econometrics

MDPI

Article

Children's Health Capital Investment: Effects of U.S. Infant Breastfeeding on Teenage Obesity

Albert Okunade [1], Ahmad Reshad Osmani [2,*], Toluwalope Ayangbayi [3] and Adeyinka Kevin Okunade [4]

1 Department of Economics, The University of Memphis, Memphis, TN 38152, USA; aokunade@memphis.edu
2 College of Business and Social Sciences, University of Louisiana Monroe, Monroe, LA 71209, USA
3 ALSAC/St. Jude Children's Research Hospital, Memphis, TN 38105, USA; toluwalope.ayangbayi@alsac.stjude.org
4 Rush Health Systems, Pediatrics, Meridian, MS 39307, USA; Adeyinka.okunade@rushhealth.org
* Correspondence: osmani@ulm.edu; Tel.: +1-318-342-1768

Abstract: Obesity, as a health and social problem with rising prevalence and soaring economic cost, is increasingly drawing scholarly and public policy attention. While many studies have suggested that infant breastfeeding protects against childhood obesity, empirical evidence on this causal relationship is fragile. Using the health capital development theory, this study exploited multiple data sources from the U.S. and a three-way error components model (ECM) with a jackknife resampling plan to estimate the effect of in-hospital breastfeeding initiation and breastfeeding for durations of 3, 6, and 12 months on the prevalence of obesity during teenage years. The main finding was that a 1% rise in the in-hospital breastfeeding initiation rate reduces the teenage obesity prevalence rate by 1.7% (9.6% of a standard deviation). The magnitude of this effect declines as the infant breastfeeding duration lengthens—e.g., the 12-month infant breastfeeding duration rate is associated with a 0.53% (3.7% of a standard deviation) reduction in obesity prevalence in the teenage years (9th to 12th grades). The study findings agree with both the behavioral and physiological theories on the long-term effects of breastfeeding, and have timely implications for public policies promoting infant breastfeeding to reduce the economic burden of teenage and later adult-stage obesity prevalence rates.

Keywords: infant breastfeeding; teenage obesity; human capital formation; jackknife resampling; three-way error component model

check for updates

Citation: Okunade, Albert, Ahmad Reshad Osmani, Toluwalope Ayangbayi, and Adeyinka Kevin Okunade. 2021. Children's Health Capital Investment: Effects of U.S. Infant Breastfeeding on Teenage Obesity. *Econometrics* 9: 42. https://doi.org/10.3390/econometrics9040042

Academic Editor: Marc S. Paolella

Received: 16 June 2021
Accepted: 24 November 2021
Published: 29 November 2021

1. Introduction and Literature

Past literature has shown the significance of investment in childhood health capital on later life outcomes (Francesconi and Heckman 2016; Keane and Wolpin 1997; Walters 2015). Childhood health capital (Grossman 1972), as an integral component of human capital, plays a substantial role in regulating socioeconomic activities in the teenage period. Any health investment in childhood could potentially trigger prospects for later life outcomes such as education, social activities, occupation, and health outcomes (Currie 2009).

Past studies have indicated that the nutritional condition of the fetus and infant has long-lasting health and wealth effects (Almond and Currie 2011; Campbell et al. 2014; Victora et al. 2016). The fetal origins hypothesis (Barker 1995), for instance, shows the importance of well-balanced nutrition for the fetus, which can avoid disproportionate fetal growth and later-life chronic heart diseases. A lack of sufficient nutrients and proper living conditions during infancy and early childhood could increase the prevalence of many health conditions (Van den Berg et al. 2006). Along the same line of reasoning, a burgeoning body of research has also shown the protective effects of maternal breast milk as a healthy nutrition choice during infancy (Martin et al. 2005; Metzger and McDade 2010).

Among other morbidities, obesity, with an annual cost of 190 billion dollars to the U.S. public, is a prevalent chronic health condition that can affect all age groups, including teenagers (Currie 2009; Tremmel et al. 2017). The prevalence of severe obesity among American children and teenagers has tripled in the last three decades (Rosin 2008). Breast milk, as a form of health capital investment, could curb or reverse teenage and adult overweight and obesity, in addition to its short-term health-promoting effects (Victora et al. 2015).

Despite the widespread interest in the medical literature about understanding the health benefits of breastfeeding, limited studies exist in the applied economics literature. The most closely related studies to our work are those of Der et al. (2006); Rees and Sabia (2009), and Metzger and McDade (2010), who also explored how an infant's cognitive and health outcomes respond to breastfeeding by exploiting individual-level sibling data. Der et al. (2006) investigated the effects of breastfeeding on the health and cognitive outcomes of children using the National Longitudinal Survey of Youths (NLSY). Although their study showed the statistically significant effect of a mother's IQ on the cognitive outcomes of their child, they barely found any substantial impact of breastfeeding on their child's cognitive measures. Using the same dataset but a different identification strategy and adding more controls to the empirical model, Rees and Sabia (2009) established a significant relationship between breastfeeding and both acute and chronic health conditions. The findings of both studies should be interpreted with caution, as they solely relied on siblings and included only mother-specific controls. A mother's IQ itself could be an endogenous variable in their model. Finally, Metzger and McDade (2010) executed a fixed-effects model of obesity as a function of breastfeeding on a sibling sample of American children. They concluded that breastfeeding is correlated with a lower body mass index (BMI) in teenagers.

Our study builds on the studies above in at least three ways. First, unlike the previous works, which only evaluate a single or limited measure of breastfeeding, our study considered in-hospital breastfeeding initiation and 3-, 6- and 12-month durations of breastfeeding. This allowed us to flexibly estimate a public policy-relevant relationship between the duration of breastfeeding and obesity. Second, none of the previous studies analyzed the relationship between resampled state-level breastfeeding data and the prevalence of obesity in high schoolers conditional on individual, household, and community observables. This enabled us to construct more comprehensive state profiles that could capture the landscape of breastfeeding support as recommended by the Centers for Disease Control and Prevention (CDC). Finally, our work departs from the existing literature by exploiting a three-way error component identification strategy that explored the complexity of the data structure and tackled the issue of unobserved heterogeneity when estimating the effects of breastfeeding. We relied on a theoretical construct of human capital development (Schultz 1974) and used the Fuller and Battese (1974) empirical crossed-error component estimation technique to analyze the effects of breastfeeding on teenage obesity by generating a retrospective linkage of cross-sectional time series pyramid-shaped datasets in a discontinuous timeframe structure. In addition, we augmented our empirical applications by applying variable selection algorithms of machine learning.

The next section of this study outlines the conceptual framework associated with the empirical results. The two subsequent sections explain the data and identification strategy. The penultimate section describes the results of the empirical models. The final section discusses the findings and concludes the research.

2. Breastfeeding and Health Capital Development: Conceptual Framework

The existing body of literature indicates at least three key channels through which breastfeeding could affect the stock of health capital in the short and long terms. First, human milk contains rich nutritional and hormonal elements, such as colostrum, that are essential for the healthy growth and development of infants. Any dietary interruption during the infancy period can increase the likelihood of diseases in later life (Rosales-Rueda 2014). Specifically, the body composition of infants who have not been breastfed comprises more unhealthy fats than lean body mass. Imbalances in body composition increase the risk of developing multiple health conditions as a result of diminished health stock (Zweifel

2012). Second, unlike infant formula, breast milk is an organic and well-balanced food that lacks any preservatives. Antioxidant food additives are associated with the high prevalence of obesity in the Western world (Ciardi et al. 2012). The primary underlying cause of obesity and overweight among children is an imbalance between caloric intake and calorie burning. Past studies have suggested that newborns cannot distinguish between amniotic fluid and breast milk in the early days after birth due to their biochemical similarities. As a natural post-natal diet regulatory mechanism, breastfeeding improves the stock of health capital. Finally, breastfeeding improves the psychological bond between mother and infant (Belfield and Kelly 2012); this could enhance the physical and mental health of the child.

3. Data

The bulk of the work on breastfeeding and its effects on later life outcomes has relied on small and short pooled cross-sectional datasets with descriptive research designs. The problems associated with a lack of appropriate observable factors and long-panel datasets are well known and have been the source of inconsistencies in estimates. For the current study, we used state-level panel datasets linking multiple breastfeeding indicators with later-life outcomes, including teenage obesity and overweight for all U.S. states. We exploited multiple sources to construct the estimation sample.

The hospital breastfeeding initiation incidence and duration of breastfeeding (3, 6, and 12 months after birth), as the key variables of interest, were obtained from the Ross Products Division of Abbott Laboratories, which is a leading pharmaceutical firm and the manufacturer of the Similac nutritional formula for newborns. Ross Products have conducted annual mail surveys of a representative sample of U.S. newborns until recently.[1] The survey encompasses at least 82% of all live births in the U.S. for the years 1999–2002. Similar to major national health surveys (e.g., the Survey of Family Growth, the National Maternal and Infant Health Survey, and the National Health and Nutrition Examination Survey), the Ross Products survey maintains a high response rate. Data on the rates of teenage obesity and overweight came from the high school Youth Risk Behavior Survey (YRBS) conducted biannually by the CDC. The survey contains state-level obesity and overweight data for 9th–12th graders, along with other relevant information, including gender, participation in physical activity classes, hours of daily sleep, and choice of food consumption (e.g., vegetables or fruit) for the years 2011–2017. The RMS data were further supplemented with state-level data from the Behavioral Risks Factor Surveillance System (BRFFS) on the percentage of U.S. adult obesity to capture the potential associations between family adult BMI pathology and the persistent rise in teenage obesity.

Unlike the measure in adults, overweight and obese children fall in the 85th and 95th percentiles of the growth chart distribution, respectively, based on gender- and age-specific groups of observed children (Cawley 2010). We retrospectively linked state-level panels of breastfeeding incidence and duration with the obesity and overweight rates for teenagers aged 14–17 years. Furthermore, we used other data sources, including the Behavioral Risks Factor Surveillance System (BRFFS), the American Community Survey (ACS), the U.S. Census Bureau, the Bureau of Labor Statistics (BLS), and the Kaiser Family Foundation (KFF), to construct rich sets of state-level observables. We confined our attention to the observables associated with students in the 9th to 12th grades, adult characteristics, and socioeconomic factors. Table 1 provides a summary of the observables. The dependent variable, the current state-level prevalence of teenage obesity, is the percentage of high schoolers who were at or above the 95th percentile for BMI based on gender- and age-specific reference data from the CDC growth charts. This information was collected by school grade for 9th–12th graders for each U.S. state participating in the YRBS. With the assumption of an average age of 14–17 years for high schoolers, the independent variable of interest was the in-hospital breastfeeding and durations (3, 6, and 12 months) during the year of birth (i.e., 1998–2001) for each grade of high school.

Table 1. Summary statistics of the data.

	Variable	Definition	Source	Mean (SD)
Dependent variable	Teenage obesity	Percentage, high schoolers who were in the ≥95th percentile for body mass index, based on gender- and age-specific reference data from the 2000 CDC growth charts	CDC, High School Youth Risk Behavior Survey 2015[2]	14.02 (2.72)
Individual factors	In-hospital breastfeeding initiation rate	Percentage, newborns breastfed in hospital at birth during the year of birth	Ross Product[3] Division, Abbott Laboratories	68.50 (10.9)
	Physical activity	Percentage of high schoolers who attended physical education classes on all 5 days in an average school week	CDC, High School Youth Risk Behavior Survey 2015	23.38 (13.69)
	Vegetable consumption	Percentage of high schoolers who ate vegetables during the 7 days before the survey	CDC, High School Youth Risk Behavior Survey 2015 (Note 1)	92.64 (2.3)
	Fruit consumption	Percentage, high schoolers who ate fruit or drank 100% fruit juices during the 7 days before survey	CDC, High School Youth Risk Behavior Survey 2015	92.31 (7.97)
	Fast food restaurants	Fast food restaurants per 10,000 population	Dun and Bradstreet[4]	3.15 (0.68)
Household factors	Adult obesity	Percentage, adults aged ≥18 years with a body mass index ≥ 30 based on reported height and weight (pre-2011 BRFSS methodology)	CDC, Behavioral Risk Factor Surveillance System[5]	29.23 (3.39)
	Single-parent household	Percentage, of children under 18 years who live with their own single parent either in a family or subfamily	U.S. Census Bureau 2002–2015 American Community Survey (ACS)	35.12 (7.06)
	Log of median household income (short-run income)	Dollar amount that divides the household income distribution into two equal groups	America's Health Rankings[6]	10.94 (0.16)
	Log of median home value (long-run income or asset)	Median home value in 2010–2012	American Community Survey Briefs[7]	12.09 (0.38)

Table 1. *Cont.*

	Variable	Definition	Source	Mean (SD)
	Hispanic	Percentage, Hispanic population in the state	Kaiser Family Foundation (Note 4)	11.82 (9.9)
Community factors	Crime rate	Number of murders, rapes, robberies and aggravated assaults per 100,000 population	Federal Bureau of Investigation[8]	379.88 (185.27)
	White	Percentage, White American population in the state	Kaiser Family Foundation[9]	68.43 (16.37)
	Black	Percentage, Black American population in the state	Kaiser Family Foundation	11.18 (10.61)

In-hospital breastfeeding was used as a proxy for the intention and motivation to breastfeed, although it is recognized that breastfeeding may cease following discharge from hospital. Moreover, breastfeeding initiation in hospital increases the chances of a newborn consuming colostrum—the immuno-protective low-fat, protein-rich (including growth factors) precursor of breast milk (Stevens et al. 2006). The independent variables for the model were identified from the economic literature on obesity and grouped into individual, household, and community factors (Courtemanche et al. 2016; Rosin 2008). The variable selection algorithm of machine learning is applied to identify the relevant set of variables.

These factors are physical activity (Slentz et al. 2005), sleep, fruit and vegetable consumption, and fast food restaurants per 10,000 population. These (except for fast food restaurants per 10,000 population) were observed at the grade and state levels. In-hospital breastfeeding initiation was expected to correlate negatively with teenage obesity due to the balanced nutritional properties of breast milk. Physical activity, sleep (Zaqout et al. 2018), and fruit and vegetable consumption (Sauder et al. 2019) are healthy options; thus, higher rates of engagement in physical activity, sleep, and fruit and vegetable consumption was expected to correlate negatively with the prevalence of obesity. Consumption of low-priced refined carbohydrate meals and sugary drinks from fast food restaurants has been reported as a contributory factor to the increase in the prevalence of obesity and was expected to correlate positively with teenage obesity. Household variables at the state level included the prevalence of adult obesity, the percentage of children in single-parent households, the median household income, and the median home value in each state.

Adult obesity represents a genetic link between teenagers and adults in a household and, depending on the contribution of genetics to obesity, a positive correlation is expected between teenage and adult obesity. A high prevalence of teenage obesity may also correlate positively with the rate of single-parent households if the single parent has to be absent from home to go to work, exposing the children in such a household to the consumption of fast foods and reduced time engaging in physical activities (due to the need to keep indoors, especially in areas with a high crime rate).

Household income and wealth are expected to correlate negatively with teenage obesity due to a negative wealth gradient from better access to parks and recreational activities in wealthier neighborhoods, parental ability to provide exercise equipment and to support activities that promote physical exercise, etc. Community factors included the proportion of White, Black, and Hispanic ethnic groups, as well as the crime rate in each state.

The proportion of each ethnic group in the community also controlled for racial differences associated with genetic and cultural variations in factors that correlate with obesity or health outcomes (Frank et al. 2007) and how obesogenic the environment is (Finkelstein et al. 2005). For instance, a high crime rate in a community may discourage children and youths from engaging in outdoor physical activities, including in public parks designed and equipped for fitness activities. Therefore, a high crime rate is expected to be associated with a higher prevalence of teenage obesity (Fish et al. 2010).

4. Econometric Model and the Jackknife Resampling Plan

The empirical model of this work was motivated by the unique retrospective linkage of cross-sectional time series pyramid-shaped datasets in a discontinuous timeframe structure. The Fuller and Battese (1974) empirical crossed-error components estimation technique[10] was used to obtain the effects of breastfeeding initiation and time-phased breastfeeding durations on teenage obesity using the following equation:

$$y_{it} = x_{it}\beta + z_{jt}\gamma + p_i\pi + q_j\varphi + \alpha_i + \theta_j + \phi_t + \varepsilon_{it}. \tag{1}$$

In Equation (1), high schoolers are indexed as $i = (9, 10, 11, 12)$ and observed once per period $t = 1, \ldots, T$ in the state $j = 1, \ldots, J$. Furthermore, y_{it} is the obesity rate, x_{it} and p_i indicate the vectors observable at the i level, and z_{jt} and q_j are the vectors observable at the j level. If one assumes that teenagers can enter and exit the panel, the data could be unbalanced. In Equation (1), the unobserved heterogeneities (error components) consist of α_1 for high schoolers, θ_j for states, and ϕ_t for time effects. Throughout the estimation, we treated ϕ_t as the time-fixed effects. Moreover, a correlation between the error components (α_i and θ_j) and the observables was assumed.[11]

The main approach taken in the previous literature has been to estimate Equation (1) using the ordinary least squares (OLS) framework. We might obtain unbiased estimates of the variable of interest by OLS if we observe and measure all of the variables in Equation (1) that are correlated with teenage obesity. Nevertheless, there are many omitted variables that are difficult to measure and are likely to potentially suffer from heterogeneity, given the unique componential data structure. Therefore, to the extent that unobserved factors potentially exist, OLS estimation would yield biased results.

Following previous work (Borjas 2003; Donohue et al. 2001), we began with the classic heteroskedasticity-based argument for the weighted least square (WLS)[12] model when the outcome of interest was a group average and the averages for heterogeneous groups were based on widely varying within-group sample sizes. Under exogenous sampling and correct specification of the conditional mean of the outcome of interest in Equation (1), WLS is consistent for estimating the regression coefficients. On the contrary, under either endogenous sampling or model misspecification, OLS and WLS generally follow different probability limits. As indicated by Dumouchel and Duncan (1983), the statistically significant variation between OLS and WLS estimates can be used as a diagnostic for model misspecification or endogenous sampling.[13]

Given the susceptibility of the dependent variable data for heteroskedasticity and other distributional features that complicate analysis, we took one further step and implemented the classic Box–Cox transformation (Box and Cox 1964)[14]. Finally, the linear mixed-effects models are another appropriate class of models that could possibly fit the estimation samples. These models are also known as multilevel models (Al-Amin and Hossain 2019). The overall error distribution of the linear mixed-effects model is assumed to be Gaussian, and heteroskedasticity and correlations within lowest-level groups also may be modeled. The multilevel mixed model estimates are reported in Table A1. These estimates are based upon a 3-level hierarchical model. This model acknowledges the multiple sources of variability and its assignment to the appropriate level. As a result, it can accurately and precisely capture the effects of included independent variables. The variance component of multi-level mixed model is an indicator of intraclass or intra-level correlation coefficient

(ICC). The latter mentioned is a statistic that quantifies the degree to which data at the lower level are correlated.[15] The ICC varies between 0 and 1. An ICC value close to 0 is suggestive of variation in the outcome variable at lower-level units while a value close to 1 is indicative of variability at upper levels. Among all the estimated comparable models in Table A1, the multilevel mixed model has the highest value of likelihood ratio (LR) value. Basically, the LR test compares the one-level ordinary linear regressions with multi-level models.

The natural grouping structure of the panel dataset, on the one hand, and the heterogeneous distribution of the groups, on the other, provide enough evidence to execute the three-way error component model (ECM) for unbalanced panels (Fuller and Battese 1974). The ECM exploits the times series cross-sectional dimensions in the data. The errors in the model are associated with error correlations across state cross-sections ξ_i and schooling grade years within each state τ_i, and the third component e_{ijt} combines both dimensions as follows:

$$\varepsilon_{ijt} = \xi_i + \tau_i + e_{ijt} \tag{2}$$

Furthermore, sample reuse strategies are useful for obtaining unbiased estimates and correct interval estimations for the coefficients of interest (Efron 1982; Okunade et al. 2004; Quenouille 1956). Specifically, the bootstrapping and jackknife methods are widely used, but there are various designs within each method (e.g., one-delete, n-delete, random jackknife, and the weighted and unweighted bootstrap strategies). These strategies exploit the full sample or a specified data-generating process (such as a coin toss), which is a model of the process of interest to generate a new sample of simulated data and to evaluate the results using these samples. A well-known type of weighted least square jackknife variance estimator was proposed by Beran (1986). This estimator is achieved by deleting any fixed number of respondents at a time. This method provides estimates that are unbiased for homoskedastic and heteroskedastic residuals. We implemented this method of resampling in three steps, as proposed by Simon and Bruce (1991). First, a simulated process of randomizing, which was compositionally similar to the process whose behavior we wished to evaluate, was constructed. Second, we identified a procedure that generates a pseudo-sample that mimics the real sample of interest. Finally, the probability of the outcome of interest from the outcomes of the resampling plan trials was acquired.

The one-delete jackknife method generates a jackknife dataset of $(n-1)$ by deleting the first observation from the entire dataset. As a result, estimation of the coefficient of interest is then performed on the reduced dataset; this provides an estimated value of the parameter of interest, say, η_{j1}. Thereafter, the first deleted unit is replaced and another one (that is, the second observation) is deleted, providing a new measured value of the estimates, say, η_{j2}. This technique is repeated for each observation in the sample of interest. This generates a set of estimated values $\left(\eta_{ji},\ i = 1, \ldots, K\right)$. For this approach, the standard error, where η is the jackknifing result of the complete sample, was calculated as the square root of the jackknife variance, as follows:

$$\sigma_\eta^2 = \frac{(n-1)\sum_{i=1}^{n}\left(\eta_{ji} - \eta\right)^2}{n} \tag{3}$$

Finally, to calculate the confidence interval for the parameter of interest, we subtracted (or added) the product of the critical values of the t-distribution with $n-1$ degrees of freedom and the bootstrap estimates for the standard error.

5. Results and Discussion

The descriptive summary of the study presented in Table 1 shows that the mean prevalence of teenage obesity among 9th–12th graders in the U.S. is 14.02%. The mean in-hospital breastfeeding initiation rate from 1998 to 2001 was 68.5%. Approximately 23.38% of the teenagers in our study sample participated in physical education classes on all five days of an average school week, while 26.37% slept for at least eight hours at night in an average school week. Approximately 92.64% and 92.31% of the 9th–12th graders in our sample ate vegetables and fruits, respectively, during the seven-day period prior to the survey. There were approximately 3.15 fast food restaurants per 10,000 population and the mean prevalence of adult obesity was 29.23%. There was a preponderance of the White ethnic group (68.43%) in the study sample, while the Black and Hispanic ethnic groups accounted for 11% each.

The results of the jackknife sample reuse procedure applied to the three-way ECM, presented in Table 2, revealed that in-hospital breastfeeding initiation is statistically significantly associated with a 9.1% reduction in the current teenage obesity rate. Physical activity was associated with a 2.3% reduction in the teenage obesity rate. Moreover, vegetable consumption reduced the teenage obesity rate by some 20.6%, and a rise in median income was associated with a 2.8% reduction in teenage obesity. A percentage increase in the population proportion of Whites, Blacks, and Hispanics was associated with a 6%, 4%, and 4.4% reduction in teenage obesity, respectively. For comparison with our preferred one-delete jackknife resampling plan estimation in the context of the three-way ECM reported in Table 2, the results of the (less preferred) alternative estimation techniques are presented in the Appendix B.

Table 2. Jackknife three-way error component (ECM) regression model results of hospital breastfeeding initiation.

Variable	Jackknife 3-Way ECM	Minimum Jackknife Estimates	Maximum Jackknife Estimates
In-hospital breastfeeding	–0.0964 ***	–0.104	–0.079
	(0.037)		
Physical activity	–0.023 *	–0.028	–0.019
	(0.012)		
Hours of sleep	0.012	–0.002	0.021
	(0.025)		
Vegetable consumption	–0.206 ***	–0.232	–0.183
	(0.08)		
Fruit consumption	–0.042	–0.082	0.031
	(0.081)		
Fast food consumption	0.131	–0.128	0.386
	(0.538)		
Adult obesity	0.244 **	0.175	0.321
	(0.103)		
Living with a single parent	–0.071 ***	–0.093	–0.048
	(0.017)		
Log of income	–2.809	–4.542	–1.912
	(2.539)		
Log of home value	–0.613	–1.651	0.321
	(1.759)		
Non-Hispanic White	–0.06 *	–0.105	–0.047
	(0.032)		
Non-Hispanic Black	–0.039	–0.079	–0.013
	(0.044)		
Hispanic	–0.044	–0.096	–0.026
	(0.041)		
Neighborhood crime rate	0.003	0.0008	0.003
	(0.002)		
Observations	10,200	10,200	10,200
R^2	0.306	0.310	0.441
RMSE	1.258	1.225	1.265

Standard errors are in parentheses, with significance noted as *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. RMSE (Root Mean Square Error).

Our overall study findings (using the estimation results in Table 2) are generally consistent with the theoretical expectation of an inverse relationship between in-hospital breastfeeding initiation and later teenage obesity rates. Specifically, our main variable of interest, the in-hospital initiation breastfeeding rate, was negatively correlated with the prevalence of teenage obesity. A rise in physical activity also had a significant negative effect on teenage obesity, a finding consistent with (Chou et al. 2004) behavioral model of obesity, in which physical activity is a determinant of caloric expenditure. Moreover, the effects of vegetable consumption and an increasing population of the White ethnic group had the appropriate sign and were statistically significant.

Used as controls, the consumption of fruit or 100% juice had the expected negative sign but had no statistically significant effect on the prevalence of teenage obesity. The density of fast food serving restaurants also had the expected positive sign but was not statistically significant. The insignificant effect of fruit and fast-food consumption may be due to the statistically significant effect of vegetable consumption. The controls for the Black and Hispanic ethnic population groups indicate diminishing effects on teenage obesity and were statistically significant. The adult obesity rate, included as a proxy for genetics and other obesogenic factors in the household environment, had the expected positive sign and was statistically significant. The median home value, included in the model as a proxy for permanent household income, had a negative but an insignificant effect on the teenage obesity rate. As expected, we detected that a higher crime rate correlates positively (but insignificantly) with an increased teenage obesity rate, as residential areas with high crime rates discourage the participation of teenagers in outdoor activities and public parks.

Detailed results of breastfeeding duration are provided in Table 3. The obesity reduction effect of breastfeeding falls (but remains statistically significant) as breastfeeding duration lengthens, so that a 12-month breastfeeding duration is associated with a 0.53% (3.7% of a standard deviation) reduction in the average teenage obesity prevalence. Overall, the effects of breastfeeding duration on teenage obesity decrease at an increasing rate.

Table 3. Jackknife three-way error component (ECM) regression model results of breastfeeding durations.

Dependent Variable: Teenage Obesity	Three Months	Six Months	Twelve Months
Breastfeeding duration	–0.0761 ***	–0.0581 ***	–0.0369 **
	(0.0147)	(0.0154)	(0.0146)
Physical activity	0.0206	0.0309	0.0179
	(0.0334)	(0.0310)	(0.0342)
Hours of sleep	–0.0711 **	–0.0657 **	–0.0757 ***
	(0.0282)	(0.0287)	(0.0273)
Hours of watching TV	0.164 ***	0.163 ***	0.179 ***
	(0.0366)	(0.0360)	(0.0330)
School grade	0.0438	0.0730	0.0328
	(0.136)	(0.139)	(0.137)
Non-Hispanic White	–0.00613	–0.00295	–0.00732
	(0.0118)	(0.0111)	(0.0116)
Non-Hispanic Black	0.0126	0.0201	0.0143
	(0.0245)	(0.0250)	(0.0251)
Hispanic	–0.0121	–0.0178	–0.0164
	(0.0266)	(0.0259)	(0.0243)
Access to primary care physician	–0.0103 *	–0.0104 *	–0.0106 *
	(0.00577)	(0.00588)	(0.00561)
Living with a single parent	0.503 **	–0.0498	0.447 *
	(0.248)	0.162	(0.237)
Fast food consumption	0.257	0.530 **	0.000158
	(0.187)	(0.244)	(0.00115)
Neighborhood crime rate	0.106 ***	–0.000382	0.104 ***
	(0.0286)	(0.00129)	(0.0262)
Unemployment rate	–0.189	0.373 **	–0.333
	(0.379)	(0.189)	(0.361)

Table 3. *Cont.*

Dependent Variable: Teenage Obesity	Three Months	Six Months	Twelve Months
Youth smoking	23.64	0.104 ***	32.04 *
	(18.88)	(0.0316)	(17.70)
Teenagers' gender (male or female)	−0.764	−0.0502	−0.06
	0.664	(0.364)	(0.01)
Jackknife resample size	7721	8354	6231
Mean dep. var. (obesity prevalence)	20.20	20.43	20.82
Sample mean of breastfeeding rates	51.79	48.44	20.67
State/time-fixed effects	Yes	Yes	Yes
Flexible controls	No	Yes	No

Standard errors are in parentheses, with significance noted as *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

6. Conclusions

This study used U.S. data on the in-hospital initiation (at birth) and duration of infant breastfeeding as an independent variable of interest in a regression model estimation of the determinants of the prevalence rate of teenage obesity. Data from the 50 U.S. states, including DC, on breastfeeding trends from 1998 through to 2002 were obtained from the Ross Products Division, a subsidiary of Abbot Labs (a major manufacturer of infant nutrition formula), which conducts nationally representative surveys of new mothers and how they feed their newborns. We augmented the dataset with the behavioral risk factor data from the Centers for Disease Control and Prevention's High School Youth Risk Behavioral Survey (YRBS) and multiple other sources, and linked the prevalence of in-hospital breastfeeding initiation and breastfeeding duration (3, 6, and 12 months) of the birth year with the current prevalence of teenage obesity among 9th–12th graders in each of the U.S. states.

By controlling for other confounding factors and using a one-delete jackknife resampling data reuse plan, we estimated a three-way ECM using the Fuller and Battese (1974) method, which combined the errors across the state cross-sections, across grades in each state, and a combination of the two.

Our results suggest that the state-level prevalence of infant breastfeeding initiation at birth correlates negatively with later teenage obesity. Our findings further suggest that teenage obesity is negatively influenced by physical activity, vegetable consumption, short-run household income, and ethnicity. The economic benefits of breastfeeding, measured as averted adulthood obesity-related illnesses and untimely mortality (Rollins et al. 2016), would easily justify instituting and monitoring the progress of policies that promote optimal breastfeeding of newborn babies in the hospital and post-discharge.

Our findings support the U.S. government's policy provisions in the Affordable Care Act (Drago et al. 2010); Healthy People 2020, and the Department of Health and the U.S. Human Services' goals of raising long-term U.S. population health through increased in-hospital infant breastfeeding initiation. Furthermore, the spectacular historical rise in the U.S. life expectancy at birth of 77.6 years (80.1 years for women and 74.8 years for men) appears to be nearing an end, according to (Olshansky et al. 2005), who projected a three- to five-year decrease in U.S. life expectancy at birth attributable to the childhood obesity epidemic.

Besides its negative health effects, obesity is responsible for one-fifth of total health expenditure in the U.S. (Tremmel et al. 2017), and the cost associated with teenage obesity is estimated to be approximately USD 14 billion annually. If the obesity rate (children and adults) was to be maintained at its 2010 rate, the U.S. could save USD 550 billion within the next two years (Biener et al. 2020). Instituting state-level policies to stem the tide of obesity could also significantly reduce the staggering cost of unemployment, job disruption (work absenteeism), and disability among the obese population (Biener et al. 2018). The CDC projects that 13.7 million children are obese, with annual direct medical costs of USD 14 billion. Using our estimates to perform a back-of-the-envelope calculation of annual cost savings from curbing teenage obesity, a 1% reduction in teenage obesity could save USD 140 million in medical expenses. Thus, achieving the Healthy People 2020 targets for the U.S. population by focusing on achieving significantly greater in-hospital breastfeeding rates and post-discharge breastfeeding duration could save the U.S. economy millions.

Author Contributions: Conceptualization, A.O., A.R.O., T.A. and A.K.O.; methodology, A.O. and A.R.O.; software, A.R.O. and T.A.; validation, A.R.O., T.A. and A.O.; formal analysis, A.R.O., T.A. and A.O.; investigation, A.O., A.R.O. and T.A.; resources, A.O., A.R.O. and T.A.; data curation, A.O., A.R.O. and T.A.; writing—original draft preparation, A.O., A.R.O., T.A. and A.K.O.; writing—review and editing, A.O., A.R.O. and T.A.; visualization, A.O., A.R.O., T.A. and A.K.O.; supervision, A.O.; project administration, A.O. All authors read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All of the datasets utilized in this study can be obtained through publicly available sources.

Acknowledgments: The authors thank the participants at the annual Midwest Economics Association (MEA) conference and at the Biennial Congress of the International Health Economics Association (iHEA) for their helpful comments. They acknowledge the help of George Relyea (Statistical Consulting Group) for providing the PROC TSCSREG codes in SAS (Statistical Analysis System, Cary, NC) for estimation of the three-way error components model (ECM) using the jackknife data resampling algorithm. The standard disclaimer applies, however.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

The econometric model of this study is derived from a class of linear estimation methods in which the error term is componential. In Equation (1), ξ_i, τ_i, e_{ijt} are independently distributed as $\xi_i = N_p(O,\ Z)$, $\tau_i = N_p = (0,\ \Theta)$, $e_{ijt} = N_p(0,\ \Gamma)$. Here, $Z,\ \Theta$ is positive semi-definite, while Γ is positive definite, all of the order q (Davis 2002) and with variances $\sigma^2_\xi, \sigma^2_\tau, \sigma^2_e$. Furthermore, it is assumed that $E[\xi_i \xi_s] = 0$ for $i \neq s$, that $E[\tau_t \tau_n] = 0$ for $t \neq n$, and that $E\left[e_{ijt} e_{tsk}\right] = 0$ for $ijt = tsk$. For convenience and without loss of generality, we can write Equation (1) in matrix notation with the assumption that each column vector of the x matrix sums to zero:

$$Obesity_{ijt} = \begin{bmatrix} BF_{111} \\ BF_{112} \\ BF_{NJT} \end{bmatrix}, (1,x) = \begin{bmatrix} 1 & x_{111} & x_{p11} \\ 1 & x_{112} & x_{p12} \\ 1 & x_{1NT} & x_{pNT} \end{bmatrix}, \ \varepsilon = \begin{bmatrix} \xi_1 + \tau_1 + e_{111} \\ \xi_2 + \tau_2 + e_{222} \\ \xi_N + \tau_T + e_{NJT} \end{bmatrix} \tag{A1}$$

Taking Equation (A1), we can rewrite Equation (1) as:

$$obesity = (1,x)\begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \varepsilon \tag{A2}$$

where *obesity* is an $NJTx1$ matrix, x is $NJTx\ p$, α is a scalar term, β is the $px1$ vector, and ε is an $NJTx1$ vector of random disturbances. Based upon assumptions about the error vectors, the $NJTxNJT$ variance–covariance matrix Ω can be written as:

$$E\left[\varepsilon, \varepsilon'\right] = \Omega = \sigma^2_\xi I_{NJT} + \sigma^2_t A + \sigma^2_e B \tag{A3}$$

In Equation (A3), I_{NJT} is an identity matrix and A and B are $NJTxNJT$ matrices.[16] Given the distributional assumption provided earlier, it is known that the normal equation that provides the best linear unbiased estimates of the coefficients of interest is:

$$\begin{pmatrix} 1' \\ x' \end{pmatrix}\Omega^{-1}(1,x)\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} 1' \\ x' \end{pmatrix}\Omega^{-1} obesity \tag{A4}$$

The form of Ω^{-1} is the following $NJT\ x\ NJT$ matrix:

$$\Omega^{-1} = \frac{1}{\sigma_w^2}\left(I_{NJT} - \rho_1 A - \rho_2 B + \rho_3 L_{NJT}\right) \tag{A5}$$

L_{NJT} is an $NJT\ x\ NJT$ matrix with ones everywhere, and ρ_i indicates the functions of the variance components that are estimated using Fuller and Battese (1974) method and are provided in the empirical result of this study as indicated below:

$$\rho_1 = \frac{\sigma_\xi^2}{\left(\sigma_w^2 + T\sigma_\xi^2\right)} \tag{A6}$$

$$\rho_2 = \frac{\sigma_\tau^2}{\left(\sigma_w^2 \ + N\sigma_\tau^2\right)} \tag{A7}$$

$$\rho_3 = \frac{\sigma_\xi^2\sigma_\tau^2}{\left(\sigma_w^2 + T\sigma_\xi^2\right)\left(\sigma_w^2 + T\sigma_\tau^2\right)}\left[\frac{2\sigma_\tau^2 + N\sigma_\tau^2 + T\sigma_\xi^2}{\sigma_\tau^2 + N\sigma_\tau^2 + T\sigma_\xi^2}\right] \tag{A8}$$

Hence, all of the coefficients in the second term of Equation (A2), including hospital breastfeeding initiation and breastfeeding duration, which are recursively linked to teenage obesity, are asymptomatically normal with a smaller generalized variance for both finite and repeated samples.

Appendix B

Table A1. Estimates of alternative models. Of breastfeeding in-hospital initiation.

Variable	OLS	WLS	Classical Box–Cox	Box–Cox Elasticity	GLS	Multi-Level Mixed Model
In-hospital breastfeeding	–0.125 ***	–0.244 ***	–0.011 ***	–0.699	–0.132 ***	−0.096
	(0.0234)	(0.0250)	(0.002)		(0.222)	(0.169)
Physical activity	–0.0922 *	–0.0815 *	–0.002 **	–0.05	–0.032 **	0.236
	(0.0487)	(0.0451)	(0.001)		(0.013)	(0.205)
Sleep	0.0607	0.0634	0.005 *	0.118	0.042 *	−0.057 ***
	(0.0399)	(0.0387)	(0.003)		(0.023)	(0.012)
Vegetable consumption	–0.183 **	–0.378 *	–0.015 **	–1.319	–0.221 **	−0.0215 **
	(0.0867)	(0.0960)	(0.007)		(0.086)	(0.0102)
Fruit consumption	0.0716	0.0565	0.008	0.697	0.117	0.120 ***
	(0.0907)	(0.0865)	(0.007)		(0.088)	(0.027)
Fast food	0.0214	0.212	–0.007	–0.019	0.135	−0.021 **
	(0.292)	(0.329)	(0.025)		(0.278)	(0.010)
Adult obesity	0.243 ***	0.103 **	0.016 **	0.429	0.243 ***	0.120 ***
	(0.0765)	(0.0785)	(0.006)		(0.074)	(0.027)
Single-parent household	–0.106 ***	–0.0948 ***	–0.009 ***	–0.284	–0.091 ***	−0.150
	(0.0233)	(0.0287)	(0.003)		(0.024)	(0.65)
Log of income	–4.020 ***	–1.594 *	–0.218 **	–2.264	–3.276 **	1.88
	(1.455)	(0.864)	(0.112)		(1.393)	(0.131)

Table A1. *Cont.*

Variable	OLS	WLS	Classical Box–Cox	Box–Cox Elasticity	GLS	Multi-Level Mixed Model
Log of home value	–0.420	–0.0767 ***	–0.065	–0.746	–0.384	0.062
	(0.992)	(0.0150)	(0.075)		(0.991)	(0.084)
White	–0.0699 ***	–0.0668 ***	–0.006 ***	–0.394	–0.07 ***	0.050
	(0.0134)	(0.0241)	(0.001)		(0.013)	(0.022)
Black	–0.0547 **	–0.0468 **	–0.004 **	–0.046	–0.053 **	−0.006
	(0.0228)	(0.0201)	(0.002)		(0.022)	(0.0001)
Hispanic	–0.0454 **	0.00361 ***	–0.004 **	–0.044	–0.046 **	−0.018
	(0.0218)	(0.000903)	(0.002)		(0.02)	(0.1433)
Crime rate	0.00321 ***	0.514	0.0003 ***	0.098	0.003 ***	0.085
	(0.000812)	(0.438)	(0.00007)		(0.008)	(0.087)
10th grade	0.522	0.552	0.027	0.006	–	0.267
	(0.413)	(0.580)	(0.034)			0.251
11th grade	0.541	0.427	0.028	0.007	–	0.754 ***
	(0.548)	(0.682)	(0.044)			(0.026)
12th grade	0.432	0.212	0.017	0.004	–	0.076
	(0.658)	(0.329)	(0.05)			(0.013)
Constant	83.62 ***	56.32 ***	7.442 ***	10.429	73.267 ***	
	(16.26)	(15.17)	(1.321)		(15.42)	
Observations	10,200	10,200	10,200	10,200	10,200	10,200
R^2	0.639	0.60	0.632		0.639	
Mean VIF	3.94	–	–	–	–	–
Ramsay RESET (Pr)	0.86					
LR test ($\chi^2{}_1$)	276.6 ***	119.2 ***	–	–	–	474.85 ***
RMSE	7.01					
F	20.00 ***	11.53 ***	18.81 ***	-	23.863 ***	–
Box–Cox λ	–		0.02			

Standard errors are in parentheses, with significance noted as *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$. OLS (Ordinary Least Square), GLS (Generalized Least Square), WLS (Weighted Least Square).

Notes

1. Until recently, Ross Products were the primary source of data on breastfeeding for the U.S.
2. https://nccd.cdc.gov/youthonline/App/Results.aspx?LID=XX (accessed on 29 May 2021).
3. http://web.archive.org/web/20070101180255/http://www.ross.com:80/images/library/BF_Trends_2002.pdf (accessed on 29 May 2021).
4. http://www.washingtonexaminer.com/what-is-your-states-fast-food-ranking/article/2565698 (accessed on 29 May 2021).
5. https://nccd.cdc.gov/NPAO_DTM/IndicatorSummary.aspx?category=28&indicator=29 (accessed on 29 May 2021).
6. http://www.americashealthrankings.org/explore/2016-annual-report/measure/Medianincome/state/AL (accessed on 29 May 2021).
7. https://www.census.gov/prod/2013pubs/acsbr12-20.pdf (accessed on 29 May 2021).
8. https://ucr.fbi.gov/crime-in-the-u.s/2015/crime-in-the-u.s.-2015/tables/table-5 (accessed on 29 May 2021).
9. http://kff.org/other/state-indicator/distribution-by-raceethnicity/?currentTimeframe=0&selectedDistributions=white--black--hispanic--asian (accessed on 29 May 2021).
10. Technical explanations of this method are provided in (Fuller and Battese 1974).
11. The technical explanation of the model is provided in Appendix A.
12. Weights were applied to the median household income and the median home value.
13. We provide the results for the alternative models in the Appendix B (Table A1). The results clearly indicate that the estimates are different in each model given the specific underlying assumptions.
14. The estimates of the classic Box–Cox model are presented in the Appendix for comparison purposes.
15. For detailed description of Multi-level modeling, see (Yang and Schmidt 2021).
16. For a detailed explanation, see (Wallace and Hussain 1969).

References

Al-Amin, A. K. M. Abdullah, and Md Jamal Hossain. 2019. Impact of non-farm income on welfare in rural Bangladesh: Multilevel mixed-effects regression approach. *World Development Perspectives* 13: 95–102. [CrossRef]

Almond, Douglas, and Janet Currie. 2011. Human Capital Development before Age Five. *Handbook of Labor Economics* 4: 1315–486. [CrossRef]

Barker, David J. 1995. Fetal origins of coronary heart disease. *BMJ: British Medical Journal* 311: 171. [CrossRef] [PubMed]

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

Beran, Rudolf. 1986. Jackknife, Bootstrap and Other Resampling Methods in Regression-Analysis—Discussion. *Annals of Statistics* 14: 1295–98. [CrossRef]

Biener, Adam, John Cawley, and Chad Meyerhoefer. 2018. The Impact of Obesity on Medical Care Costs and Labor Market Outcomes in the US. *Clinical Chemistry* 64: 108–17. [CrossRef] [PubMed]

Biener, Adam I., John Cawley, and Chad Meyerhoefer. 2020. The medical care costs of obesity and severe obesity in youth: An instrumental variables approach. *Health Economics* 29: 624–39. [CrossRef]

Borjas, George J. 2003. The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *Quarterly Journal of Economics* 118: 1335–74. [CrossRef]

Box, George E. P., and David R. Cox. 1964. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)* 26: 211–52. [CrossRef]

Campbell, Frances, Gabriella Conti, James J. Heckman, Seong Hyeok Moon, Rodrigo Pinto, Elizabeth Pungello, and Yi Pan. 2014. Early childhood investments substantially boost adult health. *Science* 343: 1478–85. [CrossRef] [PubMed]

Cawley, John. 2010. The economics of childhood obesity. *Health Affairs* 29: 364–71. [CrossRef]

Chou, Shin-Yi, Michael Grossman, and Henry Saffer. 2004. An economic analysis of adult obesity: Results from the Behavioral Risk Factor Surveillance System. *Journal of Health Economics* 23: 565–87. [CrossRef] [PubMed]

Ciardi, Christian, Marcel Jenny, Alexander Tschoner, Florian Ueberall, Josef Patsch, Michael Pedrini, Christoph Ebenbichler, and Dietmar Fuchs. 2012. Food additives such as sodium sulphite, sodium benzoate and curcumin inhibit leptin release in lipopolysaccharide-treated murine adipocytes in vitro. *British Journal of Nutrition* 107: 826–33. [CrossRef] [PubMed]

Courtemanche, Charles J., Joshua C. Pinkston, Christopher J. Ruhm, and George L. Wehby. 2016. Can changing economic factors explain the rise in obesity? *Southern Economic Journal* 82: 1266–310. [CrossRef]

Currie, Janet. 2009. Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature* 47: 87–122. [CrossRef]

Davis, Peter. 2002. Estimating multi-way error components models with unbalanced data structures. *Journal of Econometrics* 106: 67–95. [CrossRef]

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: 945. [CrossRef]

Donohue, John J., III, and Steven D. Levitt. 2001. The impact of legalized abortion on crime. *The Quarterly Journal of Economics* 116: 379–420. [CrossRef]

Drago, Robert, Jeffrey Hayes, and Youngmin Yi. 2010. *Better Health for Mothers and Children: Breastfeeding Accommodations under the Affordable Care Act*. Washington, DC: Institute for Women's Policy Research.

Dumouchel, William H., and Greg J. Duncan. 1983. Using Sample Survey Weights in Multiple-Regression Analyses of Stratified Samples. *Journal of the American Statistical Association* 78: 535–43. [CrossRef]

Efron, Bradley. 1982. *The Jackknife, the Bootstrap and Other Re-Sampling Plans*. CBMS-NSF Regional Conference Series in Applied Mathematics 38. Philadelphia: Society for Industrial and Applied Mathematics (SIAM).

Finkelstein, Eric A., Christopher J. Ruhm, and Katherine M. Kosa. 2005. Economic causes and consequences of obesity. *Annual Review of Public Health* 26: 239–57. [CrossRef] [PubMed]

Fish, Jason S., Susan Ettner, Alfonso Ang, and Arleen F. Brown. 2010. Association of perceived neighborhood safety on body mass index. *American Journal of Public Health* 100: 2296–303. [CrossRef]

Francesconi, Marco, and James J. Heckman. 2016. Symposium on Child Development and Parental Investment: Introduction. *Economic Journal* 126: F1–F27. [CrossRef]

Frank, Lawrence Douglas, Brian E. Saelens, Ken E. Powell, and James E. Chapman. 2007. Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? *Social Science & Medicine* 65: 1898–914.

Fuller, Wayne A., and George E. Battese. 1974. Estimation of linear models with crossed-error structure. *Journal of Econometrics* 2: 67–78. [CrossRef]

Grossman, Herschel I. 1972. A choice-theoretic model of an income-investment accelerator. *The American Economic Review* 62: 630–41.

Keane, Michael P., and Kenneth I. Wolpin. 1997. The career decisions of young men. *Journal of Political Economy* 105: 473–522. [CrossRef]

Martin, Richard M., David Gunnell, Christopher G. Owen, and George Davey Smith. 2005. Breast-feeding and childhood cancer: A systematic review with metaanalysis. *International Journal of Cancer* 117: 1020–31. [CrossRef] [PubMed]

Metzger, Molly W., and Thomas W. McDade. 2010. Breastfeeding as obesity prevention in the United States: A sibling difference model. *American Journal of Human Biology* 22: 291–96. [CrossRef] [PubMed]

Okunade, Albert A., Mustafa C. Karakus, and Charle Okeke. 2004. Determinants of health expenditure growth of the OECD countries: Jackknife resampling plan estimates. *Health Care Management Science* 7: 173–83. [CrossRef]

Olshansky, S. Jay, Douglas J. Passaro, Ronald C. Hershow, Jennifer Layden, Bruce A. Carnes, Jacob Brody, Leonard Hayflick, Robert N. Butler, David B. Allison, and David S. Ludwig. 2005. A potential decline in life expectancy in the United States in the 21st century. *New England Journal of Medicine* 352: 1138–45. [CrossRef]

Quenouille, Maurice H. 1956. Notes on bias in estimation. *Biometrika* 43: 353–60. [CrossRef]

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: 43–72. [CrossRef]

Rollins, Nigel C., Nita Bhandari, Nemat Hajeebhoy, Susan Horton, Chessa K. Lutter, Jose C. Martines, Ellen G. Piwoz, Linda M. Richter, and Cesar G. Victora. 2016. Why invest, and what it will take to improve breastfeeding practices? *The Lancet* 387: 491–504. [CrossRef]

Rosales-Rueda, Maria Fernanda. 2014. *Impact of Early Life Shocks on Human Capital Formation: Evidence from el nino Floods in Ecuador*. University of California Irvine Working Paper. Irvine: University of California Irvine.

Rosin, Odelia. 2008. The economic causes of obesity: A survey. *Journal of Economic Surveys* 22: 617–47. [CrossRef]

Sauder, Katherine A., Traci A. Bekelman, Kylie K. Harrall, Deborah H. Glueck, and Dana Dabelea. 2019. Gestational diabetes exposure and adiposity outcomes in childhood and adolescence: An analysis of effect modification by breastfeeding, diet quality, and physical activity in the EPOCH study. *Pediatric Obesity* 14: e12562. [CrossRef]

Schultz, Theodore W. 1974. Marriage, Family, Human Capital, and Fertility. *Journal of Political Economy* 82: 2. [CrossRef]

Simon, Julian L., and Peter Bruce. 1991. Resampling: A tool for everyday statistical work. *Chance* 4: 22–32. [CrossRef]

Slentz, Cris A., Lori B. Aiken, Joseph A. Houmard, Connie W. Bales, Johanna L. Johnson, Charles J. Tanner, Brian D. Duscha, and William E. Kraus. 2005. Inactivity, exercise, and visceral fat. STRRIDE: A randomized, controlled study of exercise intensity and amount. *Journal of Applied Physiology* 11: 38–39. [CrossRef]

Stevens, Bonnie, Denise Guerriere, Patricia McKeever, Ruth Croxford, Karen-Lee Miller, Jo Watson-MacDonell, Sharyn Gibbins, Michael Dunn, Arne Ohlsson, and Karen Ray. 2006. Economics of home vs. hospital breastfeeding support for newborns. *Journal of Advanced Nursing* 53: 233–43. [CrossRef]

Tremmel, Maximillian, Ulf-G. Gerdtham, Peter M. Nilsson, and Sanjib Saha. 2017. Economic Burden of Obesity: A Systematic Literature Review. *International Journal of Environmental Research and Public Health* 14: 435. [CrossRef] [PubMed]

Van den Berg, Gerard J., Maarten Lindeboom, and France Portrait. 2006. Economic conditions early in life and individual mortality. *American Economic Review* 96: 290–302. [CrossRef]

Victora, Cesar G, Bernardo Lessa Horta, Christian Loret De Mola, Luciana Quevedo, Ricardo Tavares Pinheiro, Denise P. Gigante, Helen Gonçalves, and Fernando C. Barros. 2015. Association between breastfeeding and intelligence, educational attainment, and income at 30 years of age: A prospective birth cohort study from Brazil. *The Lancet Global Health* 3: e199–e205. [CrossRef]

Victora, Cesar G., Rajiv Bahl, Aluisio J. Barros, Giovanny V. Franca, Susan Horton, Julia Krasevec, Simon Murch, Mari Jeeva Sankar, Neff Walker, Nigel C. Rollins, and et al. 2016. Breastfeeding in the 21st century: Epidemiology, mechanisms, and lifelong effect. *Lancet* 387: 475–90. [CrossRef]

Wallace, Thomas D., and Ashiq Hussain. 1969. The use of error components models in combining cross section with time series data. *Econometrica: Journal of the Econometric Society* 37: 55–72. [CrossRef]

Walters, Christopher R. 2015. Inputs in the production of early childhood human capital: Evidence from Head Start. *American Economic Journal: Applied Economics* 7: 76–102. [CrossRef]

Yang, Yimin, and Peter Schmidt. 2021. An econometric approach to the estimation of multi-level models. *Journal of Econometrics* 220: 532–43. [CrossRef]

Zaqout, Mahmoud, Nathalie Michels, Wolfgang Ahrens, Claudia Börnhorst, Dénes Molnár, Luis A. Moreno, Gabriele Eiben, Alfonso Siani, Stalo Papoutsou, and Toomas Veidebaum. 2018. Associations between exclusive breastfeeding and physical fitness during childhood. *European Journal of Nutrition* 57: 545–55. [CrossRef]

Zweifel, Peter. 2012. The Grossman model after 40 years. *The European Journal of Health Economics* 13: 677–82. [CrossRef] [PubMed]