The University Of Sheffield.

School of Health And Related Research

Health Economics & Decision Science (HEDS) Discussion Paper Series

The Effects of Breastfeeding on Childhood BMI: A Propensity Score Matching Approach

Laura A Gibson, Monica Hernandez Alava, Michael P Kelly, Michael J Campbell

Corresponding Author: Laura Gibson
ScHARR, University of Sheffield, Regent Court, Sheffield, S1 4DA
Telephone: +44 (0)114 222 0885
Fax: +44 (0)114 272 0749
Email: l.a.gibson@sheffield.ac.uk

No. 15.04

The Effects of Breastfeeding on Childhood BMI:

A Propensity Score Matching Approach

Laura A Gibson[1]*, Monica Hernandez Alava[1], Michael P Kelly[2], Michael J Campbell[3]

1. Health Economics and Decision Science, School of Health and Related Research, University of Sheffield
2. Institute of Public Health, University of Cambridge
3. Design, Trials and Statistics, School of Health and Related Research, University of Sheffield
   *author for correspondence

Key words: Breastfeeding; Childhood BMI; Childhood obesity; Endogenous treatment; Propensity score matching

October 2015

ABSTRACT

This paper investigates the effect of breastfeeding on childhood body mass index (BMI). We use data from the Millennium Cohort Study (MCS), a nationally representative UK cohort survey, containing detailed infant feeding information, which allow us to explore the effects of a range of breastfeeding variables on the mean BMI of children breastfed for different durations and for exclusive and partial breastfeeding.

Using propensity score matching, we find statistically significant influences of breastfeeding on childhood BMI, particularly in older children and when breastfeeding is prolonged and exclusive. The effects of breastfeeding on BMI are small in magnitude but large relative to the mean BMIs of children this age. At this young age, there is not a large difference in BMI between children who are identified as obese and those who are identified as normal weight, so even a small difference in BMI could mean the difference between children being overweight or a healthy weight. We suggest that breastfeeding should be encouraged as part of wider lifestyle interventions in order to help reduce BMI as well as improve other childhood outcomes. This could be important evidence for public health bodies when creating public health guidelines and recommendations.

1. INTRODUCTION

The effects of breastfeeding on childhood obesity have been debated in an extensive but inconclusive literature (Armstrong and Reilly, 2002; Bergmann *et al*., 2003; Beyerlein *et al*., 2008; McCrory and Layte, 2012; Oddy and Sherriff, 2003; Reilly *et al*., 2005; Salsberry and Reagan, 2005; Jiang and Foster, 2012; Kramer *et al*., 2007). There is little doubt that childhood obesity has been worsening over recent years and obese children are more likely to become obese adults who will suffer from a number of co-morbidities (Serdula et al. 1993; Power et al. 1997). As a result, early life factors could play a large role in determining levels of childhood body mass index (BMI) and therefore future obesity levels in adults. This potentially has important policy implications given that early life interventions could help reduce comorbidities in later life, allowing the NHS to reallocate limited resources to alternative priority areas. Breastfeeding is known to have a number of benefits to both mothers and infants. Policies to promote breastfeeding are already well established and breastfeeding should be encouraged regardless of its effects on childhood BMI (Renfrew et al. 2007) and both breastfeeding and childhood obesity are of increasing interest to bodies such as NICE, the Department of Health and the NHS. However, if breastfeeding can be conclusively linked to a reduction in childhood BMI then efforts to encourage breastfeeding should become an increased policy priority. Breastfeeding is not expected to be the solution to the current obesity epidemic; however, if it is found to reduce childhood BMI it could be one part of a wider early life solution.

There are a number of theories suggesting the mechanisms by which breastfeeding might influence BMI. The 'early protein hypothesis' (Günther et al. 2007) suggests that formula fed infants consume higher levels of protein than breastfed infants inducing hormone responses which cause high levels of insulin and lead to weight gain. The 'growth-accelerating theory' (Kramer et al. 2004) suggests that formula fed children experience

accelerated growth during infancy which leads to obesity in later life (Koletzko et al. 2009). The 'self-regulation theory' (Li et al. 2010) suggests that breastfed infants learn, at an early age, to stop feeding once satisfied, whilst bottle fed infants are often encouraged to finish any milk they are given despite how much they might need. This self-regulation is thought to persist into childhood and thus prevent overeating and unnecessary weight gain. In this study, we aim to identify any causal effect of breastfeeding on childhood BMI rather than to identify the reasons that this potential relationship might occur.

For ethical reasons, randomised controlled trials (RCTs) cannot be used to randomise breastfeeding behaviour. Additionally, RCTs might influence the normal behaviour of mothers to differ from how they would behave in the absence of a trial (Duflo *et al*., 2007). For these reasons, RCTs are not an appropriate way to determine the causal effect of breastfeeding and so observational data is an alternative (Iacovou & Sevilla-Sanz 2010). However, using observational data leads to the potential for selection bias. Selection bias occurs when the characteristics that have an independent effect on the outcome (BMI) differ between the groups of children who were breastfed and those who were not. If those characteristics can be observed then we have selection on observables, if they are not observed we have selection on unobservables.

Existing studies have generally used regression models, most commonly a linear or logistic regression, depending on the measurement of childhood obesity in their data (Armstrong and Reilly, 2002; Bergmann *et al*., 2003; Bogen *et al*., 2004; Burdette and Whitaker, 2007; Burke *et al*., 2005; Gillman *et al*., 2001; Grummer-Strawn and Mei, 2004; Hediger *et al*., 2001; Liese *et al*., 2001; Mayer-Davis *et al*., 2006; Oddy and Sherriff, 2003; Reilly *et al*., 2005; Salsberry and Reagan, 2005; von Kries *et al*., 1999). However, these regression models make a number of assumptions which have been criticised within the literature. For example, linear relationships are often assumed between breastfeeding and BMI which may not be

appropriate (Jiang & Foster, 2012; Beyerlein *et al*., 2008). Propensity score matching (PSM) is one potential technique which can be used to estimate the causal effects of breastfeeding on BMI. The same as a regression, it deals with selection on observables but it avoids imposing a linear or other specific type of relationship between breastfeeding and BMI. We investigate a range of breastfeeding 'treatments' in order to determine the difference in BMI between the treated and non-treated groups[1].

This study contributes to existing literature in several ways. We use a large nationally representative UK dataset, the Millennium Cohort Study (MCS), in order to produce conclusions which are meaningful at a population level. The data contains detailed information on breastfeeding and other infant feeding behaviours as well as other early life characteristics, including childhood BMI at each period of observation. It also over-represents children from disadvantaged backgrounds, which are less likely to be breastfed. To our knowledge, this is the first paper to use propensity score matching in this setting, to identify the causal influences of breastfeeding on childhood BMI accounting for a number of confounding factors which influence both a mother's propensity to breastfeed and her child's BMI. We also check the validity of the assumption of selection on observables by estimating an endogenous treatment regression model.

The remainder of this paper is structured as follows. Section 2 describes the methods and the data used, Section 3 presents the empirical results and Section 4 discusses the key findings and provides concluding remarks.

[1] By treatment we mean that a child has received a certain duration of breastfeeding. These 'breastfeeding treatments' are discussed later.

2. METHODS

2.1 Data

The Millennium Cohort Study (MCS) contains a rich set of information for a sample of 19,517 children born in and around the year 2000. Cohort members were recruited using child benefit records which, at the time, was a universal benefit, keeping sample bias to a minimum. The cohort members' carers were interviewed when the infant was approximately nine months old and detailed information on breastfeeding and other infant feeding behaviours were recorded. The same carers and children have since been interviewed when the infants were three, five and seven years old[2] (see Hansen, 2012 for a guide to these datasets). During each of these subsequent interviews, data on height and weight of the children were collected, amongst other adiposity measures, allowing BMI to be calculated. Only observations for which the cohort member's natural mother is the main carer are used in this analysis due to a lack of information and accuracy on breastfeeding variables from other carers. The data also contains detailed information on a wide variety of socioeconomic and demographic variables allowing a wide range of potential confounding factors to be accounted for.

Outcome Variable

The dependent variable analysed in this study is childhood BMI measured at ages three, five and seven years. The MCS contains data on children's height and weight from which BMI is calculated;

[2] There is a further wave of the MCS which is now available and provides data on the cohort at 11 years of age but this most recent wave is not included in this study.

$$\text{BMI} = \frac{\text{weight(kg)}}{\text{height(m)}^2}. \quad (1)$$

Summary statistics for BMI in children from the MCS sample used in this study are displayed in Table 1 along with the proportion of children considered to be overweight and obese. These summary statistics show a dip in BMI when the children are five years old illustrating the presence of the adiposity rebound. Age and gender specific overweight and obesity are defined using the international obesity task force (IOTF) BMI thresholds (Cole et al. 2000)[3]. The proportion of obese children in the sample increased with age, whereas the proportion of overweight children decreased as the children got older.

Table 1: Summary Statistics for Adiposity Variables

Variable	**3 Years**	**5 Years**	**7 Years**
BMI (kg/m^2) mean	16.78	16.31	16.60
(standard deviation)	(1.561)	(1.679)	(2.224)
BMI (kg/m^2) median	16.70	16.08	16.16
Overweight* (%)	23.34	21.03	20.16
Obesity* (%)	4.98	5.16	5.39
N	11,200	11,744	10,707

Source: Millennium Cohort Study. Notes: Mean with standard deviation in parentheses. *Overweight and obesity are defined using the International Obesity Task Force (IOTF) thresholds which vary by gender and age.

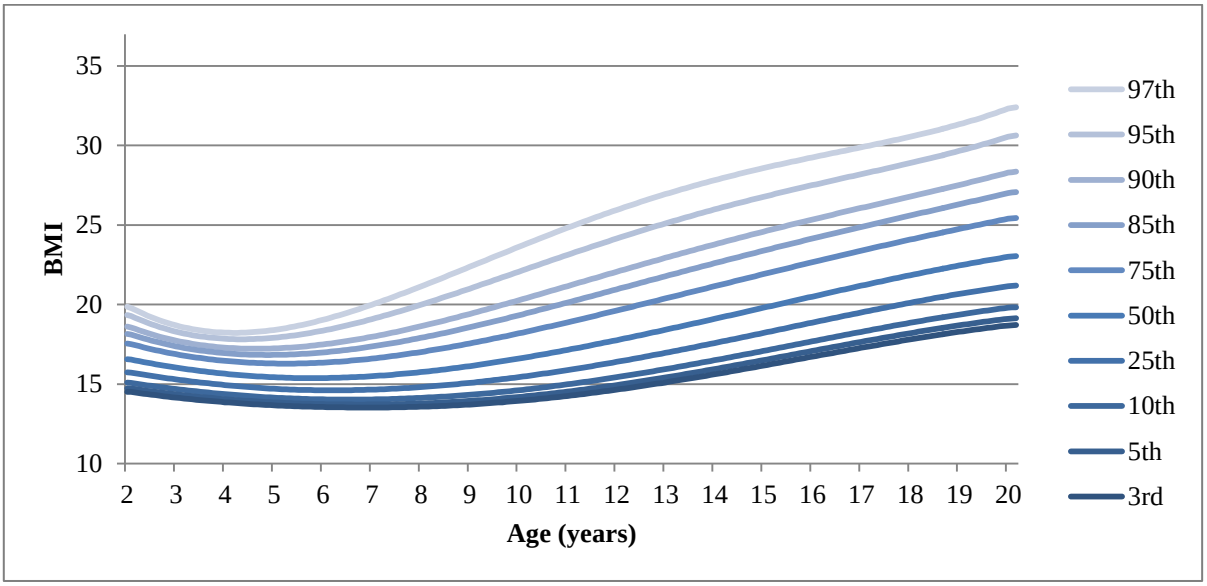
Although childhood BMI is defined using the same calculation as in adults, the distribution of BMI fluctuates during childhood and it is not possible to use the same thresholds of BMI to define obesity and overweight as it is in adults. The classifications of childhood obesity and overweight are more complex; there are many different definitions, most of which vary by age and gender. These fluctuations depend on age and gender making it impossible to classify all children over a single threshold BMI as overweight or obese. The adiposity rebound, a term established by Rolland-Cachera *et al*. (1984), occurs in children around the

[3] These thresholds are calculated for males and females every six months throughout childhood until the age of 18 and displayed in (Cole et al. 2000).

age of five years when they begin to experience an increase in BMI, after a drop in BMI during early childhood. After the adiposity rebound there is a steady increase in average BMI throughout childhood and adolescence until adult definitions can be used.

Figure 1 illustrates how BMI fluctuates during childhood in male children[4]. It shows the average BMI throughout childhood for boys on different percentiles of the BMI distribution. The figure shows that the distribution of BMI throughout childhood does not follow a normal or symmetric distribution. Those on the 97th percentile were much further from the median than those on the 3rd percentile showing just how far an 'obese' child was from the average BMI. The data from the MCS give a similar pattern, although the medians are slightly higher from the MCS, probably because the MCS over-represents children from disadvantaged backgrounds.

Figure 1: CDC & NCHS (2001) chart showing US childhood BMI percentiles



Source: Centre for Disease Control and Prevention, National Centre for Health Statistics, US.

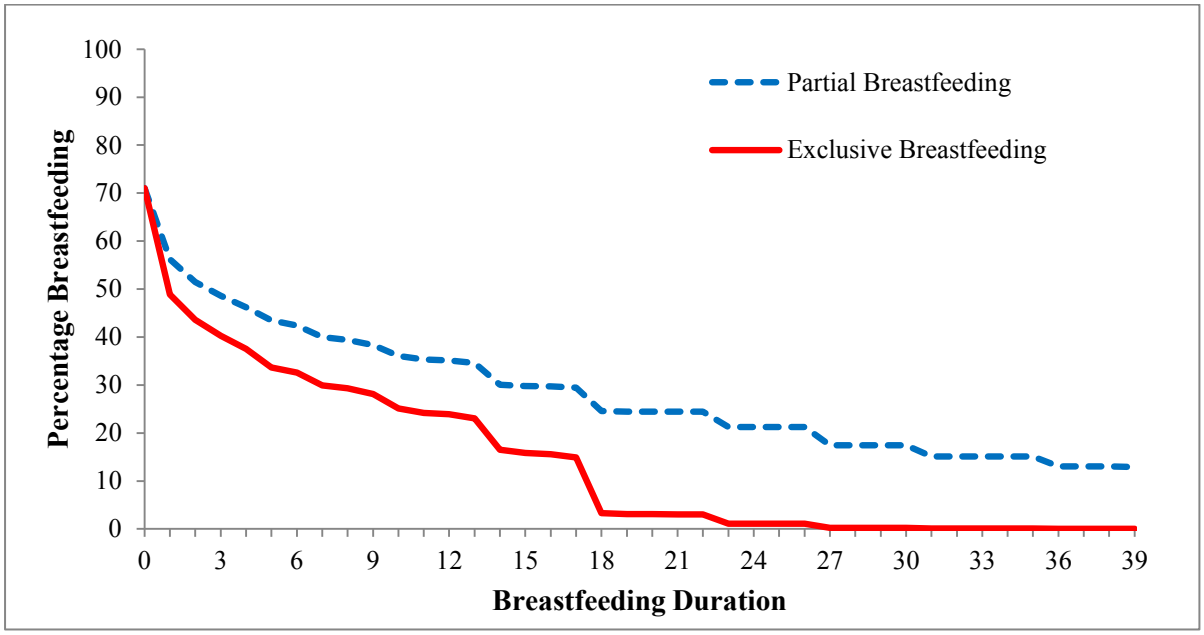
[4] The fluctuations of BMI throughout childhood are very similar in girls. These charts are from the Centers for Disease Control and Provention (CDC) & the National Center for Health Statistics (NCHS) (2001) giving data on US children between the ages of two and twenty years during 2001. The US and UK follow similar patterns and these US percentiles are often applied to UK data.

A healthy childhood BMI differs with age so comparing children of different ages could give misleading results. As children get older, the standard deviation of the BMI distribution widens and the extreme upper percentiles, representing the most obese children, move further away from the median BMI. This suggests that early childhood could be the most effective time to intervene. If a child is obese or overweight in later childhood, their BMI is further away from a 'normal' BMI suggesting that obesity in older children is more problematic, could be more difficult to reverse and might also be more likely to persist into adulthood. A universal BMI threshold to define childhood obesity or overweight would fail to identify obesity in younger children and inappropriately identify many older children as obese.

Treatment Variables

In this study, we explore a range of breastfeeding treatments based on duration and exclusivity. Breastfeeding was initiated in 71% of cohort members. Figure 2 shows the percentages of cohort members still being breastfed, exclusively and partially, by duration. At four weeks, less than 50% of cohort members were partially breastfed and less than 40% were exclusively breastfed. By sixteen weeks, these numbers drop to 30% and 16%, respectively. This sharp drop is most likely due to the World Health Organisations recommendations at the time, that weaning should start at sixteen weeks.

Figure 2: Percentage of Cohort Members Still Breastfeeding by Durations



Source: Millennium Cohort Study, UK data archive. Notes 17,385 observations.

This paper investigates a number of binary breastfeeding 'treatments' on childhood BMI at different ages. Firstly, a binary variable indicating whether or not breastfeeding was initiated is investigated. Next, two variables indicating a minimum of four and sixteen weeks of partial breastfeeding, respectively, are used. Similarly, two variables indicating a minimum of four and sixteen weeks of exclusive breastfeeding are investigated. For each of these binary variables, infants satisfying the required criteria were considered as 'treated'. They were then matched using propensity score matching with those who were 'untreated'. Untreated infants are those who were never breastfed, allowing a comparison between two groups; observations which are neither 'treated' nor 'untreated' are removed from the analysis. This means that the control groups are consistent for all binary treatments[5]. Using these binary breastfeeding variables allows a range of breastfeeding behaviours to be

[5] The results are robust to different definitions of breastfeeding treatments, including binary variables for which every observation is either treated or untreated.

investigated by identifying both length and exclusivity of breastfeeding without imposing a parametric relationship and so allows discontinuities to be investigated.

Control Variables

The analysis also includes a number of variables which are considered to confound the relationship between breastfeeding behaviours and childhood BMI. Variables recorded as close to the time of birth as possible are used because subsequent characteristics cannot causally influence breastfeeding behaviours.

These variables include socioeconomic variables; these are high and low maternal education, high and low socioeconomic status and home ownership/tenancy. We also include demographic variables, gender and ethnicity and parental variables, living with both natural parents, maternal marital status, maternal obesity, mother in care as a child and maternal longstanding illness. Pregnancy and birth variables are also included; these are whether a pregnancy was planned, maternal age at birth, maternal smoking during each trimester of pregnancy, alcohol consumption during pregnancy, birth weight, prematurity and the logged length of hospital stay.

We exclude some observations from our analysis. In the second wave, 692 families (699 children) entered the MCS because they were not identified in the initial wave. These observations are removed due to a lack of information on breastfeeding and early life variables. In accordance with Oddy & Sherriff (2003) and Burke *et al*. (2005), we remove children from multiple births due to the different breastfeeding experiences they are likely to have had and the potential influences that being from a multiple birth could have on BMI. We also exclude children who had a birth weight less than 2.5kg, those who remained in hospital immediately after birth for over fourteen days and those with a gestational period less than 196 days who are considered to be 'extremely preterm' by WHO (2012).

Observables are also removed in accordance with the WHO recommendations for biologically implausible values (BIVs); these include childhood height, weight and BMI as well as maternal height, weight and BMI[6]. We also removed observations with missing values and assume that missing data are missing at random. The number of observations excluded from the sample in each wave of the data can be found in Table 4 in the appendix.

Due to missing data and attrition in the MCS, suitable data was available for a sample of 11,200, 11,744 and 10,707 children at ages three, five and seven, respectively[7]. More details on the MCS including information on response rates can be found in a report by Plewis (2007). Further discussion of these independent variables as well as their summary statistics (Table 5) are available in the appendix.

2.2 Econometric Analysis: Propensity Score Matching (PSM)

The aim of PSM in this context is to emulate a randomised controlled trial (RCT) in a setting where, due to ethnical reasons, randomisation of breastfeeding treatments is not possible. By using PSM we create treatment and control groups, similar to those in an RCT. Observations are then matched to observations with similar characteristics in the other group. This is done using a propensity score to identify treated and non-treated observations which are similar in observable characteristics. Due to the semi-parametric nature of this technique, there is no functional form imposed on the relationship between the outcome and the treatment or any of the covariates. Standard regression techniques impose a functional form on these relationships which, if incorrect, could bias the results.

[6] Who suggest that any plausible height must lie between -5 and +3 z-scores from the mean, any plausible weight must lie between -5 and +5 z-scores and any plausible BMI values must be between -4 and +5 z-scores. These BIVs were developed using data from the NCHS and WHO growth charts from 1977.
[7] Attrition and item-non-response is assumed to be missing at random.

Treated observations are those which have been breastfed for the required duration stated in each of the binary breastfeeding variables discussed in the previous section. The untreated observations are those which have never been breastfed. Propensity scores are estimated using probit models with a number of control variables which are thought to influence maternal breastfeeding decisions as well as the outcome, BMI. We use a nearest neighbour algorithm with a calliper to restrict the difference in propensity score between matched observations. The callipers were chosen using trial and error in order to find the best balance between bias and variance. More extensive discussions of PSM can be found in the econometric literature (Rosenbaum and Rubin, 1983; Sianesi, 2006; Caliendo and Kopeinig, 2008).

PSM imposes a number of assumptions. The conditional independence assumption states that once all control variables have been accounted for, there is no remaining confounding effect by unobservable characteristics. A common support is imposed around a range of propensity scores for which there are both treated and untreated observations. Observations which do not fall within this common support are removed from the analysis. We check for bias in the matched samples by ensuring that each of the control variables do not significantly differ in mean between the treated and untreated groups.

PSM can provide estimates for the average treatment effect on the treated (ATT), the average treatment effect on the untreated (ATU) and the average treatment effect for the population (ATE). We are most interested in the ATE because any population wide policies aiming to reduce childhood BMI through breastfeeding interventions require the expected treatment effect on a random member of the population. It is also the most comparable with those estimated by the regression models in the existing literature. However, to estimate the ATE

stronger assumptions than those for the ATT and ATU are needed[8]. The ATU which estimates the benefit for those who are not currently breastfed but whose mothers might be induced to breastfeed by an intervention might also be of interest for policy makers, as might the differences between the ATT and ATU. The ATT and ATU are not discussed here but are displayed in the appendix.

Analysis was carried out using Stata 13 and the user written *psmatch2*[9] command (Leuven & Sianesi 2012) along with the *pstest* command to perform post estimation checks.

3. RESULTS

3.1 Propensity Score Matching Results

We find that at least 80% of eligible observations lie within the common support for matching for each of the binary breastfeeding treatment variables. It is assumed that this is a sufficient number of observations to estimate the treatment effects and is higher than in similar studies (Iacovou & Sevilla-Sanz 2010). We find that each of the covariates used to estimate the propensity score have similar means in both the treated and untreated groups when matching on each of the treatments. Using t-tests and a 95% significance level, the majority of covariates are balanced between treated and untreated groups and all are balanced at a 90% significance level.

[8] These assumptions are stronger versions of the same assumptions required for the estimation of other treatment effects.

[9] The *psmatch2* command provides estimates for the ATE, ATT and ATU. However, the standard errors for these estimates are calculated with the assumption that the propensity scores are known rather than estimated and therefore provide inaccurate standard errors. In addition, the standard errors for the ATE and the ATU are estimated using the Stata *bootstrap* command which has shown to produce standard errors which are not robust. Abadie and Imbens (2009) found that the true, adjusted standard errors of the ATE were consistently lower than the standard errors which assumed that the propensity score was known. We are predominantly interested in the ATE so this should not detrimentally affect our findings; the ATE parameters estimated in this study would only be more significant if the true standard errors were known.

The probit models used to estimate the propensity scores show similar results for the samples used for each age group suggesting that attrition is not significantly changing the sample. This is expected because the same control variables from the first wave of the MCS are used in estimating BMI for children at each age. This suggests that the observations which are lost to follow up in the MCS do not considerably change the estimation of the propensity scores. The probit models estimating the propensity scores for three year olds are available in Table 6 in the appendix; probit models for children at ages five and seven years were very similar. Using link tests, we find no evidence of misspecification in these probit models.

Table 2 presents the ATEs of breastfeeding on childhood BMI estimated using PSM. The table shows the average treatment effects on BMI for different breastfeeding behaviours: they are *ever breastfed, partially breastfed for four weeks, partially breastfed for sixteen weeks, exclusively breastfed for four weeks* and *exclusively breastfed for sixteen weeks*. The appendix presents more detailed results estimating the ATE, ATT and ATU. Although breastfeeding initiation appears to reduce childhood BMI in all waves, its effect is small and statistically insignificant. Simply initiating breastfeeding is not enough to significantly reduce childhood BMI. Details of the number of neighbours and size of callipers used in each set of matching are in Table 7 in the appendix.

Table 2: Average Treatment Effects using Propensity Score Matching

	BMI (outcome equations)				
	(1)	(2)	(3)	(4)	(5)
Age 3	-0.0392	-0.0333	-0.0086	-0.0602	-0.1592**
	(0.0419)	(0.0470)	(0.0077)	(0.0421)	(0.0785)
N	9,330	7,877	6,949	7,451	5,183
Age 5	-0.0782	-0.1086**	-0.1772**	-0.1401***	-0.2031**
	(0.0456)	(0.0535)	(0.0686)	(0.0484)	(0.0824)
N	9,996	6,858	4,841	7,829	5,423
Age 7	-0.1591**	-0.1665**	-0.2416***	-0.2072***	-0.2762**
	(0.0672)	(0.0767)	(0.0761)	(0.0743)	(0.1077)
N	8,372	6,168	6,534	7,167	4,948

Source: Millennium Cohort Study. Notes: Bootstrapped standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. PSM varying by breastfeeding treatment; these binary treatments are (1) ever breastfed, (2) partially breastfed for four weeks, (3) partially breastfed for sixteen weeks, (4) exclusively breastfed for four weeks, (5) exclusively breastfed for sixteen weeks. Control group is never breastfed.

Partial breastfeeding for both four and sixteen weeks has a more noticeable effect on BMI than simply initiating breastfeeding. However, these effects are still only significant in five and seven year old children. The effects of partial breastfeeding get larger and increasingly significant as the age of the children increases.

Exclusive breastfeeding generally has a larger effect on childhood BMI than partial breastfeeding, both in magnitude and statistical significance. Exclusive breastfeeding continued for at least four weeks has a statistically significant effect on childhood BMI in five and seven years old. If the exclusive breastfeeding is continued for at least sixteen weeks then reductions in BMI are also detected in three year olds. The effects of exclusive breastfeeding increase in magnitude as the children get older.

3.2 Robustness Checks

The results outlined above are robust to other matching algorithms; as well as nearest neighbour matching, we performed radius matching and Kernel matching, which produced very similar results. These results are also robust to other measures of childhood adiposity; using weight categories defined using the IOTF BMI thresholds for childhood obesity and overweight produced similar results.

Like the majority of studies in the existing literature, PSM assumes that there are no unobservable characteristics which have an effect on both breastfeeding and BMI. If there is selection on unobservables, then both PSM and standard regression models would be inappropriate and exclusion restrictions would be required. For this reason, we also investigate the relationship using a restricted version of the Roy model (Roy 1951) to test for the endogeneity of breastfeeding. This restricted version is explained in more detail by Maddala (1983). The endogenous treatment regression model simultaneously estimates the outcome (BMI) using a linear model and the binary treatment (breastfeeding) using a probit

model using maximum likelihood. The model allows for structured correlation between unobservable characteristics which affect both the outcome and the treatment. This allows us to test for endogeneity of the treatment by testing whether the error terms in the two equations are correlated using a likelihood ratio test. However, this method does impose a potentially restrictive linear functional form, assumes that the error terms of the two equations are multivariate normal and restricts the ATU to be equal to the ATT. We use the Stata command for a regression with an endogenous treatment effect, *etregress*, in Stata 13.

Table 3 presents the ATEs estimated using the restricted Roy models for the same binary breastfeeding treatments as the PSM[10]. Here, the effects of breastfeeding initiation on BMI are positive but, like the PSM analysis, they are small and insignificant, suggesting that simply initiating breastfeeding has no influence on later BMI. Similarly, the results of the Roy model suggest that prolonged and exclusive breastfeeding will produce a significant reduction in the BMI of five and seven year old children. The results shown in table three are for the same sample used in the PSM, restricted by common support. This is to ensure that there is no endogeneity in the sample that is used in this study. Results did not significantly change when all possible observations were included.

[10] The same independent variables which were included in the estimation of the propensity scores are used in both the treatment and outcome equations in the Roy model, with the exception of delivery by Caesarean section which is not included in the outcome equation. Caesarean sections are not expected to influence BMI except through their effects on breastfeeding and as a result are used as an instrument in order to more strongly identify the Roy model. Delivery by Caesarean sections were also used by Denny and Doyle (2008) as an instrumental variable in estimating the effects of breastfeeding on cognitive ability. When caesarean sections are included in the OLS model estimating BMI, it is also found not to have a significant influence. This suggests that it is a good instrument, because any influence that Caesarean sections on BMI is through its effect on breastfeeding.

Table 3: Average Treatment Effects using Linear Model with Endogenous Treatment Effect

	BMI (outcome equations)				
	(1)	(2)	(3)	(4)	(5)
Age 3	0.1009	0.0477	-0.4191	-0.0010	-0.5868**
	(0.1914)	(0.2589)	(0.2602)	(0.2905)	(0.2354)
N	9,330	7,877	6,949	7,451	5,183
Age 5	-0.0443	-0.2906	-0.6068**	-0.2425	-0.6933***
	(0.1996)	(0.2956)	(0.2523)	(0.2612)	(0.2287)
N	9,996	6,858	4,841	7,829	5,423
Age 7	0.0104	-0.1975	-0.6285*	-0.1959	-0.9697***
	(0.2295)	(0.3071)	(0.3410)	(0.3175)	(0.3030)
N	8,372	6,168	6,534	7,167	4,948

Source: Millennium Cohort Study. Notes: Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Restricted Roy model varying by breastfeeding treatment; these binary treatments are (1) ever breastfed, (2) partially breastfed for four weeks, (3) partially breastfed for sixteen weeks, (4) exclusively breastfed for four weeks, (5) exclusively breastfed for sixteen weeks. Control group is never breastfed.

The likelihood ratio tests for endogeneity of the binary breastfeeding treatments, show that for each breastfeeding treatment and at each age group, there is no evidence to reject the null hypothesis of exogeneity[11] with a 95% confidence level. Therefore, we conclude that selection on observables is a reasonable assumption and thus PSM is an appropriate and less restrictive model because it does not impose a functional form between the outcome and the covariates, including the treatment.

[11] The null hypothesis is that there is no correlation between the errors of the outcomes and treatment equations.

4. DISCUSSION

This study differs from those in the existing literature in that it acknowledges the underlying assumptions that are imposed when estimating the treatment effects of breastfeeding on BMI. The two methods used in this study each relax one of the key assumptions that many studies in the existing literature have relied on. We use PSM to relax the assumptions of a functional form and a model with endogenous treatment effect to relax the assumption of selection only on observables. We find no evidence of selection on unobservables once confounding factors are accounted for and we focus on the results estimated using PSM.

The results of the empirical analysis show that although some breastfeeding behaviours produce a significant reduction in BMI, the effects are small. The effects appear to increase as children get older and are larger and more significant when breastfeeding is exclusive and continued for longer durations.

The statistically significant effects found here, challenge findings from a number of studies which detected no influence of breastfeeding on childhood adiposity (Jiang & Foster 2012; Kramer et al. 2007; Oddy & Sherriff 2003; Beyerlein et al. 2008) and those which observed a significant effect in young children but became insignificant by the age of eight years (Burke et al. 2005). The results support previous studies which noted that a reduction in BMI as a result of breastfeeding is not apparent during infancy (Bergmann et al. 2003). The findings also support studies which have pointed out that the relationship between breastfeeding and childhood adiposity is largely attenuated by confounding factors (Liese et al. 2001).

Even though many of the effects found in this study are statistically significant, they are small in magnitude suggesting that breastfeeding policies will help in the fight against the obesity problem but might have a limited influence when used in isolation. They should be part of a wider effort to reduce childhood obesity. The small effects on BMI become modestly larger

as children get older which could be because the reductions in BMI accumulate throughout early childhood and take time to be identified. If these reductions in childhood BMI continue to become larger and more significant as children get older, then there could be substantial differences in BMI as a result of breastfeeding by the time a child reaches adolescence. Further research into the effects of breastfeeding on adolescent BMI could provide important information to policy makers because it has previously been found that obese children and in particular obese adolescents are more likely to remain obese throughout adulthood than their non-obese equivalents (Deckelbaum & Williams 2001).

Results from this study support the current WHO recommendations for six months of exclusive breastfeeding. There is no reason to dispute the recommendations for partial breastfeeding continued until a child is two years old although further research into longer durations of partial breastfeeding could provide more evidence in relation to this recommendation.

It is worth noting that children born today may experience different treatment effects to those born at the start of the Millennium due to improvements in formula milk and attitudes towards breastfeeding. Similarly, the increased prevalence of childhood obesity since members of the Millennium Cohort were born, suggest that it is possible that effects on BMI might be visible at a younger age in childhood born in more recent cohorts. The results from this study are also limited by the data and future research could investigate the effects of infant feeding on childhood BMI in cohort born in later years or in children older than seven years of age. Maternal recall on breastfeeding duration might also effect results. However, the recall of breastfeeding duration has been found to be valid and reliable (Li *et al*., 2005).

It is possible that the reductions in BMI found in this study become more apparent as children get older, possibly due to an increasing number of influences affecting BMI as children get

older. For example, when children start school there are a number of additional influences which could potentially influence BMI, such as school dinners, physical education lessons and influences from other children. Further research into how childhood obesity develops over time in relationship a range of lifestyle behaviours and influences could help us to further understand the dynamics of childhood BMI.

Even when applying methods that account for confounding factors, this study found that the causal influence of breastfeeding on childhood BMI is insufficient to prevent childhood obesity in isolation. Breastfeeding policies alone cannot solve the obesity epidemic; however, the small significant effects that breastfeeding is found to have on childhood BMI shows that breastfeeding could one part of a wider early-life solution to the problem of childhood obesity.

5. REFERENCES

Abadie, A. & Imbens, G.W., 2009. *Matching on the Estimated Propensity Score*, (No. w15301) National Bureau of Economic Research.

Armstrong, J. & Reilly, J.J., 2002. Breastfeeding and Lowering the Risk of Childhood Obesity. *Lancet*, 359, pp.2003–2004.

Austin, P.C., 2011. An Introduction to Propensity Score Methods for Reducing the Effects of Confounding in Observational Studies. *Multivariate Behavioral Research*, 46, pp.399–424.

Bergmann, K. et al., 2003. Early Determinants of Childhood Overweight and Adiposity in a Birth Cohort Study: Role of Breast-Feeding. *International Journal of Obesity*, 27, pp.162–172.

Beyerlein, A., Toschke, A.M. & von Kries, R., 2008. Breastfeeding and Childhood Obesity: Shift of the Entire BMI Distribution or Only the Upper Parts? *Obesity*, 16(12), pp.2730–2733.

Bogen, D.L., Hanusa, B.H. & Whitaker, R.C., 2004. The Effect of Breast-Feeding with and without Formula Use on the Risk of Obesity at 4 Years of Age. *Obesity Research*, 12(9), pp.1527–1535.

Burdette, H.L. & Whitaker, R.C., 2007. Differences by Race and Ethnicity in the Relationship between Breastfeeding and Obesity in Preschool Children. *Ethnicity and Disease*, 17(3), pp.467–470.

Burke, V. et al., 2005. Breastfeeding and Overweight: Longitudinal Analysis in an Australian Birth Cohort. *Journal of Pediatrics*, 147(1), pp.56–61.

Caliendo, M. & Kopeinig, S., 2008. Some Practical Guidance for the Implemenation of Propensity Score Matching. *Journal of Economic Surveys*, 22(1), pp.31–72.

Centers for Disease Control and Prevention (CDC) & National Center for Health Statstics (NCHS), 2001. Growth Charts - Data Table of BMI-for-age Charts. Available at: http://www.cdc.gov/growthcharts/html_charts/bmiagerev.htm [Accessed October 16, 2012].

Cole, T.J. et al., 2000. Establishing a Standard Definition for Child Overweight and Obesity Worldwide: International Survey. *British Medical Journal*, 320, pp.1240–1245.

Deckelbaum, R.J. & Williams, C.L., 2001. Childhood obesity: the health issue. *Obesity Research*, 9(supplement 4), p.239s–243s.

Denny, K. & Doyle, O., 2008. The Causal Effect of Breastfeeding on Children's Cognitive Development: A Quasi-Experimental Design. *Archives of General Psychiatry*, 65(5), pp.578–584.

Duflo, E., Glennerster, R. & Kremer, M., 2007. Using Randomization in Development Economics Research: A Toolkit. *Handbook of Development Economics*, 4, pp.3895–3962.

Gillman, M.W. et al., 2001. Risk of Overweight Among Adolescents who were Breastfed as Infants. *Journal of the American Medical Association*, 285(19), pp.2461–2467.

Grummer-Strawn, L.M. & Mei, Z., 2004. Does Breastfeeding Protect Against Pediatric Overweight? Analysis of Longitudinal Data From the Centers for Disease Control and Prevention Pediatric Nutrition Surveillance System. *Pediatrics*, 113(2), pp.e81–e86.

Günther, A.L.B. et al., 2007. Early protein intake and later obesity risk: which protein sources at which time points throughout infancy and childhood are important for body mass index and body fat percentage at 7 y of age? *The American Journal of Clinical Nutrition*, 86(6), pp.1765–1772.

Hansen, K., 2012. *Millennium Cohort Study: First, Second, Thurd and Fourth Surveys: A Guide to the Datasets (Seventh Edition)*, London.

Hediger, M.L. et al., 2001. Association between infant breastfeeding and overweight in young children. *Journal of the American Medical Association*, 285(19), pp.2453–2460.

Iacovou, M. & Sevilla-Sanz, A., 2010. *The Effect of Breastfeeding on Children's Cognitive Development*, (No. 2010-40) ISER, Uni. Essex.

Jiang, M. & Foster, E.M., 2012. Duration of Breastfeeding and Childhood Obesity: A Generalized Propensity Score Approach. *Health Services Research*, 48(2), pp.628–651.

Jones, A.M., 2007. Indetification of Treatment effects in Health Economics. *Health Economics*, 16(2007), pp.1127–1131.

Koletzko, B. et al., 2009. Can infant feeding choices modulate later obesity risk? *American Journal of Clinical Nutrition*, 89(2), pp.1502–1509.

Kramer, M.S. et al., 2007. Effects of Prolonged and Exclusive Breastfeeding on Child Height, Weight, Adiposity, and Blood Pressure at Age 6.5 y: Evidence from a Large Randomized Trial. *American Journal of Clinical Nutrition*, 86(6), pp.1717–1721.

Kramer, M.S. et al., 2004. Feeding effects on growth during infancy. *The Journal of Pediatrics*, 145(5), pp.600–605.

Von Kries, R. et al., 1999. Breast Feeding and Obesity: Cross Sectional Study. *British Medical Journal*, 319, pp.147–150.

Leuven, E. & Sianesi, B., 2012. Stata Module to Perform Full Mahalanobis and Propensity Score Matching, Common Support Graphing, and Covariate Imbalance Testing. Statistical Software Components.

Li, R., Fein, S.B. & Grummer-Strawn, L.M., 2010. Do infants fed from bottles lack self-regulation of milk intake compared with directly breastfed infants? *Pediatrics*, 125(6), pp.e1386–1393.

Li, R., Scanlon, K.S. & Serdula, M.K., 2005. The validity and reliability of maternal recall of breastfeeding practice. *Nutrition reviews*, 63(4), pp.103–110.

Liese, A.D. et al., 2001. Inverse Association of Overweight and Breast Feeding in 9 to 10-Y-Old Children in Germany. *International Journal of Obesity*, 25(11), pp.1644–1650.

Maddala, G.S., 1983. Models with Self-selectivity. In *Limited-dependent and Qualitative Variables in Econometrics*. Cambridge University Press, pp. 257–290.

Mayer-Davis, E.J. et al., 2006. Breast-Feeding and Risk for Childhood Obesity: Does Maternal Diabetes or Obesity Status Matter? *Diabetes Care*, 29(10), pp.2231–2237.

McCrory, C. & Layte, R., 2012. Breastfeeding and Risk of Overweight and Obesity at Nine-Years of Age. *Social Science & Medicine*, 75(2), pp.323–330.

Oddy, W.H. & Sherriff, J.L., 2003. Breastfeeding, Body Mass Index, Asthma and Atopy in Children. *Asia-Pacific Journal of Public Health*, 15(supplement), pp.s15–s17.

Plewis, I., 2007. Non-Response in a Birth Cohort Study: The Case of the Millennium Cohort Study. *International Journal of Social Research Methodology*, 10(5), pp.325–334.

Power, C., Lake, J.K. & Cole, T.J., 1997. Body Mass Index and Height from Childhood to Adulthood in the 1958 British Born Cohort. *The American Journal of Clinical Nutrition*, 66, pp.1094–1101.

Reilly, J.J. et al., 2005. Early life risk factors for obesity in childhood: cohort study. *British Medical Journal*, 330, pp.1357–1363.

Renfrew, M.J. et al., 2007. Rethinking Research in Breast-Feeding: a Critique of the Evidence Base Identified in a Systematic Review of Interventions to Promote and Support Breast-Feeding. *Public Health Nutrition*, 10(7), pp.726–732.

Rolland-Cachera, M.F. et al., 1984. Adiposity Rebound in Children: a Simple Indicator for Predicting Obesity. *American Journal of Clinical Nutrition*, 39(1), pp.129–135.

Rosenbaum, P.R. & Rubin, D.B., 1983. The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrica*, 70(1), pp.41–55.

Roy, A.D., 1951. Some Thoughts on the Distribution of Earnings. *Oxford Economic Papers*, 3(2), pp.135–146.

Salsberry, P.J. & Reagan, P.B., 2005. Dynamics of Early Childhood Overweight. *Journal of Pedicatrics*, 116(6), pp.1329–1338.

Serdula, M.K. et al., 1993. Do Obese Children Become Obese Adults? A Review of the Literature. *Preventive Medicine*, 22(2), pp.167–177.

Sianesi, B., 2006. Propensity Score Matching. In *Fiscal Studies*. pp. 1–14.

Swinburn, B. et al., 2007. Diet, nutrition and the prevention of excess weight gain and obesity. *Public Health Nutrition*, 7(1a), pp.123–146.

World Health Organization (WHO), 2012. WHO| Preterm birth: Fact Sheet No. 363. , p.Fact Sheet No363. Available at: http://www.who.int/mediacentre/factsheets/fs363/en/index.html [Accessed July 17, 2012].

6. APPENDIX

8.1 Independent variables

Maternal education, measured on a five point scale, indicates whether a mother has 'no qualifications', 'GCSEs (grade A*-C)', 'A-levels', 'Degree level' and 'higher than degree level' or equivalent level qualifications. Two dummy variables indicating high and low maternal education levels are derived. 'Highly educated' is defined as having at least one degree, or equivalent, and 'low education' is considered to be those who did no further qualifications post compulsory education. Family socioeconomic status (SES) is defined as the highest SES of a cohort member's parents. In the MCS, SES is provided using a five point National Statistics Socioeconomic Classification (NS-SEC) scale. This scale consists of 'managerial or professional', 'intermediate', 'small employer or self-employed' and 'semi-routine or routine'. A further category to indicate 'long-term unemployed' was also added to reduce the number of missing values. Two dummy variables indicating high and low SES are created, defining high SES as 'managerial or professional' and low SES as 'semi-routine, routine or long-term unemployed'.

Binary variables were included to indicate whether a child was living with both natural parents or not and whether their mother was married or not during the first wave of the MCS. In addition, a dummy variable indicating whether or not the cohort member lived in a house owned, outright or with a mortgage, by his/her parent(s) is included. Dummy variables are also used to indicate maternal pre-pregnancy obesity, whether the mother had a longstanding illness soon after the birth of her child and whether a mother was in care at the age of leaving school. A mother's age at the birth of her child is also included. Three dummy variables are used to indicate whether a mother smoked in each trimester of her pregnancy. Additionally, the number of units of alcohol consumed by a mother whilst pregnant on an average day when she consumed alcohol is also included.

Birth weight, in kilograms, and the logged number of days hospital stay after birth are included in predicting breastfeeding behaviours, as well as dummy variables indicating prematurity (less than 37 weeks gestation time) and whether a pregnancy was planned or not. Gender and ethnicity of the cohort member are also included.

Table 4: Excluded Observations

Variable	9 Months	3 Years	5 Years	7 Years
Original Sample	18,552	15,808	15,460	14,043
Late entry	0	699	573	500
Multiple birth	256	413	409	351
Mother's BMI*	819	80	666	698
Birth weight	846	690	747	673
Hospital stay	459	362	428	369
Gestation length	834	679	734	664
Child's BMI*	-	669	768	683
Number after exclusions	16,219	13,979	13,700	12,494
(% removed)	(12.58)	(11.57)	(11.38)	(11.03)
Missing observations	1,151	2,779	1,956	1,787
# observations in sample	15,068	11,200	11,744	10,707
(% of original obs.)	(81.22)	(70.85)	(75.96)	(76.24)

Source: Millennium Cohort Study. Notes: Values are for number of children, not families. Observations can be missing in more than one variable. *implausible or missing height, weight or BMI. Child's BMI is not measured at nine months.

Table 4 shows the number of observations which are excluded from each of the wave due to missing data or removed observations. The variables described above are used in the PSM as the independent variables which influence the likelihood of treatment; they are used to estimate the propensity scores. They are also used in the restricted Roy model to predict BMI. The Roy model also includes an additional variable to estimate the endogenous treatment effect, in order to improve the strength of model identification. We use a binary variable indicating whether an infant was delivered by caesarean section or not. Caesarean sections have previously been used by Denny and Doyle (2008) as an instrument for breastfeeding when estimating the effects of breastfeeding on cognitive development. In the same way as an instrumental variable, this variable will be included in the treatment equation of the Roy model but not in the outcome equation. Summary statistics of all independent variables in the sample during each wave are presented in Table 5.

Table 5: Mean and Standard Deviation of Independent Variables at 9 Months

Variable	9 Months	3 Years	5 Years	7 Years
High education*	0.2989	0.3262	0.3252	0.3332
	(0.4578)	(0.4688)	(0.4685)	(0.4714)
Low education*	0.5551	0.5233	0.5249	0.5146
	(0.4970)	(0.4995)	(0.4994)	(0.4998)
High SES*	0.1849	0.2028	0.1992	0.2090
	(0.3882)	(0.4021)	(0.3994)	(0.4066)
Low SES*	0.5330	0.4968	0.5019	0.4874
	(0.4989)	(0.5000)	(0.5000)	(0.4999)
Male*	0.5145	0.5024	0.5077	0.5023
	(0.4998)	(0.5000)	(0.5000)	(0.5000)
Black*	0.0296	0.0236	0.0246	0.0242
	(0.1695)	(0.1517)	(0.1549)	(0.1536)
Asian*	0.0898	0.0829	0.0827	0.0805
	(0.2859)	(0.2758)	(0.2754)	(0.2721)
Other*	0.0350	0.0320	0.0320	0.0305
	(0.1837)	(0.1759)	(0.1761)	(0.1721)
Home Owner*	0.5927	0.6354	0.6238	0.6391
	(0.4913)	(0.4814)	(0.4844)	(0.4803)
Private Renter*	0.0901	0.0812	0.0840	0.0816
	(0.2863)	(0.2731)	(0.2773)	(0.2738)
Natural Parents*	0.8239	0.8483	0.8428	0.8495
	(0.3809)	(0.3587)	(0.3640)	(0.3575)
Birth weight	3.367	3.379	3.375	3.381
	(0.5535)	(0.5493)	(0.5536)	(0.5494)
Premature*	0.0664	0.0635	0.0652	0.0639
	(0.2490)	(0.2438)	(0.2469)	(0.2446)
Log Hospital Stay	1.126	1.125	1.126	1.128
	(0.6070)	(0.6083)	(0.6071)	(0.6082)
Planned Pregnancy*	0.5438	0.5679	0.5617	0.5706
	(0.4981)	(0.4954)	(0.4962)	(0.4950)
Mother married*	0.5771	0.6051	0.5969	0.6051
	(0.4940)	(0.4889)	(0.4905)	(0.4888)
Mother obese*	0.0688	0.0792	0.0730	0.0707
	(0.2530)	(0.2701)	(0.2601)	(0.2563)
Mother age at birth	3.367	28.50	28.41	28.55
	(0.5535)	(5.764)	(5.768)	(5.753)
Smoking 1st Trimester*	0.2534	0.2470	0.2457	0.2433
	(0.4350)	(0.4313)	(0.4305)	(0.4291)
Smoking 2nd Trimester*	0.0280	0.0252	0.0250	0.0255
	(0.1650)	(0.1567)	(0.1562)	(0.1576)
Smoking 3rd Trimester*	0.0755	0.0685	0.0707	0.0669
	(0.2642)	(0.2526)	(0.2563)	(0.2498)
Alcohol units a day	0.2902	0.2960	0.2949	0.2940
	(0.9826)	(0.9929)	(0.9945)	(0.9854)
Mother in Care when leaving school*	0.0096	0.0090	0.0083	0.0072
	(0.0973)	(0.0945)	(0.0905)	(0.0845)
Illness*	0.2069	0.2143	0.2125	0.2137
	(0.4051)	(0.4103)	(0.4091)	(0.4099)
Caesarean Section*	0.2098	0.2139	0.2079	0.2095
	(0.4072)	(0.4101)	(0.4058)	(0.4070)
N	15,068	11,200	11,744	10,707

Source: Variables available in, or created from responses in the Millennium Cohort Study. Notes: Mean with standard deviation in parentheses. *Binary variable.

8.2 Extended Results

Table 6: Estimation of Propensity Scores at Age 3 Years

	Probit model estimating Breastfeeding				
	(1)	(2)	(3)	(4)	(5)
Age	-0.00180	-0.00231	-0.000925	-0.00191	-0.00143
	(0.00124)	(0.00141)	(0.00165)	(0.00148)	(0.00198)
Sex	0.0348	0.0357	0.0249	0.0265	-0.0474
	(0.0268)	(0.0301)	(0.0350)	(0.0314)	(0.0417)
Black	1.246***	1.483***	1.637***	1.244***	1.428***
	(0.126)	(0.133)	(0.148)	(0.147)	(0.181)
Asian	0.681***	0.787***	0.852***	0.656***	0.821***
	(0.0572)	(0.0624)	(0.0705)	(0.0667)	(0.0818)
Other	0.756***	0.864***	0.956***	0.788***	0.889***
	(0.0888)	(0.0976)	(0.111)	(0.103)	(0.131)
high education	0.339***	0.365***	0.399***	0.357***	0.426***
	(0.0454)	(0.0493)	(0.0553)	(0.0512)	(0.0655)
low education	-0.254***	-0.330***	-0.406***	-0.334***	-0.366***
	(0.0384)	(0.0430)	(0.0495)	(0.0447)	(0.0596)
high SES	0.257***	0.308***	0.340***	0.321***	0.356***
	(0.0458)	(0.0490)	(0.0539)	(0.0508)	(0.0624)
low SES	-0.274***	-0.304***	-0.366***	-0.293***	-0.343***
	(0.0325)	(0.0365)	(0.0422)	(0.0380)	(0.0503)
live with both natural parents	0.276***	0.288***	0.333***	0.263***	0.325***
	(0.0429)	(0.0505)	(0.0625)	(0.0528)	(0.0779)
mother married	0.0319	0.0561	0.0470	0.0633	0.110*
	(0.0346)	(0.0388)	(0.0451)	(0.0405)	(0.0538)
home owners	0.0947*	0.0948*	0.0726	0.0972*	0.0527
	(0.0376)	(0.0430)	(0.0509)	(0.0451)	(0.0614)
private renters	0.180***	0.220***	0.270***	0.223***	0.219*
	(0.0517)	(0.0595)	(0.0707)	(0.0622)	(0.0871)
birth weight	-0.0110	-0.00594	0.0301	-0.0108	0.0179
	(0.0276)	(0.0311)	(0.0367)	(0.0328)	(0.0439)
hospital stay (log)	0.129***	0.0948**	0.0641	0.0864**	0.0442
	(0.0258)	(0.0290)	(0.0340)	(0.0304)	(0.0410)
planned pregnancy	0.0939**	0.108**	0.0974*	0.0995**	0.0583
	(0.0299)	(0.0335)	(0.0388)	(0.0349)	(0.0460)
Premature	-0.0807	-0.0992	-0.245**	-0.162*	-0.266**
	(0.0601)	(0.0684)	(0.0830)	(0.0726)	(0.0995)
mother obese	-0.0273	-0.110	-0.282***	-0.139*	-0.379***
	(0.0488)	(0.0560)	(0.0685)	(0.0592)	(0.0858)
mother age at birth	0.0117***	0.0247***	0.0358***	0.0256***	0.0433***
	(0.00270)	(0.00306)	(0.00359)	(0.00319)	(0.00429)
smoker 1st trimester	-0.0790*	-0.168***	-0.344***	-0.183***	-0.353***
	(0.0335)	(0.0384)	(0.0457)	(0.0400)	(0.0551)
smoker 2nd trimester	-0.335***	-0.415***	-0.454***	-0.371***	-0.577***
	(0.0826)	(0.0981)	(0.119)	(0.100)	(0.158)
smoker 3rd trimester	-0.341***	-0.454***	-0.652***	-0.474***	-0.741***
	(0.0532)	(0.0633)	(0.0807)	(0.0664)	(0.104)
alcohol during pregnancy	-0.000174	-0.00106	0.0148	0.000330	0.00984
	(0.0129)	(0.0151)	(0.0169)	(0.0155)	(0.0228)
mother in care at 16 years	-0.0299	-0.116	-0.146	-0.126	0.123
	(0.132)	(0.162)	(0.210)	(0.171)	(0.233)
maternal longstanding illness	0.0522	0.0138	-0.0245	-0.0118	-0.120*
	(0.0326)	(0.0371)	(0.0435)	(0.0389)	(0.0531)
Caesarean Section delivery	-0.118**	-0.138**	-0.169***	-0.178***	-0.168**
	(0.0382)	(0.0430)	(0.0502)	(0.0455)	(0.0603)
Constant	0.122	-0.372	-1.213***	-0.479	-1.632***
	(0.247)	(0.281)	(0.331)	(0.294)	(0.395)
N	11200	8845	6949	7885	5290

Source: Millennium Cohort Study. Notes: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Probit model varying by breastfeeding treatment; these binary treatments are (1) ever breastfed, (2) partially breastfed for four weeks, (3) partially breastfed for sixteen weeks, (4) exclusively breastfed for four weeks, (5) exclusively breastfed for sixteen weeks.

Table 7: Propensity Score Matching Results

Treatment	# NN (calliper)	ATT (s.e.^)	ATT sample size (com. support)	ATU (s.e.)	ATU sample size (com. support)	ATE (s.e.$^{\epsilon}$ ^)	ATE sample size (com. support)	ATE 95% CI
Age 3								
Ever breastfed	1	-0.0448	6,196	-0.0282	3,134	-0.0392	9,330	(-0.1214,
	(0.00024)	(0.0518)	(79.9%)	(0.0223)	(90.9%)	(0.0419)	(83.3%)	0.0430)
Partial Breastfeeding								
> 4 weeks compared to	2	-0.0174	4,724	-0.0570*	3,153	-0.0333	7,877	(-0.1254,
never breastfed	(0.0005)	(0.0585)	(87.5%)	(0.0298)	(91.5%)	(0.0470)	(89.1%)	0.0589)
> 16 weeks compared to	3	-0.0088	2,835	-0.0083	2,700	-0.0086	5,602	(-0.2291,
never breastfed	(0.0005)	(0.0087)	(80.9%)	(0.0068)	(78.4%)	(0.0077)	(80.2%)	-0.0013)
Exclusive breastfeeding								
> 4 weeks compared to	3	-0.0512	4,178	-0.0231	3,279	-0.0388	7,457	(-0.1446,
never breastfed	(0.001)	(0.0580)	(94.1%)	(0.0318)	(95.2%)	(0.0465)	(94.6%)	0.0204)
> 16 weeks compared to	3	-0.1310*	1,822	-0.1746**	3,361	-0.1592**	5,183	(-0.3131,
never breastfed	(0.01)	(0.0790)	(98.8%)	(0.0768)	(97.5%)	(0.0785)	(98.0%)	-0.0054)
Age 5								
Ever breastfed	1	-0.0837	6,726	-0.0669**	3,270	-0.0782	9,996	(-0.1675,
	(0.00025)	(0.0535)	(82.8%)	(0.0294)	(90.4%)	(0.0456)	(85.1%)	0.0112)
Partial Breastfeeding								
> 4 weeks compared to	2	-0.0977*	4,080	-0.1246*	2,778	-0.1086**	6,858	(-0.2135,
never breastfed	(0.00025)	(0.0569)	(72.0%)	(0.0749)	(76.8%)	(0.0535)	(73.9%)	0.0036)
> 16 weeks compared to	3	-0.1809***	2,439	-0.1735**	2,402	-0.1772**	4,841	(-0.3117,
never breastfed	(0.0003)	(0.0651)	(66.6%)	(0.0722)	(66.4%)	(0.0686)	(66.5%)	-0.0428)
Exclusive breastfeeding								
> 4 weeks compared to	3	-0.1623***	4,363	-0.1121***	3,466	-0.1401***	7,829	(-0.2349,
never breastfed	(0.0009)	(0.0597)	(94.0%)	(0.0342)	(95.8%)	(0.0484)	(94.8%)	-0.0453)
> 16 weeks compared to	3	-0.2176***	1,883	-0.1954**	3,540	-0.2031**	5,423	(-0.3646,
never breastfed	(0.01)	(0.0794)	(97.9%)	(0.0840)	(97.9%)	(0.0824)	(97.9%)	-0.0415)
Age 7								
Ever breastfed	1	-0.1880**	5,565	-0.1019**	2,807	-0.1591**	8,372	(-0.2908,
	(0.0002)	(0.0773)	(74.4%)	(0.0472)	(86.9%)	(0.0672)	(78.2%)	0.0274)
Partial Breastfeeding								
> 4 weeks compared to	2	-0.1542*	3,697	-0.1850***	2,471	-0.1665**	6,168	(-0.3168,
never breastfed	(0.00025)	(0.0841)	(70.1%)	(0.0656)	(76.5%)	(0.0767)	(8,474%)	-0.0162)
> 16 weeks compared to	3	-0.2139**	3,360	-0.2709***	3,174	-0.2416***	6,534	(-0.3908,
never breastfed	(0.0003)	(0.1019)	(98.5%)	(0.0488)	(98.2%)	(0.0761)	(98.4%)	-0.0924)
Exclusive breastfeeding								
> 4 weeks compared to	3	-0.1845**	4,062	-0.2370***	3,105	-0.2072***	7,167	(-0.3528,
never breastfed	(0.001)	(0.0867)	(94.2%)	(0.0581)	(96.1%)	(0.0743)	(95.0%)	-0.0616)
> 16 weeks compared to	3	-0.3674***	1,762	-0.2258**	3,186	-0.2762**	4,948	(-0.4873,
never breastfed	(0.01)	(0.1131)	(98.2%)	(0.1047)	(98.6%)	(0.1077)	(98.4%)	-0.0652)

Source: Millennium Cohort Study. Notes: * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. $^{\epsilon}$bootstrap standard error (500 repetitions). ^Standard errors assume propensity score is known.

8.3 Restricted Roy Model

The Roy model accounts and tests for selection on unobservables by allowing for an endogenous breastfeeding treatment effect. This method, however, does impose a potentially restrictive linear functional form. The most restrictive assumption is probably multivariate normality in this case. If unobserved influences do exist then both PSM and standard regression models would be inappropriate and a model which accounts for selection on unobservables would be required.

A linear regression model with an endogenous treatment effect is estimated in order to investigate the relationship between breastfeeding and childhood adiposity under the assumption that breastfeeding is endogenous. The model allows the outcome, BMI to be estimated simultaneously with the binary breastfeeding treatment, $\boldsymbol{d}$. This model is described in further detail by Maddala (1983) and is a special case of the switching model developed by Roy (1951).

We start with a linear model estimating childhood BMI,

$$\boldsymbol{BMI = X'\beta + \delta d + u} \qquad (2)$$

where $\boldsymbol{X}$ is a vector of independent characteristics, $\boldsymbol{\beta}$ is a vector of corresponding coefficients and **d** is the same binary treatment indicator used in the propensity score analysis. Here, we assume an unobserved continuous latent variable,

$$\boldsymbol{d^* = w'\gamma + \epsilon} \qquad (3)$$

where $\boldsymbol{w}$ is another vector of independent characteristics, which includes both $\boldsymbol{X}$ and the additional variables, caesarean section and $\boldsymbol{\gamma}$ is a vector of corresponding coefficients. Equations 3 and 4 are estimated simultaneously using maximum likelihood. Error terms $\boldsymbol{u}$ and $\boldsymbol{\epsilon}$ are assumed to be bivariate normal with mean zero and covariance matrix

$$\begin{bmatrix} \sigma^2 & \rho\sigma \\ \rho\sigma & 1 \end{bmatrix} \quad (4)$$

where ρ is the correlation between the two error terms.

The observed binary variable $\boldsymbol{d}$ is defined as

$$\boldsymbol{d} = \begin{cases} 1 & if\ \boldsymbol{d}^* > 0 \\ 0 & if\ \boldsymbol{d}^* \leq 0 \end{cases} \quad (5)$$

to indicate treatment.

This model allows us to test for endogeneity of the treatment by testing whether the error terms in the two equations are correlated. The model allows for structured correlation between unobservable characteristics which affect both the outcome and the treatment. A likelihood ratio test can be used to test the null hypothesis that there is no correlation between the outcome and treatment errors, $\boldsymbol{u}$ and $\boldsymbol{\epsilon}$, respectively:

$$H_0: \rho = 0 \quad (6)$$

and therefore determines whether or not the treatment is endogenous.