Breastfeeding and Children's Development

Emla Fitzsimons[*]
Marcos Vera-Hernández [†]

April 2012

Abstract: There is a large gradient in breastfeeding rates across education groups, with rates in the developed world considerably higher amongst the relatively more educated. This may a contributory factor in the intergenerational transmission of human capital. In this paper, we estimate the causal effects of breastfeeding on children's development. We provide strong evidence to show that babies born just before or during the weekend are significantly less likely to be breastfed, most likely because hospitals dedicate fewer resources to non-essential services – such as breastfeeding support - at times when they are more costly, namely at weekends. We use variation in the timing/day of birth to estimate the effect of breastfeeding on children's subsequent cognitive and non-cognitive outcomes. With the exception of planned Caesarean sections, which we exclude from the analysis, we argue that timing of birth is random. We find that breastfeeding has large and significant effects on the developmental outcomes of the children of relatively less educated mothers. These effects remain until at least the age of 7. A Monte Carlo analysis shows that our Instrumental Variable estimates are conservative and our confidence intervals have the right coverage.

Keywords: Breastfeeding; weekend birth; instrumental variables; cognitive and non-cognitive development

JEL classification: I14, I18, J13

Acknowledgements: We thank Manuel Arellano, Richard Blundell, Olympia Bover, and Adam Rosen for helpful discussions, as well as participants at the "Early Life Interventions and Intergenerational Mobility" conference, University of Bristol October 2011 and the IFS internal seminar. All errors are the responsibility of the authors.

Correspondence: emla_f@ifs.org.uk; m.vera@ucl.ac.uk.

* Institute for Fiscal Studies, London

[†] University College London and Institute for Fiscal Studies, London

1. Introduction

In this paper we estimate the causal effects of breastfeeding on children's cognitive and non-cognitive development in the UK. This is important for at least three reasons. First, breastfeeding is an investment made very early on in life, a critical period in shaping the long-term development of individuals. Though there is fairly robust evidence to suggest that it reduces childhood illnesses and chronic disease (León-Cava et al, 2002), there is much less known about its effects on children's cognitive and non-cognitive development. Second, the paper contributes to opening the black box of why children from better off backgrounds achieve better outcomes compared to those from relatively less well off backgrounds: breastfeeding rates amongst the high educated are almost double those of the low educated, so a natural question is whether breastfeeding contributes to the gap in children's development across the socio-economic spectrum. Third, the question is of considerable policy interest - in 2001, the World Health Organization (WHO) put forward for consideration a global recommendation that infants should be exclusively breast fed for six months. Many Western countries, including around two thirds of European member states and the United States, elected not to follow this recommendation fully, or at all.[1] There are clear and important implications for maternity leave policies – for instance the UK offers paid time off for breastfeeding for women who choose to return to work in the first six months, and statutory maternity leave entitlement of one year.

We use an Instrumental Variable methodology to estimate the causal effects of breastfeeding. Our identification strategy uses exogenous differences in breastfeeding support immediately after birth as a source of natural variation in breastfeeding.[2] These differences in support are driven by day of the week of birth, which we show affects strongly the chances that a baby is breastfed, with those born just before or during the weekend significantly less likely to be breastfed compared to those born during the week. The reason behind this is that a hospital manager cuts back on the provision of non-essential services, such as feeding support, at the weekend, when they are more expensive. And without this support early on, particularly in the first *48 hours* after delivery, it is considerably more difficult for successful breastfeeding to be established. Crucial to our identification strategy, we argue and provide evidence that

[1] The UK has complied with the recommendation since 2003.
[2] Exploiting breastfeeding support as a predictor of breastfeeding is in line with the largest ever lactation randomised trial conducted (Kramer et al, 2008).

day of the week of birth, which is random (note, we exclude planned Caesarean sections for this reason[3]) only affects outcomes through affecting breastfeeding. We use this exogenous variation to measure the effects of breastfeeding on children's cognitive and non-cognitive development. Using a sample of children born in the early part of this century, we show that breastfeeding affects cognitive and non-cognitive development, with large and significant effects observed at ages 3, 5 and 7. The magnitude of our estimates is comparable to those obtained by Kramer et al (2008), the largest randomised trial ever conducted in the area of lactation. The trial focused on the randomisation of health care worker assistance for initiating and maintaining breastfeeding and lactation and postnatal breastfeeding support: in focusing on exogenously shifting breastfeeding support, our identification strategy shares common ground with this trial.

Our work fits into the large and growing literature on the importance of the very early stages of life in shaping the development of individuals (see for instance Carneiro and Heckman, 2003; Cunha et al. 2010; Almond and Currie 2011, in terms of social development, intellectual development, and health. Neurological research shows that brain development occurs during the first years of life, and that this is a function of the quality and range of early experiences and interactions (Thompson and Nelson, 2001). There are many studies that focus specifically on the period before or just after birth, such as those that document the importance of prenatal conditions (for instance Currie and Hyson, 1999), of birth weight (for instance, Behrman and Rosenzweig, 2004; Currie and Moretti 2007; Black, Devereux and Salvanes 2007; Royer 2009), and economic conditions around the time of birth (Berg et al. 2006; Cutler et al 2007; Banerjee et al 2009); other studies look at the early childhood environment, particularly the importance of early childhood shocks. There is more of a gap in the literature around investments in the early post-natal period, right after birth, where in terms of nutrition, the type of milk to provide is the key decision faced by parents. Of those that do consider this period, they tend to consider child health shocks and maternal health/depression (for instance Case and Paxson (2008a,b, 2010a,b); Frank and Meara, 2009). As Almond and Currie (2011) note"one of the more effective ways to improve children's

[3] Planned C-sections are those that are scheduled for medical reasons. Elective Caesarean sections (i.e. by parents) are not an issue in the UK NHS as women who cannot afford to pay private doctors for their baby's delivery have been allowed to have C-sections only if there are health concerns for mother or baby. This is set to change however.

long term outcomes might be to target women of child bearing age in addition to focusing on children after birth".

Breastfeeding falls quite naturally into this period of life, yet the literature contains relatively little on its effects. It has been hypothesised to affect children's development through at least two channels. One relates to the 'superior' constituents of breast milk, particularly higher concentrations of essential long-chain polyunsaturated fatty acids, which are believed to play an important role in cell division and brain maturation (Innis, 2004; Petryk et al, 2007), and the presence of insulinlike growth factor I, which is contained in higher concentration in breast milk than in formula (Nagashima et al, 1990) and has been shown to be absorbed intact across the newborn infant's gastrointestinal tract (Phillips et al, 1995). The second channel relates to the physical interaction which is an integral part of breastfeeding, which might lead to permanent physiologic changes that accelerate neurocognitive development: increased skin-to-skin contact has been shown to be associated with secure attachment (Briton et al, 2006), and increased maternal-infant contact in breastfeeding could increase verbal interaction between mother and infant, which might also stimulate cognitive development. Whilst this paper will not be able to disentangle the precise channels through which breastfeeding may affect outcomes, we will provide some evidence later on that the mother-child attachment is no stronger in breastfed than in non-breastfed children.

However the inherent problem in estimating the causal effects of breastfeeding is that it is endogenous, and there are likely to be other unobserved differences between women who choose to breastfeed and those who do not, which are also relevant for children's development. There is just one randomised control trial (RCT), that of Kramer et al (2008), which randomises the provision of facilities to promote breastfeeding and assesses whether exclusive breastfeeding for three months improves children's cognitive development. It finds large effects, of between one fifth and one half of a standard deviation, on verbal IQ, performance IQ, and full-scale IQ, at age 6.5 years.

Apart from this, related studies are observational, and use different methods to control for selection bias. For instance, the literature that uses sibling pairs to control for unobserved family background characteristics generally finds no effects of breastfeeding on children's cognitive development (Rothstein, 2011; Belfield and Kelly, 2010; Der et al 2006; Evenhouse and Reilly, 2005).[4] Another set of studies uses Instrumental Variables. For instance, Rothstein (2011) uses State breastfeeding rates and laws about breastfeeding in public as instruments; Del Bono and Rabe (2011) use whether the hospital where the child is born participated in a (non-randomly allocated) breastfeeding promotion programme. Del Bono and Rabe report significantly positive effects of breastfeeding on cognitive development and no effects on non-cognitive outcomes. The third strategy widely used is to control for a wide range of background variables (Rothstein 2011; Quigley et al. 2011; Belfield and Kelly 2010), and this literature generally finds small effects on cognitive measures, and little to no effects on non-cognitive measures.

2. Background and Data

2.1 Background: Maternity Care in the UK

The UK National Health Service (NHS) is a unique state-run health care system.[5] The vast majority of UK deliveries occur in NHS hospitals. In our data (described in next section), just under 2% of women reported that they had had a home birth; we do not know what proportion of children in our data were born in private hospitals but the England-wide proportion in 2001-2002, the period covering our data, was just 0.5%.[6]

The UK has a well-developed midwifery profession that provides maternity care to the majority of women.[7] Only when an instrumental delivery or a surgical birth is necessary will an obstetrical specialist be called upon, making it unusual for obstetricians to attend vaginal births. This is in contrast to North America, where maternity care is obstetric-led, using highly trained specialists (i.e., physicians) to attend nearly all births as the primary caregiver (Clarke et al, 2003; Conrad and Leiter, 2004). Indeed in 1999, 69.4% of live births in UK

[4] Note there is a large literature on the health benefits of breastfeeding in the medical literature, which suggests that the greatest and most obvious benefits of breastfeeding are for the immediate health and survival of infant. With the exception of the Kramer et al (2001) RCT, the studies are mainly observational. León-Cava et al (2002) consider the evidence to be fairly compelling "*As the epidemiological evidence favouring breastfeeding is generally derived from multiple studies in a variety of situations, the evidence is in sum, convincing*".

[5] Although funded centrally from national taxation, NHS services in England, Northern Ireland, Scotland and Wales are managed separately. While some differences have emerged between these systems in recent years, they remain similar in most respects and are considered as belonging to a single, unified system. The Department of Health is responsible for the NHS. The Secretary of State for Health is the head of the Department of Health and reports to the Prime Minister.

[6] It is likely to be even lower in our data, the sample design of which allowed for a disproportionate representation of families living in areas of child poverty.

[7] In 1992, the Winterton Report by the House of Commons Health Select Committee released a watershed report on maternity care in Britain, which said that: "Maternity Care should no longer be based on the medical model of care".

attended by midwives compared to 7.65% in US. More generally, the dynamics of private markets and competition, together with the lack of national health insurance (or other systems of compulsory health insurance) frame the issue of access to maternity care in the US in a different way from how it is framed in the UK where access to health care is secured through such arrangements (Woodhandler & Himmelstein 2007).

Women giving birth in the UK have limited choice about length of stay after delivery, and it is mainly supply driven and judged by the midwife. The median length of stay for first-borns with normal deliveries is 48 hours, and 24 hours for second-borns. After discharge, postnatal care is transferred to a Community midwife/health visitor, a nurse who has had extra training in child development and health promotion and who works in the community. Health visitors make home visits in the early days (up to ~10 days after discharge) and after that, community midwifery care in is provided at GP surgeries, 'drop-in' sessions at shopping or community centres and via Sure Start schemes (DoH, 2007).

2.2 Data

We use three waves of data from the UK Millennium Cohort Study (MCS), a panel data set which follows a sample of nearly 18,500 babies born at the beginning of the noughties.[8] The study's overarching objective is to create a new longitudinal dataset, describing the diversity of backgrounds from which children born in the new century are setting out on life. Four surveys of MCS members have been conducted so far: MCS1 at 9 months (2000/2001), MCS2 at three years (2004/05), MCS3 at five years (2006), and MCS4 at seven years (2008). The next study is planned for 2012.

The sample design allowed for disproportionate representation of families living in areas of child poverty, in the smaller countries of the UK and in areas with high ethnic minority populations in England. The first survey recorded in detail the circumstances of pregnancy and birth, as well as those of the early months of life, and the social and economic background of the family into which the children have been born such as parental characteristics, employment, childcare, income, attitudes. Subsequent surveys collected information on parenting activities, child health, child cognitive and non-cognitive development, early education, housing, employment and income.

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

We use two measures of cognitive ability, based on age-appropriate tests administered to children themselves. The first is the British Ability Scales (BAS), which is measured directly from the child at ages 3, 5 and 7 (MCS2,3,4). Six different BAS tests have been administered across the MCS sweep. The BAS Naming Vocabulary test is a verbal scale which assesses spoken vocabulary (MCS2,3). Children are shown a series of coloured pictures of objects one at a time which they are asked to name. The scale measures the children's expressive language ability. In the BAS Pattern Construction Test, the child constructs a design by putting together flat squares or solid cubes with black and yellow patterns on each side (MCS3,4). The child's score is based on both speed and accuracy in the task. The BAS Picture Similarity Test assesses pictorial reasoning (MCS3). The BAS Word Reading Test the child reads aloud a series of words presented on a card (MCS4).

The second measure of cognitive ability is the Bracken School Readiness Assessment. This is used to assess the conceptual development of young children across a wide range of categories, each in separate subtests (Bracken, 2002). MCS2 employs six of the subtests which specifically evaluate: colours, letters, numbers/counting, sizes, comparisons, and shapes. The test result used is a composite score based on the total number of correct answers across all six subtests.

We measure the behavioural development of children using with the Strengths and Difficulties Questionnaire (SDQ). This is a widely used behavioural screening questionnaire for 3 to 16-year-olds (Goodman, 1997, 2001; Goodman, Meltzer and Bailey, 1998). It consists of 25 items which generate scores for five subscales measuring: conduct problems; hyperactivity; emotional symptoms; peer problems; and pro-social behaviour. The child's behaviour is reported by a parent, normally the mother, in the computer assisted self-completion module of the questionnaire.

We also observe a measure of the relationship between the mother and child at three years from a self-reported instrument completed by mothers that assesses her perceptions of her relationship with her child (Pianta, 1992).

3. Identification Strategy

At the heart of our identification strategy is recognition of the fact that breastfeeding is a skill that requires learning at an early stage of the baby's life in order to be successful. Many women experience difficulties with breastfeeding - for instance, Hamlyn et al (2002) report that majority of women (87%) who discontinue breastfeeding would have preferred to feed for longer but did not, mainly due to problems with breastfeeding rather than by choice. An extensive body of research has demonstrated that mothers require active support for establishing and sustaining appropriate breastfeeding practices, particularly in the very early days of life. There are many studies on the importance of hospital policies and time immediately after birth as key determinants of breastfeeding success - for instance, skin-to-skin contact straight after birth (e.g. Renfrew et al, 2009; Bolling et al, 2005); increased "Baby-Friendly" hospital practices, and several other maternity-care practices (Di Girolamo et al, 2008; Merten et al, 2005); whether the mother can independently attach baby on discharge and whether or not artificial baby milk administered in hospital (McAllister et al, 2009); individualised breastfeeding support and consistency (Backstrom et al, 2010); extra professional support (Sikorski et al, 2002).

Against this backdrop, we argue in this paper that this type of feeding advice and support differs greatly in hospitals by day of the week, which subsequently affects breastfeeding.[9] In particular, the provision of feeding support is much lower at weekends. The reason for this is that weekend services are more expensive, so a hospital manager allocating resources will protect essential services at weekends, and cut back on non-essential services such as infant feeding support, and thus employ less experienced staff.[10] In particular, as the median hospital stay after the birth of a baby is 48 hours, we would expect those born on Fridays to be most exposed to weekend services, followed by those born on Saturdays.

This suggests that mothers of babies born early on in the weekend are likely to have less breastfeeding support at hospital, which is likely to affect breastfeeding. This is supported by evidence from the UK Maternity Users Survey (2007), a postal survey conducted on a sample of around 26,000 mothers around three months after giving birth, and covering 148 NHS trusts in England. The survey covered each of the three main stages in maternity care: antenatal care, labour and delivery, and postnatal care. Of particular interest here, it asked

[9] Note that in our sample, 98% of births take place in a hospital.
[10] In a questionnaire completed by random sample of 3,000 nurses - *Workloads, Pay and Morale of Qualified Nurses in 1994* – the proportion in receipt of "Special Duty Payments" (enhanced weekend/public holiday) is much higher for lower grade nurses.

respondents whether “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?” The responses are listed in Table 1 below, by day of week of birth, and separately for low and high educated mothers. Two interesting points emerge from the table. First, looking at the left hand panel, feeding support is generally considered by mothers to be significantly worse on a Friday, and to a lesser extent a Saturday, compared to Tuesday (the omitted category). This Friday effect is likely due to the fact that these mothers are most exposed to weekend services, compared to all other days (median stay in hospital after birth is 48 hours). Second, this pattern is observed for the low educated only: for the high educated, there is no discernible different in perceptions of feeding support by weekday/weekend.

We next show graphical evidence that breastfeeding varies by day of the week of birth. Note before proceeding that throughout the paper, Caesarean sections and babies placed in intensive care after delivery are excluded: for planned Caesareans, birth delivery in this way, and its timing, may reflect a choice on the part of parents and thus not be exogenous; for emergency Caesareans, we exclude them in order to alleviate concerns that results are driven by children born at weekends suffering from adverse events that might impede their future development; it is for the latter reason that we also exclude children who went into intensive care after delivery.[11]

Data used in Figure 1 are from the UK Millennium Cohort Study, a panel data set which follows a sample of nearly 18,500 babies born at the beginning of the noughties and described more fully in section 3. Note that throughout the paper, our sample excludes multiple births, those who were not born in a hospital, those in special care after delivery, and planned Caesarean sections. Northern Ireland is also excluded. We see from the figure that breastfeeding is noticeably lower on Fridays and at the weekend. Again, it is interesting to note that this pattern is particularly discernible for the relatively less educated (upper panel). The evidence presented here leads us to focus the remainder of the analysis on the relatively less educated, whose breastfeeding decisions seem to be considerably more sensitive to birth timing.

[11] Another point to make is that the issues around successful initiation and continuation of breastfeeding are different for babies born through normal delivery, and those born medically through Caesarean section/those separated from the mother and placed in intensive care after delivery.

[FIGURE 1 HERE]

Interestingly, when we look at the relationship by birth parity, we see that the weekend link holds for both, but in slightly different ways. For first-borns, Friday appears to be associated with lowest breastfeeding; Saturday for second-borns. This is in line with what one would expect: given that the median length of stay for first-(second) borns is 48 (24) hours, those born on a Friday (Saturday) are likely to be most exposed to weekend services. In the analysis, we pool birth parities to boost sample size and precision, and control for birth order throughout.

[FIGURE 2 HERE]

We next depict the relationship between breastfeeding and hour born, in Figure 3, for relatively less educated mothers. In particular the figure displays breastfeeding (y axis) against the number of hours since Sunday (00:01am onwards) that the baby was born (x axis), denoted “elapsed”. The main point to take from the figure is that the relationship is non-linear: breastfeeding rates increase as Tuesday approaches, when they peak, and then taper off right through Saturday. The reason behind this pattern is likely that Monday services are affected by a potential backlog of issues from the weekend; Tuesday births are those least likely to be exposed to the weekend; from Wednesday onwards the probability of being exposed to the weekend gradually increases, and hence we observe breastfeeding rates gradually decreasing. In the estimation, and in particular our construction of the instrumental variable, we take account of this non-linear pattern.

[FIGURE 3 HERE]

Finally, in Table 2 we compare a rich set of characteristics of those born during the week (Mon-Thurs) and at the weekend (Fri-Sun). The table first shows that aspects to do with labour and delivery are very similar, including delivery type, complications, pain relief used, birth weight of baby. One exception is emergency Caesarean sections, which are slightly more likely to occur during the week rather than at weekends.[12] Note that we exclude all C-

[12] Recall that we exclude planned Caesarean sections throughout the analysis, as they might not be exogenous.

sections from the analysis to alleviate concerns that results are driven by children born at weekends suffering from adverse events that might impede their future development.[13] The table further shows socio-economic characteristics of mothers, such as type of housing, employment, receipt of benefits, age. Again, the table leaves no cause for concern in terms of balancedness of weekend and weekday births, suggesting strongly that the event is indeed random.

[TABLE 2 HERE]

Though it is reassuring to note that the hospital birth experiences and socio-economic characteristics are very similar across weekend and weekday deliveries, we control for all of these variables throughout the estimation. It is particularly important to control for all measures relating to labour and delivery, which one might argue affect children's subsequent development, due to affecting maternal ability to care for the baby in the time immediately after birth, for instance.[14] In this way, we are confident that we are isolating the effects of breastfeeding on children's development, from the effects of hospital birth experiences more generally.

4. Estimation and Results

4.1 Methodology

In order to estimate the causal effects of breastfeeding on child's outcomes, we estimate the following linear model

$$Outcome_i = \alpha_0 + \alpha_1 Breastfed\_90_i + \alpha_2 X_i + \varepsilon_i \qquad (1)$$

where $Outcome_i$ is the outcome variable of chid i (cognitive/non-cognitive development), $Breastfed\_90_i$ is a binary variable that takes the value 1 if child i has been breastfed for the first 90 days of life and 0 otherwise[15], X_i is a vector of covariates (including a rich set of variables associated with the characteristics of the birth, as shown in Table 1), and ε_i is an error term which includes unobserved characteristics relevant for the child's development. The parameter α_1 measures the effect of being breastfed for 90 days on child i's outcomes.

[13] Another reason for excluding C-sections is that the instrument is less likely to bite for them as median hospital stay tend to be longer.
[14] Recall that we exclude babies who were put in special care after delivery, approximately 9% of the sample, as they are likely to be very different from the rest of the sample, particularly in relation to breastfeeding and development.
[15] Note that we do not distinguish between exclusive and non-exclusive breastfeeding due to data constraints.

Estimation of equation (1) may be biased because children who have been breastfed are likely to have different unobserved characteristics compared to children who have not been breastfed, resulting in a correlation between the breastfeeding variable, $Breastfed\_90_i$, and the error term, ε_i. For instance, children who are breastfed might have, on average, mothers who are more willing to obtain information on child's development, or who might be more inclined to invest in their child's development.

In order to address the potential endogeneity of $Breastfed\_90_i$, we construct an instrumental variable for it, based on the timing of birth as discussed in section 3. The variable, $Elapsed_i$, is constructed as

$$Elapsed_i=24*DayBirth_i+HourBirth_i$$

where $DayBirth_i$ is day of the week of birth that child i was born, where 0 is Sunday and 6 is Saturday, and $HourBirth_i$ is the hour of birth of child i in 24 hour format, and $Elapsed_i$ is the number of hours that have elapsed since Sunday 00:01am and the birth of child i.

For the estimation, we follow Wooldridge (2002, p. 623) and Angrist and Pischke (2008, p. 191), and use a non-linear two-stage estimator (NTSLS hereon) where we first estimate a Probit model of $Breastfed\_90_i$ over X_i and a cubic polynomial in $Elapsed_i$ (non-linear first stage):

$$Breastfed\_90_i = \beta_0 + P(\beta_1, Elapsed_i) + \beta_2 X_i + u_{i,}, \qquad (2)$$

where P(.) represents a cubic polynomial in $Elapsed_i$, in order to accommodate the non-linear shape shown in Figure 2. We denote by $\hat{B}_i$ the fitted probabilities defined as:

$$\hat{B}_i = \Phi\left[\hat{\beta}_0 + P(\hat{\beta}_1, Elapsed_i) + \hat{\beta}_2 X\mathrm{i}\right]$$

where $\hat{\beta}_{0,}\,\hat{\beta}_{1,}\,\hat{\beta}_2$ are the Probit estimates from equation (2), and $\Phi[.]$ is the cumulative distribution function of the standardized normal.

We next estimate equation (1) using Instrumental Variables, using X_i and $\hat{B}_i$ as instruments, in order to identify the causal effects of breastfeeding on outcomes.[16] There are several

[16] See Windmeijer and Santos Silva (1997) for an early application of this estimator within a count data model. See Attanasio et al (2012), Mogstad (2012) for other applications of this method.

advantages to using this two-stage method instead of standard Two Stage Least Squares (TSLS) which uses a linear rather than a non-linear first stage. The most important is that if the predictions from the first-stage Probit model (equation (2) above) provide a better approximation to $Breastfed\_90_i$ than a linear model, the resulting IV estimates are more efficient than those that use a linear first stage model (Newey, 1990).[17] A second advantage is that the consistency of the estimator does not depend on the Probit model being correct, and the IV standard errors do not need to be corrected (Wooldridge, 2002). Clearly, this procedure using nonlinear fitted values as instruments implicitly uses nonlinearities in the first stage as a source of identifying information (Angrist and Pischke, 2008). However in our case, we will show that most of the point estimates of α_1 obtained using $\hat{B}_i$ as instruments are close to those obtained using TSLS (which uses a linear first stage), which implies that our results are not driven by the nonlinearities of the first stage.

4.2 First stage estimation

Table 3 shows the results of Probit and OLS regressions of $Breastfed\_90_i$ over a cubic polynomial in the *Elapsed* variable and the set of covariates, X, estimated over the sample of low educated mothers. The relevant coefficients are those of the cubic polynomial of the *Elapsed* variable as this polynomial is later excluded from the second stage regression and hence drives the identification of the instrumental variables estimates. The models show a pattern as that of Figure 2: essentially a U-shape relation but the curve is flatter for high values of *Elapsed* (as driven by the cubic term of the polynomial). Modelling the non-linear relationship is important as a simple linear term would yield hardly any regression between *Elapsed* and the probability of breastfeeding for at least 90 days.

The three terms of the *Elapsed* polynomial are significant at either 5% or 10%, and the P-value of the joint significance of the terms of the polynomial is 0.029. While our preferred estimation method is NTSLS which uses the Probit model as first stage, we also report the estimates obtained using OLS as customary in most studies using instrumental regression (Staiger and Stock, 1997). The F-test of joint significance of the terms of the polynomial is 3, which is below the critical values reported in Stock and Yogo (2005). However, it has been recognized that the use of first stage F-statistic to assess the quality of the instruments has

[17] The Monte Carlo analysis in section 5 will show that these gains are very large.

important limitations (Hahn and Hausman 2003; Cruz and Moreira, 2005; Angrist and Pischke 2008, p.215). In particular, the size and power of the tests are not only sensitive to the explanatory power of the instruments, but they are also sensitive to other parameters such as the degree of endogeneity of the explanatory variable (Hall, Rudebusch, and Wilcox 1996). For example, Cruz and Moreira (2005) can obtain meaningful inferences with F-statistics as low as 1.93. Moreover, the critical values in Stock and Yogo (2005) are obtained for TSLS for a continous endogenous variable while we use NTSLS which has some optimality properties for discrete endogenous variables as indicated above. In section 5 we report the results of a Montecarlo analysis with our sample which indicates that our point estimates are conservative, and the confidence intervals have the right coverage.

4.3 Results

Estimates of the effects of breastfeeding on children's cognitive and noncognitive development are shown in Table 4, at ages 3 (top panel), 5 (middle panel) and 7 (bottom panel). As discussed in section 2, the measures of cognitive development at age 3 are based on the British Ability Scales (BAS) and the Bracken School Readiness test. We find large and significant effects of breastfeeding on both measures, 0.42 of a standard deviation for BAS, and 0.47 of a standard deviation for the Bracken test. The third column of the upper panel shows the effects on noncogitive development, as measured using the Strengths and Difficulties questionnaire described in section 2. The effects on this are also large, at 0.58 of a standard deviation. The final column shows the mother-to-child relationship, as measured according to the Pianta scale, a self-reported instrument completed by mothers that assesses her perception of their relationships with their sons and daughters. We detect no effect of breastfeeding on this outcome.

[TABLE 4 HERE]

We next consider the effects at age 5, shown in the middle panel of Table 4. Whilst the point estimates for the three cognitive outcomes (BAS pictures, BAS vocabulary, BAS patterns) are still large, they are less precisely estimated and not statistically distinguishable from zero. Interestingly, the point estimate for the noncognitive outcome in column (4) has dropped considerably to 0.15 of a standard deviation. By age 7 (bottom panel), the effects on development have petered out and we detect no evidence of effects at this age.

In Table 5 we show corresponding estimates using standard TSLS (with a linear first stage) for comparison. The first thing to note is that the standard errors of TSLS are more than twice the ones of NTSLS. The gain in precision of NTSLS was anticipated in section 4.1 when we discussed the optimality properties of NTSLS (and it will also be clear from a montecarlo exercise that will follow). The point estimates of NTSLS are roughly similar to those of TSLS, especially in the top panel (given that TSLS exhibit very large standard errors, one should also expect some substantial differences for some outcome variables). This means that the identification of the parameter of interest is not driven by the non-linearities embedded in the first stage Probit model, but the variation embedded in *Elapsed*.

[TABLE 5 HERE]

Finally, note that the OLS estimates, shown in Table 6, are in most cases statistically significant, and always considerably lower than the IV estimates. This finding of the local average treatment effect (LATE) parameter estimated by IV exceeding the OLS parameter is common (and is in line with, for instance, the returns to education literature). We provide further discussion of this in section 6. The fact that the NTSLS estimates are substaintially different from the OLS inform us that a weak instrument problem is not prevalent here.

[TABLE 6 HERE]

5. Montecarlo simulation

The NTSLS is relatively new in empirical practice and standard IV specification tests (i.e. weak instrument tests) do not apply to it. Despite the optimality properties of the NTSLS as referred to above, there is not much evidence on its finite sample properties. In this section we report the results of a montecarlo exercise to learn about the finite sample properties of the NTSLS. To keep results informative for our study, we use exactly the same sample and variables ($Elapsed_i$, X_i) used in the previous section. In particular, this allows us to explore whether our instrument is strong enough to provide us with confidence intervals or the right coverage and unbiased point estimates, and if not, in what direction the bias goes.

The montecarlo design keeps the sample of (N=4730) individuals and X_i and $Elapsed_i$ variables fixed. In each of the 500 montecarlo replicas, the following steps are taken:[18]

Step 1: $\{\tilde{\varepsilon}_i, \tilde{\vartheta}_i\}_{i=1}^{N}$ draws of the bivariate normal distribution with variances (σ_ε^2,1) correlation coefficient ρ are obtained.

Step 2: Using the values of $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2$ estimated using the first stage Probit model in section 4.2 (Table 3, column 1), we simulate *Breastfed_90*$_i$ as

$$\widetilde{Breastfed}_{90i}=1\left[\hat{\beta}_0 + P(\hat{\beta}_1, Elapsed_i) + \hat{\beta}_2 X_i + \tilde{\vartheta}_i > 0\right]$$

Step 3: Using the values of $\hat{\alpha}_0$, $\hat{\alpha}_1$, $\hat{\alpha}_2$, and $\hat{\sigma}_\varepsilon^2$, estimated using the NTSLS (second stage) in section 4.3 (Table 4, column 1, top panel), we simulate *Outcome*$_i$ as

$$\widetilde{Outcome}_i = \hat{\alpha}_0 + \hat{\alpha}_1 \widetilde{Breastfed}\_90_i + \hat{\alpha}_2 X_i + \tilde{\varepsilon}_i$$

Step 4: Using the 4730 observations of $Elapsed_i$, X_i , $\widetilde{Outcome}_i$, and $\widetilde{\text{Breastfed\_90}}_i$, equation (1) is estimated using NTSLS and TSLS to obtain $\bar{\bar{\alpha}}_1^{NTSLS}$and $\bar{\bar{\alpha}}_1^{TSLS}$. The variable $Covered^{NTSLS}$ ($Covered^{TSLS}$) takes value 1 if $\bar{\bar{\alpha}}_1^{NTSLS}$ ($\bar{\bar{\alpha}}_1^{TSLS}$) is not statistically different from $\hat{\alpha}_1$ at 95% confidence. In other words, $Covered^{NTSLS}$ ($Covered^{TSLS}$) takes value 1 if the 95% confidence interval for $\bar{\bar{\alpha}}_1^{NTSLS}$ ($\bar{\bar{\alpha}}_1^{TSLS}$) includes the true parameter: $\hat{\alpha}_1$. The values of $\bar{\bar{\alpha}}_1^{NTSLS}$,$\bar{\bar{\alpha}}_1^{TSLS}$, $Covered^{NTSLS}$, and $Covered^{TSLS}$ are saved. In this step, we also compute the OLS estimator of equation (1) and keep $\bar{\bar{\alpha}}_1^{OLS}$.

After the 500 montecarlo replicas, the averages of $Covered^{NTSLS}$, and $Covered^{TSLS}$ are obtained to compute the coverage of the confidence intervals of the NTLS and TSLS estimators, as well as the mean and variance of the 500 values obtained of $\bar{\bar{\alpha}}_1^{NTSLS}$,$\bar{\bar{\alpha}}_1^{TSLS}$to compute the bias and mean square error of both estimators. The results are contingent on the value of ρ used in Step 1. We try different values of ρ and choose the one (ρ =-0.24) for which the average of $\bar{\bar{\alpha}}_1^{OLS}$across the 500 montecarlo replicas is close to the OLS estimator of α1 in equation (1). So, the average correlation between $\widetilde{Breastfed}\_90_i$ and $\widetilde{Outcome}_i$ (simulated data) should be the same as in the actual data. Table 7 shows that the main moments of the simulated variables are very similar to the moments of the actual data.

[18] For ease of notation, we omit the montecarlo replica subscript.

Table 8 reports the mean square error, bias, variance and confidence interval coverage of $\bar{\bar{\alpha}}_1^{NTSLS}$ and $\bar{\bar{\alpha}}_1^{TSLS}$ when we choose British Ability Scale (at age 3) $Outcome_i$. The Mean Square Error of the NTSLS is about 10% of the Mean Square Error of TSLS. This very large difference in the Mean Square Error comes from a very large difference in the variance. As our discussion in section 4.1 indicated, NTSLS is more efficient than TSLS, and this translates into a much lower variance as well as lower estimated standard errors as it is clear when we compare Tables 4 and 5. The NTSLS seems to be more biased than TSLS (the bias of NTSLS is -0.089 =0.3340-0.423, while the bias of TSLS is -0.0187= 0.4043-0.423) at but this difference is bias in not enough to compensate for the much larger variance of TSLS. Consequently the Mean Square Error, the standard criterion to chose among estimators, favours NTSL over TSLS by large. More importantly, even if NTSLS is biased, it is biased towards zero, which means that our NTSLS point estimates are conservative.

[TABLES 7, 8 HERE]

It is known that the coverage of confidence intervals can be affected by weak instrument problems. In our case, the montecarlo simulation that the confidence interval obtained using NTSLS attains almost the nominal level of coverage (0.95). In other words, the 95% NTSLS confidence intervals include the true parameter almost 95% of the times (precisely, 93.8%). So, this does not seem to be a concern for our case.

6. Robustness

A possible criticism of our identification strategy is that if it is the case that less experienced staff was more likely to work over the weekend, and mothers had more traumatic deliveries, this could affect the mother to child relationship, and consequently their future interaction and investments. However, Table 4 shows that the point estimate of the Mother-to-Child relationship score is quite small (0.08) and not statistically different from zero.

7. Discussion

In this paper, we have estimated fairly sizeable impacts of breastfeeding on children's cognitive and non-cognitive development at age 3. What we have identified are local average treatment effects (LATE), relating to the set of individuals affected by the instrument ("compliers" – see Angrist, 2004). The IV estimates are larger than the OLS ones, which likely reflects the fact that the marginal individuals affected by the instrument are those who breastfeed only if they receive adequate support while at hospital, and who are unlikely to make tradeoffs so as to receive support from other sources as the child grows. So it is plausible to expect that children of these mothers will not receive as many investments in the future compared to children of non-marginal mothers. Thus breastfeeding represents a very important input into these children. In line with this, the observed effects are large, in the region of half a standard deviation for both cognitive and noncognitive development. However, despite being large, they are in line with estimates obtained by Kramer et al (2008), who find effects at age 6.5 years in the regions of 0.5 of a standard deviation (sd) for verbal IQ, 0.2 of a sd for performance IQ, and 0.5 if a sd for full-scale IQ. Their study involved randomising a breastfeeding promotion intervention that increased hospital support in Belarus, so their compliers are mothers who breastfeed only if adequate support is obtained, and who thus share features with ours.

8. Conclusion

This paper has estimated the effects of breastfeeding (for at least the first three months of life) on children's cognitive and noncognitive development up to the age of 7. It has used variation in timing of birth, which we argue is random (with the exception of planned Caesarean sections, which are excluded) as a predictor of breastfeeding. The reason for this is that births that occur at or just before the weekend benefit from fewer non-essential services, particularly breastfeeding support, and this affects some mothers decision and/or ability to breastfeed - in particular, mothers with lower access to support and information regarding child development. This marginal group represents those less likely to make other investments into child development in the future, so those for whom breastfeeding may represent a particularly important input. Using UK data, we find that breastfeeding has large effects, in the region of half a standard deviation, on children's cognitive and noncognitive development at age 3. There is some evidence that these effects are sustained until age 5 though they are less precisely estimated, but by age 7 they have petered out.

References

Almond, D. and J. Currie (2011). "Human capital development before age five", Chapter 15 in "Handbook of Labor Economics", Volume 4: Orley Ashenfelter and David Card, eds.

Angrist J. (2004) Treatment Effect Heterogeneity in Theory and Practice. *The Economic Journal,* 114(494): C52-C83

Angrist, Joshua D., and Jörn-Steffen Pischke. (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Backstrom, C., E. Wahn, A. Ekstrom (2010) "Two sides of breastfeeding support: experiences of women and midwives", International Breastfeeding Journal 5:20.

Banerjee, A., E. Duflo, G. Postel-Vinay, T. Watts (2009). "Long run health impacts of income shocks: wine and phylloxera in 19^{th} century France". *Review of Economics and Statistics*, forthcoming.

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

Belfield, C. and I. Kelly (2010). "The benefits of breastfeeding across the early years of childhood". NBER Working Paper 16496.

Black, S., P. Devereux and K. Salvanes (2007). "From the cradle to the labor market? The effect of birth weight on adult outcomes", *Quarterly Journal of Economics*, 122(1): 409-439.

Bolling K, C. Grant, B. Hamlyn, A. Thornton (2005). Infant Feeding Survey 2005. The Information Centre, 2007.

Carneiro, Pedro Manuel, and James J. Heckman. (2003). "Human Capital Policy." SSRN eLibrary (July). http://papers.ssrn.com/sol3/papers.cfm?abstract_id=434544.

Case, Anne, Paxson, Christina. (2008a). "Height, health, and cognitive function in old age". *American Economic Review: Papers & Proceedings* 98 (2), 463–467.

Case, Anne, Paxson, Christina. (2008b). “Stature and status: height, ability, and labor market outcomes”. *Journal of Political Economy* 116 (3), 499–532.

Case, Anne, Paxson, Christina. (2010a). “Causes and consequences of early life health”. Working Paper 15637, National Bureau of Economic Research, January.

Case, Anne, Paxson, Christina. (2010b). “The long reach of childhood health and circumstance: evidence from the Whitehall II study”. Working Paper 15640, National Bureau of Economic Research, January.

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.” *The Journal of Human Resources* 40 (2) (April 1): 393–410.

Cunha, F., J. Heckman and S. Schennach (2010). “Estimating the technology of cognitive and noncognitive skill formation”, NBER Working Paper No. 15664.

Currie, J. and R. Hyson (1999). “Is the impact of shocks cushioned by socioeconomic status? The case of low birth weight”, *American Economic Review*, 89(2):245–250.

Currie, J. and E. Moretti (2007). “Biology as destiny? Short and long-run determinants of intergenerational transmission of birth weight”, *Journal of Labor Economics*, 25(2): 231-264.

Currie, J. and D. Almond (2011). “Human Capital Development before Age Five,” In (D. Card and O.Ashenfelter, eds.) *Handbook of Labor Economics*, Elsevier.

Cutler, D., G. Miller and D. Norton (2007). “Evidence on early life income and late-life health from America’s Dust Bowl era”, *Proceedings of the National Academy of Sciences,* 104 (33): 13244-13249.

Del Bono, E. and B. Rabe (2011). “Estimating the causal effects of breastfeeding on early child outcomes”, Institute for Social and Economic Research, mimeo.

Der, G., G. Batty and I. Deary (2006). “Effect of breastfeeding on intelligence in children: prospective study, sibling pairs analysis, and meta-analysis”, *British Medical Journal*, 333(7575).

DiGirolamo, A., L. Grummer-Strawn, S. Fein (2008). “Effect of maternity-care practices on breastfeeding”, *Pediatrics* 122:s43-9.

Evenhouse, E. and S. Reilly (2005). “Improved estimates of the benefits of breastfeeding using sibling comparisons to reduce selection bias”, Health Services Research 40(6): 1781-1802.

Frank, R. and E. Meara (2009). “The effect of maternal depression and substance abuse on child human capital development”. Working Paper 15314, National Bureau of Economic Research, September.

Hall, Alastair R., Glenn D. Rudebusch, and David W. Wilcox. 1996. “Judging Instrument Relevance in Instrumental Variables Estimation.” *International Economic Review* 37 (2) (May 1): 283–298. doi:10.2307/2527324.

Hamlyn B, Brooker S, Oleinikova K,Wands S. (2002). Infant feeding 2000. TSO: London.

Hahn, Jinyong, and Jerry Hausman. (2003). “Weak Instruments: Diagnosis and Cures in Empirical Econometrics.” *American Economic Review* 93 (2) (May): 118–125. doi:10.1257/000282803321946912.

Kramer, M S, B Chalmers, E D Hodnett, Z Sevkovskaya, I Dzikovich, S Shapiro, J P Collet, et al. (2001). “Promotion of Breastfeeding Intervention Trial (PROBIT): a Randomized Trial in the Republic of Belarus.” *JAMA: The Journal of the American Medical Association* 285 (4) (January 24): 413–420.

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) (March): e435-440. doi:10.1542/peds.2007-1248.

León-Cava N, Lutter C, Ross J, Martin L. (2002). "Quantifying the benefits of breastfeeding: a summary of the evidence". Washington DC: Pan American Health Organization, 2002.

McAllister H, S. Bradshaw, G. Ross-Adjie G (2009). "A study of in-hospital midwifery practices that affect breastfeeding outcomes", *Breastfeed Rev.* 17(3):11-5.

Merten S, J. Dratva J, U. Ackermann-Liebrich (2005). "Do baby-friendly hospitals influence breastfeeding duration on a national level?" *Pediatrics,* 116(5):e702-8.

Murray, Michael P. 2006. "Avoiding Invalid Instruments and Coping with Weak Instruments." The Journal of Economic Perspectives 20 (4) (October 1): 111–132.

Newey, W. (1990). "Semiparametric Efficiency Bounds" Journal of Applied Econometrics 5(2): 99-135.

Newey, W. (1990). "Efficient Instrumental Variables Estimation of Nonlinear Models" Econometrica 58(4): 809-837.

Quigley, M., C. Hockley, C. Carson, Y. Kelly, J. Renfrew, A. Sacker (2011). "Breastfeeding is associated with improved child cognitive development: a population-based cohort study", Journal of Pediatrics, forthcoming.

Renfrew M, D. Craig, L. Dyson, F. McCormick, S. Rice, S. King, K. Misso, E. Stenhouse, A. Williams (2009)."Breastfeeding promotion for infants in neonatal units: a systematic review and economic analysis", *Health Technology Assessment*, 13(40):1-146.

Rothstein, D. (2011). "Breastfeeding and children's early cognitive outcomes". US Bureau of Labor Statistics, mimeo.

Royer, H. (2009). "Separated at Girth: Estimating the Long-Run and Intergenerational Effects of Birthweight using Twins", *American Economic Journal: Applied Economics:* 1(1): 49-85.

Sikorski J, Renfrew MJ, Pindoria S, Wade A. "Support for breastfeeding mothers", *Cochrane Database Syst Rev.* 2002;(1):CD001141.

Staiger, Douglas, and James H. Stock. (1997). "Instrumental Variables Regression with Weak Instruments." *Econometrica* 65 (3) (May 1): 557–586. doi:10.2307/2171753.

Thompson, R. and C. Nelson (2001). "Developmental science and the media: Early brain development", *American Psychologist,* 56: 5–15.

Van der Berg, G., M. Lindeboom, F. Portrait (2006). "Economic conditions early in life and individual mortality", *American Economic Review*, 96(1): 290-302.

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. John Wiley & Sons, Ltd. http://ideas.repec.org/a/jae/japmet/v12y1997i3p281-94.html.

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

Table 1. Feeding support at hospital, by day of week of birth.

Day of week of birth	Low Educated Mothers			High educated Mothers		
	Consistent advice	Practical help	Active support and encouragement	Consistent advice	Practical help	Active support and encouragement
Sunday	0.0251	0.00654	0.00685	-0.00655	0.0149	0.00644
	[0.0214]	[0.0230]	[0.0226]	[0.0140]	[0.0139]	[0.0135]
Monday	0.0220	0.0213	0.0242	0.00697	0.0128	0.00611
	[0.0219]	[0.0235]	[0.0232]	[0.0137]	[0.0135]	[0.0132]
Wednesday	0.0140	0.0150	0.00509	0.0157	0.00871	0.00912
	[0.0213]	[0.0230]	[0.0225]	[0.0136]	[0.0135]	[0.0132]
Thursday	0.0151	0.0103	0.00313	0.000959	0.00366	-0.00517
	[0.0218]	[0.0235]	[0.0230]	[0.0136]	[0.0134]	[0.0131]
Friday	-0.0750***	-0.0636***	-0.0625***	-0.00151	0.00844	0.00391
	[0.0215]	[0.0231]	[0.0227]	[0.0135]	[0.0134]	[0.0131]
Saturday	-0.00668	-0.0456**	-0.0286	0.0145	0.0196	0.0125
	[0.0215]	[0.0231]	[0.0227]	[0.0137]	[0.0135]	[0.0132]
Sample	4,914	4,772	4,813	12,946	12,580	12,820

Notes to table: Source Maternity Users Survey 2007. The low educated sample includes those who left full-time education at age 16 years or less; the high educated sample includes those who left full-time education at age 17 or above. Note that Caesarean sections (both emergency and planned) are excluded throughout. Sample covers England, and relates to 148 NHS trusts. Coefficients reported from an OLS regression of the variable listed in the left hand column, on day of the week of birth (Tuesday omitted category).

Table 2. Sample Balance, Low Educated

	Fri-Sun	Mon-Thurs	t-stat diff
Labour induced? (Y/N)	0.304	0.309	-0.331
Type Delivery:			
Forceps	0.040	0.037	0.408
Vacuum	0.064	0.065	-0.103
Emerg C-section			
Other	0.009	0.009	-0.158
Pain relief:			
Gas and air	0.808	0.798	0.786
Pethidine	0.370	0.367	0.216
Epidural	0.201	0.198	0.221
General anaesthetic	0.003	0.002	1.078
TENS	0.084	0.079	0.522
Complication:			
None	0.748	0.751	-0.212
Breech	0.003	0.004	-0.927
Other abnormal	0.020	0.021	-0.262
V long lab	0.051	0.047	0.503
V rapid lab	0.031	0.025	1.139
Foetal distress (heart)	0.080	0.075	0.621
Foetal distress (meconium)	0.035	0.043	-1.460
Female	0.523	0.496	1.766
Birth weight (kg)	3.373	3.364	0.634
Premature	0.046	0.042	0.575
First ante-natal was before:			
0-11 weeks	0.422	0.402	1.296
12-13 weeks	0.326	0.345	-1.308
≥ 14 weeks	0.177	0.188	-0.874
Mother's age	26.815	26.837	-0.118
Live in house	0.843	0.850	-0.624
# rooms	5.066	5.094	-0.669
Own outright	0.029	0.027	0.399
Rent from Local Authority	0.274	0.275	-0.069
Rent from Housing Association	0.094	0.099	-0.603
Rent privately	0.097	0.082	1.677
Live with parents	0.056	0.050	0.871
Live rent free	0.012	0.017	-1.218
DV open fire	0.036	0.035	0.101
DV gas/elec fire	0.307	0.305	0.159
DV central heating	0.879	0.904	-2.531
DV no heating	0.012	0.007	1.716
Own computer	0.421	0.412	0.592
Own/access to car	0.770	0.747	1.689
In paid employment during pregnancy	0.506	0.533	-1.740
Self-employed during pregnancy	0.027	0.019	1.700
Receives child tax credit	0.138	0.145	-0.666
Receives WFTC	0.260	0.247	0.991
Receives income support	0.267	0.277	-0.702
Receives JSA	0.040	0.044	-0.726
Sample size		4242	

Notes to table: Source Millennium Cohort Study. Low educated include mothers with ≤ NVQ level 2, or those whose NVQ level is unknown but left school before 17.
Sample excludes Caesarean sections (either emergency or planned) and children that went into intensive care.

Table 3. First stage. Breastfed for 90 days

	(1)	(2)
	Probit coefficients	OLS coefficients
Elapsed	0.0105**	0.0027**
	(0.0049)	(0.0012)
Elapsed^2	-0.000134**	-3.62e-05**
	(6.82E-05)	(1.79E-05)
Elapsed^3	4.35E-07	1.22e-07*
	(2.68E-07)	(7.04E-08)
P-Value Joint	0.01	0.02
F-stat		3.203
Observations	4,241	4,241

Notes to table: Dependent variable is whether the child was breastfed for at least 90 days. Sample comprises Low educated include mothers (NVQ level 2 or less, or those whose NVQ level is unknown but left school before 17), but excludes children born through Caesarean sections (either emergency or planned) and children that went into intensive care. Data source: Millennium Cohort Study. All other controls are also included

*** p<0.01, ** p<0.05, * p<0.1.

Table 4. Non-Linear TSLS Estimates of Breastfeeding on Developmental Outcomes

	(1)	(2)	(3)	(4)	(5)
	Top Panel: 3 Years Old				
	British Ability Scales	School Readiness	Strength and Difficulties	Mother to Child Relationship	
Breastfed 90 days	0.549**	0.665***	0.492*	0.150	
	(0.231)	(0.254)	(0.269)	(0.269)	
Observations	4,241	4,029	4,158	4,547	
	Middle Panel: 5 Years Old				
	British Ability Scales - Pictures	British Ability Scales - Vocabulary	British Ability Scales - Patterns	Strength and Difficulties	Foundation Stage Profile
Breastfed 90 days	0.529*	0.536**	0.658**	0.0994	0.199
	(0.293)	(0.262)	(0.301)	(0.283)	(0.309)
Observations	4,384	4,378	4,364	4,243	3,558
	Bottom Panel: 7 Years Old				
	British Ability Scales - Maths	British Ability Scales - Words	British Ability Scales - Patterns	Strength and Difficulties	
Breastfed 90 days	0.390	-0.0255	0.580**	0.402	
	(0.298)	(0.275)	(0.286)	(0.265)	
Observations	3,702	3,651	3,685	3,639	

Notes to table: Dependent variable (standarized) is defined at the heading of each column-panel combination. Sample comprises Low educated include mothers (NVQ level 2 or less, or those whose NVQ level is unknown but left school before 17), and excludes children born through Caesarean sections (either emergency or planned) and children that went into intensive care. Data source: Millennium Cohort Study. All other controls are also included. *** p<0.01, ** p<0.05, * p<0.1.

Table 5. TSLS Estimates of Breastfeeding on Developmental Outcomes

	(1)	(2)	(3)	(4)	(5)
	Top Panel: 3 Years Old				
	British Ability Scales	School Readiness	Strength and Difficulties	Mother to Child Relationship	
Breastfed 90 days	0.454	0.420	-0.228	0.578	
	(0.671)	(0.681)	(0.699)	(0.663)	
Observations	4,241	4,029	4,158	4,547	
	Middle Panel: 5 Years Old				
	British Ability Scales - Pictures	British Ability Scales - Vocabulary	British Ability Scales - Patterns	Strength and Difficulties	Foundation Stage Profile
Breastfed 90 days	0.673	0.173	0.411	0.409	-0.254
	(0.747)	(0.643)	(0.709)	(0.679)	(0.721)
Observations	4,384	4,378	4,364	4,243	3,558
	Bottom Panel: 7 Years Old				
	British Ability Scales - Maths	British Ability Scales - Words	British Ability Scales - Patterns	Strength and Difficulties	
Breastfed 90 days	0.119	0.117	-0.280	0.382	
	(0.667)	(0.610)	(0.667)	(0.582)	
Observations	3,920	3,869	3,903	3,850	

Notes to table: Dependent variable (standarized) is defined at the heading of each column-panel combination. Sample comprises Low educated include mothers (NVQ level 2 or less, or those whose NVQ level is unknown but left school before 17), but excludes children born through Caesarean sections (either emergency or planned) and children that went into intensive care. Data source: Millennium Cohort Study. All other controls are also included. *** p<0.01, ** p<0.05, * p<0.1.

Table 6. OLS Estimates of Breastfeeding on Developmental Outcomes

	(1)	(2)	(3)	(4)	(5)
	Top Panel: 3 Years Old				
	British Ability Scales	School Readiness	Strength and Difficulties	Mother to Child Relationship	
Breastfed 90 days	0.0890***	0.0891**	0.145***	0.00364	
	(0.0326)	(0.0355)	(0.0340)	(0.0326)	
Observations	4,241	4,029	4,158	4,547	
	Middle Panel: 5 Years Old				
	British Ability Scales - Pictures	British Ability Scales - Vocabulary	British Ability Scales - Patterns	Strength and Difficulties	Foundation Stage Profile
Breastfed 90 days	0.101***	0.0669**	0.0519	0.0802**	0.0876**
	(0.0356)	(0.0331)	(0.0356)	(0.0346)	(0.0355)
Observations	4,384	4,378	4,364	4,243	3,558
	Bottom Panel: 7 Years Old				
	British Ability Scales - Maths	British Ability Scales - Words	British Ability Scales - Patterns	Strength and Difficulties	
Breastfed 90 days	0.129***	0.0737**	0.0985**	0.114***	
	(0.0379)	(0.0374)	(0.0394)	(0.0366)	
Observations	3,920	3,869	3,903	3,850	

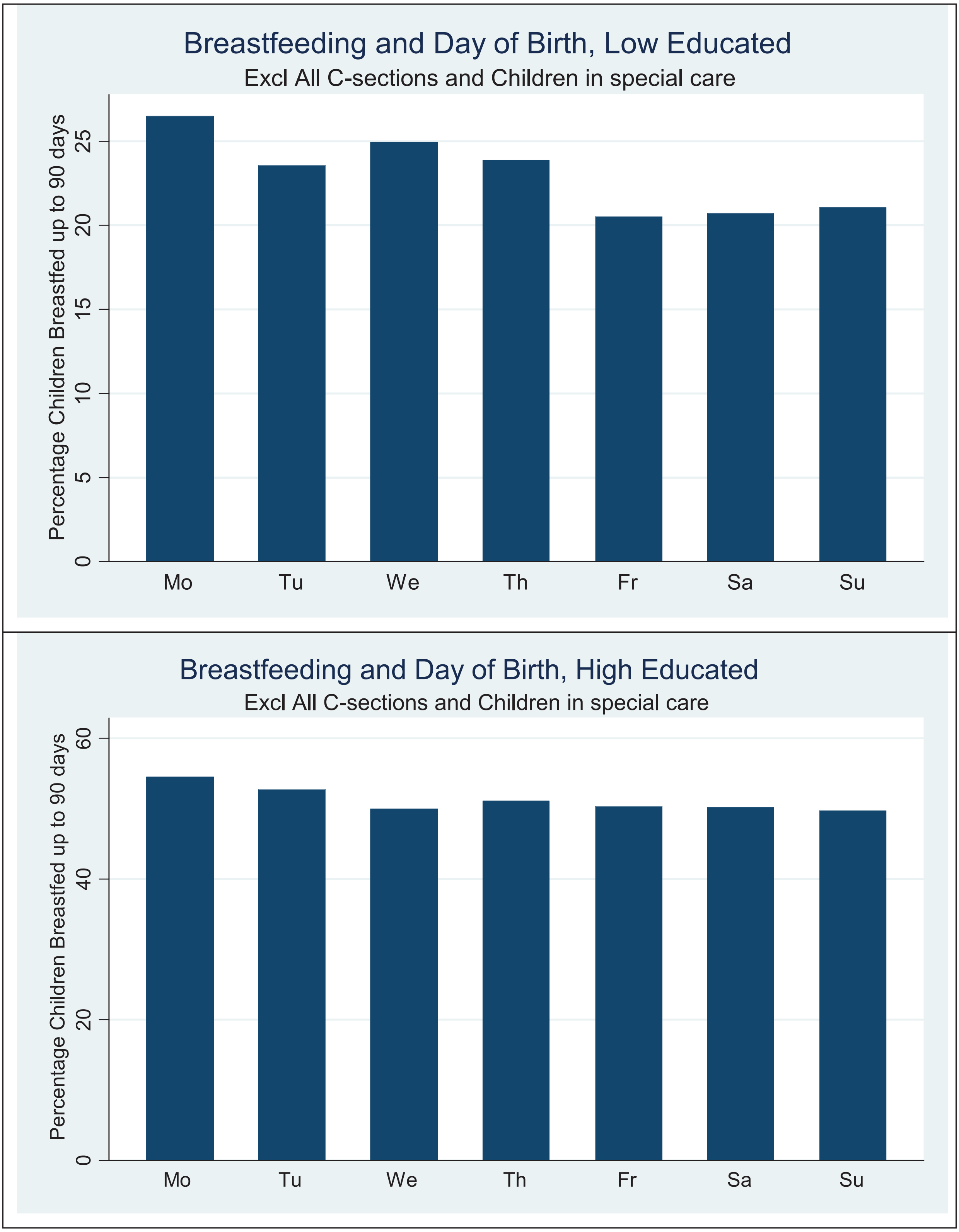
Notes to table: Dependent variable (standarized) is defined at the heading of each column-panel combination. Sample comprises Low educated include mothers (NVQ level 2 or less, or those whose NVQ level is unknown but left school before 17), but excludes children born through Caesarean sections (either emergency or planned) and children that went into intensive care. Data source: Millennium Cohort Study. All other controls are also included. *** p<0.01, ** p<0.05, * p<0.1.

Table 7. Non-Linear TSLS Estimates of Breastfeeding on Child Investments

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Spends plenty of time with child	Read to child every day	Regular Bed time	Regular Meal time	More than 3 hours TV	Home learning Environment	Home learning Environment
Breastfed 90 days	-0.0630	-0.0884	0.113	0.0905	-0.131	1.154	2.568
	(0.117)	(0.138)	(0.138)	(0.141)	(0.119)	(2.107)	(2.138)
Wave	2	2	2	2	2	3	4
Observations	4,547	4,547	4,547	4,547	4,547	4,424	3,756
Mean	0.734	0.463	0.378	0.457	0.249	24.57	21.18
Standard Deviation						7.285	7.542

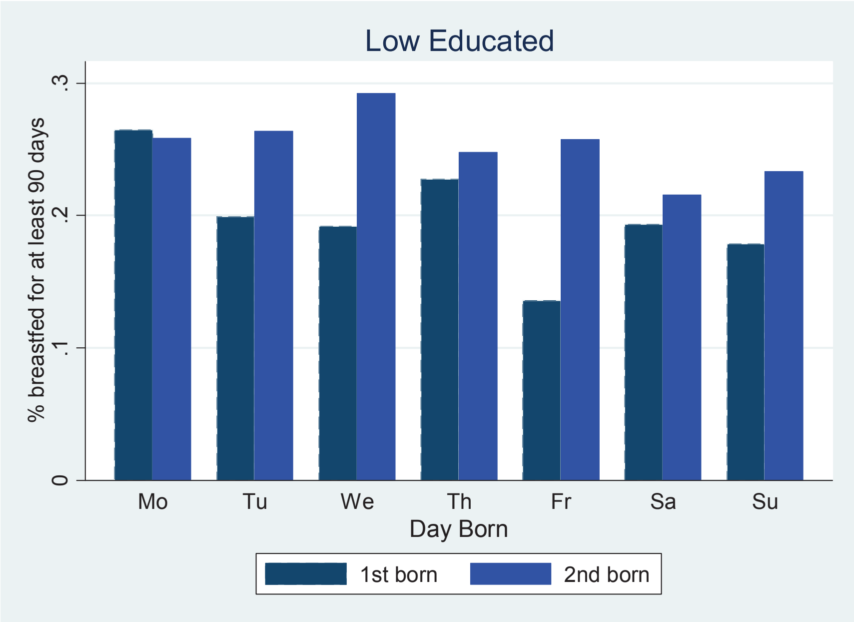
Notes to table: Dependent variable is defined at the heading of each column-panel combination. Sample comprises Low educated include mothers (NVQ level 2 or less, or those whose NVQ level is unknown but left school before 17), but excludes children born through Caesarean sections (either emergency or planned) and children that went into intensive care. Data source: Millennium Cohort Study. *** p<0.01, ** p<0.05, * p<0.1. All other controls are also included

Figure 1 Breastfeeding by day of week of birth



Notes to figure: Source Millennium Cohort Study (wave 1). Low educated include mothers with ≤ NVQ level 2, or those whose NVQ level is unknown but left school before 17. High educated include mothers with NVQ level 3 or higher. Breastfeeding refers to whether or not the baby was breastfed for at least three months.

Figure 2. Breastfeeding by day of week of birth and birth parity



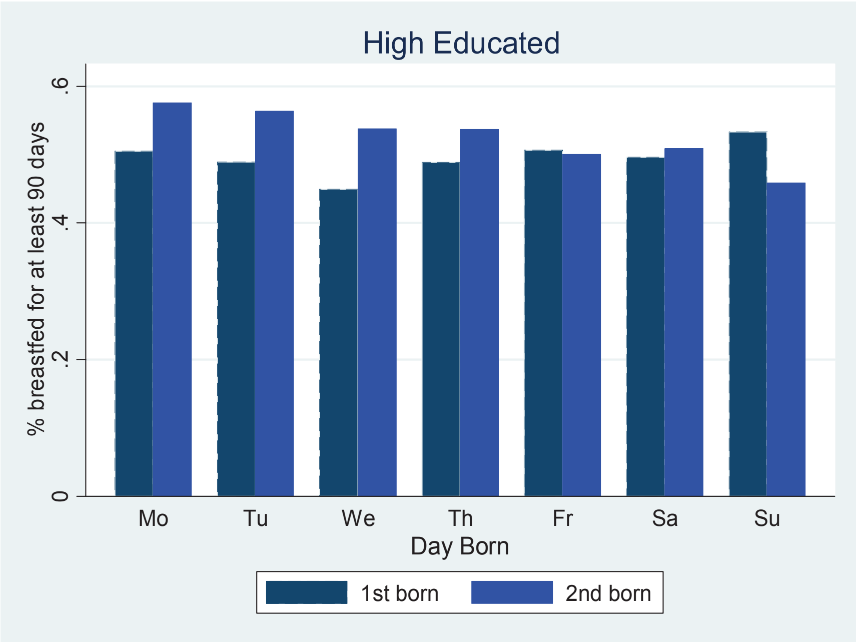
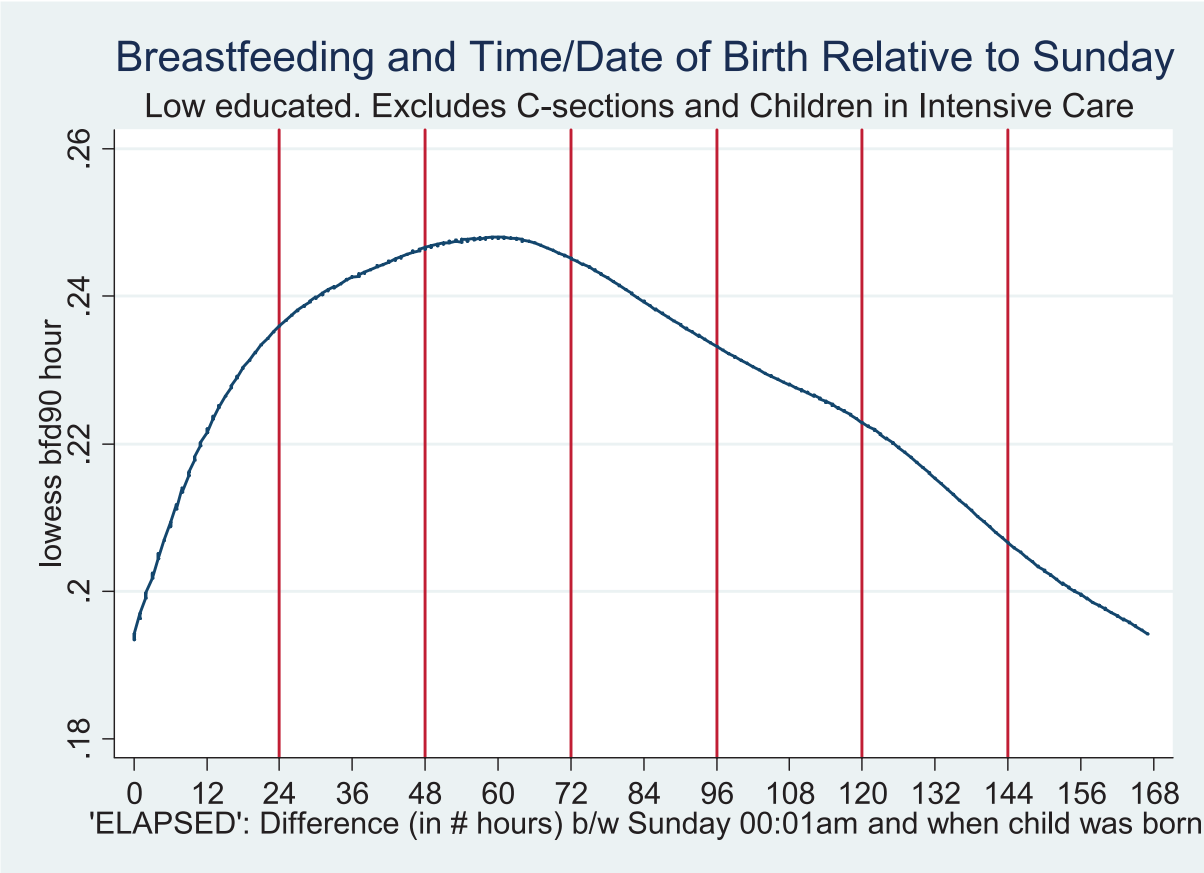


Figure 3. Breastfeeding by hour born, Low Educated



Notes to figure: Source Millennium Cohort Study. Low educated include mothers with ≤ NVQ level 2, or those whose NVQ level is unknown but left school before 17. Sample excludes planned Caesarean sections.