The Effect of Breastfeeding on Educational Attainment: Evidence from Sibling Data[*]

Daniel I. Rees
University of Colorado Denver
Department of Economics, CB 181
P.O. Box 173364
Denver, CO 80217-3364
Phone: (303) 556-3348
Fax: (303) 556-3547
E-mail: Daniel.Rees@cudenver.edu

Joseph J. Sabia
American University
Department of Public Administration & Policy
School of Public Affairs
4400 Massachusetts Ave, NW
336 Ward Circle Building
Washington, D.C. 20016
Phone: (202) 885-2377
Fax: (202) 885-2347
Email: sabia@american.edu

September 2008

[*] The authors wish to thank Julie Hotchkiss, Laura Argys, and participants at the 2008 Western Economic Association meetings in Honolulu, Hawaii for useful comments and suggestions on an earlier draft of this manuscript. This research uses data from Add Health, a program project designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris, and funded by a grant P01-HD31921 from the National Institute of Child Health and Human Development, with cooperative funding from 17 other agencies. Special acknowledgment is due Ronald R. Rindfuss and Barbara Entwisle for assistance in the original design. Persons interested in obtaining data files from Add Health should contact Add Health, Carolina Population Center, 123 W. Franklin Street, Chapel Hill, NC 27516-2524 (http://www.cpc.unc.edu/addhealth/contract.html).

The Effect of Breastfeeding on Educational Attainment: Evidence from Sibling Data

Abstract

While a few studies have examined the relationship between having been breastfed and academic achievement, none have attempted to control for unobservables by exploiting within-family variation. Using data on sibling pairs drawn from the National Longitudinal Study of Adolescent Health, we estimate the effect of having been breastfed on high school graduation, high school grades, and college attendance. Our results suggest that breastfeeding is associated with substantial increases in high school grades and in the probability of college attendance. Moreover, these associations are robust to adding controls for within-family heterogeneity such as the respondent's temperament as a child and the quality of the relationship between the respondent and his or her parents. We conclude that improvements in cognitive ability and adolescent health may be important pathways through which breastfeeding affects academic achievement.

Keywords: breastfeeding, schooling, human capital

I. INTRODUCTION

Breastfeeding rates have been on the rise since the early 1990s, and a large number of states have passed legislation designed to protect the rights of mothers who breastfeed their infant children.[1] At the beginning of this decade, the U.S. Department of Health and Human Services (HHS) declared that it would like to see seventy-five percent of mothers in the United States breastfeed their children before being discharged from the hospital. In 2005, the HHS reported that substantial progress had been made towards this goal (U.S. Department of Health and Human Services 2000, 2005).

One of the impetuses for the promotion of breastfeeding is that it reduces the incidence of childhood illnesses and chronic disease (American Academy of Pediatrics 1997), which in turn may lead to better health as an adolescent. In addition, advocates of breastfeeding claim that it yields important short- and long-run cognitive development benefits. A number of studies have, in fact, found a positive association between being breastfed and cognitive ability.[2] However, there is a growing concern that this association may be a reflection of omitted family-level variables.

Two recent, but influential, analyses have attempted to address the omitted-variables problem by using data on sibling pairs. Evenhouse and Reilly (2005) found that breastfeeding was associated with an increase in Peabody Picture Vocabulary Test (PPVT) scores. In contrast, Der et al. (2006) found no relationship between breastfeeding and Peabody Individual Achievement Test (PIAT) scores.

[1] The Centers for Disease Control and Prevention (2007) and Jacknowitz (2007) provide national data on the prevalence of breastfeeding.

[2] Anderson et al. (1999), Jain et al. (2002) and Horta et al. (2007) provide reviews of this literature. Also see Kramer et al. (2008).

The potential effects of breastfeeding on cognitive development and adolescent health suggest that it may yield important educational spillovers. However, to our knowledge, no previous study has estimated the effect of having been breastfed on schooling controlling for the influence of unobservables. In the current study we draw on data from the National Longitudinal Study of Adolescent Health to estimate the effect of breastfeeding on high school graduation, high school grades, and college attendance. Building on the work of Der et al. (2006) and Evenhouse and Reilly (2005), we exploit sibling differences in order to account for the influence of family-level unobservables. In addition, we explore the importance of two key mechanisms through which we expect having been breastfed to affect academic achievement: improvements in adolescent health and cognitive ability.

We find that having been breastfed is associated with substantial increases in high school grades and the probability of college attendance. These associations are robust to adding controls for within-family heterogeneity such as the respondent's temperament as a child, the quality of the respondent's relationship with his or her parents, and the degree of parental involvement in the respondent's education. Falsification tests using an alternative set of outcomes also provide evidence against the hypothesis that within-family heterogeneity is driving our results. When measures of adolescent health and cognitive ability are included as explanatory variables, the estimated effect of breastfeeding on academic achievement is attenuated. We conclude that adolescent health and cognitive ability may be important pathways through which breastfeeding affects academic achievement.

II. BACKGROUND

A number of studies have found evidence to support the hypothesis that breastfeeding a child leads to enhanced cognitive ability. For instance, Mortensen et al. (2002) found that

being breastfed for more than 6 months was associated with a 6 percent increase in adult IQ. These authors hypothesized that breast milk may be richer in nutrients that contribute to brain development than standard infant formula, but could not rule out the possibility that mothers who breastfeed, on average, provide a more stimulating home environment than those who do not.

In order to control for the influence of difficult-to-measure factors at the family level, Der et al. (2006) used data on 545 sibling pairs from the 1979 National Longitudinal Survey of Youth. They found that being breastfed was associated with only a small, statistically insignificant increase in PIAT scores. They concluded that, "while breastfeeding has many advantages for the child and mother, enhancement of the child's intelligence in unlikely to be among them" (p. 945).

The Der et al. (2006) study casts doubt on whether breastfeeding is causally related to cognitive development. However it is not the only study in this literature to exploit within-family variation. Evenhouse and Reilly (2005) used data on 523 sibling pairs from the National Longitudinal Study of Adolescent Health, the same data source as is used in the current study. They found that having been breastfed was associated with an increase in PPVT scores, although there was no evidence that breastfeeding impacted a variety of health measures or self-reported grades.

Aside from Evenhouse and Reilly (2005), we know of only three other previous studies that have examined the association between breastfeeding and academic achievement (as opposed to intelligence/cognitive ability): Richards et al. (2002), using data on individuals born in the United Kingdom in 1946, found that having been breastfed was associated with an increase in the probability of attaining an advanced degree; Victora et al. (2005), using data on 18-year old males in Brazil, found that having been breastfed was associated with an increase

in the number of years of schooling completed; and Harwood and Fergusson (1998), using data on individuals born in Christchurch New Zealand in 1977, found that having been breastfed was associated with higher test scores in reading and math, and a lower probability of leaving high school without a degree.

Neither Richards et al. (2002), nor Victora et al. (2005), nor Harwood and Fergusson (1998) used sibling data, and therefore one potential explanation for the positive relationship between breastfeeding and achievement documented by these authors is that it is a reflection of family-level unobservables such as maternal intelligence or the quality of the mother-infant interactions. An alternative explanation is that the relationship is causal in nature and is driven by the cognitive benefits of breastfeeding. In fact, Richards et al. (2002) found evidence that the effect of having been breastfed on attaining an advanced degree worked almost entirely through cognitive ability measured at the age of 15. Finally, the health benefits of breastfeeding may explain its relationship to achievement.[3] If improved health allows children to avoid missing school, or even to study harder or more efficiently, then it is possible that it explains the positive relationship between having been breastfed and academic achievement.

In the empirical analysis below we try to distinguish between the above hypotheses. In order to control for the influence of unobservables, we rely on within-family variation for identification. We also test the sensitivity of our identification strategy to within-family heterogeneity by controlling for child-specific observables and by conducting a set of falsification tests. Next, we examine the pathways through which breastfeeding might affect schooling. Because the Adolescent Health study administered a shortened version of the PPVT to its respondents, we are able to test whether having been breastfed impacts high school

[3] As noted in the introduction, there is strong evidence that breastfeeding provides important health benefits, such as immunization from infectious illnesses and a reduction in the likelihood of chronic illnesses (American Academy of Pediatrics 1997).

grades, the probability of high school graduation, and the probability of college attendance through cognitive ability. In addition, the Adolescent Health data contain extensive information on the health of respondents when most were between the ages of 12 and 18. This information allows us to examine whether adolescent health mediates the relationship between having been breastfed and academic achievement.

III. DATA AND BASIC MEASURES

The data used in this study come from the National Longitudinal Study of Adolescent Health, conducted by the Carolina Population Center at the University of North Carolina at Chapel Hill. The Adolescent Health data collection effort began with the identification of more than 26,000 schools in the United States that served 11th graders and had an enrollment of at least 30 students. Eighty high schools were chosen from this population with unequal probability based on their size, region of the country, level of urbanization, type (public vs. private), and racial mix. Most were then matched with a junior high or middle school from the same community, bringing the total number of participating schools to 132.

From the student rosters of these 132 schools, a core sample was randomly chosen to be administered the Adolescent Health Wave I (baseline) in-home survey, which was completed by 20,746 adolescents between April and December of 1995, and produced a nationally representative sample of students in grades 7 through 12. A follow-up survey was administered approximately one year later, and a second follow-up, the Wave III in-home survey, was administered in 2001 when respondents were between the ages of 18 and 28.[4]

[4] Further information regarding the Adolescent Health data collection effort is available from a variety of sources. See, for instance, Harris et al. (2002).

Three outcome variables were constructed from the Adolescent Health data. The first is equal to 1 if the respondent had received a high school diploma by the time of the Wave III survey in 2001, and equal to 0 if he or she dropped out. The second is equal to 1 if the respondent was attending college at the time of the Wave III survey or had completed at least one year of college prior to the survey.[5] The third outcome is the respondent's cumulative high school grade point average (GPA), which was calculated using the official transcripts made available to researchers with access to the restricted-use Adolescent Health data.

Information on whether and for how long the respondent was breastfed comes from answers to the Adolescent Health parent in-home questionnaire, administered at the time of the Wave I survey. One of the respondent's parents (typically the biological mother) was asked for how long the respondent was breastfed. Possible answers were as follows:

1. less than 3 months
2. 3 months to less than 6 months
3. 6 months to less than 9 months
4. 9 months to less than 12 months
5. 12 months to less than 24 months
6. 24 months or more
7. He/She was not breastfed
8. Don't know

The top panel of Table 1 presents the proportion of respondents who were breastfed by the outcomes under study.[6] It provides evidence that breastfeeding is associated with academic success. For instance, respondents who graduated high school were more likely to have been

[5]A small number of respondents (n = 12) were excluded from the analysis because they were still attending high school when the Wave III survey was administered in 2001. Inclusion of these respondents in the analysis did not qualitatively change the results presented below. Because 11.3% of the respondents were still teenagers when interviewed at Wave III, it is likely that some proportion subsequently graduated high school and went on to attend college. Restricting the sample to respondents who were at least 20 years of age at the time of the Wave III interview produced qualitatively similar estimates of the relationship between having been breastfed and achievement as those presented below.

[6] Respondents whose parent did not know whether they were breastfed were dropped from the analysis.

breastfed than their counterparts who dropped out, as were respondents who were attending college at the third wave of the Adolescent Health data collection effort.

Following Evenhouse and Reilly (2005), we created a continuous measure of how long the respondent was breastfed, *Duration*, based on the midpoints of the first six categories listed above (for instance, respondents in the less-than-3-months category were assigned a breastfeeding duration of 1.5 months, respondents in the 3-months-to-less-than-6-months category were assigned a breastfeeding duration of 4.5 months). The second panel of Table 1 shows the mean of *Months Breastfed* by the outcomes under study. The pattern of results is consistent with a standard dose-response relationship. That is, respondents with higher grades were typically breastfed longer. Similarly, those who graduated from high school were, on average, breastfed for more months than those who did not, and respondents who were attending college were breastfed for more months than respondents who were not attending college at Wave III. This same pattern of results is evident in the third panel of Table 1, which divides respondents into three mutually exclusive categories based on whether they were breastfed for 1 to 5 months, 6 to 12 months, or 12 or more months.

IV. EMPIRICAL MODEL

We begin by testing whether the positive relationship between breastfeeding and academic achievement documented in Table 1 is robust to controls for a standard set of observables. Specifically, we estimate the following using the full Adolescent Health sample:

$$A_i = \beta_0 + \boldsymbol{\beta'_1 X_i} + \beta_2 \, Breastfed_i + \varepsilon_i , \qquad (1)$$

where A_i represents the achievement of respondent *i*; the vector $\mathbf{X_i}$ includes controls for family background and personal characteristics; *Breastfed*, which is equal to 1 if the respondent was breastfed, and equal to 0 otherwise; and ε_i is a random error term. Our primary is on β_2, which represents the relationship between having been breastfed and educational attainment, although equation (1) can easily be modified to explore the relationship between the length of time a respondent was breastfed and his or her academic achievement by replacing *Breastfed* with *Months Breastfed*.

The regression model outlined above can generate an unbiased estimate of the effect of breastfeeding provided that the appropriate controls are included on the right-hand side. However, in practice it is often difficult to obtain information on all of the controls that might be in the vector $\mathbf{X_i}$. For instance, although we can control for the highest degree received by the respondent's parent, we have no information on the mother's cognitive ability, her parenting skills, her health endowment, or the health care services she received while pregnant.

We use two approaches to address this issue. Both exploit the fact that Adolescent Health data contain information on siblings raised in the same family. Restricting the sample to respondents whose sibling was also interviewed, we can estimate:

$$A_{ij} = \beta_0 + \boldsymbol{\beta'_1}\mathbf{X_i} + \beta_2 \, Breastfed_{ij} + \omega_{ij}, \qquad (2)$$

where the subscript *j* denotes the family of respondent *i*, ω_{ij} is the composite error term equal to $\kappa_j + \varepsilon_{ij}$, and κ_j is an unobserved family effect. Under the assumption that Cov ($Breastfed_{ij}$, κ_j) = 0 and Cov ($\mathbf{X}_i$, κ_j) = 0, the parameters in equation (2) can be estimated consistently via a random effects model. A random effects model is preferred to a pooled OLS regression if within-family unobservables are correlated. Because ω_{ij} is a function of κ_j, ω_{ij} will be

correlated across families; in a random effects model we assume Corr (ω_{ij}, ω_{sj}) = (σ^2_κ/(σ^2_κ + σ^2_ε) for $i \neq s$. Using estimates of σ^2_κ and σ^2_ε generated from pooled OLS regressions, a random effects estimate of β_2 can be obtained using feasible generalized least squares.

While the random effects model accounts for within-family correlation of unobservables, it will not produce an unbiased estimate of β_2 if κ_j correlated with $Breastfed_{ij}$. Therefore, following Der et al. (2006) and Evanhouse and Reilly (2005), we present estimates equation (2) treating κ_j as a family fixed effect:

$$A_{ij} = \beta_0 + \boldsymbol{\beta'_1 X_i} + \beta_2 Breastfed_{ij} + \kappa_j + \varepsilon_{ij,} \qquad (3)$$

where the vector $\mathbf{X_i}$ includes controls for birth weight, age (at Wave III), whether the respondent had an older sibling, family size at birth, and gender. The advantage of this estimation strategy as compared to the pooled OLS or the random effects model is that only the within-family variation is used to estimate the effect of breastfeeding on achievement. All factors common to both siblings are controlled for by the vector κ_j, eliminating the need to observe and measure a myriad of potentially important confounders.

While the estimation of (3) accounts for family-level unobservables, there are at least three drawbacks to this identification strategy that are worthy of note. First, it entails a substantial reduction in sample size and identifying variation. Second, estimates obtained from a sample of siblings may not be generalizable to only children. Finally, controlling for family fixed effects does not account for within-family unmeasured heterogeneity. For instance, if the quality of the mother-child relationship, the child's temperament, or the mother's health at the time of birth is correlated with both the decision to breastfeed and academic success, then the

result could be a biased estimate of β_2.[7] Fortunately, the Adolescent Health data contain information that can be used to minimize the problem of unmeasured within-family heterogeneity.

Because there is strong evidence that birth weight is related to the mother's consumption of prenatal care (Liu 1999) and whether she smoked, drank, or used other substances during pregnancy (Shankaran et al. 2004), all specifications include a continuous birth weight measure (in grams) as a control.[8] This measure should capture differences in maternal health investments across siblings. In addition, although the Adolescent Health data provide no information on the quality of the infant-mother relationship, a number of questions were asked with regard to the quality of the *adolescent*-mother relationship.[9] As part of a series of robustness checks performed in Section V of the paper, we use the answers to the following four questions to augment the vector of controls, $\mathbf{X_i}$:

1. How close do you feel to your mother? (=1 "not at all"; =2 "very little"; =3 "somewhat"; =4 "quite a bit; =5 "very much")

2. How much do you think [your mother] cares about you? (=1 "not at all"; =2 "very little"; =3 "somewhat"; =4 "quite a bit"; =5 "very much")

[7] The decision to work outside the home following childbirth could also affect both breastfeeding duration and the quality of infant-mother interactions. Chatterji and Frick (2003), Baker and Milligan (2008), Blau et al. (1996), Lindberg (1996), and Ryan and Martinez (1989) document evidence of a negative relationship between work and breastfeeding duration.

[8] Replacing birth weight in grams with an indicator of low birth weight produced qualitatively similar results to those reported in the paper.

[9] One possibility is that being breastfed is positively correlated with other determinants of academic achievement. Alternatively, mothers may try to equalize inputs among their children and, for instance, spend more time with the sibling who was not breastfed. If being breastfed is simply a proxy for the quality of the infant-mother relationship, then we would expect controlling for the quality of the adolescent-mother relationship to reduce our estimate of β_2. If controlling for the quality of the adolescent-mother relationship has no impact on our estimate of β_2, this could indicate that (1) the quality of the infant-mother relationship is orthogonal to the quality of the adolescent-mother relationship, or (2) that being breastfed is related to academic achievement by an alternative mechanism such as cognitive ability. In our opinion, the second scenario is more plausible.

3. Do you get along well with your child? (=1 “always”; =2 “often”; =3 “sometimes”; =4 “seldom”; =5 “never”)

4. Do you and your child make decisions about his/her life together? (=1 “always”; =2 “often”; =3 “sometimes”; =4 “seldom”; =5 “never”)[10]

We also experiment with adding a set of controls intended to measure the degree to which the respondent’s parents were involved in their child’s education at Wave I. The controls are based on answers to the following four questions:

1. Which of the following have you done with your mother in the past 4 weeks? (a) talked about your school work or grades, (b) worked on a project for school, (c) talked about other things you’re doing in school.

2. On a scale of 1 to 5, where 1 is low and 5 is high, how disappointed would [your mother] be if you did not graduate from high school?

3. On a scale of 1 to 5, where 1 is low and 5 is high, how disappointed would [your mother] be if you did not graduate from college?

4. If [your child] could be one of the following in high school, which would be most important to you? (a) a brilliant student, (ii) a leader in school activities, (iii) an athletic star, (iv) the most popular.[11]

Finally, we experiment with augmenting the vector of controls, $\mathbf{X_i}$, with measures of childhood temperament, which could be correlated with both being breastfed and achievement.[12] At

[10] The first two questions were asked as part of the Wave I in-home survey; the final two come from the parent questionnaire. Four sets of dichotomous variables were created capturing all of the possible answers to each question and missing values. Appendix Table 2 shows that respondents from the same family often provided different answers to these questions. For instance, 45.0 percent of respondents in the sibling sample reported a different degree of closeness to their biological mother than did their sibling, and 39.0 percent of parents reported differences in how well they got along with their children.

[11] The first three questions were asked as part of the Wave I in-home survey in-home; the final question comes from the parent questionnaire. Four sets of dichotomous variables were created capturing all of the possible answers to each question and missing values. Appendix Table 2 of the appendix shows that respondents from the same family often provided different answers to these questions. For instance, 41.1 percent of respondents reported differences from their sibling in whether their mother had talked about grades or school work in the previous four weeks, and 57.7 percent reported differences in the degree of disappointment their mother would feel if they did not graduate from high school.

[12] For instance, especially active infants may breastfeed for shorter durations than their more placid counterparts. Activity level may also be correlated with achievement.

Wave III, respondents were asked 17 retrospective behavioral questions with regard to whether they experienced the symptoms of Attention Deficit Hyperactivity Disorder (ADHD) when they were between the ages of 5 and 12. For example, respondents are asked how often they squirmed in their seats, felt restless, fidgeted with their hands, and did not follow through on instructions. Each response was coded on a scale of 0 to 3 based on frequency (0 = never or rarely; 1 = sometimes; 2 = often; and 3 = very often), and adding up these responses produced an ADHD score of between 0 and 51, which was included in the vector $\mathbf{X_i}$.[13] In addition, a dichotomous measure of child temperament was constructed from answers to an item on the parent survey, "Does [your child] have a bad temper?"

V. RESULTS

Full Sample Estimates

The top panel of Table 3 presents OLS estimates of the relationship between breastfeeding and the three outcome variables. The results suggest that having been breastfed is associated with a 0.118 increase in high school GPA, or an approximately 4.6 percent increase in the GPA of the typical Adolescent Health respondent. It is also associated with a .025 increase in the probability of graduating from high school, and a .069 increase in the probability of attending college.

In the second panel of Table 3, we replace the dichotomous measure, *Breastfed*, with the continuous measure, *Months Breastfed*, described in Section III. The results suggest that the length of time a respondent was breastfed is positively related to the three outcomes. Specifically, an additional month of breastfeeding is associated with a 0.011 increase in high

[13] We also experimented with a dichotomous ADHD measure. If a respondent answered "often" or "very often" on six or more of the inattention-related ADHD questions or six or more of the hyperactivity- or impulsivity-related questions (see http://www.cdc.gov/ncbddd/adhd/symptom.htm). A similar measure was employed by Fletcher and Wolfe (2008), Kollins et al. (2005), and Murphy and Barkley (1996).

school GPA. It is also associated with a .002 increase in the probability of high school graduation, and a .006 increase in the probability of college attendance.

In the third and final panel of Table 3 we show the relationship between the three duration categories introduced in Table 1 and achievement. This specification allows breastfeeding duration to have a nonlinear effect on schooling. The omitted category is composed of respondents who were never breastfed. The results confirm that breastfeeding duration is positively related to the outcomes under study. For instance, having been breastfed for less than 6 months is associated with a .021 increase in the probability of graduating high school as compared to not having been breastfed (the omitted category); having been breastfed for 6 to months is associated with a .026 increase in this probability; and having been breastfed for 12 months or more is associated with a .038 increase in this probability.

In summary, the results in Tables 3 are consistent with much of the previous literature (Harwood and Fergusson 1998; Richards et al. 2002; Victora et al. 2005) and, if naively interpreted, suggest that breastfeeding leads to substantial increases in academic achievement and provide evidence of a dose-response relationship. However, if family-level unobservables are correlated with both breastfeeding and the outcomes under study, then this interpretation may be incorrect. Below we attempt to control for unmeasured family-level factors through the examination of sibling data and the introduction of random and family fixed effects.

Estimates Based on Sibling Data

The top panel of Table 4 presents random and fixed effects estimates of the impact of having been breastfed on high school grades, the probability of graduating high school, and the probability of attending college. OLS estimates are presented for the purposes of comparison, and standard errors are corrected for clustering at the family-level. Following Der et al. (2006)

and Evenhouse and Reilly (2005), the sample is restricted to siblings who received different breastfeeding treatments as infants.[14] Means of schooling outcomes for the sibling sample are presented in Appendix Table 1.

The results in columns (1) through (3) provide strong evidence that *Breastfed* is associated with a substantial increase in high school GPA. The OLS estimate of β_2 for this restricted sample is 0.381, quite a bit larger than that presented in Table 3. Treating the unmeasured family effect as random reduces this estimate to 0.340, and controlling for family fixed effects reduces this estimate to 0.286, or a 10.4 percent GPA increase for the typical respondent in the sibling sample.

Turning to the results presented in columns (4) through (9) of Panel I, there is no evidence that siblings who were breastfed enjoyed an advantage in terms of graduating from high school or attending college. The OLS, random effects, and fixed effects estimates of β_2, although generally positive, are never significant at conventional levels.

Panels II and III of Table 4 presents OLS, random effects, and fixed effects estimates of the relationship between *Months Breastfed* and achievement using sibling data. Again, the sample is restricted to siblings who received different breastfeeding treatments as infants.[15] The results suggest that breastfeeding duration is positively related to high school grades, or in other words, the results provide evidence of a dose-response relationship. For instance, fixed effects estimates show that an increase in breastfeeding duration of one month is associated

[14] If a family contributed information on two siblings to the Adolescent Health data, this restriction means that one sibling was breastfed, while the other was not. If a family contributed three siblings, then at least one was breastfed and at least one was not. When grades are the dependent variable, the sample is composed of 126 siblings from 59 families. When high school completion or college attendance is on the left-hand side, the sample is composed of 191 siblings from 90 families.

[15] This restriction implies that if a family contributed information on two siblings to the Adolescent Health data, each was breastfed for different periods of time. If a family contributed information on three siblings, then at least one was breastfed for a different period of time than the other two. When grades are the dependent variable, the sample is composed of 333 siblings from 159 families. When high school completion or college attendance is on the left-hand side, the sample is composed of 459 siblings from 220 families.

with an increase in high school GPA of 0.017 points, or about 1 percent for the typical respondent in the sibling sample (Panel II). Similarly, having been breastfed for 1-5 months is associated with a 0.246 point increase in GPA as compared to never having been breastfed, and having been breastfed for 6-12 months is associated with a 0.401 point increase (Panel III). The latter estimate translates to 14.6 percent increase in the GPA of the typical respondent in the sibling sample.

Although the results presented in Panels II and III of Table 4 provide little evidence that breastfeeding duration is related to the probability of graduating from high school, they show a positive relationship between *Months Breastfed* and the probability of attending college. Specifically, an additional month of breastfeeding is associated with a 0.009 to a 0.014 increase in the probability of college attendance. In fact, controlling for family-level unobserved heterogeneity (column 9) produces a larger estimate than that obtained using OLS. Panel III shows that the college attendance effect appears to be largest for respondents who were breastfed for 12 or more months, while having been breastfed for less than 6 months is not associated with a statistically significant increase in the probability of attending college.

Robustness Checks

Table 4 provides evidence that, controlling for family-level unobservables, *Months Breastfed* is associated with better high school grades and an increase in the probability of attending college. Table 5 explores the extent to which these associations can be attributed to heretofore unmeasured within-family heterogeneity.

The first column of Table 5 reproduces the estimates based on sibling data originally presented in Table 4. The second column shows what happens to these estimates when measures of the quality of the adolescent-mother relationship are added to the right-hand side

of equation (3). The results provide little evidence that the quality of the adolescent-mother relationship is driving the estimated effect of breastfeeding duration on academic achievement. In fact, when these controls are added, the estimated coefficient of *Months Breastfed* actually increases from 0.017 to 0.021 in the GPA equation, and from 0.014 to 0.016 in the college attendance equation.

The third column of Table 5 shows the estimated effect of breastfeeding duration after adding controls for parental involvement in the respondent's education; the fourth column shows the estimated effect after adding the respondent's ADHD score and his or her parent's answer to the question, "Does [your child] have a bad temper?"; and the fifth column shows the estimated effect after adding each set of controls from the previous columns simultaneously. Again, there is little evidence that unmeasured within-family heterogeneity can explain the association between breastfeeding duration and academic achievement documented in Table 4. Although not shown, when the continuous measure of breastfeeding duration was replaced with the categorical measures, the same basic pattern of results was obtained.

Falsification Tests

An alternative method of exploring whether the estimates presented in Table 4 are driven by within-family unobservables is through the use of falsification tests. This requires identifying a set of outcomes that, in theory, should not be affected by breastfeeding.

Using sibling data, we examine the relationship between breastfeeding and seven outcomes measured at Wave I: an abridged version of the Rosenberg Self-Esteem Scale

(RSE)[16]; a depression indicator based on the Center for Epidemiologic Studies Depression (CES-D) Scale[17]; an indicator of binge drinking equal to 1 if the respondent reported being "drunk or very high on alcohol" at least twice in the previous year, and equal to 0 otherwise; an indicator of whether the respondent smoked cigarettes in the past month; the number of hours of television the respondent watched per week; an indicator of sports participation equal to 1 if the respondent reported playing "an active sport, such as baseball, softball, basketball, soccer, swimming, or football" at least three times during the week prior to being interviewed, and equal to 0 otherwise; and the respondent's percentile standing in the gender-specific BMI-for-age distribution.[18] A causal interpretation of the results presented in Table 4 would be called into question if the breastfeeding variables were found to be related to these seven alternative outcomes.

[16] The Adolescent Health study administered six of the ten questions typically used to derive the full RSE Scale as described by Rosenberg (1965).For instance, respondents were asked whether they had "good qualities," whether they had "a lot to be proud of," and if they liked "themselves the way they are." Responses available to respondents were: "strongly agree" (= 5), "agree" (= 4), "neither agree nor disagree" (= 3), "disagree" (= 2), or "strongly disagree" (= 1). These responses were summed to produce a score of 6 to 30, with higher scores corresponding to greater self-esteem. Other studies using the abridged RSE Scale include Nelson and Gordon-Larsen (2006) and Shrier et al. (2001).

[17] Originally developed by Radloff (1977), the CES-D Scale is a widely-used measure of depressive symptomatology. The Adolescent Health study administered 18 of the 20 items that typically comprise the CES-D Scale.[17] Specifically, respondents were instructed to indicate the frequency with they had experienced certain feelings or emotions during the past week, including how often they felt "too tired to do things," how often they felt "fearful," and how often they "talked less than usual." Possible responses were "rarely or none of the time" (= 0); "some or a little of the time" (= 1); "occasionally or a moderate amount of the time" (= 2); and "most or all of the time" (= 3). Responses to the 18 items were summed to produce a score of between 0 and 54, which was adjusted to correspond to the original 20-item CES-D Scale. Following Roberts et al. (1991) and Sabia and Rees (2008), we coded *Depress* equal to 1 if a male respondent scored above 22 on the CES-D Scale, and equal to 0 otherwise. For females a cut-point of 24 was employed. The CES-D Scale is often dichotomized in this fashion by psychologists and medical researchers.

[18] The means of each of these outcomes is presented in Appendix Table 1. A number of studies have found that being breastfed is associated with lower BMI. However, a recent review of the literature concluded that the effect is "small and is likely to be strongly influenced by publication bias and confounding factors" (Owen et al. 2005, p. 1298). Nelson et al. (2005) found that cross-section estimates of the relationship between breastfeeding and the probability of obesity are biased due to family-level unobservables. Their analysis of sibling pairs suggested that there was no effect of having been breastfed on obesity.

Table 6 shows the falsification test results. Without exception, the estimated effect of breastfeeding falls short of statistical significance at conventional levels. Moreover, the magnitude of the estimates is often quite small. For instance, having been breastfed is associated with a 0.574 reduction in the number of hours of television watched per week. For the typical adolescent in the sibling sample, this represents a small fraction (3.9%) of the total number of hours of television watched per week. Similarly, the estimated relationship between the length of time a respondent was breastfed and the seven alternative outcomes is never statistically significant and is often small in terms of magnitude. This pattern of results bolsters the case for interpreting the results in Table 4 as causal.

Exploring Likely Pathways

If neither between- nor within-family family unobservables can explain the estimated relationship between the length of time a respondent was breastfed and academic achievement, what can? There are at least two pathways through which, in theory, having been breastfed might impact academic success.

Column (1) of Table 7 reproduces the fixed effects estimates of the relationship between breastfeeding and high school grades originally presented in Table 4. In column (2) we show what happens to these estimates when the PPVT score, a measure of cognitive ability, is added to $\mathbf{X_i}$.[19] Controlling for the PPVT score reduces, but does not eliminate, the estimated effect of breastfeeding on high school grades. For instance, the estimated coefficient of *Breastfed* falls from 0.286 to 0.218, a reduction of about 24 percent. A similar pattern of results emerges when *Breastfed* is replaced by the duration variables (Panels II and III).

[19] The PPVT measures verbal comprehension and vocabulary. The respondent is read a word, and then chooses which of four illustrations best fits the word. The standard PPVT consists of 78 items (Harris and Thomas 2002). Adolescent Health respondents were administered 39 of these 78 items.

Column (5) of Table 7 reproduces the estimates of the relationship between breastfeeding and college attendance originally presented in Table 4. In column (6) we add the PPVT score as an explanatory variable. The estimated effect of having been breastfed for an additional month on the probability of college attendance falls from 0.014 to 0.013; the estimated effect of having been breastfed for 6 to 12 months on the probability of college attendance (as compared to never breastfed) falls from 0.145 to 0.126; the estimated effect of having been breastfed for more than 12 months falls from 0.229 to 0.221.

The findings discussed above provide some evidence that cognitive ability mediates the relationship between breastfeeding and academic achievement. However, it would seem that cognitive ability, at least as measured by the PPVT score, cannot account for the entire effect of breastfeeding on achievement. One interpretation of this finding is that the PPVT score does not adequately capture IQ gains due to breastfeeding. An alternative interpretation is that there exist additional mediators through which breastfeeding impacts schooling.

One such mediator may be adolescent health. There is strong evidence in the medical literature that breastfeeding protect infants from a variety of ailments. The case for long-term health benefits is weaker, but nevertheless many medical professionals argue that breastfeeding confers lifelong immunologic protection (Jackson and Nazar 2006). Case et al. (2005) provide evidence that childhood health is a strong predictor of educational attainment and adult socioeconomic status.

In order to test whether adolescent health mediates the effect of having been breastfed on academic achievement, we created two health indexes. The first was based on answers to 9 questions asked at Wave I about the respondent's general health; the second was based on answers to the same 9 questions asked at Wave II. For example, respondents were asked how frequently in the last 12 months they had had a stomach ache. If they answered "about once a

week," "almost every day," or "every day" they were coded as suffering from stomach aches. Similarly, respondents were asked about headaches, feeling hot, cold sweats, feeling physically weak, sore throats or coughs, frequent or painful urination, and feeling "very sick." Adding up the number of ailments reported by a respondent produced an adolescent health index ranging from 0 to 9. Respondents were also asked at the Wave I and II interviews about the number of times they were absent from school in the past school-year for a full day with an excuse ("for example, because you were sick or out of town"). Possible responses, which were dichotomized, included "never" (the omitted category), "1 or 2 times," "3 to 10 times," or "more than 10 times."

In columns (3) and (7) of Table 7, we introduce controls for the number of excused absences from school and the adolescent health indices described above to the basic estimating equation. The results suggest that adolescent health mediates the relationship between having been breastfed and academic performance as measured by high school grades. For instance, the estimated coefficient of *Breastfed* falls from 0.286 to 0.203 and loses its statistical significance when the adolescent health controls are added. It falls to 0.097 when controls for both cognitive ability and adolescent health are added. There is also evidence that adolescent health can explain part of the relationship between having been breastfed and college attendance. For instance, the estimated coefficient of *Months Breastfed* falls from 0.014 to 0.012 when the adolescent health controls are added, and falls to 0.011, or 27 percent, when controls for both cognitive ability and adolescent health are added.

VI. CONCLUSION

While a number of studies have found a positive relationship between breastfeeding and academic achievement, each has struggled with the issue of unobservables. If, for

instance, mothers who breastfeed, on average, provide a more stimulating environment to their children, then the standard estimates in the literature may be misleading.

The current study builds on the work of Evenhouse and Reilly (2005) and Der et al. (2006) in that we also utilize sibling data to control for family-level unobservables. Fixed effects results suggest that being breastfed is associated with substantial increases in high school grades and the probability of college attendance. However, because these estimates are potentially subject to within-family heterogeneity bias, we exploit the wealth of information available in the Adolescent Health data to control for factors that previous researchers have been forced to leave unmeasured. Specifically, we are able to include measures of the quality of the adolescent-mother relationship, parental involvement in the respondent's education, and the respondent's temperament as a child. These measures are far from perfect, but tellingly their inclusion has very little impact on our estimates of the relationship between being breastfed and academic achievement. In fact, our estimates increase in magnitude with their inclusion. It is difficult to believe that more detailed measures of, for instance, the quality of the respondent's relationship as an infant with his or her mother would have the *opposite* effect.

The case for interpreting the fixed effects estimates of the relationship between being breastfed and academic achievement as casual is bolstered by a series of falsification tests. We can think of no reason why being breastfed should be associated with outcomes such as self-esteem, depression, drinking, smoking, the amount of television watched, sports participation or BMI except through the influence of unobservables. In fact, regressions based on sibling data provide little evidence that being breastfed is related to these outcomes, a pattern of results that suggests unmeasured within-family heterogeneity is unlikely to be an issue. If an

unobservable is driving our results, it would have to be related to academic achievement, but unrelated to these alternative outcomes.

If within-family family heterogeneity does not explain the relationship between breastfeeding and academic achievement, are there potential causal mechanisms that might? We argue that the two obvious conduits are cognitive ability and health. In fact, these factors when simultaneously entered as controls explain almost two-thirds of the estimated effect of length of time breastfed on high school grades. In contrast, they explain only about one-quarter of the effect of breastfeeding on college attendance.

Although more work must be done to identify the precise mechanisms through which having been breastfed is related to academic achievement, this research confirms that the basic results of Harwood and Fergusson (1998), Richards et al. (2002), and Victora et al. (2005) are robust to the use of sibling data as well as controls for within-family heterogeneity.

The magnitudes of the estimated effects are substantial. For instance, we find that having been breastfed leads to a 0.2 to 0.3-point increase in cumulative high school GPA, an effect that is large enough to improve an adolescent's chances of admittance to high-quality colleges and increase his or her earnings as an adult (Manski and Wise 1983; Brewer and Ehrenberg 1996; Brewer et al. 1999). Thus, recent increases in the percentage of U.S. mothers who breastfeed can be expected to have substantial benefits in terms of the human capital acquired by the next generation.

References

American Academy of Pediatrics. 1997. "Breastfeeding and the Use of Human Milk," *Pediatrics*, Vol. 100, No. 6, pp.1035-1039.

Anderson, J.W., B.M. Johnstone, and D.T. Remley. 1999. "Breastfeeding and Cognitive Development: A Meta-Analysis." *American Journal of Clinical Nutrition*, Vol. 70, pp. 525-535.

Baker, Michael and Kevin Milligan. 2008. "Maternal Employment, Breastfeeding, and Health: Evidence from maternity leave mandates," *Journal of Health Economics* 27: 871-887.

Blau, D.M., Guilkey, D.K. and Popkin, B.M. 1996. "Infant health and the labor supply of mothers," *Journal of Human Resources* 31: 90–139.

Brewer, D., & Ehrenberg, R. 1996. "Does it pay to attend an elite private college? Evidence from the senior class of 1980," *Research in Labor Economics*, 15, 239-272.

Brewer, D., Eide, E., & Ehrenberg, R. 1999. "Does it pay to attend an elite private college? Cross cohort evidence on the effects of college type on earnings." *Journal of Human Resources*, 34(1), 104-123

Case, Anne, Angela Fertig, Christina Paxson. 2005. "The Lasting Impact of Childhood Health and Circumstance." *Journal of Health Economics*, Vol. 22, No. 2, pp. 365-389.

Centers for Disease Control and Prevention. 2007. "Breastfeeding Trends and Updated National Health Objectives for Exclusive Breastfeeding -- United States, Birth Years 2000—2004." *Morbidity and Mortality Weekly Report*, Vol. 56, No. 30, pp. 760-763.

Chatterji, Pinka and Kevin Frick. 2003. "Does Returning to Work After Childbirth Affect Breastfeeding Patterns?" NBER Working Paper No. 9630.

Der, G., G.D. Batty, and IJ. Deary. 2006. "Effect of Breast Feeding on Intelligence in Children: Prospective Study, Sibling Pairs, and Meta-Analysis." *British Medical Journal*, Vol. 333, No. 7575, pp. 945-951.

Evenhouse, E. and S. Reilly. 2005. "Improved Estimates of the Benefits of Breastfeeding Using Sibling Comparisons to reduce Selection Bias." *Health Services Research*, Vol. 40 , No. 6, pp.1781-1802.

Fletcher, J.M. and Wolfe, B.L. 2008. Child Mental Health and Human Capital Accumulation: The Case of ADHD Revisited. *Journal of Health Economics*, in press.

Golding, J., I.S. Rogers, and P.M. Emmett. 1997. "Association between Breastfeeding, Child Development, and Behavior." *Early Human Development*, Vol. 49, pp.177-184.

Harris, Kathleen M., Greg J. Duncan, Johanne Boisjoly. 2002. "Evaluating the Role of 'Nothing to Lose' Attitudes on Risky Behavior in Adolescence." *Social Forces*, Vol. 80, No. 3, pp. 1005-1039.

Harris, David R. and Justin L. Thomas. 2002. "The Educational Costs of Being Multiracial: Evidence from a National Survey of Adolescents," Research Report No. 02-521, Population Studies Center at the Institute of Social Research, University of Michigan.

Horwood, John L. and David M. Fergusson. 1998. "Breastfeeding and Later Cognitive and Academic Outcomes" *Pediatrics*, Vol. 101 No. 1, p. e9

Horta, Bernando L., Rajiv Bahl, Jose C. Martinez, and Cesar G. Victora. 2007. Evidence on the Long-Term Effects of Breastfeeding: Systematic Review and Meta-analyses. Geneva, Switzerland: World Health Organization.

Jacknowitz, A. 2007. "Increasing Breastfeeding Rates: Do Changing Demographics Explain Them?" *Women's Health Issues*, Vol. 17, No. 2, pp. 84-92.

Jackson, Kelly M. and Andrea M. Nazar. 2006. "Breastfeeding, the Immune Response, and Long-term Health." The Journal of the American Osteopathic Association, Vol. 106, No. 4, pp. 203-207.

Jacobson, S.W., I.M. Chiodo, and J.L. Jacobson. 1999. "Breastfeeding Effects on Intelligence Quotient in 4- and 11-Year-Old Children." *Pediatrics*, Vol. 103, p. 71.

Jain, Anjali, John Concato, and John M. Leventhal. 2002. "How Good Is the Evidence Linking Breastfeeding and Intelligence?" Pediatrics Vol. 109, No. 6, pp. 1044-1053.

Kollins, S.H.F., Joseph, M., and F. Bernard. 2005. "Association between smoking and attention deficit/hyperactivity disorder symptoms in a population-based sample of young adults," *Archives of General Psychiatry* 62: 1142-1147.

Kramer, M.S., F. Aboud, E. Mironova, I. Vanilovich, R.W. Platt, L. Matush; S. Igumnov, E. Fombonne, N. Bogdanovich, T. Ducruet, J. Collet, B. Chalmers, E. Hodnett, S. Davidovsky, O. Skugarevsky, O. Trofimovich, L. Kozlova, S. Shapiro, For the Promotion of Breastfeeding Intervention Trial (PROBIT) Study Group. 2008. "Breastfeeding and Child Cognitive Development: New Evidence from a Large Randomized Trial." *Archives of General Psychiatry*, Vol. 65, No. 5, pp. 578-584.

Lindberg, Laura D. 1996. "Women's Decisions about Breastfeeding and Maternal Employment," *Journal of Marriage and the Family* 58(1): 239-251.

Liu, G.G., 1998. "Birth Outcomes and the Effectiveness of Prenatal Care" *Health Services Research*, Vol. 32, No. 6, pp. 805–823.

Manski, Charles F. and David A. Wise. 1983. *College Choice in America*. Cambridge, MA: Harvard University Press.

Mortensen, E.L., KF. Michaelsen, S.A. Sanders, and J.M. Reinisch. 2002. "The Association between Duration of Breastfeeding and Adult Intelligence." *Journal of the American Medical Association*, Vol. 287, No. 18, pp. 2365-2371.

Murphy, K. and Barkley, R. 1996. "Prevalence of DSM-IV symptoms of ADHD in adult licensed drivers: implications for clinical diagnosis," *Journal of Attention Disorders* 1: 47-161.

Nelson, Melissa C., Penny Gordon-Larsen, and Linda S. Adair. 2005. "Are Adolescents Who Were Breast-fed Less Likely to Be Overweight?: Analyses of Sibling Pairs to Reduce Confounding," *Epidemiology* 16(2):247-253.

Nelson, Melissa C. and Penny Gordon-Larsen. 2006. "Physical Activity and Sedentary Behavior Patterns Are Associated with Selected Adolescent Health Risk Behaviors," *Pediatrics*, Vol. 117, No. 4, pp.1281-90.

Owen, Christopher G., Richard M. Martin, Peter H. Whincup, George Davey-Smith, Matthew W. Gillman and Derek G. Cook. 2005**.** "The Effect of Breastfeeding on Mean Body Mass Index Throughout Life: A Quantitative Review of Published and Unpublished Observational Evidence" *American Journal of Clinical Nutrition*, Vol. 82, No. 6, pp. 1298-1307.

Radloff, L. 1977. "The CES-D Scale: A Self-Report Depression Scale for Research in the General Population." *Applied Psychological Measurement*, Vol. 1, No. 3, pp. 385-401.

Roberts, RE, Lewinsohn PM, Seeley JR. 1991. "Screening for Adolescent Depression: A Comparison of Depression Scales." *Journal of the American Academy of Child and Adolescent Psychiatry*, Vol. 30, pp. 58-66.

Rosenberg, M. 1965. Society and the Adolescent Self-Image. Princeton, NJ: Princeton University Press.

Richards, Marcus and Rebecca Hardy. 2002. "Long-term Effects of Breast-Feeding in a National Birth Cohort: Educational Attainment and Midlife Cognitive Function." *Public Health Nutrition*, Vol. 5, No. 5, pp. 631-635.

Ryan, A.S. and Martinez, G.A., 1989. "Breast-feeding and the working mother: a profile," *Pediatrics* 83: 524–531.

Sabia, Joseph J. and Daniel I. Rees. 2008. "The Effect of Adolescent Virginity Status on Psychological Well-being," *Journal of Health Economics* (in press).

Shankaran, Seetha, Abhik Das, Charles R. Bauer, Henrietta S. Bada, Barry Lester, Linda L. Wright, and Vincent Smeriglio. 2004. "Association Between Patterns of Maternal Substance Use and Infant Birth Weight, Length, and Head Circumference." *Pediatrics*, Vol. 114, No. 2, pp. e226-e234.

Shrier, L.A., S.K. Harris, M. Sternberg, and W.R. Beardslee. 2001. “Associations of Depression, Self-Esteem, and Substance Use with Sexual Risk among Adolescents.” *Preventative Medicine*, Vol. 33, No. 3, pp.179-189.

US Department of Health and Human Services. 2005. Healthy People 2010 Midcourse Review. Washington, DC: U.S. Government Printing Office. Available at http://www.healthypeople.gov/data/midcourse

US Department of Health and Human Services. 2000. Healthy People 2010 (2nd ed). Washington, DC: U.S. Government Printing Office.

Victora, Cesar G., Fernando C. Barros, Bernardo L. Horta, Rosângela C. Lima. 2005. “Breastfeeding and School Achievement in Brazilian Adolescents.” Acta Paediatrica, Vol. 94, No.11, pp: 1656-60.

Table 1. Mean Breastfeeding Rates and Duration of Breastfeeding by Child's Educational Attainment

	Cumulative HS GPA			*HS Graduation*		*College Attendance*	
	Lower third	Middle third	Upper third	Dropout	HS Grad	No College	College
	(1)	(2)	(3)	(4)	(5)	(7)	(8)
	Panel I: Proportion Breastfed						
Breastfed	0.357	0.454	0.555	0.354	0.462	0.347	0.508
	(0.479)	(0.498)	(0.497)	(0.478)	(0.499)	(0.476)	(0.500)
	Panel II: Duration of Breastfeeding						
Months Breastfed	2.28	3.05	4.20	2.28	3.27	2.16	3.72
	(4.60)	(5.24)	(5.94)	(4.67)	(5.41)	(4.46)	(5.72)
	Panel III: Duration Categories						
1 Month ≤ Breastfed < 6 Months	0.217	0.269	0.283	0.215	0.255	0.214	0.271
	(0.412)	(0.443)	(0.450)	(0.411)	(0.436)	(0.411)	(0.444)
6 Months ≤ Breastfed < 12 Months	0.093	0.116	0.178	0.092	0.134	0.090	0.152
	(0.290)	(0.320)	(0.383)	(0.289)	(0.341)	(0.286)	(0.359)
Breastfed ≥ 12 Months	0.047	0.069	0.094	0.047	0.072	0.042	0.085
	(0.212)	(0.254)	(0.291)	(0.212)	(0.259)	(0.201)	(0.279)
N	3,367	3,395	3,439	2,189	10,471	5,049	7,602

*** Statistically significant at the 1% level; ** at the 5% level; * at the 10% level

Notes: Standard deviations are in parentheses. Means are unweighted and based on data drawn from Waves I and III of the National Longitudinal Study of Adolescent Health.

Table 2. Means of Dependent and Independent Variables

Dependent Variables		*Control Variables*		*Control Variables*	
Cumulative High School GPA (Transcript)[1]	2.59 (0.834)	Parent Post-College Education	0.096 (0.294)	Asian	0.056 (0.230)
Received High School Diploma (Excluding GED)	0.827 (0.378)	Single Parent	0.052 (0.223)	Indian	0.016 (0.125)
Attend College	0.601 (0.490)	Divorced	0.141 (0.348)	Hispanic/Other	0.161 (0.368)
Breastfeeding Variables		Separated	0.049 (0.215)	Class Size	26.4 (5.65)
Breastfed	0.443 (0.497)	Widowed	0.034 (0.182)	Public School	0.923 (0.266)
Months Breastfed	3.10 (5.31)	Rural	0.182 (0.386)	% Enrolled in college courses	47.4 (31.2)
1 Month ≤ Breastfed < 6 Months	0.248 (0.432)	Suburban	0.535 (0.499)	Female	0.53 (0.499)
6 Months ≤ Breastfed < 12 Months	0.127 (0.333)	West	0.236 (0.425)	Small School Size	0.161 (0.368)
Breastfed ≥ 12 Months	0.068 (0.252)	Midwest	0.260 (0.438)	Medium School Size	0.377 (0.485)
Control Variables		South	0.370 (0.483)	Birthweight	3311 (571.1)
Age at Wave 3	21.8 (1.74)	Catholic	0.261 (0.439)	Older sibling	0.499 (0.500)
Log Household Income	10.5 (0.811)	Baptist or Methodist	0.382 (0.486)	Family size at birth	1.59 (1.41)
Parent Completed High School	0.290 (0.454)	Other Christian	0.198 (0.398)		
Parent Trade School	0.095 (0.294)	Non-Christian Relig	0.041 (0.198)		
Parent Some College	0.200 (0.400)	Black	0.213 (0.409)		
Parent College Ed	0.147 (0.355)				

Note: Standard deviations appear in parentheses; N = 12,660. Means are unweighted.

[1]Sample restricted to those with non-missing transcript information on cumulative GPA; N = 10,201.

Table 3. OLS Estimates of Effect of Breastfeeding Duration on Educational Attainment

	Cumulative HS GPA	*HS Graduation*	*College Attendance*
	(1)	(2)	(3)
	Panel I: Breastfed (yes/no)		
Breastfed	0.118***	0.025***	0.069***
	(0.017)	(0.008)	(0.013)
	Panel II: Duration of Breastfeeding		
Months Breastfed	0.011***	0.002***	0.006***
	(0.001)	(0.001)	(0.001)
	Panel III: Duration Categories[1]		
1 Month ≤ Breastfed < 6 Months	0.088***	0.021**	0.057***
	(0.018)	(0.009)	(0.013)
6 Months ≤ Breastfed < 12 Months	0.159***	0.026**	0.074***
	(0.025)	(0.010)	(0.015)
Breastfed ≥ 12 Months	0.170***	0.038**	0.112***
	(0.029)	(0.013)	(0.017)
N	10,201	12,660	12,660

*** Statistically significant at the 1% level; ** at the 5% level; * at the 10% level

Notes: Standard errors corrected for clustering at the family level are in parentheses. Estimates are from unweighted OLS regressions based on data drawn from Waves I and III of the National Longitudinal Study of Adolescent Health. The full set of controls given are in Table 2. Regressions also include indicators for missing information for the controls.

[1]Omitted category consists of respondents who were not breastfed.

Table 4. OLS, Random Effects, and Fixed Effects Estimates of Effect of Breastfeeding Duration on Educational Attainment[1]

	Cumulative HS GPA			*HS Graduation*			*College Attendance*		
	OLS	RE	FE	OLS	RE	FE	OLS	RE	FE
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Panel I: Breastfed (yes/no)								
Breastfed	0.381***	0.340***	0.286**	0.011	0.002	0.025	0.050	0.042	0.073
	(0.145)	(0.141)	(0.110)	(0.046)	(0.043)	(0.040)	(0.069)	(0.068)	(0.059)
	[126]	[126]	[126]	[191]	[191]	[191]	[191]	[191]	[191]
	{59}	{59}	{59}	{90}	{90}	{90}	{90}	{90}	{90}
	Panel II: Duration of Breastfeeding								
Months Breastfed	0.014*	0.016**	0.017**	-0.000	-0.000	-0.001	0.009***	0.010***	0.014***
	(0.008)	(0.007)	(0.009)	(0.003)	(0.003)	(0.003)	(0.004)	(0.003)	(0.005)
	[333]	[333]	[333]	[459]	[459]	[459]	[459]	[459]	[459]
	{159}	{159}	{159}	{220}	{220}	{220}	{220}	{220}	{220}
	Panel III: Duration Categories[2]								
1 Month ≤ Breastfed < 6 Months	0.268**	0.260**	0.246**	-0.021	-0.017	0.021	0.073	0.034	0.012
	(0.113)	(0.109)	(0.111)	(0.041)	(0.039)	(0.044)	(0.060)	(0.058)	(0.060)
6 Months ≤ Breastfed < 12 Months	0.356***	0.401***	0.401***	0.013	0.016	0.032	0.110*	0.105	0.145**
	(0.130)	(0.121)	(0.124)	(0.048)	(0.044)	(0.052)	(0.067)	(0.065)	(0.072)
Breastfed ≥ 12 Months	0.298***	0.263*	0.238	-0.028	-0.028	-0.029	0.166**	0.171**	0.229***
	(0.146)	(0.145)	(0.163)	(0.059)	(0.058)	(0.065)	(0.075)	(0.072)	(0.084)
	[333]	[333]	[333]	[459]	[459]	[459]	[459]	[459]	[459]
	{159}	{159}	{159}	{220}	{220}	{220}	{220}	{220}	{220}

*** Statistically significant at the 1% level; ** at the 5% level; * at the 10% level

Notes: Sample sizes are in brackets and unique families are in braces. Standard errors corrected for clustering at the family level are in parentheses. Estimates are from unweighted OLS regressions based on data drawn from Waves I and III of the National Longitudinal Study of Adolescent Health. The full set of controls given in Table 2 are included in the ordinary least squares (OLS) and random effects (RE) models. Family Fixed effects (FE) models include controls for age of the respondent at Wave III, family size at the time of the respondent's birth, whether the respondent has an older sibling, birth weight of the respondent, and sex of the respondent.

[1]The sample is restricted to siblings with non-missing information on the educational outcome of interest, breastfeeding duration, and sex. In Panel I, the sample is limited to sibling pairs where one sibling is breastfed and one is not. In Panels II and III, the sample also includes siblings with different breastfeeding durations.
[2]Omitted category consists of respondents who were not breastfed.

Table 5. Sensitivity of Fixed Effects Estimates to Controls for Sibling Heterogeneity[1]

	(1)	(2)	(3)	(4)	(5)
	Panel I: Cumulative HS Grades				
Months Breastfed	0.017**	0.021**	0.019**	0.017**	0.018**
	(0.009)	(0.009)	(0.009)	(0.008)	(0.009)
	[333]	[333]	[333]	[333]	[333]
	{159}	{159}	{159}	{159}	{159}
	Panel II: HS Graduation				
Months Breastfed	-0.001	0.002	-0.002	-0.002	0.001
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)
	[459]	[459]	[459]	[459]	[459]
	{220}	{220}	{220}	{220}	{220}
	Panel III: College Attendance				
Months Breastfed	0.014***	0.016***	0.014***	0.014***	0.014***
	(0.005)	(0.004)	(0.005)	(0.004)	(0.004)
	[459]	[459]	[459]	[459]	[459]
	{220}	{220}	{220}	{220}	{220}

*** Statistically significant at the 1% level; ** at the 5% level; * at the 10% level

Notes: Sample sizes are in brackets and unique families are in braces. Standard errors corrected for clustering at the family level are in parentheses. Estimates are from unweighted OLS regressions based on data drawn from Waves I and III of the National Longitudinal Study of Adolescent Health. All models include controls for age of the respondent at Wave III, family size at the time of the respondent's birth, whether the respondent has an older sibling, birth weight of the respondent, and sex of the respondent. Model (2) adds controls for the quality of the mother-child relationship; Model (3) includes controls for parental involvement in the child's education; Model (4) adds controls for the respondent's retrospective ADHD score and parent's report of the child having a temper; and Model (5) includes all of the above controls.

[1]The sample is restricted to siblings with non-missing information on the educational outcome of interest, breastfeeding duration, and sex. In all panels, the sample is limited to siblings with different breastfeeding durations.

Table 6. Falsification Tests[1]

	RSE Scale	Depress	Drunk ≥ 3X/Year	Smoked last 30 days	TV Hours per Week	Sports ≥ 3X/Year	BMI-for-age-sex Percentile
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Panel I: Breastfed (yes/no)						
Breastfed	0.656	0.012	0.026	0.046	-0.574	0.010	-1.16
	(0.427)	(0.038)	(0.041)	(0.044)	(1.44)	(0.049)	(3.05)
	[263]	[263]	[263]	[263]	[263]	[263]	[250]
	{125}	{125}	{125}	{125}	{125}	{125}	{119}
	Panel II: Duration of Breastfeeding						
Months Breastfed	0.044	0.000	0.004	-0.000	-0.070	-0.005	0.137
	(0.034)	(0.003)	(0.003)	(0.004)	(0.140)	(0.005)	(0.241)
	[602]	[602]	[602]	[602]	[602]	[602]	[572]
	{286}	{286}	{286}	{286}	{286}	{286}	{273}
	Panel III: Duration Categories [2]						
1 Month ≤ Breastfed	0.623	0.026	0.013	0.060	-0.367	0.001	-0.226
< 6 Months	(0.444)	(0.039)	(0.044)	(0.048)	(1.46)	(0.053)	(3.22)
6 Months ≤ Breastfed	0.324	-0.004	0.014	-0.014	-1.13	0.053	1.16
< 12 Months	(0.503)	(0.045)	(0.051)	(0.058)	(1.75)	(0.067)	(3.95)
Breastfed ≥ 12 Months	1.01	0.045	0.075	0.033	-1.98	-0.041	1.35
	(0.667)	(0.055)	(0.061)	(0.067)	(2.85)	(0.091)	(4.73)
	[602]	[602]	[602]	[602]	[602]	[602]	[572]
	{286}	{286}	{286}	{286}	{286}	{286}	{273}

*** Statistically significant at the 1% level; ** at the 5% level; * at the 10% level
Notes: Sample sizes are in brackets and unique families are in braces. Standard errors corrected for clustering at the family level are in parentheses. Estimates are from unweighted family fixed effects regressions based on data drawn from Waves I and III of the National Longitudinal Study of Adolescent Health. All models include controls for age of the respondent at Wave III, family size at the time of the respondent's birth, whether the respondent has an older sibling, birthweight of the respondent, and sex of the respondent.

[1]The sample is restricted to siblings with non-missing information on the educational outcome of interest, breastfeeding duration, and sex. In Panel I, the sample is limited to sibling pairs where one sibling is breastfed and one is not. In Panels II and III, the sample also includes siblings with different breastfeeding durations.

[2]Omitted category consists of respondents who were not breastfed.

Table 7. Examining Factors that Mediate Fixed Effects Estimates of Relationship Between Breastfeeding and Educational Attainment[1]

	Cumulative HS GPA				*College Attendance*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel I: Breastfed (yes/no)							
Breastfed	0.286**	0.218*	0.203	0.097	0.073	0.060	0.067	0.055
	(0.110)	(0.115)	(0.136)	(0.132)	(0.059)	(0.059)	(0.068)	(0.067)
Wave I PVT Score		0.019***		0.025***		0.004		0.005
		(0.007)		(0.007)		(0.005)		(0.005)
1-2 Absences Wave I			-0.397*	-0.294			0.103	0.116
			(0.223)	(0.219)			(0.140)	(0.139)
3-9 Absences Wave I			-0.237	-0.313			0.043	0.063
			(0.241)	(0.228)			(0.169)	(0.170)
10+ Absences Wave I			0.217	0.142			0.104	0.073
			(0.344)	(0.361)			(0.187)	(0.188)
1-2 Absences Wave II			-0.508**	-0.407*			-0.115	-0.094
			(0.237)	(0.218)			(0.140)	(0.160)
3-9 Absences Wave II			-0.248	-0.293			-0.246*	-0.244
			(0.196)	(0.207)			(0.138)	(0.165)
10+ Absences Wave II			-0.422	-0.449			-0.133	-0.135
			(0.423)	(0.412)			(0.211)	(0.223)
Illness Index Wave I			-0.091	-0.097			0.011	0.004
			(0.080)	(0.082)			(0.030)	(0.028)
Illness Index Wave II			0.014	0.060			-0.008	-0.002
			(0.099)	(0.090)			(0.040)	(0.039)
	[126]	[126]	[126]	[126]	[191]	[191]	[191]	[191]
	{59}	{59}	{59}	{59}	{90}	{90}	{90}	{90}

Table 7 Continued

	Cumulative HS GPA				*College Attendance*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel II: Duration of Breastfeeding							
Months Breastfed	0.017**	0.014*	0.009	0.006	0.014***	0.013***	0.012***	0.011**
	(0.009)	(0.008)	(0.008)	(0.008)	(0.005)	(0.005)	(0.005)	(0.005)
Wave I PPVT Score		0.015***		0.015***		0.007**		0.007**
		(0.004)		(0.004)		(0.003)		(0.005)
1-2 Absences Wave I			-0.145	-0.128			0.110	0.116
			(0.156)	(0.144)			(0.101)	(0.095)
3-9 Absences Wave I			-0.119	-0.125			0.136	0.128
			(0.150)	(0.144)			(0.103)	(0.099)
10+ Absences Wave I			0.176	-0.154			0.223*	0.222
			(0.243)	(0.213)			(0.127)	(0.134)
1-2 Absences Wave II			-0.400***	-0.360**			-0.121	0.108
			(0.154)	(0.146)			(0.080)	(0.080)
3-9 Absences Wave II			-0.390***	-0.402***			-0.159*	-0.167**
			(0.137)	(0.135)			(0.082)	(0.081)
10+ Absences Wave II			-0.678***	-0.668***			-0.265**	-0.264**
			(0.205)	(0.205)			(0.109)	(0.106)
Illness Index Wave I			-0.045	-0.039			-0.009	-0.004
			(0.038)	(0.037)			(0.020)	(0.020)
Illness Index Wave II			-0.051	-0.047			-0.007	-0.006
			(0.051)	(0.050)			(0.024)	(0.024)
	[333]	[333]	[333]	[333]	[459]	[459]	[459]	[459]
	{159}	{159}	{159}	{159}	{220}	{220}	{220}	{220}

Table 7 Continued

	Cumulative HS GPA				*College Attendance*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel III: Duration Categories[2]							
1 Month ≤ Breastfed < 6 Months	0.246**	0.204*	0.192*	0.138	0.012	0.000	-0.006	-0.023
	(0.111)	(0.108)	(0.116)	(0.112)	(0.060)	(0.060)	(0.061)	(0.061)
6 Months ≤ Breastfed < 12 Months	0.401***	0.332***	0.262*	0.183	0.145**	0.126*	0.099	0.074
	(0.124)	(0.124)	(0.134)	(0.130)	(0.072)	(0.073)	(0.072)	(0.072)
Breastfed ≥ 12 Months	0.238	0.208	0.115	0.078	0.229***	0.221**	0.198**	0.187**
	(0.163)	(0.162)	(0.158)	(0.158)	(0.084)	(0.083)	(0.084)	(0.083)
Wave I PPVT Score		0.014***		0.014***		0.007**		0.007**
		(0.002)		(0.004)		(0.003)		(0.003)
1-2 Absences Wave I			-0.129	-0.118			0.113	0.118
			(0.155)	(0.144)			(0.102)	(0.097)
3-9 Absences Wave I			-0.114	-0.121			0.141	0.133
			(0.148)	(0.142)			(0.105)	(0.101)
10+ Absences Wave I			0.153	-0.141			0.217	0.219*
			(0.206)	(0.207)			(0.139)	(0.127)
1-2 Absences Wave II			-0.385***	-0.354**			-0.116	-0.104
			(0.151)	(0.145)			(0.082)	(0.081)
3-9 Absences Wave II			-0.348**	-0.372***			-0.151*	-0.163**
			(0.135)	(0.134)			(0.014)	(0.082)
10+ Absences Wave II			-0.628***	-0.635***			-0.250**	-0.254**
			(0.202)	(0.202)			(0.114)	(0.111)
Illness Index Wave I			-0.050	-0.044			-0.006	-0.001
			(0.040)	(0.035)			(0.020)	(0.020)
Illness Index Wave II			-0.055	-0.050			-0.010	-0.008
			(0.054)	(0.053)			(0.024)	(0.024)

Table 7 Continued

	Cumulative HS GPA				*College Attendance*			
	(1)	(2)	(3)	(4)	(5)	(6)	(5)	(6)
	[333]	[333]	[459]	[459]	[459]	[459]	[459]	[459]
	{159}	{159}	{220}	{220}	{220}	{220}	{220}	{220}

*** Statistically significant at the 1% level; ** at the 5% level; * at the 10% level
Notes: Sample sizes are in brackets and unique families are in braces. Standard errors corrected for clustering at the family level are in parentheses. Estimates are from unweighted family fixed effects regressions based on data drawn from Waves I and III of the National Longitudinal Study of Adolescent Health. All models include controls for age of the respondent at Wave III, family size at the time of the respondent's birth, whether the respondent has an older sibling, birthweight of the respondent, and sex of the respondent.

[1]The sample is restricted to siblings with non-missing information on the educational outcome of interest, breastfeeding duration, and sex. In Panel I, the sample is limited to sibling pairs where one sibling is breastfed and one is not. In Panels II and III, the sample also includes siblings with different breastfeeding durations.

[2]Omitted category consists of respondents who were not breastfed.

Appendix Table 1. Means of Education and Falsification Outcome Variables for Sibling Sample

	Any Breastfeeding Sibling Sample (Panel I Sample)	Months Breastfeeding Sibling Sample (Panels II, III Sample)
Cumulative High School GPA (Transcript)	2.66 (0.880)	2.77 (0.849)
Received High School Diploma (Excluding GED)	0.838 (0.370)	0.834 (0.372)
Attend College	0.586 (0.494)	0.649 (0.478)
RSE Scale	24.7 (3.70)	24.6 (3.58)
Depress	0.091 (0.289)	0.086 (0.281)
Drunk ≥ 3X per Year	0.175 (0.381)	0.173 (0.378)
Smoked last 30 days	0.255 (0.357)	0.269 (0.444)
TV Hours per Week	14.8 (12.3)	14.6 (13.2)
Sports ≥ 3X per Week	0.548 (0.499)	0.527 (0.500)
BMI-for-age-sex percentile	60.3 (28.8)	56.5 (29.1)

Note: Standard deviations appear in parentheses. Means are unweighted. Sample sizes correspond to those reported in Tables 4 and 6.

Appendix Table 2. Mean proportion of siblings who differ on quality of the parent-child relationship, schooling involvement, and child temperament

	Any Breastfeeding Sibling Sample (Panel I Sample)	Months Breastfeeding Sibling Sample (Panels II, III Sample)
Perceived closeness of child to biological mother	0.450 (0.498)	0.460 (0.499)
Perceived belief that biological mother cares for child	0.162 (0.369)	0.179 (0.383)
Parent reports getting along well with child	0.390 (0.489)	0.418 (0.494)
Parent reports that she and child make decisions about life together	0.523 (0.500)	0.490 (0.500)
Parent talked with child about school grades	0.411 (0.493)	0.423 (0.495)
Paerent worked with child on project in school	0.213 (0.411)	0.205 (0.404)
Parent talked with child about other school issues	0.393 (0.489)	0.423 (0.495)
Parent disappointed if child does not complete HS	0.577 (0.495)	0.593 (0.492)
Parent disappointed if child does not attend college	0.192 (0.395)	0.233 (0.423)
Parent's view on child identity in HS (brilliance, leader, etc.)	0.252 (0.453)	0.266 (0.442)
ADHD Scale	0.970 (0.171)	0.952 (0.214)
Parent's report of child having bad temper	0.339 (0.475)	0.344 (0.476)

Note: Standard deviations appear in parentheses. Means are unweighted. Sample sizes correspond to those reported in Tables 4 and 6.