Breastfeed, If You Choose: Parental Context and the Long-Term Legacy of Lactation

Mike Cassidy
Rutgers University

June 23, 2020

Abstract

Despite consensus among medical authorities about the desirability of breastfeeding, causal evidence about its effects is surprisingly scant. Using a thorough collection of empirical approaches and detailed longitudinal data spanning five decades, I investigate a comprehensive set of outcomes with greater breadth and continuity than previous work. On average (per OLS), breastfeeding is associated with modest and persistent cognitive advantages from childhood through young adulthood—even after controlling for an extensive set of confounding forces. Accounting for breastfeeding duration strengthens these relationships and uncovers favorable labor market and fertility linkages as well. But there is no evidence for enduring health benefits. At the same time, a novel extended family fixed effects analysis comparing differentially breastfed siblings and cousins finds little association between breastfeeding and any outcome. I argue these findings are not mutually exclusive by providing evidence that, contrary to conventional wisdom, the divergent estimates are the consequence of considerable negative selection into the subset of families contributing to fixed effects identification. (*JEL J13, I12, I18, I21*)

Acknowledgments: Generous funding has been provided by the National Academy of Education and the National Academy of Education/Spencer Dissertation Fellowship Program. Carolyn Moehling, Jennifer Hunt, Roger Klein, and Brendan O'Flaherty, and Rosanne Altshuler have offered invaluable mentorship and guidance. I would also like to thank my wife, Molly, and son, Max, whose tenure breastfeeding coincided with this study. Without their love, support, and inspiration, this paper would not have been possible. Whatever the long-term effects of breastfeeding may be, Max is a fan.

1 Introduction

Breastfeeding is nature's way of nourishing infants, so it is hardly surprising that it is associated with an array of benefits. The medical literature is replete with studies finding breastfed babies are healthier—quicker to acquire immunity and less susceptible to the infectious respiratory and gastrointestinal maladies of infancy. As children and young adults, they exhibit augmented intelligence, lower likelihoods of being overweight, and attenuated incidences of cardiovascular and metabolic diseases. Their mothers experience improved birth spacing and reduced propensities for reproductive cancers[1]. The evidence is seen as sufficiently strong that, since 1997, the American Academy of Pediatrics has decreed six months of exclusive breastfeeding as a public health goal—a sentiment echoed by the World Health Organization and U.S. Department of Health and Human Services[2].

Despite this consensus, the causal evidence to support these claims is surprisingly weak. Virtually the entire literature is observational, and many studies fail to convincingly address confounding factors. A fundamental challenge complicates evaluation: breastfeeding is the outcome of a parental optimization problem and not randomly assigned[3]. Consequently, favorable correlations between breastfeeding choices and wellbeing may be the artifact of unobserved heterogeneity. Several recent systematic meta-analyses underscore this concern: as a rule, the more rigorous the study, the more modest, if any, are breastfeeding's salubrious associations (see, e.g., Ip et al. (2007); Horta et al. (2013); Victora et al. (2016)). Beyond health advantages during infancy—for which both empirical documentation and biological plausibility is strong—the claims on breastfeeding's behalf consist mostly of informed speculation. Further complicating matters is generalizability: breastfeeding is studied in settings rich and poor, with scopes micro and macro, and across margins intensive and extensive, raising oft-neglected issues of external validity.

Work by economists and other social scientists placing greater priority on causal inference emphasizes the theme of apparent advantage obscured by endogeneity. With cognitive performance the preferred topic of inquiry[4], multivariate regression and propensity score matching approaches almost always document favorable associations between breastfeeding various measures of achievement, from childhood intelligence tests[5] to educational attain-

[1]There is an extensive medical literature studying breastfeeding. Horta et al. (2007); Ip et al. (2007); Salone, Vann Jr and Dee (2013); Horta et al. (2013); Dieterich et al. (2013); Horta, Loret De Mola and Victora (2015); Victora et al. (2016) offer perhaps the best systematic reviews to date; the many studies cited therein comprise the evidence discussed in these first three paragraphs.

[2]See Eidelman and Schanler (2012); US Department of Health and Human Services et al. (2011); U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion (2020); World Health Organization (2018).

[3]Their is a rich literature on the determinants of breastfeeding. See, e.g., Rollins et al. (2016); Dennis (2002); Dieterich et al. (2013).

[4]For a helpful summary of research in economics on breastfeeding and cognitive performance through 2013, see Rothstein (2013).

[5]Evenhouse and Reilly (2005); Der, Batty and Deary (2006); Denny and Doyle (2010); Jiang, Foster

ment[6] and earnings[7]. Nearly as often, these studies find advantages attenuate in the face of mother fixed effects, or simply adjusting for correlated parenting behaviors[8]. The same is true of health and noncognitive outcomes evaluated by any method: there appear few impacts enduring beyond infancy[9]. These findings are in keeping with the one randomized control trial of breastfeeding, outside of economics, which documents small cognitive gains but few health advantages beyond infancy (Kramer et al., 2001, 2007, 2008). While breastfeeding's purported benefits have deemed additional RCT's unethical, economists have contributed quasi-experimental evidence in the form of instrumental variable approaches exploiting such things as breastfeeding promotion programs, non-elective C-sections, and hospital staffing patterns[10]. While raising questions of validity and power, these results generally tend towards null effects.

As a theoretical matter, it is not obvious breastfeeding would be unambiguously preferred. On one hand, the medical literature posits several hypotheses as to why breastfeeding may be advantageous. One is the biochemical properties of breastmilk. In addition to its macronutrient composition—evolutionarily attuned to adapt to infants' ever-evolving nutritional needs—human milk is abundant in immune-boosting antibodies, awash in essential minerals, vitamins, hormones, and enzymes, and rich in long-chain fatty acids, which are thought to promote neural development[11]. But more important than covalent bonding may be that between mother and child: intimate time together may catalyze the gains for which milk is credited. Indeed, the most important changes may occur in the mother, with breastfeeding stimulating nurturing instincts[12].

On the other hand, the scientifically-engineered alternative—infant formula—has progressed to the point where, at least from a nutritional standpoint, it is virtually equivalent to mother's milk[13]. Likewise, lactation is neither a necessary nor a sufficient condition for conscientious parenting or quality time together. In addition, breastfeeding imposes nontrivial physical and time costs on women, potentially compromising other elements of child wellbeing[14].

In this paper, I bring the causal question to the fore. Using a expansive set of empirical approaches in the context of rich longitudinal data, I investigate a comprehensive collection

and Gibson-Davis (2011); Belfield and Kelly (2012); Del Bono and Rabe (2012); Borra, Iacovou and Sevilla (2012); Rothstein (2013); Fitzsimons and Vera-Hernández (2013); Onda et al. (2016).

[6]Rees and Sabia (2009).

[7]Cesur et al. (2017).

[8]Evenhouse and Reilly (2005); Der, Batty and Deary (2006); Gibson-Davis and Brooks-Gunn (2006); Rees and Sabia (2009); Rothstein (2013); Gibbs and Forste (2014); Colen and Ramey (2014); Cesur et al. (2017).

[9]Baker and Milligan (2008); Del Bono and Rabe (2012); Fitzsimons and Vera-Hernández (2013).

[10]Denny and Doyle (2010); Del Bono and Rabe (2012); Fitzsimons and Vera-Hernández (2013).

[11]For details, see Lessen and Kavanagh (2015); Martin, Ling and Blackburn (2016); Victora et al. (2016).

[12]World Health Organization (2009); Rothstein (2013).

[13]See Stevens, Patrick and Pickler (2009); Martin, Ling and Blackburn (2016).

[14]Rollins et al. (2016).

of educational, health, labor, and behavioral outcomes from childhood through young adulthood, offering a perspective novel in its scope and temporal integrity. My contribution is twofold. First, where prior research has settled for snapshots—infants, children, adolescents, *or* young adults—of disjoint domains—cognition, education, labor, *or* health—I provide a synthesis, evaluating a broad spectrum of subjects among the same individuals from ages 5 through 25 years, imparting continuity and coherence that has been lacking. Second, I harmonize previously contradictory results by introducing a new identification strategy to the study of breastfeeding while characterizing the extant biases of observational econometric estimates.

Good data makes this possible. I use the 1979 National Longitudinal Survey of Youth Child and Young Adult cohort (NLSY-CYA), which, through the 2016 cycle, follows 11,530 individuals from birth through adulthood in the United States. Featuring contemporaneous queries on infant feeding and extraordinary detail describing children and their mothers, accumulated through biennial surveys spanning five decades, the NLSY-CYA is especially well-suited to the study of breastfeeding in the rich country context.

Given this extensive detail and following convention, I begin with OLS linear regression. I augment standard demographic and socioeconomic controls with covariates encompassing maternal intelligence, employment, health, perinatal behaviors, and home environments, as well as children's birth circumstances—a collection rarely, if ever, matched in the breastfeeding literature. This helps lay claim to more a credible selection-on-observables story than is usually possible. Per OLS, breastfeeding is associated with early gains in cognitive achievement—by about 0.1 of a standard deviation (SD) on math, reading, and vocabulary intelligence tests—that translate into 2–3 percentage point (pp) gains in the probabilities of completing high school and attending college by age 25.

These results, which are in line with prior estimates by economists, pertain to breastfeeding initiation, which is the simplest treatment definition and the measure most commonly used in the literature. Other gains appear on the intensive margin. Individuals breastfed at least three months are 3 pp less likely to receive public benefits (welfare, Food Stamps, or medical insurance) at age 21 and are 6 pp more likely to be employed at 25—advantages not apparent on the extensive margin. They are also 4 pp less likely to have a child before marriage. Early cognitive gains exhibit dose-response as well, peaking near 0.2 SD's among those breastfeed between six and twelve months.

In contrast, there is little evidence for enduring health benefits, and, indeed, a positive association between having been breastfed and reporting health problems in childhood and young adulthood. Given breastfeeding's known benefits during infancy, this is puzzling.

These results assume no omitted variables. My second identification strategy, mother fixed effects, is also common in the breastfeeding literature. By narrowing the unit of comparison to siblings with varied breastfeeding experiences, this approach implicitly differences out family-level unobservables invariant among offspring. In keeping with the existing literature, breastfeeding appears to matter little among siblings, with precise nulls estimated for nearly all outcomes.

As with OLS, mother fixed effects relies on strong assumptions: mothers who, by choice or necessity, breastfeed one child and not another might compensate in other ways, potentially creating biases worse than the omissions of OLS. To address this concern, I introduce a new identification approach to the study of breastfeeding: extended family fixed effects. Made possible because the NLSY includes linkages between mothers and their sisters, this strategy includes cousins as well as siblings in the within-family comparison set, thus rendering more explicable contrasting breastfeeding choices while retaining some of the shared genetic and parental background features that make family fixed effects attractive[15]. It confirms the nuclear family findings: among kin, breastfeeding appears to make little difference in cognitive or other outcomes.

It is tempting to view fixed effects results as a corrective of OLS. I provide evidence—previously lacking in the breastfeeding literature (see, e.g., (Colen and Ramey, 2014; Rees and Sabia, 2009; Rothstein, 2013))—that this is a mistake. The subset of families contributing to fixed effects identification is significantly negatively selected. Among other things, mothers from inconsistently-breastfeeding families have markedly lower educational attainment, income, and intelligence scores than their counterparts from families with uniform feeding behaviors. Their children have worse outcomes by almost all measures; for example, they score about 0.2 SD's worse on cognitive assessments and are 5 pp less likely to have college experience by age 25. In addition, when these mothers breastfeed, they do so 30 percent shorter, on average, than their breastfeeding-consistent peers. Excluding these incongruously fed kin from the OLS analysis boosts many coefficients by between a fifth and a half, rendering some newly statistically significant.

Put differently, null fixed effects findings are, at least in part, the consequence non-intensive breastfeeding in disadvantaged environments. This turns the conventional wisdom on its head: it would appear downward bias in fixed effects is at least as large a concern as the inverse bias in OLS.

My preferred interpretation is that breastfeeding is, on average, associated with modest, and persistent, intellectual advantages. More complex outcomes effectuated during the young adult years—labor and fertility behaviors derivative of cognitive antecedents—require more sustained doses. At the same time, there are reasons to believe at least a portion of these gains are attributable to parenting quality in general, rather than breastfeeding per se. Breastfeeding's abundance of auspicious associations makes it difficult to ensure all else is equal; to this point, dose-response relationships may reflect augmented selectivity

[15]While new to breastfeeding, the so-called "cousin" fixed effects approach has been used in other settings (e.g., (Barclay, Lyngstad and Conley, 2018; Duncan et al., 2018*a*)).

among more intensively-breastfeeding mothers (e.g., those with access to generous parental leave). Even more plainly, the fixed effects analysis demonstrates that parents and their circumstances can compensate for infant feeding in other ways. Among the NLSY cohort, this remuneration appears generally disadvantageous, but there of course remains scope for salutary redress as well. To the extent there is a true effect of breastfeeding, it likely lies somewhere in the middle: smaller than covariate-adjusted linear associations but greater than fixed effect nulls.

The takeaway for policymakers, physicians, and public health professionals is that breastfeeding, at least in the rich country context, is desirable, but not dispositive. For most children, it is likely to be an advantage, but for others, accumulating these assets may be achievable via alternative routes—or stymied by roadblocks. Encouraging breastfeeding among families least likely to pursue (or sustain) it own their own may be especially important, as these are likely the children with the most to gain. However, in so doing, it is essential to acknowledge the ex ante costs of breastfeeding for these families may outweigh its benefits, potentially implicating deleterious trade-offs (e.g., employment) without adequate supports (e.g., paid family leave). Policies to promote breastfeeding must explicitly address opportunity costs and constraints, putting in place necessary foundations to ensure that, especially for disadvantaged populations, breastfeeding does not compromise other aspects of child or parent wellbeing.

2 Data

My data consists of the 1979 National Longitudinal Survey of Youth Child and Young Adult cohort (NLSY-CYA), which studies the 11,530 biological children of the original female members of the NLSY79. Sponsored by the U.S. Bureau of Labor Statistics (BLS), the NLSY79 thoroughly chronicles the characteristics and experiences of Americans born between 1957 and 1964 (and aged 14–22 in 1979). The NLSY-CYA charts the analogous attributes of their children, detailing their cognitive, physical, social, and behavioral development, as well as their educational, health, labor market, and familial experiences[16]. Born between 1970 and 2014, these children (and their mothers) have been interviewed biennially since 1986 and most recently in 2016[17]. Data is collected through four survey instruments: a child assessment administered to those 14 years of age and younger, a young adult questionnaire for those 15 and older, and a mother supplement augmenting the NLSY proper.

[16]The Center for Human Resource Research (CHRR) at The Ohio State University manages the NLSY79 and NLSY-CYA and the National Opinion Research Center (NORC) at the University of Chicago performs the actual interviews. The National Institute for Child Health and Human Development co-directs the NLSY-CYA along with BLS. For more, see Bureau of Labor Statistics, U.S. Department of Labor, and National Institute for Child Health and Human Development (2019).

[17]Interview rounds occur in even-numbered years.

I transform the raw NLSY-CYA data, for which the unit of observation is the child[18], by age-aggregating survey round responses into pooled cross sections corresponding to the five "outcome age" groups I study: ages 5, 10, 13, 21, and 25. Child i's age a outcome Y_{ia} comes from the earliest survey round r for which the child's age in years satisfies the relation $0 \leq age_{ir} - a \leq 2$ and the child was successfully interviewed. That is, outcomes are derived from the interviews most proximate to the pivotal age without going younger, given crossing thresholds such as 5 or 21 years can confer discontinuous social and legal changes. Correspondingly, time-varying covariates are culled from the interviews most proximate to, or immediately preceding, a child's birth, in order to avoid endogenous controls.

Table 1 illustrates the scope of the NLSY-CYA data. Each row describes the universe of individuals relevant to identification for an estimation strategy (to be described below). Cells give sample sizes, with all children in Column 2, those with breastfeeding data in Col 3, and interview participation at outcome ages enumerated in the ensuing columns. Response rates at my outcome ages of interest start at about 70 percent at age five and decline to about half by age 25 (with some of the reduction attributable individuals too young to have yet responded). Only 37 percent of children participate across all five age checkpoints. The numbers of individuals contributing identifying variation to breastfeeding in the fixed effects models (i.e., siblings or cousins with different breastfeeding statuses) are about a third of the sample.

In the following subsections, I focus on the aspects of the data most important to studying the long-term effects of breastfeeding[19].

2.1 Breastfeeding

The NLSY is especially well-suited to answer questions about the long-term effects of breastfeeding. Mothers who have given birth since the date of the last interview are asked two questions. The first is binary—"Was child breastfed?"—and the second records breastfeeding duration in weeks. Temporal proximity enhances recall and is a feature not always present in other datasets commonly used to study breastfeeding[20].

[18]For simplicity, I frequently refer to respondents as "children" even subsequent to attaining "young adult" status.

[19]For a thorough discussion of the data, including much more detail about variable definitions, see the extensive documentation BLS makes available at `https://www.nlsinfo.org/`, as well as the the Stata code for this paper.

[20]For example, the National Longitudinal Survey of Adolescent Health, in contrast, asks breastfeeding questions retrospectively of adolescents. Nevertheless, the NLSY-CYA is not exclusively contemporaneous. Births prior to 1979 (11 percent), when the mother interviews began, and prior to 1986 (51 percent), when child assessments commenced, rely partly on recall and the associated loss in precision.

2.2 Outcomes

The first three age-outcome groups concern children; the latter two correspond to young adults. Table A.1 details the distribution of age-outcome ages. About 95 percent of age-outcome ages are within one year of the pivotal age, and split evenly between them (e.g., for outcome-age 25, 50.3 percent of respondents are 25 at the time of their outcome measurement, 46.2 percent are 26 years, and 3.5 percent are 27 years). I include indicators for these ages in all regressions to control for micro-age discrepancies[21].

The following subsections provide an overview of the outcomes I assess in the main text. Detailed definitions of these outcomes, as well as a discussion and analysis of several alternative outcome measures, are available in Appendices B and D.

2.2.1 Child Outcomes

During childhood, I focus on three domains: cognitive performance, behavior, and health. Cognitive outcomes consist of four assessments: Peabody Individual Achievement Tests (PIAT) for math, reading recognition, and reading comprehension, as well as the Peabody Picture Vocabulary Test (PPVT). These tests are among the most common measures intelligence and are known to have high reliability and validity. I used normed scores of mean zero and standard deviation one, so that regression results can be interpreted in standard deviation units, with the caveat that these norms are in reference to external norming samples with somewhat lower average performance than the NLSY-CYA cohort. I evaluate behavioral outcomes using the Behavioral Problems Index (BPI), which is also standard in the NLSY-CYA literature and analogously scaled as a standard normal. Higher scores indicate worse behaviors; the within-NLSY mean is somewhat above zero. I use two measures of health. The first is broad a binary indicator for health problem, which include school-limiting conditions, illnesses requiring medical attention, and subjective fair or poor health. The second is an binary indicator for overweight, defined as a body mass index (BMI) of 21 or greater.

2.2.2 Young Adult Outcomes

The young adult outcomes span an analogous four domains: educational attainment, labor market experiences, health, and behavior. For education, I consider binary indicators for high school graduation and college experience, as well as a count of years of education. For labor, I assess indicators for employed, either in school or working, and public benefit receipt (public assistance, Food Stamps, and/or public health insurance), as well as log earned income (with earnings in 2019 dollars). For health, I define health problem and overweight

[21] I analogously assign "age-outcome years" to the interview round corresponding with the "age-outcome" age, and include it as a covariate to control for year-of-interview effects.

indicators symmetric with the child versions. For behavior, I consider an indicator for premarital childbearing.

2.3 Covariates

Throughout the analysis, I control for a rich set of factors jointly relevant to breastfeeding decisions and child outcomes. With the exception of home environment scores measured at ages 0–2, all covariates are defined at the time of, or in the year prior to, a child's birth. Given it is common for individuals to have missing data on at least one control, I append each covariate with a dummy reflecting unknown values.

These covariates cover three broad subjects, and are comprised as follows (with mutually exclusive indicators in parentheses and respective "unknown" categories left implicit):

- **Mother Characteristics**: education (less than high school, high school graduate, some college or more), race (Hispanic, White, Black), age (under 21, 21–24, 25–29, 30–50), year of birth ($1957, 1958, \ldots, 1965$), region (Northeast, North Central, South, West), marital status (never married, married, other), nativity (U.S. native, foreign-born), Armed Forces Qualifying Test (AFQT) quartile, employment (no, yes; during 12 months prior to birth), family income quartile (12 months prior to birth)), age 21 BMI quartile, and having a sister in the NLSY (yes, no).

- **Mother Pregnancy Behaviors**: prenatal doctor's visit (no, yes), prenatal vitamin consumption (no, yes), smoking (no, yes; during 12 months prior to birth), and alcohol consumption (no, yes; during 12 months prior to birth).

- **Child Characteristics and Birth Circumstance**s: birth order (1, 2, 3+), female sex (no, yes), birth weight quartile, cesarean delivery (no, yes), preterm birth (no, yes; defined as < 37 weeks gestation), postnatal hospital stay longer than mother (no, yes), and Home Observation Measurement of the Environment (HOME) quality quartile (measured at age 0–2), as well as a quadratic in child's birth year and indicators for birth month.

I also define indicators for children's ages, in years, during the survey rounds their responses for particular outcome ages occur (to control for short-term age effects), as well as indicators for the interview years themselves (to control for response-year effects). Henceforth I refer to this collection of controls as "full" covariates.

3 Empirical Approach

The central causal challenge is that women who choose to breastfeed—or who choose to breastfeed certain children and not others—may be different than those who do not. I pursue two identification strategies[22].

3.1 OLS

My baseline estimation approach is multivariate linear regression estimated via ordinary least squares (OLS). For individual i, outcome Y is estimated as:

$$Y_i = \mathbf{X}_i\boldsymbol{\beta} + \tau^{OLS}B_i + \boldsymbol{\varepsilon}_i \tag{1}$$

where $\mathbf{X}$ denotes the full set of mother, pregnancy behavior, and child covariates described in Section 2, $\boldsymbol{\varepsilon}$ subsumes unobservables, and B represents breastfeeding status[23]. For reasons of simplicity and accuracy of recall, I focus on the $B_i \in \{0, 1\}$, an indicator equal to one if child i is ever breastfed and zero otherwise (i.e., breastfeeding initiation). However, for robustness and to investigate dose-response effects, I additionally consider indicators for breastfeeding durations of 3+, 6+, and 12+ months. I cluster standard errors at the mother level to allow for correlation in unobservables among siblings. All regressions are weighted using NLSY custom longitudinal weights for the outcome age sample in question. To avoid incidentally truncating the sample while guarding against confounding, each covariate contains a dummy reflecting unknown values; as shown in Section 5, this choice does not affect my conclusions.

My parameter of interest is τ^{OLS}, which approximates the causal effect of breastfeeding if outcomes are unconfounded by unobservables.

3.2 Fixed Effects

The NLSY's abundant detail lays claim to plausible independence of breastfeeding and potential outcomes. Nevertheless, omitted factors—notably paternal characteristics—remain. To remove sibling-invariant family-level unobservables which may be correlated with B, I add mother fixed effects[24], ϕ_f, to Equation 1:

$$Y_{if} = \mathbf{X}_{if}\boldsymbol{\beta} + \tau^{FE}B_{if} + \phi_f + \boldsymbol{\varepsilon}_{if} \tag{2}$$

with Equation 1 augmented to include indicators, ϕ_f, for nuclear family membership.

But siblings with differential breastfeeding experiences begs an equally obvious criticism: why would a mother choose to breastfeed one child and not another? If breastfeeding statuses change, there may be other child-varying unobservables as well—and this sort of granular

[22] In Appendix A, I sketch a theoretical model to inform the empirical design.

[23] Implicitly, outcomes Y_i are indexed to specific ages, $a \in \{5, 10, 13, 21, 25\}$, obviating the need for explicit subscripts.

[24] Also popular in the breastfeedng literature, this is sometimes confusingly referred to as "sibling" fixed effects. I prefer the "mother" label, as it makes the level of clustering clear.

heterogeneity is exactly the blind spot of mother fixed effects. From this standpoint, sibling comparisons could be biased worse than OLS (e.g., the breastfed sibling may have benefited from a stay-at-home mother, while formula feeding may be indicative of financial hardships necessitating a quicker return to work). One's stance on this issue depends upon whether one views between- or within-family unobservables as the greater threat[25].

Given these downsides, I introduce an extension: extended family fixed effects. I do so by exploiting sister linkages in the NLSY79. Included in the NLSY proper are sisters who were both aged 14–21 as of December 31, 1978 and resident in the same household[26]. 3,279 children in the NLSY-CYA have an NLSY79 aunt.

The idea is that grouping women with their sisters retains the advantages of controlling for family backgrounds (including genetics) while expanding the comparison pool in a manner that makes the exogeneity of time-varying breastfeeding decisions more plausible. Cousins are more comparable than randomly-selected children; at the same time, mother-varying observables are plentiful in the data. In the interest of retaining as many children contributing to identification as possible, I define the extended family for children without cousins to be the nuclear family. Although so-called "cousin" fixed effects have been used in other settings[27], this approach is novel in the breastfeeding literature[28].

In the estimation, I modify Equation 2 such that ϕ_f refers to mothers and their sisters. As usual, consistency relies upon the assumption that cousin-varying (where, here, the term "cousin" encompasses siblings) unobservables are unrelated to breastfeeding status and outcomes.

4 Results

4.1 Descriptive Statistics

Tables 2 and 3 provide descriptive statistics for covariates and outcomes, respectively. The first three columns of Table 2 provide mean contrasts by breastfeeding status among those successfully interviewed at age five; the succeeding columns give means among respondents at each outcome age[29].

Most striking is the considerable positive selection among breastfed children. By virtually every measure, breastfed children have advantages compared with their formula-fed peers[30].

[25]For a recent perspective, see: Miller, Shenhav and Grosz (2019).

[26]The NLSY79's defines "sister" based on respondents' self-reported relationship statuses.

[27]Duncan et al. (2018*b*); Barclay, Lyngstad and Conley (2018).

[28]I prefer "extended family" to "cousin" as the label because the family is the child-invariant "within" level of clustering. Another way of thinking about this grouping is as maternal grandparent fixed effects.

[29]Since covariates are measured at, or prior to, birth, differences at different ages reflect the changing composition of respondents, not changes for any single respondent.

[30]One exception is alcohol consumption, which is 8 pp more common among breastfeeding mothers, a fact likely explicable by education and income.

The mothers of breastfed children are 29 percentage points (pp) more likely to be college educated, 20 pp more likely to be White, 2.6 years older, and 18 pp more likely to be married. They have higher rates of employment (7 pp), more income ($38k), and lower BMI (1.3 points). Perhaps most notably, they score 21 percentiles higher on the AFQT. Pregnancy behaviors follow suit. Breastfeeding moms have a higher probability of taking prenatal vitamins (4 pp), but lower probabilities of smoking (15 pp), preterm births (4 pp), and having children with long neonatal hospital stays (5 pp). Breastfed children weigh a third of a pound more at birth. They are also 5 pp more likely to be first-born and have better home environments following birth (by a third of a standard deviation). Figure 1 visually emphasizes this selection bias theme: there are pronounced gaps in breastfeeding by race, education, and cognitive test scores. At the same time, breastfeeding rates among the NLSY cohort evolved markedly during their childbearing epoch, as it did for the U.S. generally[31].

Table 3 gives a corresponding overview of outcomes. The basic point is clear: having been breastfed is associated with demonstrably better outcomes in childhood and young adulthood.

4.2 Main Results

Breastfeeding's unequivocal affiliation with *both* propitious circumstances and auspicious outcomes makes plain the difficulty isolating the effects of infant feeding. In this section, I present my main results for the relationship between ever having been breastfed and summary measures of intellect, health, and behavior—after adjusting for a wide array of confounding forces.

4.2.1 Child Outcomes

Table 4 analyzes childhood outcomes. Rows index outcomes. Supercolumns denote outcome ages (5, 10, and 13), with one column within each group for each of my three identification strategies: OLS, mother fixed effects, and extended family fixed effects. Each cell gives the treatment effect coefficient from a separate (weighted) regression of the row-enumerated outcome on an indicator for ever having been breastfed and full covariates. Mother-clustered standard errors are in parentheses. Numbers of observations are in braces[32].

OLS results (Cols 1, 4, and 7) demonstrate that the strong association between breastfeeding and cognitive performance persists even after controlling for maternal characteristics, pregnancy behaviors, and children's birth circumstances. At age 5, breastfed children per-

[31] Table A.2 gives birth counts and breastfeeding rates by year.

[32] For fixed effects, I give the *effective* numbers of observations contributing to identification—that is, the counts of siblings/cousins with different breastfeeding statuses. The models, nevertheless, are estimated on the full sample.

form 0.1 standard deviations (SD's) better in math, 0.07 SD's better in reading recognition, and 0.13 SD's better in vocabulary. These age 5 findings are in line with the previous literature on cognitive effects. What is notable, however, is that these advantages persist—and even grow stronger—during childhood. Breastfeeding-related gains at age 10 are 0.11 SD's for math, 0.07 SD's for reading recognition, 0.11 SD's for reading comprehension, and 0.16 SD's for vocabulary. At age 13, all three PIAT scores remain 0.11–0.12 SD's higher for breastfed children (the PPVT is not common at that age).

A tenth of a standard deviation is economically meaningful. Across all the cognitive assessments, the ranges between the 25th and 75th percentiles are 1–1.5 standard deviations, so the breastfeeding-related boost is worth several percentiles at the least—and as much as a decile in the middles of the distributions.

While a rich conditioning set cannot guarantee the absence of omitted variable bias warping these associations, the more detailed the controls, the more plausible the argument. Especially valuable in this regard are typically elusive traits, such as maternal intelligence. Excluding AFQT performance from the age 5 PIAT Math regression, for example, raises the coefficient on breastfeeding to to 0.131—an increase of 34 percent.

OLS results for health and behavior are more ambiguous. Breastfed children are no more likely to report health or behavioral issues at ages 5 or 10, though they are somewhat less prone to be overweight (by 0.7 pp at age 5 and 1.7 pp at age 10). By age 13, the weight difference disappears, but gaps emerge for health issues and behavior—in the opposite directions as expected. As early teens, breastfeed children are 2.7 pp more likely to have a notable health issue and are rated 0.06 SD's worse behaviorally. Whether these gaps reflect true disadvantages or more conscientious reporting by their mothers is unclear.

In contrast to OLS, mother (Cols 2, 5, 8) and extended family (Cols 3, 6, 9) fixed effects models suggest the cognitive benefits of breastfeeding may be smaller and transitory. While sibling comparisons at age 5 yield similar estimates to OLS (0.12 SD's for math and 0.28 SD's for reading comprehension), including cousins in the comparison attenuates these coefficients and renders them statistically insignificant. What's more, there is little evidence of continued gains in either model at ages 10 and 13; coefficients are uniformly smaller than 0.07 in magnitude and statistically insignificant, with standard errors that are relatively tight. Nevertheless, the samples of children contributing to fixed effects identification—which include only those families whose offspring have differential feeding experiences—are small, generally on the order of 1,000–2,000 observations, which may preclude the necessary power to detect effects. Health and behavioral effects in both FE models are also precise zeros, with two exceptions. Compared with siblings and cousins who are not breastfed, breastfed kin have 0.12 fewer behavioral problems at age 5 (per extended families), but are 6 pp more likely to have health issues at age 13 (per both FE models).

The many outcomes I assess are instructive in their breadth but may yield RCT-style concerns about multiple hypothesis testing[33]. There are well-known methods of adjusting significance levels to account for multiple comparisons (Romano and Wolf, 2005; Romano, Shaikh and Wolf, 2010). However, I do not pursue them here, as my objective is not to stake my claims on any one significant result, but instead to infer conclusions from the themes presented by the preponderance of the evidence, whether or not this evidence falls precisely on one side or the other of conventional significance levels.

By this standard, a reasonable interpretation of the foregoing analysis is that breastfeeding—and the interactions that go along with it—is indicative of modest, but persistent, cognitive gains throughout childhood, on average, but such advantages can also be conferred by other means, as the largely null fixed effects results suggest. On the other hand, there is little evidence for noncognitive benefits, and indeed some to the contrary: increased prevalence of health issues at age 13 is the lone result significant across all three models.

4.3 Young Adult Outcomes

An important question is whether these patterns persist into the young adult years. Table 5 provides some answers.

Similar in setup to Tables 4, Table 5 is split into two horizontal panels, with the first three columns describing age 21 outcomes and the latter trio studying age 25. Based on the OLS results for childhood cognition, one would expect to see better educational attainment among breastfed young adults, and indeed this is what OLS finds. At age 25 (Col 4), breastfed YA's are 2.7 pp more likely to have graduated high school (Row 1) and 3.1 pp more likely to have attended at least some college (Row 2). Much of these gains come early: by age 21 (Col 1), they have 0.14 additional years of education (Row 3), are 2.5 pp more likely to have completed high school, and have 4.6 pp greater probability of some college experience. The fixed effects educational results (Cols 2–3 and 5–6) are also in keeping with childhood: precise zeros for all outcomes.

The labor outcomes assessed in the next four rows reinforce this theme of punctilious nulls. In this case, even OLS finds breastfed young adults are no more likely to be employed or in school at ages 21 or 25; nor do they earn more or receive fewer public benefits. The same is generally true of health and behavior (bottom three rows). With the exception of an OLS-estimated 4.6 pp reduction in the probability of being overweight at 25, there are no significant results; indeed, the point estimates for the likelihood of health issues remain positive, as is true at age 13, though imprecisely so.

To summarize, OLS suggests the modest childhood gains in cognitive performance associated with breastfeeding continue through the young adult years, but fail to translate into

[33] As outcomes proliferate, the probability of a Type I error (erroneous rejection of a true null hypothesis) becomes increasingly likely for conventional hypothesis tests and confidence intervals. With 20 independent outcomes at the standard $\alpha = 0.05$ level, the probability of at least one false positive is 0.64.

labor market advantages. As in childhood, there is no evidence of noncognitive benefits, as well as reason—thoroughly null fixed effects results—to believe family context looms large.

5 Discussion, Robustness, and Extensions

The analysis thus far begs questions of sensitivity and substance. In this section, I demonstrate the robustness of my main findings while explaining how treatment intensity and quantifiable biases render comprehensible the (potentially) discordant implications of discernible cognitive benefits estimated via OLS and the null findings across other domains and fixed effects models.

5.1 Robustness: Alternative Outcomes, Covariates, and Weights

The outcomes I assess are comprehensive and conventional, but each is, nevertheless, subject to a series of subjective definitional decisions. In Tables A.3 (children) and A.4A–A.4B (young adults), I test the sensitivity of my results to alternative outcome measures (with associated descriptive statistics given in Tables A.5 and A.6). These measures are variants or constituents of my primary outcomes; full definitions are available in the Appendix. My main results are confirmed.

Tables 6 (children) and 7 (young adults) repeat the main analysis using conventional covariates. That is, individuals with missing data on any control variable are omitted. I also use continuous (rather than quartile) versions of covariates that are natively so (BMI, birthweight, income, AFQT), and exclude (potentially endogenous) age 0–2 HOME scores. The main results are again confirmed, albeit with somewhat less precision due to smaller sample sizes.

Tables A.7 (children) and A.8 (young adults) repeat the main analysis without using the NLSY-CYA longitudinal custom weights to weight the regressions. The main results are again largely confirmed, with the exception that somewhat more precise young adult OLS estimates detect statistically significant higher probabilities of health issues (by 2.2 pp at 21 and 2.7 pp at 25) and lower probabilities of premartial childbearing (by 2.3 pp at 21 and 2.8 pp at 25).

5.2 Survey Nonresponse

Another concern is bias from differential survey nonresponse. In Appendix C (and Tables A.9–A.13), I investigate this issue. I distill two facts: (1) breastfed individuals are about 4 percent more likely to respond, and (2) those who consistently respond to surveys have better outcomes. Together, these patterns impart mild accentuation bias in age 5–21 outcomes, as well as mild attenuation bias in age 25 outcomes, helping explain the observed temporal diminution of breastfeeding associations. Nevertheless, my main conclusions remain unchanged.

5.3 Treatment Intensity: Breastfeeding Duration

Prior research has frequently found breastfeeding and its benefits to exhibit a dose-response relationship: the longer breastfeeding continues, the larger the advantages, at least up until a point. In this section, I refine my main results by investigating breastfeeding duration.

Table 8 presents my childhood duration OLS results. I assess three (not mutually exclusive) treatment definitions: indicators for breastfeeding durations of at least 3, 6, and 12 months, given in columns. Extended breastfeeding is also somewhat rare: while 53 percent of NLSY-CYA respondents are ever breastfed, this rate drops to 28 percent at three months, 17 percent at six months, and 6 percent at twelve months. Among those breastfed, the mean duration is 23 weeks. As in the main analysis, supercolumns group results by outcome ages and rows index outcomes, with each cell giving estimated average treatment effects from a separate regression.

The big picture point of cognitive gains and noncognitive nulls is reinforced. In the case of the latter, the dose-response question is moot: an intensely null treatment remains null. For cognitive outcomes, the refinement is illuminating. At age 5, I find little evidence of a dose-response relationship. With the exception of vocabulary scores, which peak among those breastfed at least six months, OLS estimates indicate little changes beyond three months. Ages 10 and 13, however, suggest 6 months may be a sweet spot. There are pronounced 6-month peaks in almost all test scores; continuing beyond 12 months does not appear to offer additional gains, and may implicate modest losses.

Table 9 gives analogous childhood duration results for extended family fixed effects (I omit mother fixed effects for parsimony). The findings are simple to summarize: uniformly precise nulls, though cognitive point estimates are in keeping with a six-month sweet spot.

The childhood duration analysis mostly reinforces the main results, while suggesting a Goldilocks rule of sorts for maximizing breastfeeding's benefits. By contrast, the young adult OLS duration results in Table 10 are as much corrective as they are confirmatory. The education estimates in the first three rows are consistent with the six-month childhood peak in cognition. Those breastfed at least six months are 2.1 pp more likely to graduate high school, 6.2 pp more likely to go to college, and attain 0.28 more years of education by age 25 (Col 5). The effects are somewhat less when including those breastfeed between three and six months (Col 4) and disappear altogether when limited to those breastfed more than a year. The gap is less pronounced, but still apparent, at 21, when individuals breastfed 6+ months are 4.5 pp more likely to have attended college.

The big differences from the main results on breastfeeding initiation are labor and behavior. At age 21, those breastfed at least a year are 8 pp more likely to be employed and earn 173 percent more (Col 3); the earnings (but not employment) bump also remains visible in consideration of those breastfed at least six months (Col 2). In addition, the probability of receiving public benefits at 21 is significantly reduced with as little three months of breastfeeding (-2.7 pp), with the effect strengthening at six (-3.2 pp) and twelve (-5.6 pp) months. By age 25, however, there is a bit of a reversal. While three months of breastfeeding is associated with boosts of 6.1 pp in employment and 3.4 pp in the probability of being engaged in work or school (Col 4), those breastfed longer show no such gains. In addition, there is no evidence for improvements in earnings or reductions in public benefit use. Somewhat unexpectedly, one outcome that is consistent across both ages is behavioral: premarital childbearing. With the exception of 12+ months of breastfeeding at age 25, all durations are associated with reductions of 2.6–4.5 pp in the probability of early fertility. There is also some, albeit uneven, evidence of health impacts, with a 4.9 pp increase in the probability of age 21 health issues and a 4 pp reduction in the probability of being overweight at age 25, both among those breastfed six months or more.

The extended family fixed effects duration analysis for young adults (Table 11) provides no evidence for dose-response, or, for that matter, benefits of any sort. There is one exception: a positive, and statistically significant, relation between breastfeeding duration and the probability of overweight—by 7 pp at age 21 and 10–12 pp at age 25 among those breastfed at least six months.

Taken together, the dose-response results indicate young adult outcomes are more sensitive to duration that are those in childhood. Effects on labor and fertility behavior that are absent from treatment defined as breastfeeding initiation emerge among individuals breastfed at least three months. Though employment impacts attenuate somewhat by age 25, educational effects strengthen, and more so for those breastfed at least six months. At the same time, while any amount of breastfeeding has positive associations for child cognition, gains are stronger among those breastfed for prolonged periods.

These relationships should be construed cautiously: it is not clear whether they reflect true intensity effects or, rather, accentuated selectivity among families able to breastfeed longer (e.g., due to higher incomes or more generous family leave). As with the main analysis, the fixed effects results make plain that breastfeeding should not be interpreted outside the overall parental context. And, once again, there is little evidence of enduring health impacts.

5.4 Fixed Effects Selection

The null fixed effects results in the main analysis could reflect heterogeneous circumstances—genuinely different breastfeeding impacts for some families—or they could be seen as evidence of omitted variables biasing OLS. My preferred interpretation is the former, and here I provide two pieces of evidence in its favor.

The first is breastfeeding intensity. Extended families with diverse breastfeeding experiences also breastfeed shorter. Mean breastfeeding duration among families with unitary breastfeeding outcomes is 25.3 weeks. Among the subsample contributing to extended family identification, it drops to 17.7 weeks—a 30 percent reduction (see Table 12). Given the evidence for dose-response, this diminished duration is one explanation for smaller effects.

The second piece of evidence is sample composition. The extended family (identification) subsample is negatively selected. Tables 12 (covariates) and 13 (outcomes) provide descriptive statistics comparing these families with the larger sample of families with uniform infant feeding experiences. In comparison to those from feeding-consistent families, children from breastfeeding-variant ones have consistently worse cognitive outcomes throughout childhood (by 0.1–0.2 SD's), though noncognitive contrasts are minor and gaps appear to close by young adulthood. What's more, the feeding-variant subsample is generally disadvantaged, with mothers, who, among other things, have lower AFQT scores, income, HOME scores, and probabilities of attending college. In other words, breastfeeding may have less impact for these families because it is fighting stronger headwinds.

These facts carry an important implication countercurrent to the conventional wisdom: the assumed upward bias of OLS is arguably less important than the definite downward bias of fixed effects. Moreover, it appears not that breastfeeding-variant mothers "make up" for not breastfeeding some children, but instead the reverse: the circumstances that make it less likely to breastfeed result in disadvantages that dominate any beneficial effects of breastfeeding.

Tables A.14 (children) and A.15 (young adults) confirm these intuitions: excluding children from breastfeeding-variant families from the OLS analysis consistently bolsters breastfeeding coefficients by between about a fifth to a half, and renders some newly significant.

Whether this negative selection among inconsistently breastfeeding families is a general phenomenon or specific to the NLSY-CYA (whose mothers were co-resident 14–22-year-old sisters in 1979) is an open question.

5.5 External Validity

The majority of this discussion has addressed questions of internal validity. But there is also a point of caution regarding generalizability. The NLSY-CYA's greatest strength—its extraordinary longitudinality—is also its greatest weakness: it follows the offspring of a specific cohort of women (those born between 1957 and 1964) during a particular period in American history (that closing decades of the 20th century and the early years of the 21st)—a period not only of rapidly evolving breastfeeding practices, but also one of unprecedented prosperity and many changes of other sorts. While the NLSY-CYA is representative of this generation of children, it is not assured that results apply with equal force to children situated in other settings—though the timeliness of the data bolsters the case for relevance to affluent countries today. On the other hand, my results should not be extrapolated to developing countries, where concerns about sanitation, malnutrition, and infectious disease burden are paramount and make breastfeeding correspondingly more essential.

6 Conclusion

The current consensus among public health authorities is six months of exclusive breastfeeding for all children. My results suggest that this edict is reasonable for most families.

Across a broader set of outcomes spanning a longer trajectory of ages and controlling for more characteristics than previously assessed, I find that, on average (per OLS), breastfeeding is associated with modest but persistent cognitive benefits, ranging from higher test scores during childhood to greater educational attainment by young adulthood. Nevertheless, I cannot claim causality. While the detail of the NLSY-CYA renders the scope for unobserved influences comparatively small, omitted variables remain (notably paternal characteristics and maternal intangibles). Similarly, gains in young adult labor outcomes and fertility that appear conditional upon sustained breastfeeding durations of at least three months, as well as childhood cognitive outcomes that are strongest among those breastfed 6–12 months, may reflect genuine dose-response—or, instead, more pronounced sample selection among families able to breastfeed longer.

In contrast, I find little evidence of health benefits (and some results to the contrary). This is surprising given the well-established linkages between breastfeeding and infant health—and especially in light of the economics literature on the enduring importance of early-life experiences[34]. One intriguing possibility is that the most important consequence of infant health is cognitive development.

At the same time, not everyone is average, and all else is not always equal. Family fixed effects results indicate breastfeeding is less impactful among kin with differential breastfeeding experiences: within-family comparisons yield consistently null contrasts. I find that part of the explanation is context. Families with diverse breastfeeding are worse off than average and breastfeed less intensively.

My preferred interpretation is that the OLS and fixed effects findings are not mutually exclusive. Allowing for the effects of breastfeeding to be heterogeneous (which any reasonable understanding of parenting choices of any sort must permit), responses will vary. It is entirely possible OLS accurately describes an average treatment effect of breastfeeding that truly

[34] Almond and Currie (2011*a*,*b*); Currie and Rossin-Slater (2015) and Almond, Currie and Duque (2018) summarize this work.

confers modest intellectual advantages among the population at large, while, simultaneously, fixed effects characterizes an equally genuine observation that women who, by choice or circumstance, breastfeed some children and not others can (dis)compensate human capital investment in other ways. Breastfeeding can be an advantage without being a panacea.

Nor need breastfeeding's impacts be strictly literal. Breastfeeding entails a variety of interactions, time allocations, and related behaviors beyond the physical contents of mother's milk; precisely deciphering the relative contributions of each component is less important than recognizing the effects of the complete package.

The lesson for policymakers and medical professionals is that breastfeeding should not be one-size-fits all. Exclusive breastfeeding for six months appears optimal for many—and perhaps most—families. However, as the fixed effects results suggest, the overall environment into which breastfeeding is introduced cannot be ignored. Formula should not be a four-letter-word; the focus instead ought be on formulating strategies that replicate breastfeeding's most essential elements—nutrients, nurturing, and a general predisposition for conscientious parenting—in manners tailored to families' preferences, constraints, and resource endowments.

More research needed in this regard: little is known about the heterogeneity of breastfeeding itself. Variates of interest, in addition to supplementation and longevity, include frequency and quantity of feeds, manner of administration (e.g., breast or bottle), identity of feeder (e.g., mother or other), introduction of solid foods, and individual-level variation in the composition of breastmilk. Perhaps more important, it is essential to understand what extant omitted variables, if any, may account for breastfeeding's positive associations—and what may be done to promote such behaviors. The reverse question is also valid: to what extent does breastfeeding encourage other favorable changes in parents? What seems clear, however, is that breastfeeding is but one element of propitious parenting—an element whose importance ought not be overstated and must be interpreted in the context of the full body of work.

Internal validity aside, there is a point of caution to bear in mind when generalizing these findings. The NLSY follows a very specific cohort of American women (those born 1957–1964) during a period of time that not only witnessed widespread peace and prosperity, but also in which breastfeeding practices in the U.S. were evolving rapidly[35]—conditions that may not be replicable. In particular, my results have limited relevance for low- and middle-income countries, where sanitary deficiencies, infectious disease burden, and risk of malnutrition make breastfeeding considerably more essential[36].

Breastfeeding does not occur in a vacuum. Context matters. Being the type of mother who would breastfeed—or who conditionally would not—is probably more important than

[35] See Baker (2016) for details.
[36] Victora et al. (2016).

its realization. This does not mean breastfeeding does not have a true effect. It only suggests there are multiple means to achieving identical ends, and that impacts depend on environment. Far more important than the what of infant nutrition is the who and how of parenting—the constellation of behaviors, interactions, and values that go into raising a child.

7 References

Almond, Douglas, and Janet Currie. 2011*a*. "Human Capital Development before Age Five." In *Handbook of labor economics*. Vol. 4, 1315–1486. Elsevier.

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

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

Baker, Lindsay Gartman. 2016. "Breastfeeding in the United States: Economic Analyses of Trends and Policies." PhD diss. University of Michigan, Department of Economics.

Baker, Michael, and Kevin Milligan. 2008. "Maternal Employment, Breastfeeding, and Health: Evidence from Maternity Leave Mandates." *Journal of Health Economics*, 27(4): 871–887.

Barclay, Kieron, Torkild Lyngstad, and Dalton Conley. 2018. "The Production of Inequalities within Families and across Generations: The Intergenerational Effects of Birth Order and Family Size on Educational Attainment." National Bureau of Economic Research.

Belfield, Clive R, and Inas Rashad Kelly. 2012. "The Benefits of Breast Feeding across the Early Years of Childhood." *Journal of Human Capital*, 6(3): 251–277.

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

Bureau of Labor Statistics, U.S. Department of Labor, and National Institute for Child Health and Human Development. 2019. "Children of the NLSY79, 1979-2016." Produced and distributed by the Center for Human Resource Research, The Ohio State University., Columbus, OH.

Cesur, Resul, Joseph J Sabia, Inas Rashad Kelly, and Muzhe Yang. 2017. "The Effect of Breastfeeding on Young Adult Wages: New Evidence from the Add Health." *Review of Economics of the Household*, 15(1): 25–51.

Colen, Cynthia G, and David M Ramey. 2014. "Is Breast Truly Best? Estimating the Effects of Breastfeeding on Long-term Child Health and Wellbeing in the United States Using Sibling Comparisons." *Social Science & Medicine*, 109: 55–65.

Currie, Janet, and Maya Rossin-Slater. 2015. "Early-life Origins of Life-Cycle Well-being: Research and Policy Implications." *Journal of Policy Analysis and Management*, 34(1): 208–242.

Del Bono, Emilia, and Birgitta Rabe. 2012. "Breastfeeding and Child Cognitive Outcomes: Evidence from a Hospital-based Breastfeeding Support Policy." ISER Working Paper Series.

Dennis, Cindy-Lee. 2002. "Breastfeeding Initiation and Duration: A 1990-2000 Literature Review." *Journal of Obstetric, Gynecologic, & Neonatal Nursing*, 31(1): 12–32.

Denny, Kevin, and Orla Doyle. 2010. "The Causal Effect of Breastfeeding on Children's Cognitive Development: A Quasi-experimental Design."

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

Dieterich, Christine M, Julia P Felice, Elizabeth O'Sullivan, and Kathleen M Rasmussen. 2013. "Breastfeeding and Health Outcomes for the Mother-infant Dyad." *Pediatric Clinics of North America*, 60(1): 31.

Duncan, Greg J, Kenneth TH Lee, Maria Rosales-Rueda, and Ariel Kalil. 2018*a*. "Maternal Age and Child Development." *Demography*, 55(6): 2229–2255.

Duncan, Greg J, Kenneth TH Lee, Maria Rosales-Rueda, and Ariel Kalil. 2018*b*. "Maternal Age and Child Development." *Demography*, 55(6): 2229–2255.

Eidelman, Arthur I, and Richard J Schanler. 2012. "Breastfeeding and the Use of Human Milk." *Pediatrics*.

Evenhouse, Eirik, and Siobhan Reilly. 2005. "Improved Estimates of the Benefits of Breastfeeding Using Sibling Comparisons to Reduce Selection Bias." *Health Services Research*, 40(6p1): 1781–1802.

Fitzsimons, Emla, and Marcos Vera-Hernández. 2013. "Food for Thought? Breastfeeding and Child Development." IFS Working Papers.

Gibbs, Benjamin G, and Renata Forste. 2014. "Breastfeeding, Parenting, and Early Cognitive Development." *The Journal of Pediatrics*, 164(3): 487–493.

Gibson-Davis, Christina M, and Jeanne Brooks-Gunn. 2006. "Breastfeeding and Verbal Ability of 3-Year-Olds in a Multicity Sample." *Pediatrics*, 118(5): e1444–e1451.

Horta, Bernardo, Cesar Victora, World Health Organization, et al. 2013. "Long-term Effects of Breastfeeding: A Systematic Review."

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

Horta, Bernardo L, Rajiv Bahl, José Carlos Martinés, Cesar G Victora, World Health Organization, et al. 2007. "Evidence on the Long-term Effects of Breastfeeding: Systematic Review and Meta-analyses."

Ip, Stanley, Mei Chung, Gowri Raman, Priscilla Chew, Nombulelo Magula, Thomas Trikalinos, and Joseph Lau. 2007. "Breastfeeding and Maternal and Infant Health Outcomes in Developed Countries. Evidence Report/Technology Assessment No. 153 (Prepared by Tufts-New England Medical Center Evidence-based Practice Center, under Contract No. 290-02-0022)." *AHRQ Publication No. 07-E007. Rockville, MD: Agency for Healthcare Research and Quality.*

Jiang, Miao, E Michael Foster, and Christina M Gibson-Davis. 2011. "Breastfeeding and the Child Cognitive Outcomes: A Propensity Score Matching Approach." *Maternal and Child Health Journal*, 15(8): 1296–1307.

Kramer, Michael S, Beverley Chalmers, Ellen D Hodnett, Zinaida Sevkovskaya, Irina Dzikovich, Stanley Shapiro, Jean-Paul Collet, Irina Vanilovich, Irina Mezen, Thierry Ducruet, et al. 2001. "Promotion of Breastfeeding Intervention Trial (PROBIT): a randomized trial in the Republic of Belarus." *Jama*, 285(4): 413–420.

Kramer, Michael S, Frances Aboud, Elena Mironova, Irina Vanilovich, Robert W Platt, Lidia Matush, Sergei Igumnov, Eric Fombonne, Natalia Bogdanovich, Thierry Ducruet, et al. 2008. "Breastfeeding and Child Cognitive Development: New Evidence from a Large Randomized Trial." *Archives of General Psychiatry*, 65(5): 578–584.

Kramer, Michael S, Lidia Matush, Irina Vanilovich, Robert W Platt, Natalia Bogdanovich, Zinaida Sevkovskaya, Irina Dzikovich, Gyorgy Shishko, Jean-Paul Collet, Richard M Martin, 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." *The American journal of clinical nutrition*, 86(6): 1717–1721.

Lessen, Rachelle, and Katherine Kavanagh. 2015. "Position of the Academy of Nutrition and Dietetics: Promoting and Supporting Breastfeeding." *Journal of the Academy of Nutrition and Dietetics*, 115(3): 444–449.

Martin, Camilia, Pei-Ra Ling, and George Blackburn. 2016. "Review of Infant Feeding: Key Features of Breast Milk and Infant Formula." *Nutrients*, 8(5): 279.

Miller, Douglas L, Na'ama Shenhav, and Michel Z Grosz. 2019. "Selection into Identification in Fixed Effects Models, with Application to Head Start." National Bureau of Economic Research.

Onda, Masayuki, et al. 2016. "Breastfeeding and Early Childhood Outcomes: Is There a Causal Relationship?"

Rees, Daniel I, and Joseph J Sabia. 2009. "The Effect of Breast Feeding on Educational Attainment: Evidence from Sibling Data." *Journal of Human Capital*, 3(1): 43–72.

Rollins, Nigel C, Nita Bhandari, Nemat Hajeebhoy, Susan Horton, Chessa K Lutter, Jose C Martines, Ellen G Piwoz, Linda M Richter, Cesar G Victora, and The Lancet Breastfeeding Series Group. 2016. "Why Invest, and What It Will Take to Improve Breastfeeding Practices?" *The Lancet*, 387(10017): 491–504.

Romano, Joseph P, and Michael Wolf. 2005. "Exact and approximate stepdown methods for multiple hypothesis testing." *Journal of the American Statistical Association*, 100(469): 94–108.

Romano, Joseph P, Azeem M Shaikh, and Michael Wolf. 2010. "Hypothesis testing in econometrics." *Annu. Rev. Econ.*, 2(1): 75–104.

Rothstein, Donna S. 2013. "Breastfeeding and Children's Early Cognitive Outcomes." *Review of Economics and Statistics*, 95(3): 919–931.

Salone, Lindsey Rennick, William F Vann Jr, and Deborah L Dee. 2013. "Breastfeeding: An Overview of Oral and General Health Benefits." *The Journal of the American Dental Association*, 144(2): 143–151.

Stevens, Emily E, Thelma Patrick, and Rita Pickler. 2009. "A History of Infant Feeding." *The Journal of Perinatal Education*, 18: 32–39.

U.S. Department of Health and Human Services, Office of Disease Prevention and Health Promotion. 2020. "Maternal, Infant, and Child Health." *Healthy People 2020.*

US Department of Health and Human Services, et al. 2011. "The Surgeon General's Call to Action to Support Breastfeeding."

Victora, Cesar G, Rajiv Bahl, Aluísio JD Barros, Giovanny VA França, Susan Horton, Julia Krasevec, Simon Murch, Mari Jeeva Sankar, Neff Walker, Nigel C Rollins, et al. 2016. "Breastfeeding in the 21st Century: Epidemiology, Mechanisms, and Lifelong Effect." *The Lancet*, 387(10017): 475–490.

World Health Organization. 2009. "Infant and Young Child Feeding: Model Chapter for Textbooks for Medical Students and Allied Health Professionals."

World Health Organization. 2018. "Infant and Young Child Feeding." *Seventy-first World Health Assembly, Wha71.9.*

8 Tables

Table 1: National Longitudinal Survey of Youth Child and Young Adults, 1986-2016, Sample Overview

			Interviews Among Breastfed					
	All	BF Sample	Age 5	Age 10	Age 13	Age 21	Age 25	All
Ordinary Least Squares (OLS)	11,530	10,842	7,790	7,874	7,490	6,354	5,587	4,020
Mother Fixed Effects	3,543	3,066	2,245	2,281	2,206	1,923	1,716	1,260
Extended Family Fixed Effects	4,116	3,684	2,743	2,763	2,657	2,306	2,029	1,503

This table describes age-interview responses for the National Longitudinal Survey of Youth: Child and Young Adult 1986-2016 dataset. Rows enumerate identification strategies. Column 1 includes all observations. Column 2 includes observations with breastfeeding data. Columns 3–8 indicate age-interview responses among those with breastfeeding data. Each cell gives a count of observations. Unit of observation is individual child.

Table 2: Descriptive Statistics, NLSY-CYA 1986-2016 Breastfeeding Sample

	Age 5 Mean Contrast			Means by Age				
Variable	BF	Not BF	Diff	Age 5	Age 10	Age 13	Age 21	Age 25
Mom's Birth Year	1961.4	1961.5	-0.0	1961.4	1961.2	1961.2	1961.0	1960.8
Mom's Education: Less than HS	0.09	0.24	-0.15**	0.16	0.19	0.19	0.21	0.24
Mom's Education: High School	0.36	0.50	-0.13**	0.42	0.42	0.42	0.44	0.45
Mom's Education: Some College+	0.54	0.25	0.29**	0.41	0.38	0.37	0.33	0.28
Mom's Age	28.55	25.98	2.57**	27.45	26.48	26.33	24.97	23.70
Mom's Race: Hispanic	0.07	0.09	-0.01**	0.08	0.08	0.08	0.08	0.08
Mom's Race: White	0.86	0.66	0.20**	0.77	0.76	0.77	0.76	0.75
Mom's Race: Black	0.07	0.26	-0.19**	0.15	0.16	0.16	0.16	0.17
Mom Foreign-Born	0.05	0.04	0.01	0.05	0.05	0.04	0.04	0.05
Mom's Age 21 BMI	21.93	23.20	-1.27**	22.47	22.58	22.63	22.71	22.82
Mom AFQT Quartile	54.94	33.64	21.31**	45.83	44.87	44.91	43.36	41.29
Mom Employed	0.80	0.73	0.07**	0.77	0.77	0.77	0.79	0.79
Mom's Income	97901	59778	38123**	82194	79964	79433	68891	63005
Mom Region: Northeast	0.16	0.18	-0.02	0.17	0.16	0.15	0.15	0.13
Mom Region: North Central	0.28	0.26	0.02	0.27	0.25	0.25	0.27	0.26
Mom Region: South	0.25	0.36	-0.10**	0.30	0.27	0.27	0.28	0.27
Mom Region: West	0.21	0.11	0.10**	0.17	0.15	0.15	0.15	0.15
Mom Marital Status: Never Married	0.15	0.30	-0.15**	0.21	0.21	0.20	0.22	0.24
Mom Marital Status: Married	0.69	0.51	0.18**	0.61	0.56	0.55	0.56	0.51
Mom Marital Status: Other	0.07	0.10	-0.02**	0.08	0.08	0.07	0.08	0.07
Mom Prenatal Visit	0.99	0.99	0.01**	0.99	0.99	0.99	0.99	0.99
Mom Drink	0.53	0.45	0.08**	0.49	0.49	0.50	0.50	0.49
Mom Smoke	0.25	0.40	-0.15**	0.31	0.32	0.32	0.35	0.37
Mom Prenatal Vitamins	0.98	0.94	0.04**	0.96	0.96	0.96	0.95	0.95
C-Section	0.23	0.24	-0.01	0.24	0.22	0.22	0.21	0.21
Preterm Birth	0.11	0.15	-0.04**	0.12	0.12	0.12	0.12	0.12
Long Hospital Stay	0.06	0.11	-0.05**	0.08	0.08	0.08	0.08	0.08
Female	0.49	0.48	0.01	0.49	0.49	0.49	0.49	0.48
Birth Parity: 1	0.41	0.37	0.05**	0.39	0.43	0.43	0.45	0.49
Birth Parity: 2	0.34	0.37	-0.02**	0.36	0.34	0.34	0.34	0.33
Birth Parity: 3+	0.24	0.26	-0.02*	0.25	0.23	0.23	0.21	0.18
Birthweight	7.57	7.24	0.33**	7.43	7.41	7.41	7.41	7.37
Birth Month	6.54	6.56	-0.02	6.55	6.58	6.60	6.66	6.66
Birth Year	1989.7	1987.2	2.5**	1988.6	1987.4	1987.2	1985.7	1984.2
HOME Score Age 0-2	0.13	-0.17	0.30**	0.01	0.01	0.01	0.03	0.01
Has Aunt	0.32	0.33	-0.01	0.33	0.31	0.31	0.30	0.28
Obs.	3,784	4,006	7,790	7,790	7,874	7,490	6,354	5,587

This table summarizes covariates for the 1986–2016 NLSY-CYA. Columns 1–3 describe the differences in means by breastfeeding status for the age 5 interview sample, obtained from separate bivariate OLS regressions of each characteristic on the treatment indicator. Columns 4–8 give overall means for each age-interview sample. Unit of observation is individual child. All statistics are weighted using NLSY-CYA longitudinal custom weights for the corresponding age-interview sample. Categorical indicators do not sum to one due to missing values. See Appendix for additional covariate detail. * $p < 0.10$, ** $p < 0.05$

Table 3: Descriptive Outcome Statistics, NLSY-CYA 1986-2016 Breastfeeding Sample

	Overall			Breastfeeding Comparison				
	Mean	SD	Obs	Yes	No	Diff	SE	T-Stat
A. Age 5 Outcomes								
PIAT Math (standard deviation units)	0.15	0.94	5,710	0.32	-0.07	0.39**	0.03	11.88
PIAT Reading Recognition (std. dev. units)	0.54	0.92	5,600	0.69	0.33	0.36**	0.03	11.35
PIAT Reading Comprehension (std. dev. units)	0.82	0.81	1,964	0.90	0.72	0.19**	0.04	4.16
PPVT Vocabulary (std. dev. units)	-0.39	1.37	4,237	-0.14	-0.74	0.60**	0.05	11.14
Health Problem	0.43	0.50	7,759	0.46	0.39	0.07**	0.02	4.63
Behavior Problems Index (std. dev. units)	0.18	0.99	7,347	0.08	0.30	-0.21**	0.03	-6.31
Overweight	0.01	0.12	7,790	0.01	0.02	-0.01**	0.00	-2.66
B. Age 10 Outcomes								
PIAT Math (standard deviation units)	0.32	0.99	7,105	0.55	0.03	0.53**	0.03	16.88
PIAT Reading Recognition (std. dev. units)	0.42	1.01	7,101	0.62	0.18	0.43**	0.03	12.94
PIAT Reading Comprehension (std. dev. units)	0.16	0.92	6,989	0.34	-0.07	0.40**	0.03	13.52
PPVT Vocabulary (std. dev. units)	-0.12	1.29	6,024	0.20	-0.49	0.69**	0.04	15.92
Health Problem	0.33	0.47	7,806	0.34	0.31	0.03**	0.01	2.30
Behavior Problems Index (std. dev. units)	0.31	0.99	7,376	0.22	0.42	-0.20**	0.03	-6.26
Overweight	0.09	0.29	7,874	0.07	0.11	-0.04**	0.01	-4.78
C. Age 13 Outcomes								
PIAT Math (standard deviation units)	0.22	0.98	6,315	0.45	-0.06	0.51**	0.03	15.47
PIAT Reading Recognition (std. dev. units)	0.39	1.06	6,323	0.61	0.13	0.48**	0.04	13.33
PIAT Reading Comprehension (std. dev. units)	-0.06	0.88	6,269	0.14	-0.29	0.42**	0.03	14.29
PPVT Vocabulary (std. dev. units)	-0.42	1.16	857	-0.21	-0.57	0.36**	0.11	3.27
Health Problem	0.30	0.46	7,123	0.32	0.28	0.04**	0.01	2.81
Behavior Problems Index (std. dev. units)	0.36	0.99	6,764	0.30	0.42	-0.12**	0.03	-3.59
Overweight	0.16	0.37	7,490	0.14	0.19	-0.05**	0.01	-5.02
D. Age 21 Outcomes								
High School Grad+	0.87	0.34	5,751	0.92	0.81	0.10**	0.01	9.25
Some College+	0.54	0.50	5,751	0.64	0.43	0.21**	0.02	11.89
Years of Education	12.95	1.75	5,751	13.30	12.55	0.75**	0.06	11.81
Employed	0.65	0.48	5,927	0.68	0.63	0.04**	0.02	2.80
In School or Working	0.80	0.40	6,140	0.85	0.75	0.11**	0.01	8.34
Log Earnings (2019 Dollars)	7.94	3.34	5,294	8.10	7.76	0.34**	0.11	3.04
Public Benefits	0.16	0.37	6,349	0.12	0.21	-0.09**	0.01	-7.95
Health Problem	0.27	0.44	6,337	0.29	0.24	0.04**	0.01	2.96
Overweight	0.36	0.48	6,353	0.32	0.41	-0.09**	0.02	-6.03
Premarital Child	0.17	0.37	6,354	0.11	0.23	-0.13**	0.01	-10.95
E. Age 25 Outcomes								
High School Grad+	0.88	0.32	5,267	0.92	0.84	0.08**	0.01	7.25
Some College+	0.56	0.50	5,267	0.65	0.47	0.17**	0.02	9.48
Years of Education	13.49	2.29	5,267	13.94	13.04	0.90**	0.09	10.03
Employed	0.74	0.44	4,922	0.78	0.71	0.06**	0.02	4.13
In School or Working	0.79	0.41	5,014	0.83	0.76	0.07**	0.01	4.76
Log Earnings (2019 Dollars)	8.86	3.43	4,946	9.08	8.65	0.43**	0.12	3.56
Public Benefits	0.21	0.41	5,579	0.17	0.26	-0.09**	0.01	-6.96
Health Problem	0.32	0.47	5,583	0.34	0.31	0.03*	0.02	1.86
Overweight	0.41	0.49	5,587	0.35	0.46	-0.12**	0.02	-6.99
Premarital Child	0.28	0.45	5,587	0.20	0.36	-0.16**	0.02	-10.62

This table summarizes outcomes for the 1986–2016 NLSY-CYA sample with reported breastfeeding statuses. Columns 1–3 give overall means, standard deviations, and sample sizes. Cols 4–8 give mean comparisons by breastfeeding status, with respective group means for breastfed and not breastfed in Cols 4 and 5, point estimates for the mean differences in Col 6, and associated standard errors and test statistics in Cols 7 and 8. Results are obtained from separate bivariate OLS regressions of each outcome on the breastfeeding treatment indicator. Unit of observation is individual child. All statistics are weighted using NLSY-CYA longitudinal custom weights for the corresponding age-outcome sample. * $p < 0.10$, ** $p < 0.05$

Table 4: Child Outcomes

	Age 5			Age 10			Age 13		
	(1) OLS	(2) MOM	(3) EXT	(4) OLS	(5) MOM	(6) EXT	(7) OLS	(8) MOM	(9) EXT
Math	0.098**	0.121*	0.082	0.114**	-0.040	-0.028	0.116**	-0.004	0.019
	(0.033)	(0.071)	(0.072)	(0.031)	(0.055)	(0.059)	(0.032)	(0.058)	(0.062)
	{5,710}	{906}	{1,319}	{7,105}	{1,448}	{1,903}	{6,315}	{1,256}	{1,679}
Reading Recog.	0.070**	0.040	-0.029	0.073**	0.065	0.042	0.108**	0.026	0.009
	(0.031)	(0.065)	(0.066)	(0.034)	(0.056)	(0.061)	(0.037)	(0.061)	(0.067)
	{5,600}	{881}	{1,293}	{7,101}	{1,442}	{1,896}	{6,323}	{1,255}	{1,675}
Reading Comp.	0.020	0.284**	0.122	0.107**	-0.009	-0.002	0.111**	-0.039	-0.039
	(0.043)	(0.130)	(0.158)	(0.030)	(0.051)	(0.054)	(0.030)	(0.052)	(0.056)
	{1,964}	{120}	{184}	{6,989}	{1,409}	{1,857}	{6,269}	{1,227}	{1,636}
Vocabulary	0.129**	0.008	0.072	0.157**	0.056	0.012	-0.129	-0.347	-0.553
	(0.050)	(0.120)	(0.113)	(0.041)	(0.075)	(0.081)	(0.094)	(0.252)	(0.406)
	{4,237}	{518}	{801}	{6,024}	{1,076}	{1,444}	{857}	{40}	{45}
Health Problem	0.013	-0.029	-0.003	-0.005	0.005	0.012	0.027*	0.060*	0.055*
	(0.017)	(0.031)	(0.032)	(0.016)	(0.029)	(0.031)	(0.016)	(0.031)	(0.032)
	{7,759}	{1,505}	{2,114}	{7,806}	{1,629}	{2,164}	{7,123}	{1,458}	{1,953}
Behavior (BPI)	-0.037	-0.077	-0.117**	-0.018	0.056	0.008	0.064*	0.014	0.001
	(0.034)	(0.051)	(0.055)	(0.034)	(0.049)	(0.053)	(0.035)	(0.054)	(0.059)
	{7,347}	{1,337}	{1,910}	{7,376}	{1,475}	{1,988}	{6,764}	{1,337}	{1,793}
Overweight	-0.007*	-0.004	-0.005	-0.017*	0.006	0.001	-0.011	0.013	0.009
	(0.004)	(0.006)	(0.007)	(0.009)	(0.018)	(0.019)	(0.012)	(0.023)	(0.023)
	{7,790}	{1,513}	{2,130}	{7,874}	{1,659}	{2,204}	{7,490}	{1,601}	{2,144}

Rows enumerate outcomes, supercolumns denote outcome ages, and columns indicate estimation methods. Each cell reports the coefficient on an indicator for breastfeeding from a separate regression, with columns, respectively, for OLS, mother fixed effects, and extended family fixed effects. All regressions control for maternal characteristics (categorical indicators for mother's education, race, age, year of birth, region, marital status, AFTQ quartile, family income quartile, age 21 BMI quartile, foreign nativity, and having a sister), maternal pregnancy behaviors (categorical indicators for prenatal visits, prenatal vitamins, smoking, and alcohol consumption), and child's birth circumstances (categorical indicators for child's birth order, sex, birthweight quartile, C-section delivery, preterm birth, hospital stay longer than mother's, and age 0–2 HOME quartile), as well as a quadratic in child's birth year and indicators for birth month. All covariates include a category for missing values, and are defined at the time of child's birth or during preceding year, unless otherwise noted. In subsequent tables, this collection of controls is referred to as covariates. Data source is NLSY-CYA, 1986–2016. Unit of observation is individual child. Sample includes all births with nonmissing breastfeeding status and outcomes. All regressions are weighted using NLSY-CYA longitudinal custom weights for the age-interview sample. Standard errors clustered by mother in parentheses. Number of observations in braces; for fixed effects specifications, counts refer to observations contributing to breastfeeding identification.
* $p < 0.10$, ** $p < 0.05$

Table 5: Young Adult Outcomes

	Age 21			Age 25		
	(1) OLS	(2) MOM	(3) EXT	(4) OLS	(5) MOM	(6) EXT
High School Grad+	0.025**	0.008	0.001	0.027**	-0.010	0.000
	(0.012)	(0.024)	(0.025)	(0.012)	(0.025)	(0.026)
	{5,751}	{1,069}	{1,489}	{5,267}	{965}	{1,291}
Some College+	0.046**	-0.006	0.002	0.031*	0.012	0.021
	(0.017)	(0.030)	(0.032)	(0.018)	(0.032)	(0.034)
	{5,751}	{1,069}	{1,489}	{5,267}	{965}	{1,291}
Years of Education	0.141**	0.068	0.062	0.102	0.074	0.095
	(0.062)	(0.132)	(0.131)	(0.082)	(0.138)	(0.149)
	{5,751}	{1,072}	{1,492}	{5,267}	{965}	{1,291}
Employed	0.010	-0.031	-0.016	0.007	-0.044	-0.001
	(0.017)	(0.035)	(0.038)	(0.017)	(0.031)	(0.034)
	{5,927}	{1,169}	{1,569}	{4,922}	{892}	{1,174}
In School or Working	0.019	-0.026	-0.018	0.004	-0.040	-0.005
	(0.013)	(0.029)	(0.031)	(0.015)	(0.028)	(0.031)
	{6,140}	{1,210}	{1,631}	{5,014}	{917}	{1,203}
Log Earned Income	-0.034	0.122	0.119	-0.118	-0.000	0.001
	(0.125)	(0.269)	(0.292)	(0.130)	(0.263)	(0.285)
	{5,294}	{934}	{1,266}	{4,946}	{902}	{1,174}
Public Benefits	-0.011	-0.016	-0.008	-0.001	-0.028	-0.028
	(0.012)	(0.026)	(0.027)	(0.014)	(0.030)	(0.032)
	{6,349}	{1,275}	{1,723}	{5,579}	{1,087}	{1,422}
Health Problem	0.021	-0.006	-0.026	0.027	0.057	0.037
	(0.016)	(0.034)	(0.035)	(0.018)	(0.037)	(0.040)
	{6,337}	{1,271}	{1,718}	{5,583}	{1,088}	{1,423}
Overweight	-0.022	0.033	0.024	-0.046**	0.027	0.016
	(0.016)	(0.032)	(0.034)	(0.016)	(0.031)	(0.035)
	{6,353}	{1,277}	{1,726}	{5,587}	{1,088}	{1,424}
Premarital Child	-0.015	-0.006	0.003	-0.019	0.010	0.002
	(0.012)	(0.024)	(0.025)	(0.016)	(0.031)	(0.035)
	{6,354}	{1,277}	{1,726}	{5,587}	{1,088}	{1,424}

Rows enumerate outcomes, supercolumns denote outcome ages, and columns indicate estimation methods. Each cell reports the coefficient on an indicator for breastfeeding from a separate regression, with columns, respectively, for OLS, mother fixed effects, and extended family fixed effects. All regressions control for full covariates, as described in Table 4. Data source is NLSY-CYA, 1986–2016. Unit of observation is individual child. Sample includes all births with nonmissing breastfeeding status and outcomes. All regressions are weighted using NLSY-CYA longitudinal custom weights for the age-interview sample. Standard errors clustered by mother in parentheses. Number of observations in braces; for fixed effects specifications, counts refer to observations contributing to breastfeeding identification. * $p < 0.10$, ** $p < 0.05$

Table 6: Child Outcomes: Conventional Covariates

	Age 5			Age 10			Age 13		
	(1) OLS	(2) MOM	(3) EXT	(4) OLS	(5) MOM	(6) EXT	(7) OLS	(8) MOM	(9) EXT
Math	0.098**	0.084	0.076	0.081**	-0.073	-0.101	0.111**	0.118	0.091
	(0.041)	(0.112)	(0.113)	(0.041)	(0.093)	(0.096)	(0.043)	(0.108)	(0.111)
	{3,267}	{321}	{507}	{3,566}	{421}	{639}	{3,287}	{387}	{571}
Reading Recog.	0.071*	0.008	-0.029	0.058	0.084	0.043	0.075	0.091	0.033
	(0.040)	(0.094)	(0.097)	(0.044)	(0.092)	(0.103)	(0.049)	(0.108)	(0.115)
	{3,196}	{314}	{493}	{3,567}	{422}	{640}	{3,295}	{388}	{575}
Reading Comp.	0.048	-0.233**	-0.092	0.080**	-0.089	-0.036	0.084**	-0.023	-0.110
	(0.055)	(0.105)	(0.134)	(0.039)	(0.088)	(0.097)	(0.039)	(0.078)	(0.089)
	{1,030}	{39}	{69}	{3,518}	{419}	{624}	{3,264}	{382}	{563}
Vocabulary	0.052	0.045	0.131	0.123**	0.060	-0.101	0.399**	0.000	0.000
	(0.065)	(0.213)	(0.211)	(0.052)	(0.138)	(0.145)	(0.157)	(.)	(.)
	{2,201}	{158}	{281}	{3,035}	{305}	{471}	{199}		
Health Problem	0.015	0.001	0.008	-0.006	-0.009	-0.001	-0.000	0.056	0.054
	(0.021)	(0.051)	(0.053)	(0.022)	(0.055)	(0.057)	(0.022)	(0.056)	(0.058)
	{4,386}	{532}	{808}	{3,862}	{471}	{720}	{3,600}	{432}	{650}
Behavior (BPI)	-0.052	-0.112	-0.128	-0.039	-0.051	-0.046	0.070	-0.037	-0.057
	(0.044)	(0.078)	(0.092)	(0.045)	(0.098)	(0.103)	(0.047)	(0.087)	(0.091)
	{4,175}	{483}	{742}	{3,641}	{407}	{642}	{3,436}	{395}	{588}
Overweight	-0.005	0.000	-0.002	-0.004	-0.014	-0.016	-0.019	-0.008	-0.047
	(0.004)	(0.011)	(0.011)	(0.012)	(0.032)	(0.033)	(0.016)	(0.041)	(0.043)
	{4,395}	{537}	{813}	{3,874}	{477}	{725}	{3,788}	{472}	{706}

This table repeats the main analysis using conventional covariates. Rows enumerate outcomes, supercolumns denote outcome ages, and columns indicate estimation methods. Each cell reports the coefficient on an indicator for breastfeeding from a separate regression, with columns, respectively, for OLS, mother fixed effects, and extended family fixed effects. All regressions control for full covariates; the difference from the main text is that individuals with missing data for any covariate are dropped and HOME scores are omitted. Data source is NLSY-CYA, 1986–2016. Unit of observation is individual child. Sample includes all births with nonmissing breastfeeding status, outcomes, and covariates. All regressions are weighted using NLSY-CYA longitudinal custom weights for the age-interview sample. Standard errors clustered by mother in parentheses. Number of observations in braces; for fixed effects specifications, counts refer to observations contributing to breastfeeding identification. * $p < 0.10$, ** $p < 0.05$

Table 7: Young Adult Outcomes: Conventional Covariates

	Age 21			Age 25		
	(1) OLS	(2) MOM	(3) EXT	(4) OLS	(5) MOM	(6) EXT
High School Grad+	0.033**	-0.026	-0.015	0.027*	0.002	0.006
	(0.016)	(0.040)	(0.041)	(0.015)	(0.039)	(0.039)
	{3,262}	{365}	{559}	{3,051}	{353}	{512}
Some College+	0.050**	-0.032	-0.037	0.030	0.046	0.025
	(0.022)	(0.042)	(0.044)	(0.023)	(0.057)	(0.059)
	{3,262}	{365}	{559}	{3,051}	{353}	{512}
Years of Education	0.165**	0.234	0.142	0.111	0.352	0.248
	(0.083)	(0.263)	(0.247)	(0.106)	(0.235)	(0.244)
	{3,262}	{368}	{562}	{3,051}	{353}	{512}
Employed	-0.001	-0.058	-0.033	0.022	-0.064	0.004
	(0.022)	(0.059)	(0.062)	(0.022)	(0.060)	(0.063)
	{3,280}	{380}	{561}	{2,588}	{265}	{390}
In School or Working	0.020	-0.023	-0.020	0.011	-0.084	-0.034
	(0.017)	(0.046)	(0.049)	(0.021)	(0.056)	(0.060)
	{3,371}	{393}	{581}	{2,644}	{279}	{407}
Log Earned Income	-0.013	0.098	-0.098	-0.037	0.101	0.170
	(0.158)	(0.461)	(0.483)	(0.169)	(0.508)	(0.532)
	{2,873}	{292}	{443}	{2,736}	{299}	{428}
Public Benefits	-0.008	-0.022	-0.006	-0.012	-0.048	-0.051
	(0.016)	(0.045)	(0.047)	(0.019)	(0.054)	(0.057)
	{3,430}	{409}	{608}	{3,048}	{353}	{512}
Health Problem	0.023	-0.021	-0.052	0.042*	-0.035	-0.028
	(0.021)	(0.062)	(0.061)	(0.023)	(0.061)	(0.067)
	{3,430}	{409}	{607}	{3,050}	{353}	{512}
Overweight	-0.032	0.046	0.032	-0.025	0.092**	0.091*
	(0.022)	(0.057)	(0.060)	(0.020)	(0.045)	(0.048)
	{3,431}	{409}	{608}	{3,051}	{353}	{512}
Premarital Child	-0.014	0.017	0.013	-0.008	0.038	0.015
	(0.015)	(0.042)	(0.045)	(0.020)	(0.053)	(0.057)
	{3,431}	{409}	{608}	{3,051}	{353}	{512}

This table repeats the main analysis using conventional covariates. Rows enumerate outcomes, supercolumns denote outcome ages, and columns indicate estimation methods. Each cell reports the coefficient on an indicator for breastfeeding from a separate regression, with columns, respectively, for OLS, mother fixed effects, and extended family fixed effects. All regressions control for full covariates; the difference from the main text is that individuals with missing data for any covariate are dropped and HOME scores are omitted. Data source is NLSY-CYA, 1986–2016. Unit of observation is individual child. Sample includes all births with nonmissing breastfeeding status, outcomes, and covariates. All regressions are weighted using NLSY-CYA longitudinal custom weights for the age-interview sample. Standard errors clustered by mother in parentheses. Number of observations in braces; for fixed effects specifications, counts refer to observations contributing to breastfeeding identification. * $p < 0.10$, ** $p < 0.05$

Table 8: Child Outcomes: Breastfeeding Duration, OLS Estimates

	Age 5			Age 10			Age 13		
	(1) 3m+	(2) 6m+	(3) 12m+	(4) 3m+	(5) 6m+	(6) 12m+	(7) 3m+	(8) 6m+	(9) 12m+
Math	0.078**	0.060	-0.020	0.133**	0.170**	0.136**	0.150**	0.215**	0.196**
	(0.037)	(0.045)	(0.071)	(0.034)	(0.039)	(0.061)	(0.036)	(0.043)	(0.066)
	{5,560}	{5,560}	{5,560}	{6,946}	{6,946}	{6,946}	{6,172}	{6,172}	{6,172}
Reading Recog.	0.096**	0.060	0.032	0.047	0.119**	0.065	0.077**	0.115**	0.078
	(0.034)	(0.042)	(0.068)	(0.036)	(0.040)	(0.061)	(0.039)	(0.043)	(0.062)
	{5,457}	{5,457}	{5,457}	{6,942}	{6,942}	{6,942}	{6,180}	{6,180}	{6,180}
Reading Comp.	-0.033	-0.023	-0.144	0.125**	0.216**	0.236**	0.122**	0.179**	0.167**
	(0.049)	(0.064)	(0.142)	(0.034)	(0.039)	(0.058)	(0.033)	(0.038)	(0.056)
	{1,909}	{1,909}	{1,909}	{6,832}	{6,832}	{6,832}	{6,128}	{6,128}	{6,128}
Vocabulary	0.147**	0.194**	0.051	0.179**	0.211**	0.190**	-0.100	0.012	0.213
	(0.059)	(0.075)	(0.121)	(0.048)	(0.060)	(0.094)	(0.106)	(0.134)	(0.270)
	{4,131}	{4,131}	{4,131}	{5,885}	{5,885}	{5,885}	{855}	{855}	{855}
Health Problem	-0.005	0.002	-0.034	-0.021	-0.004	-0.028	0.009	0.005	-0.004
	(0.019)	(0.021)	(0.033)	(0.018)	(0.021)	(0.030)	(0.018)	(0.021)	(0.032)
	{7,565}	{7,565}	{7,565}	{7,634}	{7,634}	{7,634}	{6,963}	{6,963}	{6,963}
Behavior (BPI)	-0.063	-0.049	0.011	-0.037	-0.010	0.077	0.023	-0.005	0.051
	(0.039)	(0.043)	(0.064)	(0.039)	(0.045)	(0.070)	(0.040)	(0.044)	(0.065)
	{7,162}	{7,162}	{7,162}	{7,214}	{7,214}	{7,214}	{6,609}	{6,609}	{6,609}
Overweight	0.002	0.003	0.002	-0.011	0.000	-0.012	-0.011	0.000	-0.003
	(0.004)	(0.004)	(0.007)	(0.009)	(0.011)	(0.016)	(0.012)	(0.014)	(0.020)
	{7,596}	{7,596}	{7,596}	{7,702}	{7,702}	{7,702}	{7,321}	{7,321}	{7,321}

Rows enumerate outcomes. Supercolumns denote outcome ages. Columns correspond binary indicators for breastfeeding duration: 3 months or more (Cols 1, 4, 7), 6 months or more (Cols 2, 5, 8), and 12 months or more (Cols 3, 6, 9). Each cell reports the coefficient on column-enunumerated breastfeeding treatment indicator from a separate OLS regression. All regressions control for full covariates. Unit of observation is individual child. Sample includes all births with nonmissing breastfeeding status and outcomes. Standard errors clustered by mother in parentheses. Number of observations in braces. * $p < 0.10$, ** $p < 0.05$

Table 9: Child Outcomes: Breastfeeding Duration, Extended Family Fixed Effects Estimates

	Age 5			Age 10			Age 13		
	(1) 3m+	(2) 6m+	(3) 12m+	(4) 3m+	(5) 6m+	(6) 12m+	(7) 3m+	(8) 6m+	(9) 12m+
Math	0.039	-0.024	-0.068	-0.036	0.010	0.052	0.014	0.060	-0.019
	(0.079)	(0.081)	(0.113)	(0.056)	(0.062)	(0.093)	(0.062)	(0.072)	(0.109)
	{1,214}	{1,214}	{1,214}	{1,784}	{1,784}	{1,784}	{1,577}	{1,577}	{1,577}
Reading Recog.	0.094	-0.054	-0.009	-0.006	-0.002	-0.091	0.033	0.050	-0.075
	(0.072)	(0.080)	(0.114)	(0.060)	(0.070)	(0.101)	(0.065)	(0.080)	(0.100)
	{1,194}	{1,194}	{1,194}	{1,778}	{1,778}	{1,778}	{1,574}	{1,574}	{1,574}
Reading Comp.	0.007	0.198	0.102	0.039	0.141**	0.044	-0.003	0.045	-0.057
	(0.164)	(0.157)	(0.343)	(0.053)	(0.064)	(0.083)	(0.059)	(0.068)	(0.094)
	{171}	{171}	{171}	{1,739}	{1,739}	{1,739}	{1,535}	{1,535}	{1,535}
Vocabulary	-0.089	0.001	-0.448*	-0.018	0.139	0.021	0.012	-0.020	-1.048
	(0.104)	(0.126)	(0.244)	(0.083)	(0.099)	(0.131)	(0.369)	(0.483)	(2.044)
	{753}	{753}	{753}	{1,345}	{1,345}	{1,345}	{45}	{45}	{45}
Health Problem	-0.012	-0.012	-0.029	0.003	0.040	0.043	0.011	0.035	0.017
	(0.033)	(0.036)	(0.050)	(0.031)	(0.038)	(0.053)	(0.033)	(0.037)	(0.054)
	{1,976}	{1,976}	{1,976}	{2,029}	{2,029}	{2,029}	{1,839}	{1,839}	{1,839}
Behavior (BPI)	-0.074	-0.050	-0.002	0.004	-0.020	-0.016	0.009	0.003	0.031
	(0.056)	(0.063)	(0.087)	(0.062)	(0.065)	(0.095)	(0.065)	(0.072)	(0.100)
	{1,780}	{1,780}	{1,780}	{1,869}	{1,869}	{1,869}	{1,689}	{1,689}	{1,689}
Overweight	0.004	0.000	-0.006	-0.018	-0.002	0.004	-0.019	0.018	0.024
	(0.006)	(0.007)	(0.012)	(0.016)	(0.018)	(0.025)	(0.023)	(0.027)	(0.035)
	{1,991}	{1,991}	{1,991}	{2,069}	{2,069}	{2,069}	{2,022}	{2,022}	{2,022}

Rows enumerate outcomes. Supercolumns denote outcome ages. Columns correspond binary indicators for breastfeeding duration: 3 months or more (Cols 1, 4, 7), 6 months or more (Cols 2, 5, 8), and 12 months or more (Cols 3, 6, 9). Each cell reports the coefficient on column-enunumerated breastfeeding treatment indicator from a separate extended family fixed effects regression. All regressions control for full covariates. Unit of observation is individual child. Sample includes all births with nonmissing breastfeeding status and outcomes. Standard errors clustered by mother in parentheses. Number of observations contributing to breastfeeding identification in braces. * $p < 0.10$, ** $p < 0.05$

Table 10: Young Adult Outcomes: Breastfeeding Duration, OLS Estimates

	Age 21			Age 25		
	(1) 3m+	(2) 6m+	(3) 12+	(4) 3m+	(5) 6m+	(6) 12+
High School Grad+	-0.005	-0.008	-0.027	0.019	0.021*	-0.013
	(0.011)	(0.013)	(0.021)	(0.012)	(0.012)	(0.022)
	{5,576}	{5,576}	{5,576}	{5,116}	{5,116}	{5,116}
Some College+	0.024	0.045**	0.011	0.035*	0.062**	0.006
	(0.019)	(0.022)	(0.038)	(0.021)	(0.023)	(0.044)
	{5,576}	{5,576}	{5,576}	{5,116}	{5,116}	{5,116}
Years of Education	0.065	0.120	-0.029	0.188**	0.277**	-0.063
	(0.068)	(0.078)	(0.125)	(0.092)	(0.107)	(0.196)
	{5,576}	{5,576}	{5,576}	{5,116}	{5,116}	{5,116}
Employed	-0.012	0.003	0.080**	0.061**	0.035	0.022
	(0.020)	(0.024)	(0.039)	(0.019)	(0.024)	(0.043)
	{5,761}	{5,761}	{5,761}	{4,784}	{4,784}	{4,784}
In School or Working	-0.001	0.011	0.028	0.034*	0.015	0.001
	(0.014)	(0.016)	(0.025)	(0.017)	(0.020)	(0.036)
	{5,970}	{5,970}	{5,970}	{4,873}	{4,873}	{4,873}
Log Earned Income	0.023	0.272*	0.548**	-0.102	-0.216	-0.283
	(0.143)	(0.152)	(0.218)	(0.147)	(0.183)	(0.330)
	{5,151}	{5,151}	{5,151}	{4,809}	{4,809}	{4,809}
Public Benefits	-0.027**	-0.032**	-0.056**	-0.017	-0.015	0.003
	(0.013)	(0.014)	(0.018)	(0.016)	(0.017)	(0.029)
	{6,174}	{6,174}	{6,174}	{5,428}	{5,428}	{5,428}
Health Problem	0.026	0.049**	0.017	-0.001	0.014	-0.014
	(0.018)	(0.021)	(0.036)	(0.021)	(0.027)	(0.047)
	{6,162}	{6,162}	{6,162}	{5,432}	{5,432}	{5,432}
Overweight	-0.020	0.004	-0.018	-0.030	-0.040*	-0.026
	(0.019)	(0.021)	(0.032)	(0.020)	(0.024)	(0.037)
	{6,178}	{6,178}	{6,178}	{5,436}	{5,436}	{5,436}
Premarital Child	-0.026**	-0.029**	-0.045**	-0.042**	-0.045**	-0.036
	(0.012)	(0.013)	(0.016)	(0.016)	(0.018)	(0.029)
	{6,179}	{6,179}	{6,179}	{5,436}	{5,436}	{5,436}

Rows enumerate outcomes. Supercolumns denote outcome ages. Columns correspond binary indicators for breastfeeding duration: 3 months or more (Cols 1 and 4), 6 months or more (Cols 2 and 5), and 12 months or more (Cols 3 and 6). Each cell reports the coefficient on column-enunumerated breastfeeding treatment indicator from a separate OLS regression. All regressions control for full covariates. Unit of observation is individual child. Sample includes all births with nonmissing breastfeeding status and outcomes. Standard errors clustered by mother in parentheses. Number of observations in braces. * $p < 0.10$, ** $p < 0.05$

Table 11: Young Adult Outcomes: Breastfeeding Duration, Extended Family Fixed Effects Estimates

	Age 21			Age 25		
	(1) 3m+	(2) 6m+	(3) 12+	(4) 3m+	(5) 6m+	(6) 12+
High School Grad+	-0.038*	-0.029	-0.049	-0.016	-0.015	-0.049
	(0.021)	(0.025)	(0.035)	(0.021)	(0.025)	(0.041)
	{1,361}	{1,361}	{1,361}	{1,174}	{1,174}	{1,174}
Some College+	-0.004	0.019	-0.004	-0.041	-0.045	-0.038
	(0.036)	(0.042)	(0.065)	(0.036)	(0.045)	(0.071)
	{1,361}	{1,361}	{1,361}	{1,174}	{1,174}	{1,174}
Years of Education	-0.027	0.064	-0.180	0.013	-0.104	-0.218
	(0.113)	(0.128)	(0.177)	(0.147)	(0.187)	(0.285)
	{1,364}	{1,364}	{1,364}	{1,174}	{1,174}	{1,174}
Employed	-0.050	-0.007	0.095	0.047	0.046	0.082
	(0.043)	(0.049)	(0.085)	(0.047)	(0.054)	(0.068)
	{1,442}	{1,442}	{1,442}	{1,073}	{1,073}	{1,073}
In School or Working	-0.059*	-0.034	-0.040	0.012	0.016	0.048
	(0.032)	(0.036)	(0.048)	(0.041)	(0.046)	(0.070)
	{1,501}	{1,501}	{1,501}	{1,098}	{1,098}	{1,098}
Log Earned Income	0.055	0.312	0.817*	-0.234	-0.018	-0.146
	(0.285)	(0.320)	(0.432)	(0.353)	(0.385)	(0.735)
	{1,172}	{1,172}	{1,172}	{1,072}	{1,072}	{1,072}
Public Benefits	-0.028	-0.022	-0.054	-0.007	0.047	0.034
	(0.030)	(0.032)	(0.040)	(0.036)	(0.038)	(0.056)
	{1,590}	{1,590}	{1,590}	{1,301}	{1,301}	{1,301}
Health Problem	-0.004	0.018	-0.049	-0.021	0.047	0.088
	(0.035)	(0.044)	(0.066)	(0.044)	(0.058)	(0.091)
	{1,585}	{1,585}	{1,585}	{1,302}	{1,302}	{1,302}
Overweight	0.029	0.068*	0.073	0.064	0.095**	0.118*
	(0.036)	(0.039)	(0.056)	(0.042)	(0.047)	(0.068)
	{1,593}	{1,593}	{1,593}	{1,303}	{1,303}	{1,303}
Premarital Child	-0.007	0.010	0.001	-0.010	0.042	0.048
	(0.023)	(0.025)	(0.027)	(0.034)	(0.040)	(0.052)
	{1,593}	{1,593}	{1,593}	{1,303}	{1,303}	{1,303}

Rows enumerate outcomes. Supercolumns denote outcome ages. Columns correspond binary indicators for breastfeeding duration: 3 months or more (Cols 1 and 4), 6 months or more (Cols 2 and 5), and 12 months or more (Cols 3 and 6). Each cell reports the coefficient on column-enunumerated breastfeeding treatment indicator from a separate extended family fixed effects regression. All regressions control for full covariates. Unit of observation is individual child. Sample includes all births with nonmissing breastfeeding status and outcomes. Standard errors clustered by mother in parentheses. Number of observations contributing to breastfeeding identification in braces. * $p < 0.10$, ** $p < 0.05$

Table 12: Descriptive Covariate Statistics, Extended Family Fixed Effects Comparison

	Families		Mean Comparison			
	Non-FE	FE	Diff	SE	T-Stat	Obs
Mom's Birth Year	1961.2	1961.4	0.2	0.1	1.52	10,842
Mom's Education: Less than HS	0.18	0.22	0.04**	0.02	2.45	10,842
Mom's Education: High School	0.38	0.44	0.06**	0.02	2.86	10,842
Mom's Education: Some College+	0.42	0.32	-0.10**	0.02	-4.61	10,842
Mom's Age	27.19	26.52	-0.67**	0.23	-2.93	10,842
Mom's Race: Hispanic	0.07	0.13	0.06**	0.01	5.90	10,842
Mom's Race: White	0.75	0.67	-0.08**	0.02	-4.68	10,842
Mom's Race: Black	0.17	0.20	0.02	0.01	1.53	10,842
Mom Foreign-Born	0.04	0.07	0.03**	0.01	3.30	10,842
Mom's Age 21 BMI	22.34	22.90	0.56**	0.19	2.97	10,427
Mom AFQT	45.44	39.59	-5.85**	1.33	-4.39	10,301
Mom Employed	0.77	0.73	-0.04**	0.02	-2.65	8,581
Mom's Income	89851	67631	-22221**	5576	-3.99	7,002
Mom Region: Northeast	0.15	0.16	0.01	0.02	0.79	10,842
Mom Region: North Central	0.24	0.24	0.00	0.02	0.25	10,842
Mom Region: South	0.29	0.24	-0.05**	0.02	-3.26	10,842
Mom Region: West	0.15	0.17	0.02*	0.01	1.65	10,842
Mom Marital Status: Never Married	0.19	0.22	0.03**	0.01	2.84	10,842
Mom Marital Status: Married	0.57	0.50	-0.06**	0.02	-3.95	10,842
Mom Marital Status: Other	0.07	0.09	0.02**	0.01	2.42	10,842
Mom Prenatal Visit	0.99	0.99	-0.00	0.00	-0.65	9,919
Mom Drink	0.49	0.42	-0.07**	0.02	-3.94	9,898
Mom Smoke	0.31	0.33	0.01	0.02	0.61	9,888
Mom Prenatal Vitamins	0.95	0.96	0.01	0.01	0.84	8,774
C-Section	0.24	0.22	-0.02	0.02	-1.37	9,824
Preterm Birth	0.12	0.15	0.03**	0.01	2.33	9,730
Long Hospital Stay	0.08	0.09	0.01	0.01	1.07	9,444
Female	0.48	0.50	0.01	0.01	1.01	10,841
Birth Parity: 1	0.44	0.34	-0.10**	0.01	-12.54	10,842
Birth Parity: 2	0.34	0.33	-0.01	0.01	-1.57	10,842
Birth Parity: 3+	0.22	0.33	0.11**	0.01	8.98	10,842
Birthweight	7.38	7.31	-0.07	0.05	-1.45	10,117
Birth Month	6.57	6.54	-0.03	0.09	-0.32	10,842
Birth Year	1988.1	1987.6	-0.5*	0.3	-1.94	10,842
HOME Score Age 0-2	0.03	-0.11	-0.14**	0.04	-3.38	5,251
Has Aunt	0.24	0.47	0.23**	0.02	10.63	10,842
Breastfed	0.56	0.49	-0.07**	0.02	-4.12	10,842
Weeks Breastfed	25.28	17.72	-7.57**	1.51	-5.01	4,653

This table compares covariates between families contributing to extended family fixed effects breastfeeding identification (families with siblings or cousins with diverse breastfeeding experiences; FE) with children from families with unitary breastfeeding experiences (non-fixed effects families; Non-FE). Columns 1 and 2 give overall means. Cols 3–6 give mean comparisons, with point estimates for the mean differences in Col 3 (), and associated standard errors, test statistics, and sample sizes in Cols 4–6. Results are obtained from separate bivariate OLS regressions of each outcome on an indicator for fixed effects family membership. Unit of observation is individual child. All statistics are weighted using NLSY-CYA longitudinal custom weights for the corresponding age-outcome sample. * $p < 0.10$, ** $p < 0.05$

Table 13: Descriptive Outcome Statistics, Extended Family Fixed Effects Comparison

	Families		Mean Comparison			
	Non-FE	FE	Diff	SE	T-Stat	Obs
A. Age 5 Outcomes						
Math	0.20	0.04	-0.17**	0.04	-4.45	5,710
Reading Recog.	0.58	0.43	-0.15**	0.04	-3.94	5,600
Reading Comp.	0.84	0.78	-0.06	0.05	-1.23	1,964
Vocabulary	-0.32	-0.57	-0.26**	0.06	-4.02	4,237
Health Problem	0.44	0.41	-0.03	0.02	-1.48	7,759
Behavior (BPI)	0.17	0.20	0.03	0.04	0.72	7,347
Overweight	0.01	0.02	0.00	0.00	1.03	7,790
B. Age 10 Outcomes						
Math	0.38	0.17	-0.21**	0.04	-5.37	7,105
Reading Recog.	0.46	0.31	-0.15**	0.04	-3.77	7,101
Reading Comp.	0.19	0.07	-0.12**	0.04	-3.49	6,989
Vocabulary	-0.04	-0.30	-0.25**	0.05	-4.92	6,024
Health Problem	0.33	0.32	-0.01	0.02	-0.91	7,806
Behavior (BPI)	0.30	0.34	0.05	0.04	1.14	7,376
Overweight	0.08	0.11	0.03**	0.01	2.67	7,874
C. Age 13 Outcomes						
Math	0.27	0.09	-0.18**	0.04	-4.60	6,315
Reading Recog.	0.43	0.29	-0.15**	0.04	-3.34	6,323
Reading Comp.	-0.01	-0.18	-0.17**	0.03	-4.93	6,269
Vocabulary	-0.36	-0.57	-0.21*	0.11	-1.86	857
Health Problem	0.31	0.28	-0.03	0.02	-1.64	7,123
Behavior (BPI)	0.35	0.38	0.03	0.04	0.76	6,764
Overweight	0.15	0.19	0.04**	0.01	2.99	7,490
D. Age 21 Outcomes						
High School Grad+	0.88	0.85	-0.03**	0.01	-2.34	5,751
Some College+	0.56	0.49	-0.07**	0.02	-3.24	5,751
Years of Education	13.05	12.71	-0.33**	0.08	-4.31	5,751
Employed	0.65	0.66	0.01	0.02	0.34	5,927
In School or Working	0.81	0.78	-0.03*	0.02	-1.84	6,140
Log Earnings (2019 Dollars)	7.99	7.81	-0.18	0.14	-1.33	5,294
Public Benefits	0.15	0.19	0.05**	0.01	3.45	6,349
Health Problem	0.26	0.27	0.00	0.01	0.31	6,337
Overweight	0.36	0.39	0.03*	0.02	1.72	6,353
Premarital Child	0.16	0.20	0.04**	0.02	2.61	6,354
E. Age 25 Outcomes						
High School Grad+	0.88	0.87	-0.01	0.01	-0.69	5,267
Some College+	0.58	0.53	-0.05**	0.02	-2.34	5,267
Years of Education	13.58	13.29	-0.28**	0.10	-2.70	5,267
Employed	0.75	0.72	-0.03*	0.02	-1.65	4,922
In School or Working	0.80	0.78	-0.02	0.02	-1.41	5,014
Log Earnings (2019 Dollars)	8.90	8.76	-0.15	0.14	-1.06	4,946
Public Benefits	0.20	0.25	0.05**	0.02	2.96	5,579
Health Problem	0.31	0.36	0.05**	0.02	2.90	5,583
Overweight	0.40	0.43	0.03	0.02	1.63	5,587
Premarital Child	0.27	0.31	0.04**	0.02	2.37	5,587

This table compares outcomes between families contributing to extended family fixed effects breastfeeding identification (families with siblings or cousins with diverse breastfeeding experiences; FE) with children from families with unitary breastfeeding experiences (non-fixed effects families; Non-FE). Columns 1 and 2 give overall means. Cols 3–6 give mean comparisons, with point estimates for the mean differences in Col 3 (), and associated standard errors, test statistics, and sample sizes in Cols 4–6. Results are obtained from separate bivariate OLS regressions of each outcome on an indicator for fixed effects family membership. Unit of observation is individual child. All statistics are weighted using NLSY-CYA longitudinal custom weights for the corresponding age-outcome sample. * $p < 0.10$, ** $p < 0.05$

9 Figures

Figure 1