balance of several mothers' and infants' characteristics, as a function of timing of birth. We report the cor-

[13] Note also that infants placed in intensive care are more likely to be different from the rest of the sample in terms of their development, and may receive additional medical care that may affect their development. For instance, Bharadwaj, Løken, and Neilson (2013) show that infants who receive extra medical care at birth (surfactant therapy) have lower mortality rates and higher school attainment. In the United Kingdom, surfactant therapy is administered in the Intensive Care Unit, where babies with neonatal respiratory distress syndrome are transferred.

[14] Using potential rather than actual exposure circumvents problems of endogenous length of hospital stays (though note that women have little to no choice in this).

Table 2—Balance by Exposure to Weekend (Continuous)—Low-Educated Mothers

Variable	Correlation with exposure	p-value
Antenatal		
Received antenatal care	−0.008	0.525
First antenatal was before:		
0–11 weeks	−0.002	0.908
12–13 weeks	−0.001	0.969
≥14 weeks	−0.003	0.808
Don't know	0.003	0.842
Attended antenatal classes	0.006	0.666
Received fertility treatment	−0.002	0.878
Planned parenthood	0.003	0.818
Delivery		
Labor induced	0.054	0.000
No pain relief	−0.023	0.077
No complications during birth	−0.001	0.937
Baby		
Female	0.013	0.313
Birth weight (kg)	−0.005	0.709
Premature	0.016	0.237
Length of gestation (days)	−0.014	0.296
Present at birth		
Father	0.003	0.794
Mother's friend	−0.016	0.224
Grandmother (in law)	0.016	0.218
Someone else	0.011	0.398

Variable	Correlation with exposure to weekend	p-value
Mother's characteristics		
Age	−0.008	0.546
Expected educ. at age 16	0.010	0.447
Married	−0.015	0.266
Religion		
No religion	0.008	0.534
Catholic	0.013	0.306
Protestant	−0.008	0.526
Anglican	−0.005	0.727
Another type of Christian	0.004	0.771
Hindu	0.003	0.828
Muslim	−0.016	0.219
Other	0.008	0.536
Ethnicity		
White	0.004	0.767
Mixed	0.026	0.050
Indian	−0.008	0.541
Pakistani/Bangladeshi	−0.013	0.317
Black	−0.003	0.848
Other	0.010	0.456
Smoked during pregnancy (number of avg. cig. per day)	−0.004	0.782
Drank during pregnancy	−0.006	0.645
Longstanding illness	−0.003	0.821
Limiting longstanding illness	0.016	0.231
Worked during pregnancy	−0.006	0.645

Notes: Figures report the correlation between the variable to the left and the Exposure variable, as well as the p-value that the correlation is equal to zero. Sample comprises low-educated mothers (NVQ level 2 or less, or NVQ level unknown but left school before 17), and excludes children born through caesarean sections (either emergency or planned) and children placed in intensive care after delivery. All variables are dummy variables, with the exception of birth weight, mother's age, and smoked during pregnancy.

Source: Millennium Cohort Study

relation of mothers' and infants' characteristics with *Exposure* and the p-value of such correlations.

Table 2 shows that most mothers' and infants' characteristics are not significantly correlated with *Exposure* (online Appendix Table B2 shows an extended set of variables). To highlight some important variables, the p-value of the correlation between *Exposure* and infant's birth weight is 0.71, with mother's education is 0.45, and whether the mother worked during pregnancy is 0.65. Given the large number of variables that we test (and that they are not independent from each other), it is unsurprising that some reject the null of no difference (mother's hay fever, epilepsy, and digestive disorders), although even in these cases the value of the correlation is small (below 0.04).

We would not expect to see such a clean balance on maternal and birth characteristics by timing of birth in the United States, where there is more flexibility regarding elective C-sections and inductions (American College of Obstetricians and Gynecologists 2009; 2003) and where 50 percent of deliveries are covered by private insurance, rendering competition much more important, and with certain preference to schedule deliveries through inductions or C-sections.

Hospital Maternity Care and Timing of Birth.—It is crucial to assess whether other hospital services relevant for child development, apart from breastfeeding support, vary by timing of birth. For instance, a more complicated delivery could affect a child's development either through its effects on the child's health or on the health of the mother. Our hypothesis is that hospital managers protect all services relating to birth delivery, because of the major repercussions if mistakes do occur. Moreover, our sample is one of uncomplicated cases as we exclude births through C-sections, which we note do not vary by day of week (see online Appendix Table B1).

In this section we use a variety of data sources to provide evidence that other hospital maternity care does not differ by timing of delivery.[15] First, we use administrative health records, covering all hospital births in England, corresponding to the period September 2000–August 2001, to show that readmission to the hospital within 30 days does not vary by timing of birth. Second, we show in the MCS and MUS that a wide range of characteristics relating to labor, delivery, and postnatal care are extremely similar regardless of timing of birth. Third, we discuss the limited potential for other unobserved hospital-related factors.

Evidence from Hospital Administrative Records.— Hospital readmission within 30 days is a common measure of hospital quality (Axon and Williams 2011) because it is sensitive to both poor treatment as well as to poor detection of potential complications. We use the Hospital Episode Statistics (HES), a major administrative dataset containing details of all admissions at NHS (public) hospitals in England. We use data on all births in our sample period, totaling just over half a million, to compute the rate of readmission to hospital or in-hospital death within 30 days of birth, as an overall indicator of the quality of maternity care. Note we observe readmission to or death in *any* NHS hospital, not only in the hospital in which the birth took place.

We find that the rates of readmission to hospital, or death, by day of the week of birth are extremely similar (online Appendix Table B3). The same conclusion holds using babies' hospital outpatient visits for the 2003–2004 period (first period for which this data is available) as well as rates of readmission or death of mothers, which due to data availability, we analyze by day of admission rather than day of

[15] Recent findings on weekend excess mortality (not restricted to maternity) have been attributed to differential selection into admission in weekend versus weekdays (Meacock et al. 2017; Walker et al. 2017; Freemantle et al. 2015). Using data from a different period to us (2012–2013 versus 2000–2001), Palmer, Bottle, and Aylin (2015) compare weekend obstetric outcomes with Tuesday outcomes, finding that four out of seven adverse obstetric outcomes are more frequent on weekend than on Tuesday admissions. However, unlike us, they include C-sections, resulting in marked differences in delivery methods, birth weight, and maternal socioeconomic characteristics between weekend and Tuesday admissions. Moreover, using Wednesday as a comparison instead of Tuesday, we see that two of the four significant outcomes are much more similar to weekend outcomes. Finally, the study does not adjust for multiple hypothesis testing which may affect conclusions given the large number of outcomes tested.

birth delivery.[16] Hence, even using extremely large samples, there is no evidence that adverse events (readmission or deaths) are worse at weekends than weekdays in our study period.

Evidence from the MCS and MUS.—The MCS asks several questions of mothers about their experience at the hospital during the birth of their baby. Importantly, this covers a comprehensive set of characteristics including whether the labor was induced, duration of labor, whether forceps were used, whether an epidural was administered (which requires an anaesthetist, and is a proxy for availability of core services), and whether complications occurred. Table 2 reports the correlation between these characteristics and *Exposure*. We examine whether labor was induced or not, duration of labor, type of pain relief used during delivery, and complications during labor. Of those, the only one for which we reject a null correlation with *Exposure* at 5 percent is whether the delivery was induced or not. This may be worrying because the timing of inductions is not necessarily exogenous. In online Appendix B (Table B4), we assess this issue in more detail by estimating three regressions in which the dependent variable is labor induction and the covariates are *Exposure,* as well as those in the rest of Table 2 and online Appendix Table B2. Online Appendix Table B4 shows that the correlation between labor induction and *Exposure* is practically the same whether we include the controls or not, hence, the correlation between induction and *Exposure* does not reflect sociodemographic differences between those whose labor is induced or not. Moreover, the interactions of the controls (including a socio economic index) with *Exposure* are not significantly associated with labor induction either. These results point in the direction that the variation in *Exposure* among cases of induced labor is exogenous.

Online Appendix B also confirms the balance results of birth-related characteristics by splitting the sample between those with null and positive *Exposure*, as well as the p-value of a third order polynomial in *Hour*. It also considers other samples including high-educated mothers (online Appendix Tables B9–B14) and one that includes emergency C-sections and children in intensive care (online Appendix Tables B17–B22).

Using data from the Maternity Users Survey (see Section II), we can also examine postnatal care variables including whether the baby received a newborn health check and how staff treated the mother, as well as what the mother thought of the information she received. We do this by day of the week, as a continuous measure of timing of delivery is not available in this data source. We find that the values of all of these variables are markedly similar between weekdays and weekends (online Appendix Table B15). A good balance is also found for high-educated mothers (online Appendix Table B16).

We also examine differences in the six weeks after the birth of the baby, particularly in the help and advice received from health professionals. Because they are six weeks after the birth, differences between weekend and weekday births might be due to

[16]The 2003–2004 outpatient dataset was released "under experimental status" and might suffer from quality issues. (https://digital.nhs.uk/data-and-information/publications/statistical/hospital-outpatient-activity/outpatient-data-quality-report.)

differences in breastfeeding that we have already observed. Indeed, of the 10 variables tested, the only two significant at 10 percent are related to feeding: weekday births are 1.9 percentage points more likely to have received advice on feeding the baby, and 3.5 percentage points more likely to have last been visited by a midwife at home when the baby was 11 days or older (online Appendix Table B15). It is to be expected that mothers of breastfed babies (who are more likely to be born on weekdays) require more advice regarding feeding (and, hence, they receive more visits by midwives) because breastfed babies take longer to gain weight (Nelson et al. 1989), and mothers might face more discomfort and complications because of nursing.

Other Evidence.— While the above provides compelling evidence that hospital maternity services do not differ by timing of birth, the extent to which unobserved characteristics vary by timing of birth must be addressed. As our identification strategy relies on the fact that weekend delivery negatively affects breastfeeding only, the threat to identification is that hospital weekend services "harm" children's health. We believe this is not a concern, for several reasons.

First, we consider a sample of vaginal deliveries, and babies not placed in intensive care, for whom medical care is routine and relatively uncomplicated. Some work has shown large effects of specialized medical care on children at serious health risk (Bharadwaj, Løken, and Neilson 2013; Almond et al. 2010). This is not a concern as we exclude children who have been in intensive care units (moreover in online Appendix Table B1, we showed this is also balanced by day of week).

Second, we anticipate one of our key findings, which is that breastfeeding does not affect children's later health. This suggests strongly that there are no unobserved core hospital services that are simply better during the week than at the weekend and reinforces the belief that other unobserved hospital services are not confounding estimated impacts.

Third, it is highly unlikely that services targeting directly child's cognitive development are provided in maternity wards: according to the National Collaborating Centre for Primary Care (2006) ("Routine Post-natal Care of Women and Their Babies"),[17] postnatal services focus on three key areas maternal health, infant health, and infant feeding. There is no indication in the extensive guidelines that hospitals implement programs (apart from infant feeding support) that could affect children's development apart from those that could operate through maternal and/or child health. Indeed, the median stay in hospital is 48 hours, leaving little time for anything but essential care; moreover the mother is recovering and focused on her and her newborn baby's basic needs; and hospitals are capacity constrained (and the majority of mothers and newborns stay in communal not individual postnatal wards).

D. *Breastfeeding and Child Development by Timing of Birth*

In this section we provide semi-parametric evidence on how breastfeeding rates and child development relate to timing of birth, for our main sample—low-educated

[17] The year 2006 is the first year that the guidelines were issued. We have no reason to believe that they represented a change from prior practice, but rather a formalization of existing practice.

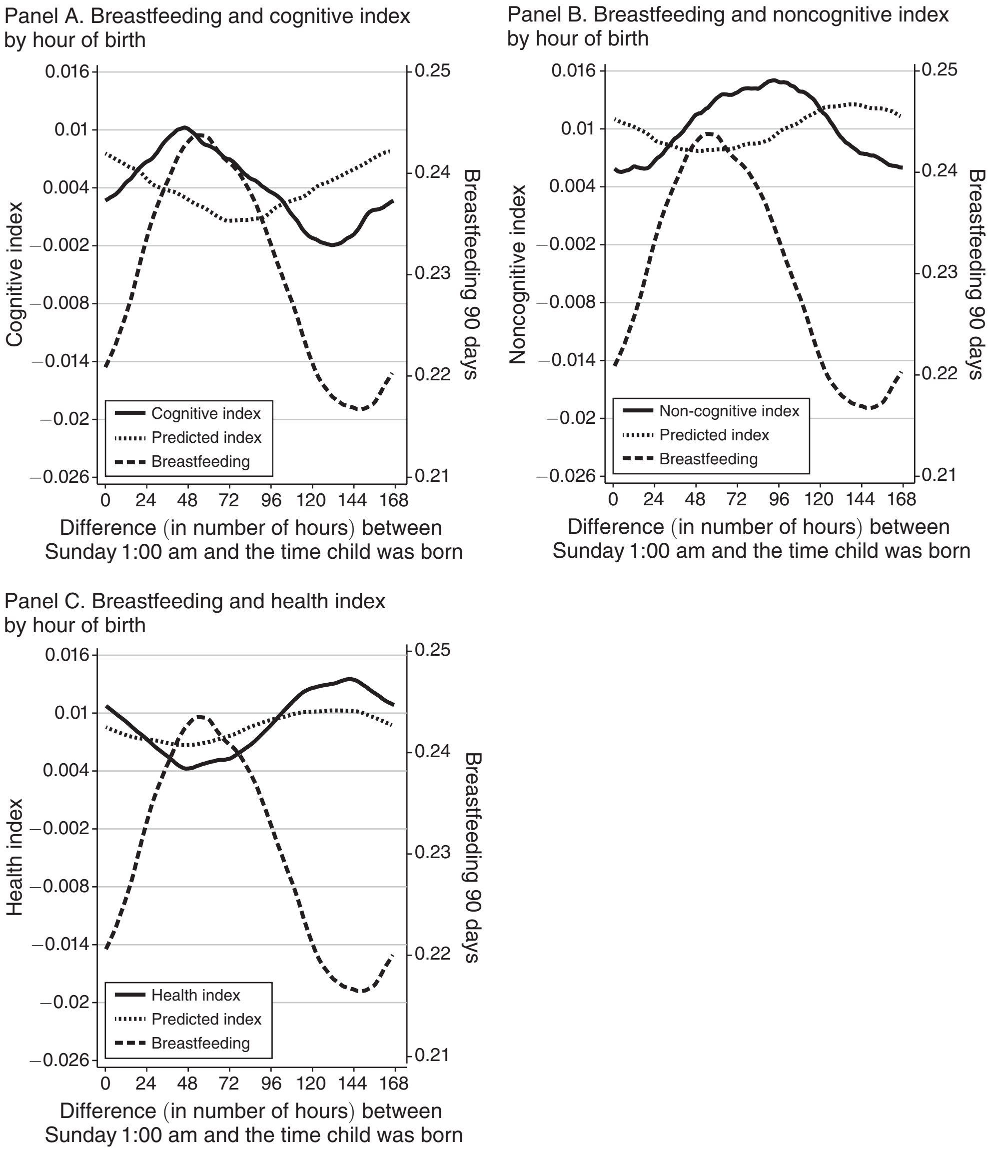

FIGURE 3. RELATIONSHIP BETWEEN BREASTFEEDING/DEVELOPMENTAL INDICES AND TIMING OF BIRTH

Notes: The horizontal axis shows the hour of birth within the week (0 corresponds to Sunday 00:01–00:59 and 163 to 23:00–23:59 on Saturday). The dashed lines is the estimate of the function $F(hour)$ on the partially linear regression defined as $Y = F(hour) + \mathbf{X}\beta + \varepsilon$, where hour is the variable in the horizontal axis, $\mathbf{X}$ is a set of control variables (those in Table 2 and online Appendix Table B2), and Y is defined as equal to 1 if the child was breastfed for at least 90 days, and 0 otherwise. $F(hour)$ is estimated following Robinson (1988) using Kernel regression (triangular Kernel with bandwidth of 72). Sample comprises low-educated mothers (NVQ level 2 or less, or unknown NVQ level but left school before age 17), but excludes children born through caesarean sections (either emergency or planned) and children placed in intensive care.

Source: Millennium Cohort Study

mothers who had normal deliveries and whose babies were not in intensive care—as a precursor to the more formal analysis in the following sections.

Figure 3, panels A, B, and C plot the relationship between breastfeeding rates and *Hour* on the right vertical axis, and the relationship between the index (cognitive, noncognitive, and health, respectively) on the left vertical axis, shown in solid lines. The figure first shows that breastfeeding rates are quite low early on into Sunday but increase quite steeply at the beginning of the week, and then taper off right through to Saturday. Although breastfeeding support is likely to be as good on Mondays as it is on Wednesdays, the later on in the week the child is born, the more likely it is that (s)he stays during the weekend (shown in online Appendix Figure F1) when the breastfeeding support will be worse.

Second, the relationship between the cognitive index and *Hour* in Figure 3, panel A tracks strikingly the relationship between breastfeeding and *Hour*. They both peak around Monday night, and they both have their minimums between Friday noon and midnight. This similarity in the patterns preempts a strong effect of breastfeeding on child cognitive development when we estimate a formal Instrumental Variables model specified in Section IV.

In Figure 3, panel B, the pattern of the relationship between the noncognitive index and *Hour* tracks less closely the breastfeeding pattern; while the overall shape is fairly similar, its peak is around one day later. This anticipates the fact that we will not find conclusive results on how breastfeeding affects noncognitive development. In Figure 3, panel C, the health index is flatter than the cognitive development index, and, if anything, the peaks and troughs are inversely related to breastfeeding. In fact, the health index appears to be slightly higher over weekends and lower on weekdays, alleviating concerns that the strong effects on cognitive outcomes are due to hospital weekend services harming children's health.

Figure 3, panels A, B, and C also plot, in the dotted lines, the prediction of the cognitive, noncognitive, and health indices as a function of an extensive set of variables (those in Table 2 and online Appendix Table B2). In all three figures the predicted indices exhibit a flatter pattern than the actual ones, and do not track the pattern in breastfeeding, confirming the comprehensive sample balance shown in Section IIIC.

IV. Estimation

In this section we describe the empirical model we estimate, show results from the First Stage estimation, and report on a Monte Carlo simulation exercise to understand the direction of potential biases.

A. *Model*

We estimate the following linear model:

$$Y_{ij} = \alpha_0 + \alpha_1 B_i + \boldsymbol{\alpha}_2 \mathbf{X_i} + \mathbf{h_j} + \varepsilon_i, \tag{2}$$

where Y_{ij} is the outcome variable of child i (cognitive development/noncognitive development/health) born in hospital j, B_i is a binary variable taking the value 1 if child i has been breastfed for at least the first 90 days of life and 0 otherwise, $\mathbf{X_i}$ is a vector of covariates (including all those shown in Table 2 and online Appendix Table B2, and in addition month of birth, month of interview, and regional dummies), $\mathbf{h_j}$ denotes hospital fixed effects, and ε_i is an error term which includes unobserved characteristics relevant to the child's development. The parameter α_1 measures the effect of being breastfed for at least 90 days on child i's outcomes.

As discussed, our identification strategy exploits timing of birth within the week. As exclusion restrictions, we use mainly $Exposure_i$, the share of hours falling in a weekend, in the interval between the infant's birth and 45 hours later (see Section IIIC). We also show some results using as an exclusion restriction a third order polynomial in $Hour_i$ (equation (1)) that captures well the relationship between breastfeeding and hour of birth (see online Appendix Figures F2 and F3). Both exclusion restrictions exploit the fact that some mothers are exposed to the weekend more than others.

For estimation, we follow Wooldridge (2002, 623) and Angrist and Pischke (2009, 191) and use a nonlinear two-stage estimator (NTSLS) where we first estimate a Probit model of breastfeeding, B_i, over $\mathbf{X_i}$ and $Exposure_i$ (equivalently for the cubic polynomial in $Hour_i$). The underlying latent variable $\dot{B}_i$ measures the propensity for child i to be breastfed:

$$\dot{B}_i = \beta_0 + \beta_1 Exposure_i + \boldsymbol{\beta}_2 \mathbf{X_i} + \vartheta_i, \tag{3}$$

where $B_i = 1 \, if \, \dot{B}_i \geq 0; B_i = 0 \, if \, \dot{B}_i < 0$, ϑ_i is standardized normal, and β_0, β_1, $\boldsymbol{\beta}_2$ are parameters to be estimated.[18] Next, we compute the fitted probabilities, $\hat{B}_i$, associated with the Probit model as

$$\hat{B}_i = \Phi\left[\hat{\beta}_0 + \hat{\beta}_1 Exposure_i + \hat{\boldsymbol{\beta}}_2 \mathbf{X_i}\right], \tag{4}$$

where $\hat{\beta}_0$, $\hat{\beta}_1$ and $\hat{\boldsymbol{\beta}}_2$ are estimates from the model specified in (3) and $\Phi[.]$ is the cumulative distribution function of the standardized normal. Finally, we use IVs to estimate the causal effect of breastfeeding on outcome Y_{ij} using $\mathbf{X_i}$ and $\hat{B}_i$ as instruments.[19]

The advantage of this NTSLS method over the standard Two Stages Least Squares (TSLS), which uses a linear First Stage, relates to the efficiency of the estimator. In general, the efficiency of an IV estimation depends on the fit of the First Stage (Angrist and Pischke 2009; Newey 1990a). Hence, if the fit of the linear First Stage is poor compared to that of the Probit model, TSLS is too inefficient, resulting in standard errors which are too large compared to the NTSLS ones (Mogstad and Wiswall 2016; Angrist and Pischke 2009; Wooldridge 2002; Newey 1990a;

[18] We do not include hospital fixed effects amongst the covariates we use to estimate the Probit model, as there are more than a hundred of them and B_i is constant in some of them.

[19] This procedure is akin to using the propensity score as an instrument in linear IV (see Carneiro, Heckman, and Vytlacil 2011; Heckman and Navarro-Lozano 2004). See also Windmeijer and Santos Silva (1997) in the context of Count Data models.

Table 3—First Stage. Breastfed for at Least 90 Days. Coefficient Estimates

	Low-educated mothers					
	PROBIT (1)	OLS (2)	OLS (3)	PROBIT (4)	OLS (5)	OLS (6)
Exposure to weekend	−0.151 (0.050)	−0.039 (0.013)	−0.036 (0.014)			
Hour				0.010 (0.004)	0.002 (0.001)	0.003 (0.001)
(Hour^2)/100				−0.012 (0.006)	−0.003 (0.001)	−0.003 (0.002)
(Hour^3)/10,000				0.004 (0.002)	0.001 (0.001)	0.001 (0.001)
p-value	0.003	0.003	0.009	0.002	0.002	0.004
F-stat		8.695	6.897		4.785	4.386
Hospital FE	No	No	Yes	No	No	Yes
Observations	5,809	5,809	5,809	5,809	5,809	5,809

Notes: Each column reports the coefficients from a regression in which the dependent variable is whether the child was breastfed for at least 90 days, and the independent variables include the exclusion restrictions listed in the first column (exposure to weekend or cubic polynomial in hour), and all of the variables listed in Tables 2 and online Appendix Table B2 (including a cubic polynomial in child's age, quadratic polynomial on mother's age and a dummy if highest qualification is missing but left school before age 17), month of birth, interview months, country dummies, and whether the baby was born on a bank holiday (137 covariates in total). The model (Probit or OLS) is noted at the top of the column. The *p*-value and *F*-stat refer to the null hypothesis that the coefficient(s) of the instrument is zero or jointly zero. Sample excludes children born through caesarean sections (either emergency or planned) and children placed in intensive care after delivery. Low-educated mothers are those with NVQ level 2 or less, or unknown NVQ level but that left school before 17. Standard errors in parentheses.

Source: Millennium Cohort Study

1990b).[20] In Section IVC we provide evidence that in our case, the Probit First Stage greatly outperforms the linear First Stage in terms of fit.

B. *First-Stage Estimation*

Table 3 shows the results of Probit and OLS regressions of breastfeeding at 90 days, B, on *Exposure* (columns 1–3) or a cubic polynomial in the *Hour* variable (columns 4–6) and the set of covariates, $\mathbf{X}$, estimated over our main sample (low-educated mothers who had a vaginal delivery and whose babies were not admitted to intensive care). Mothers with low education levels who are fully exposed to the weekend are around 3.9 percentage points less likely to breastfeed for at least 90 days (marginal effect associated with column 1).[21]

The F-tests for the hypotheses that either the coefficient on *Exposure* or the terms of the polynomial are null are between 4.4 and 8.7, which lie below the critical

[20] Moreover, the consistency of the estimator does not depend on the Probit model being correct (Kelejian 1971) and the IV standard errors do not need to be corrected (Wooldridge 2002, 623). Although NTSLS implicitly uses the nonlinearities in the First Stage as a source of identifying information, Figure 3, panel A shows that both cognitive development and breastfeeding jointly track hour quite closely, indicating that our exclusion restriction provides meaningful identifying variation.

[21] The average duration of breastfeeding for those who breastfed for less than 90 days is 8.19 days, and for those who breastfed for at least 90 days is 150 days.

values in Stock and Yogo (2005). However, their critical values are derived under the assumption of a continuous endogenous regressor[22] and might have low power. We therefore conduct a Montecarlo simulation to understand the implications of the First Stage for our results (see online Appendix C).[23]

C. *Fit of the First Stage*

In this subsection we provide three pieces of evidence to show that the fit of the Probit First Stage is considerably better than that of the linear First Stage, which is the basis for obtaining efficiency gains from NTSLS over TSLS. The first is that the linear First Stage provided negative fitted values in 9 percent of the sample.

Second, we compare the predictive performance of the Probit and linear First Stages across six strata. We determine the six strata by estimating a linear model of B over $\mathbf{X}$, and obtaining its fitted values.[24] The strata correspond to individuals with fitted values below tenth percentile, between tenth and twenty-fifth percentile, between twenty-fifth and fiftieth percentile, and so on. The first observation to note is that the covariates $\mathbf{X}$ exhibit good predictive power over B: the within strata average breastfeeding rate (column 2 of Table 4) is significantly higher for higher strata.

Columns 3 and 4 of Table 4 report the success rates at predicting B from the OLS and Probit models respectively. Both linear and Probit First Stages deliver very similar success rates in predicting B, except for the fiftieth–seventy-fifth strata for which the Probit success rate is more than ten percentage points higher. This is particularly important because individuals in the fiftieth–seventy-fifth strata are neither definite $B = 1$ nor $B = 0$, and, hence, the instrument will play an important role in determining the value of B for that strata.

Third, the marginal effects of *Exposure* over B exhibit substantial heterogeneity over $\mathbf{X}$, but they are constant in the linear First Stage. As B is a binary variable, and as the covariates have good predictive power over B, we expect the marginal effect of *Exposure* on B, $\frac{\mathrm{dProb}(B = 1|\mathbf{X}, Exposure)}{\mathrm{d}Exposure}$, to depend on the value of the covariates. To see this, it is useful to consider a simple threshold-crossing model for B:

$$B_i = 1\left[\beta_1 Exposure_i + \boldsymbol{\beta}_2 \mathbf{X_i} > \vartheta_i\right],$$

where ϑ_i is a random error term with mean zero and finite variance. Consider those individuals whose $\mathbf{X}$ values put them in strata zero–tenth and tenth–twenty-fifth of Table 4. According to their $\mathbf{X}$ values, they are very unlikely to be breastfed (see column 2), so their $\boldsymbol{\beta}_2 \mathbf{X_i}$ is extremely negative. Because B is bounded by 0, an increase in *Exposure* cannot make B negative (unlike a linear model). Moreover, a

[22] This is relevant because TSLS implicitly uses the optimal linear instrument (the conditional mean) when the endogenous regressor is continuous but not when it is discrete. Intuitively, OLS will result in a relatively poor fit (and hence relatively "low" F-statistics) if the dependent variable is discrete.

[23] Stock and Yogo (2005) indicate that the critical values could be much lower depending on the value of unknown parameters. Cruz and Moreira (2005) obtain meaningful estimates even when the First-Stage F-statistics are as low as 2, suggesting that the rule-of-thumb of F-statistic larger than 10 is far from conclusive (Angrist and Pischke 2009; Murray 2006).

[24] Note that we use *Exposure* to estimate both the Probit and linear First Stage, but we do not use it to form the strata.

Table 4—Analysis of Goodness of Fit of the First-Stage Models: OLS and Probit

Predicted probability of breastfeeding ↓	Number of observations (1)	Average probability of breastfeeding (2)	Prediction success rate (OLS) (3)	Prediction success rate (Probit) (4)	Marginal effect (Probit) (5)
<tenth percentile	580	0.045	0.955	0.955	−0.010 [0.003]
Tenth–twenty-fifth percentile	872	0.073	0.927	0.927	−0.021 [0.007]
Twenty-fifth–fiftieth percentile	1,453	0.136	0.857	0.864	−0.034 [0.011]
Fiftieth–seventy-fifth percentile	1,451	0.251	0.320	0.435	−0.048 [0.016]
Seventy-fifth–nintieth percentile	872	0.396	0.396	0.396	−0.057 [0.019]
≧Nintieth percentile	581	0.585	0.585	0.585	−0.056 [0.019]

Notes: Children are classified into six strata according to the predicted probability of being breastfed, estimated through OLS over the set of covariates listed in Table 2 and online Appendix Table B2 (including a cubic polynomial in child's age, quadratic polynomial on mother's age and a dummy variable if highest qualification is missing but left school before age 17), month of birth dummies, interview month dummies, country dummies, and whether the baby was born on a bank holiday (137 covariates in total). Column 2 reports the actual average probability of breastfeeding within each strata. Columns 3 and 4 report the success rate at predicting breastfeeding of the OLS and Probit models respectively, estimated using the same covariate set and Exposure. Column 5 reports the marginal effect of Exposure estimated using the Probit model, with standard errors computed using the delta method in parentheses.

Source: Millennium Cohort Study

decrease in *Exposure* will hardly shift *B* from 0 to 1 unless the instrument effect size is so large as to overcome the very negative value of $\beta_2 \mathbf{X_i}$. Hence, it is unlikely that for these individuals, the instrument will shift *B* from 0 to 1. Consistent with this, column 5 of Table 4 reports an average marginal effect of *Exposure* very close to zero for children in the zero–tenth and tenth–twenty-fifth strata. For higher strata, the covariate values are such that $\beta_2 \mathbf{X_i}$ take values closer to zero and, hence, *Exposure* can play a bigger role. Hence, the marginal effects for these strata are larger than for lower strata, as reflected in column 5 of Table 4. The heterogeneity of marginal effects reported in Table 4 contrasts with the linear model, for which the marginal effect is constant (−0.039) and independent of the covariate values. Note that this argument does not depend on ϑ_i being normally distributed, and that this heterogeneity of the marginal effects is not necessarily a general property, as we would not expect to observe it if the **X**s did not have had good predictive power over *B*.

D. *Finite Sample Properties*

Given the strength and goodness of fit of our First Stage, can we expect our estimator to have good finite sample properties? In online Appendix C, we provide details on the design and results of a Monte Carlo experiment to assess the finite sample properties of the estimators. We use our sample and parameter estimates (including our First Stage estimates) to simulate the Monte Carlo samples. We find that both NTSLS and TSLS are consistent if the true effect of breastfeeding is relatively small (including zero), NTSLS is biased toward zero if the true effect is large, and the standard errors are correctly estimated. This means that our estimates are conservative and, if anything, provide lower bounds. We also find that NTSLS is far more precise than TSLS.

As part of the Monte Carlo experiment, we also assess the sensitivity of the findings to departures from normality in the error term. Following Westerlund and Hjertstrand (2014), we assume that the error term that generates the breastfeeding variable is distributed following a t-distribution, a mixture of two normals, or a generalized logistic. Even if we use a Probit in the estimation of the NTSLS, we find that the standard errors are correctly estimated, that NTSLS is biased toward zero if the true effect is large, and that NTSLS is far more precise than TSLS. This is not surprising because the properties of the NTSLS do not crucially depend on the Probit being the correct model (Wooldridge 2002; Kelejian 1971).

V. Results

In this section we first describe results for child development as measured using the summary indices, and next show the results separately by age and subscale. We then consider mechanisms relating to maternal behavior, including the home environment and maternal mental health.

A. *Effects on Overall Child Development*

Measures of cognition are based on age-appropriate tests administered to the child, and noncognitive skills are based on maternal reports at ages 3, 5 and 7 (Section II and online Appendix A). We use child measured weight and maternal-reported measures of health and morbidity (at ages 9 months, 3, 5, 7 years). We consider as outcomes the indices summarizing cognitive skills, noncognitive skills and health across all ages (constructed as described in Section II).[25] All indices are coded so that larger values correspond to higher levels of development.

Results by Child Development Domain.— The main results for the three summary indices are shown in Table 5. The key finding is that, irrespective of whether we use *Exposure* or the cubic polynomial in *Hour* as exclusion restriction (columns 1 and 4), breastfeeding affects positively the overall cognitive development of children whose mothers have relatively low levels of education (in line with Figure 3 panel A), and the effect is significant at the 1 percent level. The *p*-values of the Conditional Likelihood Ratio test for our coefficient of interest are 0.0078 for cognitive development, 0.1534 for noncognitive development, and 0.459 for the health index (Andrews, Moreira, and Stock 2007; Mikusheva and Poi 2006; Moreira 2003). The key difference between NTSLS and TSLS is the precision of the estimates: the NTSLS standard errors are much smaller than those of

[25] Like Anderson (2008) and Kling, Liebman, and Katz (2007), the number of tests contributing to the index need not be constant across individuals. So we can still create the index for individuals who attrit/have some missing test measures, which we return to in Section VIA.

Table 5—Effect of Breastfeeding on Child Development

Exclusion restriction →	Exposure to weekend			Polynomial in hour		
Estimation method ↓	Cognitive index (1)	Noncognitive index (2)	Health index (3)	Cognitive index (4)	Noncognitive index (5)	Health index (6)
NTSLS	0.464	0.319	0.009	0.452	0.344	−0.010
	(0.179)	(0.224)	(0.082)	(0.169)	(0.213)	(0.078)
TSLS	0.501	0.233	−0.396	0.467	0.559	−0.303
	(0.616)	(0.803)	(0.292)	(0.421)	(0.587)	(0.202)
OLS	0.057	0.097	0.016	0.057	0.097	0.016
	(0.019)	(0.023)	(0.009)	(0.019)	(0.023)	(0.009)
F-statistic	7.023	5.701	8.636	3.728	3.094	4.737
p-value instrument	0.008	0.017	0.003	0.011	0.026	0.003
Observations	5,015	4,957	5,809	5,015	4,957	5,809

Notes: Each cell reports coefficient of breastfeeding for at least 90 days from separate regressions in which the dependent variable is listed at the top of the column and the estimation method is listed in the left hand column (NTSLS denotes nonlinear two-stage least squares; TSLS denotes two-stage least squares; OLS denotes ordinary least squares). Control variables are those listed in Table 2 and online Appendix Table B2 (including a cubic polynomial in child's age, quadratic polynomial on mother's age and a dummy variable if highest qualification is missing but left school before age 17), month of birth dummies, interview month dummies, country dummies, and whether the baby was born on a bank holiday (137 covariates in total), as well as hospital fixed effects. In columns 1 to 3 exposure to weekend is excluded from the second-stage regression, while in columns 4 to 6 the cubic polynomial in hour is excluded. *F*-statistic and *p*-value correspond to the null hypothesis that the coefficient(s) of the excluded variable(s) are zero or jointly zero, as estimated from an OLS regression where the dependent variable is breastfeeding for at least 90 days, and controls are as noted already. Sample comprises low-educated mothers (NVQ level 2 or less, or NVQ level unknown but left school before 17), and excludes children born through caesarean sections (either emergency or planned) and children placed in intensive care after delivery. Standard errors in parentheses.

Source: Millennium Cohort Study

TSLS, for reasons explained in Section IVC. As in Table 5, throughout the paper the results using *Hour* as the exclusion restriction are very similar to those using *Exposure* to weekend, hence we focus on the latter from hereon.

A key finding from Table 5 is that the effects of breastfeeding are mainly concentrated on cognitive development: we cannot reject that breastfeeding has no effect on health in this period of childhood, and the effects on noncognitive development are inconclusive (as had been anticipated from Figures 3 panels B and C). We note, however, that both health and noncognitive development are likely measured with more error as they are based on maternal report, unlike the cognitive measures, which are direct assessments.[26]

IV versus OLS Comparison.—Table 5 also reports OLS estimates, which are positive and statistically significant throughout (the health one is significant only at 10 percent). The IV estimates are markedly larger than the OLS ones. There are three potential reasons for this: misclassification error, negative selection into

[26]The sample used in the health index is larger because mothers are asked about children's health from nine months onward, but children's cognitive and noncognitive development is assessed from age three. However, as we report in Section VIA, attrition is uncorrelated with the instruments.

breastfeeding, and heterogeneous treatment effects. In what follows, we discuss the latter two.

The conventional omitted variable concern is that mothers with characteristics that facilitate an improvement in their child's cognitive development (e.g., higher socioeconomic status, higher maternal involvement) are also more likely to breastfeed for longer. In online Appendix F (Table F3), we report findings from two OLS regressions, one in which the dependent variable is breastfeeding for at least 90 days, another in which the dependent variable is the cognitive development index. If the signs of the coefficients from both regressions are the same, it indicates that selection is positive in that covariate; if the signs are different then selection is negative. Strictly speaking, these results only speak to selection on observables, but may be informative about selection on unobservables. As expected, there are several variables for which selection is positive (owning a computer, expected education attainment at age 16, income support, attending antenatal classes), but also some for which it is negative (ethnic minority, whether the mother worked during pregnancy, and use of epidural as pain relief).[27]

Another potential explanation for the IV estimate to exceed the OLS one is that when the treatment effect is heterogenous, IV identifies a Local Average Treatment Effect, LATE (Heckman and Vytlacil 2007; Heckman, Urzua, and Vytlacil 2006; Imbens and Angrist 1994). As our instrument is continuous, we follow the methodology of Card, Fenizia, and Silver (2018) to characterize the compliers' characteristics. For selected characteristics, we report the average of the characteristic in the population as well as amongst compliers (online Appendix Table F4). Findings show that among the sample of low-educated mothers, compliers are more likely to have had a complication during delivery (e.g., use of forceps or vacuum extractor). It is likely that hospital support is more important for these mothers postpartum. Compliers are also more advantaged: they are more educated, have higher socioeconomic status, the mothers are more likely to have worked during pregnancy, the mother is more likely to be in a relationship, the father is more likely to be present at birth, and the mother is more likely to have taken up early antenatal care. In our sample they are more likely to be making other complementary investments in their children, which may amplify the effects of breastfeeding, due to dynamic complementarities (Cunha and Heckman 2007; Heckman 2007; Cunha et al. 2006).

Results by Age.— Table 6 reports the results at age 3 and 5 for cognitive development using *Exposure* to weekends as the exclusion restriction (similar results are obtained using *Hour*).[28] The top panel reports results using the score as the dependent variable. Although it is customary to assess the size of the effect by the standardized effect (the effect in levels divided by the standard deviation, reported at the bottom of the table), we are concerned that the score distribution is not normal

[27] As in Rayfield, Oakley, and Quigley (2015); Santorelli et al. (2013); and Agboado et al. (2010), we find that mothers with non-White ethnicity are more likely to breastfeed (higher by around 20 percentage points compared to White ethnicity).

[28] Results are also available for age 7. However, due to the marked increase in attrition at age 7, these results are shown in online Appendix D, where we also report the results on noncognitive development and health by age.

Table 6—Effect of Breastfeeding on Cognitive Outcomes at Ages 3 and 5 Years

	3 years		5 years		
	Expressive language (1)	School readiness (2)	Expressive language (3)	Pictorial reasoning (4)	Visuo-spatial (5)
Panel A. Level dependent variable					
NTSLS	9.88	8.26	8.583	3.212	4.892
	(5.037)	(3.707)	(5.164)	(4.162)	(6.753)
TSLS	21.979	7.336	22.183	14.443	23.256
	(21.872)	(12.345)	(19.729)	(15.538)	(25.323)
OLS	2.062	1.038	1.579	1.100	1.119
	(0.623)	(0.456)	(0.544)	(0.442)	(0.727)
F-statistic	4.696	6.539	5.386	5.570	5.498
p-value	0.030	0.011	0.020	0.018	0.019
Mean	70.38	22.19	104.1	80.24	85.43
SD	17.74	12.56	15.64	11.75	19.7
Observations	4,212	4,004	4,349	4,355	4,333
Panel B. Percentile dependent variable					
NTSLS	15.399	22.177	17.847	18.194	3.772
	(8.570)	(8.485)	(9.997)	(10.665)	(10.290)
TSLS	37.374	25.784	40.348	39.727	15.662
	(36.647)	(29.502)	(37.027)	(39.437)	(33.622)
OLS	3.118	2.33	2.753	3.121	1.895
	(1.005)	(1.012)	(1.037)	(1.112)	(1.054)
F-statistic	4.696	6.539	5.386	5.570	5.498
p-value	0.030	0.011	0.020	0.018	0.019
Observations	4,212	4,004	4,349	4,355	4,333

Notes: Each cell reports coefficient of breastfeeding for at least 90 days from separate regressions in which the dependent variable is listed at the top of the column and the estimation method is listed in the left-hand column (NTSLS denotes nonlinear two-stage least squares; TSLS denotes two-stage least squares; OLS denotes ordinary least squares). Control variables are those listed in Table 2 and online Appendix Table B2 (including a cubic polynomial in child's age, quadratic polynomial on mother's age and a dummy variable if highest qualification is missing but left school before age 17), month of birth dummies, interview month dummies, country dummies, and whether the baby was born on a bank holiday (137 covariates in total), as well as hospital fixed effects. The upper panel uses as the score in levels as dependent variable, whilst the bottom panel uses the percentile in the sample distribution of the score as dependent variable. The exclusion restriction from the second-stage regressions is exposure to weekend. *F*-statistic and *p*-value correspond to the null hypothesis that the coefficient(s) on the excluded variable(s) is zero, as estimated from an OLS regression where the dependent variable is breastfeeding for at least 90 days, and controls are as noted already. Sample comprises low-educated mothers (NVQ level 2 or less, or NVQ level unknown but left school before 17), and excludes children born through caesarean sections (either emergency or planned) and children placed in intensive care after delivery. Standard errors in parentheses.

Source: Millennium Cohort Study

and, hence, the standardized effect size may give a misleading impression of the effect size.[29] For this reason, we also estimate models in which the dependent variable is the percentile of the child's score in the sample distribution (reported in

[29] The Kolmorov-Smirnov tests rejected normality for all five scores. See also Figure A1 of online Appendix A.

the bottom panel of Table 6), to represent the increase in terms of percentiles of the cognitive score distribution.

The estimates are all positive across the different measures of cognition and statistically significant for expressive language (age 3 and 5) and school readiness, but not for pictorial reasoning and visuospatial skills (age 5). The magnitude of the effects is around 55 percent SD for expressive language and 65 percent SD for school readiness. However, the standardized effect size may give an inflated impression of the effect size. For instance, if the score distribution was normal, the effect size of expressive language at age 3 (55 percent SD) would imply that an individual at the median would be shifted to the seventy-first percentile, but only to the sixty-fifth percentile according to the percentile estimates at the bottom of Table 6 $(50 + 15 = 65)$.

While the effect sizes are large, they are also imprecisely estimated, with wide confidence intervals. Our estimates are half way between previous estimates that are based on methods that rely on the selection of observables assumption (OLS, matching), and those that attempt to control for unobservables. On the former, a recent meta-analysis that summarizes the estimates from 16 different studies from high-income countries (Horta, Mola, and Victora 2015) finds that the average effect of breastfeeding on cognitive development is 25 percent of a SD (95 percent confidence interval: 16 percent–33 percent).[30] On the latter, the randomized trial in Belarus found improvements in verbal IQ, vocabulary and similarities of around 1.2 SD (Kramer, Aboud, et al. 2008).[31] Del Bono and Rabe (2012) also exploit the same UNICEF initiative as Kramer, Aboud, et al. (2008), using as an instrument the distance from the mother's home to the closest hospital that voluntarily implemented the UNICEF program, finding effects on cognitive development of between 0.7 SD and 1.5 SD, depending on the measure.

It is also interesting to note that, especially at age 5, our effects seem to be concentrated on verbal skills, rather than pictorial reasoning or visuospatial skills. Interestingly, Isaacs et al. (2011) discuss several studies linking DHA (the fatty acid component that breast milk is rich in) with verbal performance. Moreover, the results of the randomized trial by Kramer, Aboud, et al. (2008) in Belarus are also concentrated on verbal/language domains (results on performance IQ and Full Scale IQ were not statistically different from zero at 6.5 years of age).

A rich literature in health science studies the association between breastfeeding and health in developed countries. A report summarizing around 400 individual studies concluded that breastfeeding was associated with several health benefits (Ip et al. 2007). This contrasts with our results of Table 5 (and separately by each age group in online Appendix Tables D3–D6) in which we report lack of statistically significant improvements in health (subject to the caveats noted in Section VA). Interestingly, the results of the randomized trial of Kramer et al. (2001) are more in line with ours: in the first year of life, breastfeeding reduced gastrointestinal tract

[30] See table 2 of Kramer, Aboud, et al. (2008). We report their estimates divided by 15, which is the standard deviation of the Intelligence Quotient.

[31] Kramer, Aboud, et al. (2008) report intention to treat estimates of around 0.45 SD. We report Wald estimates computed using the estimates reported in their paper.

infection and atopic eczema (but did not reduce upper respiratory tract infections, otitis media, croup, wheezing, or pneumonia). At 6.5 years of age, no reductions were found in allergies, asthma, blood pressure or obesity (Kramer et al. 2009; 2007). Like ours, other papers using IV strategies have found no evidence of breastfeeding improving health outcomes (Del Bono and Rabe 2012; Baker and Milligan 2008).

There is far less evidence on the effects of breastfeeding on noncognitive skills. Kramer, Fombonne, et al. (2008) cannot reject that breastfeeding does not improve noncognitive skills in Belarus. We reach the same conclusion. We reported lack of statistically significant results of breastfeeding on the overall index of noncognitive skills (Table 5), and at ages 3, 5, and 7 separately (online Appendix Table D2). This contrasts with Del Bono and Rabe (2012) who find that breastfeeding improves child emotional development.

B. *Mechanisms*

The stark findings shown raise the question as to the underlying mechanisms through which breastfeeding may affect children's cognition. Our data lends itself to testing one of the four mechanisms discussed in Section I, that breastfeeding may improve the relationship between mother and child, due to hormonal responses that may reduce maternal stress and depression, and/or breastfeeding resulting in the mother spending more time with the baby. An improved mother-child relationship may result in an increase in interactive activities likely to increase cognitive development (such as reading/telling stories); any observed increase in such activities may also be due to their perceived returns being higher for breastfed children. Of course, the direction of the relationship could go the other way, for instance, if mothers invest more in these activities in order to compensate for not having breastfed. We here consider both the effect of breastfeeding on maternal activities with the child, and on the quality of the mother-child relationship (which could indirectly affect maternal behaviors, as the literature hypothesizes).

Maternal Investments.—We use the frequency of learning activities such as reading to the child, library visits, singing, painting (see online Appendix A) to analyze whether mothers respond to breastfeeding by altering other parental investments. The activities comprise the Home Learning Environment (HLE) index, a composite measure of the quality and quantity of stimulation and support available to a child at home (Bradley 1995). Column 1 of online Appendix Table F5 reports the overall summary index of the HLE indices at ages 3, 5, and 7 computed following Anderson (2008). The remaining columns of the upper panel focus on age 3, and the lower panel on age 5. Columns 2–7 report results for separate activities, and column 8 shows the result for the activities combined into the HLE index. Though imprecisely estimated, we cannot reject that there is no effect of breastfeeding on the learning activities that parents provide their children with.

Maternal Mental Health and Mother-child Relationship.— We find no significant effects of breastfeeding on maternal mental health measured using the Malaise Inventory or Kessler-6, either overall (column 1 of online Appendix Table F6) or at specific ages (columns 2–4). The last two columns of online Appendix Table F6 estimate whether breastfeeding affects the quality of the mother-child relationship, measured using the Pianta Scales at child age 3. We detect no effect of breastfeeding on either relationship warmth or relationship conflict.

Breastfeeding, *Fertility, and Family Size*.—It is plausible that the effects are due to smaller family size. Extended breastfeeding could reduce fertility, resulting in parents investing more resources into fewer children (Jayachandran and Kuziemko 2011; Angrist, Lavy, and Schlosser 2010; Black, Devereux, and Salvanes 2005; Rosenzweig and Wolpin 1980). This is not very likely because extended breastfeeding is not common in our sample (only 7 percent of children were breastfed beyond 9 months). Indeed, we see little difference in the average number of younger siblings across weekday- and weekend-born children (0.639 and 0.627, respectively), and it is not statistically significant ($P = 0.634$).

VI. Robustness

In this section we discuss attrition from the sample and show a battery of robustness exercises.

A. *Sample Attrition*

Online Appendix E provides a detailed analysis of attrition from the sample. We summarize its four key aspects here. First, attrition is uncorrelated with the variation we exploit for identification. Indeed, attrition at various waves is practically the same for children exposed to weekend and to those who are not (the difference ranges between –1.1 percent and +0.6 percent depending on the wave, and is not statistically different from zero in any case, see online Appendix Table E1). This balance also extends to the instruments used in the analysis *Exposure* and *Hour* (online Appendix Table E2). Second, the rich set of characteristics that we observe are well balanced between those exposed to weekend and those who are not, across ages 3, 5, and 7 (see online Appendix Tables E3–E8). Third, the sample used to obtain our main result (Table 5, column 1) is well balanced as was shown in online Appendix Tables B23–B28. Fourth, those who attrit are from more disadvantaged backgrounds (Table E9).

B. *Falsification Test*

As we previously saw, exposure to weekend does not predict breastfeeding status for the group of high-educated mothers (online Appendix Tables F1 and F2). We use that to present a falsification exercise, in which we show the reduced form because the lack of a First Stage for this group precludes us from using IVs. Similar to, for instance Blundell and Powell (2003), the reduced form is given by the expectation

Table 7—Reduced Forms

	High-educated mothers			Low-educated mothers		
Prob [B = 1| X, Exposure] ↓	Cognitive index (1)	Noncognitive index (2)	Health index (3)	Cognitive index (4)	Noncognitive index (5)	Health index (6)
$\Phi[\beta_0 + \beta_1\text{Exposure} + \beta_2\text{X}]$	0.104	0.338	0.071	0.467	0.326	0.01
	(0.311)	(0.428)	(0.152)	(0.171)	(0.229)	(0.085)
Observations	4,822	4,792	5,354	5,015	4,957	5,809

Notes: Each cell reports results from a separate OLS regression, in which the dependent variable is listed at the top of the column. The coefficient reported is that of the variable listed on the left. Control variables are those listed in Table 2 and Appendix Table B2 (including a cubic polynomial in child's age, quadratic polynomial on mother's age and education dummies), month of birth dummies, interview month dummies, country dummies, and whether the baby was born on a bank holiday, as well as hospital fixed effects. The sample in the left panel comprises higher educated mothers (NVQ 3 or higher) and the sample in the right panel comprises low-educated mothers (NVQ level 2 or less, or NVQ level unknown but left school before 17). The sample excludes children born through caesarean section (either emergency or planned) and children placed in intensive care after delivery. Standard errors in parentheses.

Source: Millennium Cohort Study

of the outcome variable, Y_i, conditional on the covariates and exclusion restriction ($Exposure_i$), so

$$\text{E}\left[Y_i|\mathbf{X_i}, Exposure_i\right] = \alpha_0 + \alpha_1 \text{E}\left[B_i|\mathbf{X_i}, Exposure_i\right] + \boldsymbol{\alpha}_2\mathbf{X_i}, \tag{5}$$

where $\text{E}\left[B_i|\mathbf{X_i},\ Exposure_i\right] = \text{Prob}\left[B_i = 1|\mathbf{X_i},\ Exposure_i\right]$, because B_i only takes values 0 or 1. In Table 7 we report the OLS reduced form estimates using that $\text{Prob}\left[B_i = 1|\mathbf{X_i}, Exposure_i\right] = \Phi\left[\hat{\beta}_0 + \hat{\beta}_1 Exposure_i + \hat{\boldsymbol{\beta}}_2\mathbf{X_i}\right]$. The left panel reports the results for the sample of high-educated mothers. We find no significant relation between any of the measures of development and $\Phi\left[\hat{\beta}_0+\hat{\beta}_1 Exposure_i + \hat{\boldsymbol{\beta}}_2\mathbf{X_i}\right]$, consistent with the notion that *Exposure* only affects children's cognitive development through its effect on breastfeeding. The results of the right panel (low-educated mothers) are in line with our IV regressions (Table 5).[32]

C. *Robustness Exercises*

We carry out a number of exercises to check robustness of our main findings to specification and sample selection. Column 1 of online Appendix Table F7 reports our main results using *Exposure* as exclusion restriction (already reported in Table 5). In column 2 we remove labor inductions from the sample; in column 3 we include emergency C-sections; and in columns 4 and 5, we condition on time of birth within the day (using either a third-order polynomial in the hour of birth

[32] If we assume that Prob $\left[B_i = 1|\mathbf{X_i}, Exposure_i\right]$ is linear in both Exposure and $\mathbf{X}$, the results have the expected sign but are not statistically significant (in accordance with the linear IV results).

defined between 0 and 23, or dummy variables for each hour of birth).[33] In all cases, the effect of breastfeeding on cognitive development remains large and statistically significant. In column 6 we impute missing values (due to attrition) in the cognitive outcomes based on the values of non-missing waves and find very similar results.[34] In column 7 we show that the estimate is somewhat smaller when we do not control for hospital fixed effects, which seems to indicate that any unobserved hospital or area level variable would result in downward bias.

As an additional robustness check, we use cut-offs different from 90 days to define the breastfeeding binary variable. Rather than trying to estimate the optimal duration of breastfeeding (for which we would need exogenous variation in the cost of breastfeeding at different ages of the child), the aim here is to show that our results apply more generally and are not an artefact of the specific 90 days threshold used in the main analysis. While online Appendix Table F8 shows that the effect of breastfeeding for at least 30 days is smaller (and not statistically significant) than the effect of breastfeeding for at least 90 days, the effects of breastfeeding for at least 60 or 120 days are extremely similar to that of breastfeeding for at least 90 days.

VII. Conclusion

In this paper, we have used exogenous variation in timing of birth to estimate the impacts of breastfeeding on children's development at different stages up to age 7. Our results, which apply to mothers with relatively low levels of education, are striking. We find strong effects of breastfeeding on children's cognitive development, the effects on noncognitive skills are inconclusive, and we find no evidence of effects on health during this period of childhood. Results from the only randomized trial to study the effect of breastfeeding on child development are quite similar to our results in several dimensions (Kramer et al. 2009; Kramer, Aboud, et al. 2008; Kramer, Fombonne, et al. 2008; Kramer et al. 2007, 2001). We also note that estimates from the IV strategy of Del Bono and Rabe (2012) are also suggestive of large effects of breastfeeding on cognition, but no effects on a number of child health outcomes; similarly, Baker and Milligan (2008) find no evidence of breastfeeding affecting infant or maternal health.

Identifying the effects of breastfeeding on child development has been a challenging research topic because it has been difficult to identify a credible exogenous source of variation. Whilst we believe that our paper makes an important contribution in this regard and advances knowledge in important ways, it also has some limitations: (1) we include labor inductions within our estimating sample. This might be problematic if expectant mothers exercise some choice in this regard. Although we believe that most mothers will follow the health professionals' recommendations, we note that the association between *Exposure* and labor induction does not vary according to other observable characteristics; (2) our confidence intervals are wide,

[33] We do this because there is a within day cycle in inductions and epidurals. Inductions are more frequent in the morning and hence children are born later in the day (epidurals follow the same pattern because epidurals are administered more frequently for induced deliveries).

[34] Robustness results (available on request) on noncognitive skills and health are also in line with the main ones.

though the estimates for cognitive development are statistically significant; (3) our data on health and noncognitive measures are mostly based on maternal reports (with the exception of child weight), unlike the cognitive development measures which are directly assessed from the child; (4) our estimates are only applicable to compliers, who are relatively better-off mothers (amongst those with relatively low education) and those who experienced some complication during delivery, and we cannot extrapolate our results to other groups of the population.

We find no effects on mother's mental health, the quality of the child-mother relationship, or parental time investments in their children. However, the same caveat, that our estimates are quite imprecise, also applies for these mechanisms, and further research is necessary.

The evidence provided suggests that breastfeeding may well contribute to the gap in children's cognitive development across the socioeconomic spectrum. Moreover, the instrument used to identify the effects suggests a specific policy focus—on hospital breastfeeding support—to help close this gap.

REFERENCES

Agboado, Gabriel, Elaine Michel, Elaine Jackson, and Arpana Verma. 2010. "Factors Associated with Breastfeeding Cessation in Nursing Mothers in a Peer Support Programme in Eastern Lancashire." *BMC Pediatrics* 10: 3.

Almond, Douglas. 2006. "Is the 1918 Influenza Pandemic Over? Long-Term Effects of *In Utero* Influenza Exposure in the Post-1940 U.S. Population." *Journal of Political Economy* 114 (4): 672–712.

Almond, Douglas, and Janet Currie. 2011. "Killing Me Softly: The Fetal Origins Hypothesis." *Journal of Economic Perspectives* 25 (3): 153–72.

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature* 56 (4): 1360–1446.

Almond, Douglas, and Joseph J. Doyle. 2011. "After Midnight: A Regression Discontinuity Design in Length of Postpartum Hospital Stays." *American Economic Journal: Economic Policy* 3 (3): 1–34.

Almond, Douglas, Joseph J. Doyle, Jr., Amanda E. Kowalski, and Heidi Williams. 2010. "Estimating Marginal Returns to Medical Care: Evidence from At-Risk Newborns." *Quarterly Journal of Economics* 125 (2): 591–634.

Almond, Douglas, and Bhashkar Mazumder. 2011. "Health Capital and the Prenatal Environment: The Effect of Ramadan Observance During Pregnancy." *American Economic Journal: Applied Economics* 3 (4): 56–85.

American Academy of Pediatrics. 2012. "Breastfeeding and the Use of Human Milk." *Pediatrics* 129 (3): e827–e841.

American College of Obstetricians and Gynecologists. 2003. "ACOG Committee on Ethics Number 289, November 2003: Surgery and Patient Choice: The Ethics of Decision Making." *Obstetrics and Gynecology* 102 (5): 1101–06.

American College of Obstetricians and Gynecologists. 2009. "ACOG Practice Bulletin No. 107: Induction of Labor." *Obstetrics and Gynecology* 114 (2): 386–97.

Ampaabeng, Samuel K., and Chih Ming Tan. 2013. "The Long-Term Cognitive Consequences of Early Childhood Malnutrition: The Case of Famine in Ghana." *Journal of Health Economics* 32 (6): 1013–27.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–95.

Andrews, Donald W.K., Marcelo J. Moreira, and James H. Stock. 2007. "Performance of Conditional Wald Tests in IV Regression with Weak Instruments." *Journal of Econometrics* 139 (1): 116–32.

Angrist, Joshua, Victor Lavy, and Analia Schlosser. 2010. "Multiple Experiments for the Causal Link between the Quantity and Quality of Children." *Journal of Labor Economics* 28 (4): 773–824.

Angrist, Joshua, and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Axon, R. Neal, and Mark V. Williams. 2011. "Hospital Readmission as an Accountability Measure." *JAMA* 305 (5): 504–05.

Baker, Michael, and Kevin Milligan. 2008. "Maternal Employment, Breastfeeding, and Health: Evidence from Maternity Leave Mandates." *Journal of Health Economics* 27 (4): 871–87.

Barham, Tania. 2012. "Enhancing Cognitive Functioning: Medium-Term Effects of a Health and Family Planning Program in Matlab." *American Economic Journal: Applied Economics* 4 (1): 245–73.

Barham, Tania, Karen Macours, and John A. Maluccio. 2013. "Boys' Cognitive Skill Formation and Physical Growth: Long-Term Experimental Evidence on Critical Ages for Early Childhood Interventions." *American Economic Review* 103 (3): 467–71.

Behrman, Jere R., and Mark R. Rosenzweig. 2004. "Returns to Birthweight." *Review of Economics and Statistics* 86 (2): 586–601.

Belfield, Clive R., and Inas Rashad Kelly. 2012. "The Benefits of Breast Feeding across the Early Years of Childhood." *Journal of Human Capital* 6 (3): 251–77.

Bharadwaj, Prashant, Katrine Vellesen Løken, and Christopher Neilson. 2013. "Early Life Health Interventions and Academic Achievement." *American Economic Review* 103 (5): 1862–91.

Black, Sandra E., and Paul J. Devereux. 2011. "Recent Developments in Intergenerational Mobility." In *Handbook of Labor Economics*, Vol. 4B, edited by David Card and Orley Ashenfelter, 1487–1541. Elsevier.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2005. "The More the Merrier? The Effect of Family Size and Birth Order on Children's Education." *Quarterly Journal of Economics* 120 (2): 669–700.

Blundell, Richard, and James L. Powell. 2003. "Endogeneity in Nonparametric and Semiparametric Regression Models." In *Advances in Economics and Econometrics*, Vol. 2, edited by Mathias Dewatripont, Lars Peter Hansen, and Stephen J. Turnovsky, 312–57. Cambridge, UK: Cambridge University Press.

Bolling, Keith, Catherine Grant, Becky Hamlyn, and Alex Thornton. 2007. *Infant Feeding Survey 2005.* Information Centre.

Borra, Cristina, Maria Iacovou, and Almudena Sevilla. 2012. "The Effect of Breastfeeding on Children's Cognitive and Noncognitive Development." *Labour Economics* 19 (4): 496–515.

Bracken, Bruce A. 2002. *Bracken School Readiness Assessment: Administration Manual*. San Antonio, TX: Psychological Corporation.

Bradley, Robert H. 1995. "Environment and Parenting." In *Handbook of Parenting, Volume 2: Biology and Ecology of Parenting*, edited by Marc H. Bornstein, 235–61. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Brion, Marie-Jo A., Debbie A. Lawlor, Alicia Matijasevich, Bernardo Horta, Luciana Anselmi, Cora L. Araújo, Ana Maria B. Menezes, Cesar G. Victora, and George Davey Smith. 2011. "What Are the Causal Effects of Breastfeeding on IQ, Obesity, and Blood Pressure? Evidence from Comparing High-Income with Middle-Income Cohorts." *International Journal of Epidemiology* 40 (3): 670–80.

Britton, Jennifer C., K. Luan Phan, Stephan F. Taylor, Robert C. Welsh, Kent C. Berridge, and I. Liberzon. 2006. "Neural Correlates of Social and Nonsocial Emotions: An fMRI Study." *NeuroImage* 31 (1): 397–409.

Brodman, Keeve, Albert J. Erdmann, Irving Lorge, and Charles P. Gershenson. 1952. "The Cornell Medical Index—Health Questionnaire. III. The Evaluation of Emotional Disturbances." *Journal of Clinical Psychology* 8 (2): 119–24.

Brodman, Keeve, Albert J. Erdmann, and Harold G. Wolff. 1949. "Cornell Medical Index Health Questionnaire." Cornell Medical Index. http://www.scalesandmeasures.net/files/files/Cornell%20Medical%20Index%20(1949).pdf.

Card, David, Alessandra Fenizia, and David Silver. 2018. "The Health Effects of Cesarean Delivery for Low-Risk First Births." NBER Working Paper 24493.

Carneiro, Pedro, James J. Heckman, and Edward J. Vytlacil. 2011. "Estimating Marginal Returns to Education." *American Economic Review* 101 (6): 2754–81.

Carneiro, Pedro, Katrine V. Løken, and Kjell G. Salvanes. 2015. "A Flying Start? Maternity Leave Benefits and Long-Run Outcomes of Children." *Journal of Political Economy* 123 (2): 365–412.

Case, Anne, Darren Lubotsky, and Christina Paxson. 2002. "Economic Status and Health in Childhood: The Origins of the Gradient." *American Economic Review* 92 (5): 1308–34.

CDC. (2020), 2018 National Immunization Survey - Child. National Center for Immunization and Respiratory Diseases. https://www.cdc.gov/breastfeeding/data/nis_data/rates-any-exclusive-bf-socio-dem-2018.html (accessed March 22, 2022).

Chay, Kenneth Y., Jonathan Guryan, and Bhashkar Mazumder. 2009. "Birth Cohort and the Black-White Achievement Gap: The Roles of Access and Health Soon After Birth." NBER Working Paper 15078.

Clandinin, M.T., J.E. Chappell, T. Heim, P.R. Swyer, and G.W. Chance. 1981. "Fatty Acid Utilization in Perinatal de Novo Synthesis of Tissues." *Early Human Development* 5 (4): 355–66.

Cruz, Luiz M., and Marcelo J. Moreira. 2005. "On the Validity of Econometric Techniques with Weak Instruments: Inference on Returns to Education Using Compulsory School Attendance Laws." *Journal of Human Resources* 40 (2): 393–410.

Cunha, Flavio, and James Heckman. 2007. "The Technology of Skill Formation." *American Economic Review* 97 (2): 31–47.

Cunha, Flavio, James J. Heckman, Lance Lochner, and Dimitriy V. Masterov. 2006. "Interpreting the Evidence on Life Cycle Skill Formation." In *Handbook of the Economics of Education*, Vol. 1, edited by E. Hanushek and F. Welch, 697–812. Amsterdam: Elsevier.

Cunha, Flavio, James J. Heckman, and Susanne M. Schennach. 2010. "Estimating the Technology of Cognitive and Noncognitive Skill Formation." *Econometrica* 78 (3): 883–931.

Currie, Janet, and Douglas Almond. 2011. "Human Capital Development before Age Five." In *Handbook of Labor Economics*, Vol. 4B, edited by David Card and Orley Ashenfelter, 1315–1486. Amsterdam: Elsevier.

Dahl, Gordon B., and Lance Lochner. 2012. "The Impact of Family Income on Child Achievement: Evidence from the Earned Income Tax Credit." *American Economic Review* 102 (5): 1927–56.

Daniels, Melissa C., and Linda S. Adair. 2005. "Breast-Feeding Influences Cognitive Development in Filipino Children." *Journal of Nutrition* 135 (11): 2589–95.

Dawson, Diane, Rowena Jacobs, Steve Martin, and Peter Smith. 2004. "Is Patient Choice an Effective Mechanism to Reduce Waiting Times?" *Applied Health Economics and Health Policy* 3 (4): 195–203.

Del Bono, Emilia, and Birgitta Rabe. 2012. "Breastfeeding and Child Cognitive Outcomes: Evidence from a Hospital-Based Breastfeeding Support Policy." Institute for Social and Economic Research Working Paper 2012-29.

Der, Geoff, G. David Batty, and Ian J. Deary. 2006. "Effect of Breast Feeding on Intelligence in Children: Prospective Study, Sibling Pairs Analysis, and Meta-Analysis." *BMJ* 333 (7575): 945.

DiGirolamo, Ann M., Laurence M. Grummer-Strawn, and Sara B. Fein. 2008. "Effect of Maternity-Care Practices on Breastfeeding." *Pediatrics* 122 (S2): S43–S49.

Evenhouse, Eirik, and Siobhan Reilly. 2005. "Improved Estimates of the Benefits of Breastfeeding Using Sibling Comparisons to Reduce Selection Bias." *Health Services Research* 40 (6): 1781–1802.

Fernald, Lia C.H., Patricia Kariger, Patrice Engle, and Abbie Raikes. 2009. *Examining Early Child Development in Low-Income Countries: A Toolkit for the Assessment of Children in the First Five Years of Life*. Washington, DC: World Bank. https://doi.org/10.1596/28107.

Fernstrom, John D. 1999. "Effects of Dietary Polyunsaturated Fatty Acids on Neuronal Function." *Lipids* 34 (2): 161–69.

Field, Erica, Omar Robles, and Maximo Torero. 2009. "Iodine Deficiency and Schooling Attainment in Tanzania." *American Economic Journal: Applied Economics* 1 (4): 140–69.

Fitzsimons, Emla, and Marcos Vera-Hernández. 2022. "Replication data for Breastfeeding and Child Development." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.38886/E135401V1.

Freemantle, Nick, Daniel Ray, David McNulty, David Rosser, Simon Bennett, Bruce E. Keogh, and Domenico Pagano. 2015. "Increased Mortality Associated with Weekend Hospital Admission: A Case for Expanded Seven Day Services?" *BMJ* 351: h4596.

Gerber Medical. 2013. "FAQ about DHA: The Skinny on the Popular Omega-3 Fatty Acid." https://medical.gerber.com/nutrition-health-topics/breastfeeding/articles/faq-about-dha.

Gertler, Paul, James Heckman, Rodrigo Pinto, Arianna Zanolini, Christel Vermeerch, Susan Walker, Susan M. Chang, and Sally Grantham-McGregor. 2014. "Labor Market Returns to an Early Childhood Stimulation Intervention in Jamaica." *Science* 344 (6187): 998–1001.

Glewwe, Paul, and Elizabeth M. King. 2001. "The Impact of Early Childhood Nutritional Status on Cognitive Development: Does the Timing of Malnutrition Matter?" *World Bank Economic Review* 15 (1): 81–113.

Goodman, Robert. 1997. "The Strengths and Difficulties Questionnaire: A Research Note." *Journal of Child Psychology and Psychiatry* 38 (5): 581–86.

Grantham-McGregor, Sally M., Lia C. Fernald, and Kavita Sethuraman. 1999. "Effects of Health and Nutrition on Cognitive and Behavioural Development in Children in the First Three Years of Life: Part 1: Low Birthweight, Breastfeeding, and Protein-Energy Malnutrition." *Food and Nutrition Bulletin* 20 (1): 53–75.

Hall, Alastair R., Glenn D. Rudebusch, and David W. Wilcox. 1996. "Judging Instrument Relevance in Instrumental Variables Estimation." *International Economic Review* 37 (2): 283–98.

Healthcare Commission, Picker Institute Europe. 2009. "Maternity Survey, 2007. 3rd Edition." UK Data Service. https://doi.org/10.5255/UKDA-SN-5785-1 (accessed August 25, 2011).

Heckman, James J. 2007. "The Economics, Technology, and Neuroscience of Human Capability Formation." *Proceedings of the National Academy of Sciences* 104 (33): 13250–55.

Heckman, James J., and Dimitriy V. Masterov. 2007. "The Productivity Argument for Investing in Young Children." *Review of Agricultural Economics* 29 (3): 446–93.

Heckman, James J., and Stefano Mosso. 2014. "The Economics of Human Development and Social Mobility." *Annual Review of Economics* 6: 689–733.

Heckman, James, and Salvador Navarro-Lozano. 2004. "Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models." *Review of Economics and Statistics* 86 (1): 30–57.

Heckman, James J., Sergio Urzua, and Edward Vytlacil. 2006. "Understanding Instrumental Variables in Models with Essential Heterogeneity." *Review of Economics and Statistics* 88 (3): 389–432.

Heckman, James J., and Edward J. Vytlacil. 2007. "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments." In *Handbook of Econometrics*, Vol. 6B, edited by James J. Heckman and Edward E. Leamer, 4875–5143. Amsterdam: Elsevier.

Horta, Bernardo L., Christian Loret de Mola, and Cesar G. Victora. 2015. "Breastfeeding and Intelligence: A Systematic Review and Meta-Analysis." *Acta Paediatrica* 104 (S467): 14–19.

Hoynes, Hilary, Diane Whitmore Schanzenbach, and Douglas Almond. 2016. "Long-Run Impacts of Childhood Access to the Safety Net." *American Economic Review* 106 (4): 903–34.

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2): 467–75.

Imbens, Guido W., and Jeffrey M. Wooldridge. 2009. "Recent Developments in the Econometrics of Program Evaluation." *Journal of Economic Literature* 47 (1): 5–86.

Ip, Stanley, Mei Chung, Gowri Raman, Priscilla Chew, Nombulelo Magula, Deirdre DeVine, Thomas Trikalinos, and Joseph Lau. 2007. *Breastfeeding and Maternal and Infant Health Outcomes in Developed Countries.* Rockville, MD: Agency for Healthcare Research and Quality.

Isaacs, Elizabeth B., Sarah Ross, Kathy Kennedy, Lawrence T. Weaver, Alan Lucas, and Mary S. Fewtrell. 2011. "10-Year Cognition in Preterms after Random Assignment to Fatty Acid Supplementation in Infancy." *Pediatrics* 128 (4): e890-898.

Jayachandran, Seema, and Ilyana Kuziemko. 2011. "Why Do Mothers Breastfeed Girls Less than Boys? Evidence and Implications for Child Health in India." *Quarterly Journal of Economics* 126 (3): 1485–1538.

Kelejian, Harry H. 1971. "Two-Stage Least Squares and Econometric Systems Linear in Parameters but Nonlinear in the Endogenous Variables." *Journal of the American Statistical Association* 66 (334): 373–74.

Kelly, Elaine. 2011. "The Scourge of Asian Flu: In Utero Exposure to Pandemic Influenza and the Development of a Cohort of British Children." *Journal of Human Resources* 46 (4): 669–94.

Kessler, Ronald C., Peggy R. Barker, Lisa J. Colpe, Joan F. Epstein, Joseph C. Gfroerer, Eva Hiripi, Mary J. Howes, et al. 2003. "Screening for Serious Mental Illness in the General Population." *Archives of General Psychiatry* 60 (2): 184–89.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.

Koletzko, Berthold, Eric Lien, Carlo Agostoni, Hansjosef Böhles, Cristina Campoy, Irene Cetin, Tamas Decsi, et al. 2008. "The Roles of Long-Chain Polyunsaturated Fatty Acids in Pregnancy, Lactation and Infancy: Review of Current Knowledge and Consensus Recommendations." *Journal of Perinatal Medicine* 36 (1): 5–14.

Kramer, Michael S., Frances Aboud, Elena Mironova, Irina Vanilovich, Robert W. Platt, Lidia Matush, Sergei Igumnov, et al. 2008. "Breastfeeding and Child Cognitive Development: New Evidence from a Large Randomized Trial." *Archives of General Psychiatry* 65 (5): 578–84.

Kramer, Michael S., Beverley Chalmers, Ellen D. Hodnett, Zinaida Sevkovskaya, Irina Dzikovich, Stanley Shapiro, Jean-Paul Collet, et al. 2001. "Promotion of Breastfeeding Intervention Trial (PROBIT): A Randomized Trial in the Republic of Belarus." *JAMA* 285 (4): 413–20.

Kramer, Michael S., Eric Fombonne, Sergei Igumnov, Irina Vanilovich, Lidia Matush, Elena Mironova, Natalia Bogdanovich, et al. 2008. "Effects of Prolonged and Exclusive Breastfeeding on Child Behavior and Maternal Adjustment: Evidence from a Large, Randomized Trial." *Pediatrics* 121 (3): e435–40.

Kramer, Michael S., Lidia Matush, Irina Vanilovich, Robert Platt, Natalia Bogdanovich, Zinaida Sevkovskaya, Irina Dzikovich, Gyorgy Shishko, and Bruce Mazer. 2007. "Effect of Prolonged and Exclusive Breast Feeding on Risk of Allergy and Asthma: Cluster Randomised Trial." *BMJ* 335 (7624): 815.

Kramer, Michael S., Lidia Matush, Irina Vanilovich, Robert W. Platt, Natalia Bogdanovich, Zinaida Sevkovskaya, Irina Dzikovich, et al. 2009. "A Randomized Breast-Feeding Promotion Intervention Did Not Reduce Child Obesity in Belarus." *Journal of Nutrition* 139 (2): 417S-421S.

Liebman, Jeffrey B., Lawrence F. Katz, and Jeffrey R. Kling. 2004. "Beyond Treatment Effects: Estimating the Relationship Between Neighborhood Poverty and Individual Outcomes in the MTO Experiment." https://dataspace.princeton.edu/bitstream/88435/dsp01g158bh29k/1/493.pdf.

Maccini, Sharon, and Dean Yang. 2009. "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall." *American Economic Review* 99 (3): 1006–26.

Maluccio, John A., John Hoddinott, Jere R. Behrman, Reynaldo Martorell, Agnes R. Quisumbing, and Aryeh D. Stein. 2009. "The Impact of Improving Nutrition During Early Childhood on Education among Guatemalan Adults." *Economic Journal* 119 (537): 734–63.

Martorell, Reynaldo, Paul Melgar, John A. Maluccio, Aryeh D. Stein, and Juan A. Rivera. 2010. "The Nutrition Intervention Improved Adult Human Capital and Economic Productivity." *Journal of Nutrition* 140 (2): 411–14.

Meacock, Rachel, Laura Anselmi, Søren Rud Kristensen, Tim Doran, and Matt Sutton. 2017. "Higher Mortality Rates amongst Emergency Patients Admitted to Hospital at Weekends Reflect a Lower Probability of Admission." *Journal of Health Services Research and Policy* 22 (1): 12–19.

Merten, Sonja, Julia Dratva, and Ursula Ackermann-Liebrich. 2005. "Do Baby-Friendly Hospitals Influence Breastfeeding Duration on a National Level?" *Pediatrics* 116 (5): e702-708.

Mikusheva, Anna, and Brian P. Poi. 2006. "Tests and Confidence Sets with Correct Size When Instruments Are Potentially Weak." *Stata Journal* 6 (3): 335–47.

Milligan, Kevin, and Mark Stabile. 2011. "Do Child Tax Benefits Affect the Well-Being of Children? Evidence from Canadian Child Benefit Expansions." *American Economic Journal: Economic Policy* 3 (3): 175–205.

Mogstad, Magne, and Matthew Wiswall. 2016. "Testing the Quantity-Quality Model of Fertility: Estimation Using Unrestricted Family Size Models." *Quantitative Economics* 7 (1): 157–92.

Moreira, Marcelo J. 2003. "A Conditional Likelihood Ratio Test for Structural Models." *Econometrica* 71 (4): 1027–48.

Murray, Michael P. 2006. "Avoiding Invalid Instruments and Coping with Weak Instruments." *Journal of Economic Perspectives* 20 (4): 111–32.

Nelson, Steven E., Ronald R. Rogers, Ekhard E. Ziegler, and Samuel J. Fomon. 1989. "Gain in Weight and Length during Early Infancy." *Early Human Development* 19 (4): 223–39.

Newey, Whitney K. 1990a. "Efficient Instrumental Variables Estimation of Nonlinear Models." *Econometrica* 58 (4): 809–37.

Newey, Whitney K. 1990b. "Semiparametric Efficiency Bounds." *Journal of Applied Econometrics* 5 (2): 99–135.

NHS Digital. n.d. "Hospital Episode Statistics (HES)". https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/hospital-episode-statistics (accessed May 9, 2021).

National Collaborating Centre for Primary Care (UK). 2006. *Postnatal Care: Routine Post-Natal Care of Women and Their Babies.* London: Royal College of General Practitioners.

Palmer, William L., A. Bottle, and P. Aylin. 2015. "Association between Day of Delivery and Obstetric Outcomes: Observational Study." *BMJ* 351: h5774.

Pianta, Robert C. 1992. "Child-Parent Relationship Scale." VA: University of Virginia. https://www.bristol.ac.uk/media-library/sites/sps/documents/c-change/cprs.pdf.

Pinheiro, Emily, Debra L. Bogen, Denada Hoxha, Jody D. Ciolino, and Katherine L. Wisner. 2015. "Sertraline and Breastfeeding: Review and Meta-Analysis." *Archives of Women's Mental Health* 18 (2): 139–46.

Quigley, Maria A., Christine Hockley, Claire Carson, Yvonne Kelly, Mary J. Renfrew, and Amanda Sacker. 2012. "Breastfeeding Is Associated with Improved Child Cognitive Development: A Population-Based Cohort Study." *Journal of Pediatrics* 160 (1): 25–32.

Rayfield, Sarah, Laura Oakley, and Maria A. Quigley. 2015. "Association between Breastfeeding Support and Breastfeeding Rates in the UK: A Comparison of Late Preterm and Term Infants." *BMJ Open* 5 (11): e009144.

Renfrew, M.J., D. Craig, L. Dyson, F. McCormick, S. Rice, S.E. King, K. Misso, E. Stenhouse, and A.F. Williams. 2009. "Breastfeeding Promotion for Infants in Neonatal Units: A Systematic Review and Economic Analysis." *Health Technology Assessment* 13 (40).

Reynolds, Ann. 2001. "Breastfeeding and Brain Development." *Pediatric Clinics of North America* 48 (1): 159–71.

Robinson, P.M. 1988. "Root-N-Consistent Semiparametric Regression." *Econometrica* 56 (4): 931–54.

Rosenzweig, Mark R., and Kenneth I. Wolpin. 1980. "Testing the Quantity-Quality Fertility Model: The Use of Twins as a Natural Experiment." *Econometrica* 48 (1): 227–40.

Rothstein, Donna S. 2013. "Breastfeeding and Children's Early Cognitive Outcomes." *Review of Economics and Statistics* 95 (3): 919–31.

Rutter, M., J. Tizard, and K. Whitmore, eds. 1970. *Education, Health and Behaviour*. Longman Group: London, UK.

Santorelli, Gillian, Emily Petherick, Dagmar Waiblinger, Baltica Cabieses, and Lesley Fairley. 2013. "Ethnic Differences in the Initiation and Duration of Breast Feeding—Results from the Born in Bradford Birth Cohort Study." *Paediatric and Perinatal Epidemiology* 27 (4): 388–92.

Scholte, Robert S., Gerard J. van den Berg, and Maarten Lindeboom. 2015. "Long-Run Effects of Gestation during the Dutch Hunger Winter Famine on Labor Market and Hospitalization Outcomes." *Journal of Health Economics* 39: 17–30.

Stock, James H., and Motohiro Yogo. 2005. "Testing for Weak Instruments in Linear IV Regression." In *Identification and Inference for Econometric Models: Essays in Honor of Thomas Rothenberg*, edited by Donald W.K. Andrews and James H. Stock, 80–108. Cambridge, UK: Cambridge University Press.

Uauy, Ricardo, and Isidora De Andraca. 1995. "Human Milk and Breast Feeding for Optimal Mental Development." *Journal of Nutrition* 125 (8S): 2278S–2280S.

Uauy, Ricardo, and Patricio Peirano. 1999. "Breast Is Best: Human Milk Is the Optimal Food for Brain Development." *American Journal of Clinical Nutrition* 70 (4): 433–34.

Renfrew, Mary J., Subhash Pokhrel, Maria Quigley, Felicia McCormic, Julia Fox-Rushby, Rosemary Dodds, Steven Duffy, Paul Trueman, and Anthony Williams. 2012. *Preventing Disease and Saving Resources: The Potential Contribution of Increasing Breastfeeding Rates in the UK*. London: UNICEF UK.

University College London. 2001. "Millennium Cohort Study: First Survey, 2001–2003." UK Data Archive. https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=4683 (accessed June 15, 2010).

University College London. 2003. "Millennium Cohort Study: Second Survey, 2003–2004." UK Data Archive. https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=5350 (accessed June 15, 2010).

University College London. 2006. "Millennium Cohort Study: Third Survey, 2006." UK Data Archive. https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=5795 (accessed June 15, 2010).

University College London. 2008. "Millennium Cohort Study: Fourth Survey, 2008." UK Data Archive. https://beta.ukdataservice.ac.uk/datacatalogue/studies/study?id=6411 (accessed June 15, 2010).

Van den Berg, Gerard J., Dorly J.H. Deeg, Maarten Lindeboom, and France Portrait. 2010. "The Role of Early-Life Conditions in the Cognitive Decline Due to Adverse Events Later in Life." *Economic Journal* 120 (548): F411–28.

Van den Berg, Gerard J., Maarten Lindeboom, and France Portrait. 2006. "Economic Conditions Early in Life and Individual Mortality." *American Economic Review* 96 (1): 290–302.

Walker, A. Sarah, Amy Mason, T. Phuong Quan, Nicola J. Fawcett, Peter Watkinson, Martin Llewelyn, Nicole Stoesser, et al. 2017. "Mortality Risks Associated with Emergency Admissions during Weekends and Public Holidays: An Analysis of Electronic Health Records." *Lancet* 390 (10089): 62–72.

Walker, Susan P., Susan M. Chang, Marcos Vera-Hernández, and Sally Grantham-McGregor. 2011. "Early Childhood Stimulation Benefits Adult Competence and Reduces Violent Behavior." *Pediatrics* 127 (5): 849–57.

Westerlund, Joakim, and Per Hjertstrand. 2014. "Indirect Estimation of Semiparametric Binary Choice Models." *Oxford Bulletin of Economics and Statistics* 76 (2): 298–314.

Windmeijer, F.A.G., and J.M.C. Santos Silva. 1997. "Endogeneity in Count Data Models: An Application to Demand for Health Care." *Journal of Applied Econometrics* 12 (3): 281–94.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.