Published in final edited form as:

Matern Child Health J. 2009 November ; 13(6): 780–794. doi:10.1007/s10995-009-0461-6.

Birth Weight Effects on Children's Mental, Motor, and Physical Development: Evidence from Twins Data

Ashlesha Datar, PhD and

RAND Corporation, 1776 Main Street, Santa Monica, CA 90407, datar@rand.org, Phone: (310) 393 0411 x 7367, Fax: (310) 260 8161

Alison Jacknowitz, PhD, MPP

American University, Department of Public Administration and Policy, 4400 Massachusetts Avenue, NW, Washington, DC 20016, jacknowi@american.edu, Phone: (202) 885 2137, Fax: (202) 885 2347

Abstract

Objective—To determine the effect of very low birth weight (VLBW; <1500g) and moderately low birth weight (MLBW; 1500–2499g) on children's mental and motor development and physical growth during the first two years of life and whether VLBW and MLBW babies catch up to normal birth weight (NBW; ≥2500g) children by age 2.

Methods—We use data on dizygotic (DZ) and monozygotic (MZ) twins and singleton births from the first two waves of the Early Childhood Longitudinal Study — Birth Cohort (ECLS-B), a nationally representative dataset of children born in the U.S. in 2001. We estimate the effects of VLBW and MLBW on children's mental and motor development scores, weight-for-age, weight-for-length, weight-for-height, and length-for-age z-scores at 9 months and 2 years. We examine whether differences in outcomes *within* twin pairs are related to differences in their birth weights. The within-twins analysis is conducted on samples of DZ and MZ twins. For comparison, we also estimate birth weight effects on child outcomes from multivariate linear regression models using the full singleton and twins sample. We also estimate the effect of being small-for-gestational age (SGA; birth weight <10th percentile for gestation) using the same set of models in order to separate out the effects of fetal growth restriction from prematurity.

Results—Evaluation of all births showed that VLBW and MLBW have large negative effects on mental development, motor development, and growth at 9 months and 2 years of age. However, results from within-twin models with DZ twins that control for shared maternal and environmental factors showed much less effect of birth weight on mental or motor development, but continued large effects on growth for the VLBW group. Within-twin models with MZ twins that control for shared maternal, environmental, *and* genetic factors showed statistically insignificant effects of birth weight on mental and motor development, but continued effects on growth. Similar patterns were found when examining the effects of SGA.

Conclusions—After controlling for the influence of maternal, environmental, and genetic factors, low birth weight has at most a small negative effect on children's mental and motor development in their first two years of life. However, low birth weight is a major risk factor for children's physical growth in the early years and there is no evidence of catch-up by age 2.

Keywords

Birth weight; Twins; Mental development; Motor development; Growth

INTRODUCTION

In 2004, 8.1 percent of all newborns in the U.S. were born at a low birth weight (LBW); conventionally defined as birth weight less than 2500 grams. This is the highest percentage since the early 1970s¹. A significant body of research finds a strong association between LBW and a wide variety of developmental and health outcomes in childhood and adulthood. LBW has been associated with adult conditions such as Type 2 diabetes, hypertension, and coronary heart disease^{2–4}. Among children, LBW has been linked to high blood pressure, cerebral palsy, deafness, blindness, asthma, lung disease, poor cognitive and physical performance, and mortality^{5–14}.

Most of this literature, however, is based on cross-sectional study designs that have limited ability to control for maternal, environmental, and genetic factors that are also correlated with birth weight and later outcomes. Therefore, it is unclear whether these associations are caused by intrauterine malnutrition *per se* or whether they are mediated by genetic and/or shared environmental effects.

Studies based on twins data have made a significant contribution to this literature because of their ability to address such concerns. Differences in birth weight within twin pairs reflect differences in fetal growth but not gestational age. Within-twins analyses based on fraternal (dizygotic; DZ) twin pairs allow matching on all observed and unobserved maternal factors (e.g., maternal height, maternal nutritional status, maternal diseases) and environmental factors that are shared by twins, and within-twins analyses based on identical (monozygotic, MZ) twin pairs additionally allow matching on genetic factors.

However, studies based on differences in birth weight within twin pairs are relatively few and are typically based on samples that have limited geographic variability and, hence, have limited generalizeability. Further, these studies have focused on the effect of being LBW on adult diseases (see Iliadou et al¹⁵ and Hubinette et al^{16,17} for an overview of these studies). In children, studies based on within twin pair differences are limited and have mixed findings. A group of studies done in the 1960s and 1970s examined whether within twin pair differences in birth weight among MZ twins were correlated with within twin pair differences in IQ^{18–21}. While most studies found the heavier twin to have higher IQ, one study found no significant difference in IQ between the heavier and lighter twin. Most recently, Newcombe et al²² examined the same question using a large sample of MZ twin pairs born in the United Kingdom and found that the heavier twin had significantly higher IQ compared to the lighter twin.

A recent national survey of a birth cohort allows examination of this issue in a sample of U.S. twins. We use data from the first 2 waves of the Early Childhood Longitudinal Study – Birth Cohort (ECLS-B) to examine the independent effect of very low birth weight (VLBW; <1500g) on children's mental and motor development and physical growth during the first 2 years of life and whether VLBW children catch up to normal birth weight (NBW; ≥2500g) children by age 2. We also examine the independent effect of being moderately low birth weight (MLBW; 1500–2499g) since less is known about children in this range of birth weight²³. We use a “within-group” estimator to examine whether differences in outcomes *within* twin pairs are related to differences in their birth weights. The within-twins analyses are conducted on samples of fraternal and identical twins. We compare the estimates from within-twin models to those from multivariate linear regression models estimated on the full sample of singleton and twin births (hereafter referred to as “cross-sectional” models). This comparison allows us to assess the contribution of unobserved maternal, environmental, and genetic factors in determining birth weight and children's developmental and growth outcomes. We also estimate the effect of being small-for-gestational age (SGA; birth weight < 10th percentile for gestation) on the same set of outcomes using the same set of models. These estimates capture the effects of low birth weight on children's developmental and growth outcomes resulting from fetal growth restriction rather than prematurity.

METHODS

Sample

We use data from the ECLS-B, a longitudinal nationally representative data set that was collected by the National Center for Education Statistics (NCES), U.S. Department of Education²⁴. The ECLS-B follows a nationally representative cohort of children born in 2001 in the U.S. from birth through kindergarten with data collection occurring when the child is approximately nine-months of age, two-years of age, four-years of age (at pre-school), and at kindergarten entry. To date, the first two waves of survey data are available and used in this study. Children were primarily sampled via registered births from the National Center for Health Statistics (NCHS) vital statistics system. Births were sampled within a set of primary sampling units, which were individual counties or groups of contiguous counties. The survey oversampled VLBW and MLBW babies, multiple births, and Asian and American Indian births. The baseline sample consists of 10,221 births that participated in the child assessments. Of the 10,221 children who participated in the first wave of the survey, 8,944 (87.5 percent) of them participated in the second wave. Children who attrited from the sample between the first two waves are relatively similar in terms of maternal education, region of residence, and child's race/ethnicity to those who remained in the sample²⁵.

Our analysis sample consists of about 6,750 singleton births, 625 twin pairs (fraternal pairs = 525, identical pairs = 100), and 50 twins and other higher-order births whose siblings were not included in the ECLS-B. The reported sample sizes for subgroups of the data have been rounded to the nearest 50 per NCES restrictions regarding disclosure of restricted use data. However, the analyses and statistics presented in the tables and text are generated using all observations in each subsample. Births were excluded from the

analysis sample using the following five criteria; the number of observations excluded for each criterion are in parentheses. First, births whose survey respondent was not their biological mother were excluded (150). Second, births that did not participate in the child assessments in each wave were excluded (1,700). Third, births that did not have complete information on birth weight, small-for-gestation, twin status, and type of twin were excluded (400). Fourth, births born with a congenital anomaly were excluded (300). Fifth, births with missing information for any of the control variables were excluded with the exception of the following four variables: first birth, adequacy of prenatal care, presence of pregnancy risk factors, and presence of labor and delivery complications (100). Each of these four variables had more than 150 missing values before the fifth sample restriction; therefore, after implementing the final data restriction we assigned the remaining missing data points the modal value. All models included variables that indicated whether the observation included a value that was mode-replaced for each of these four variables.

Measures

Identification of Twin Type—Twin type in the ECLS-B is identified using parent and interviewer reports. Opposite-sex twin pairs were classified as dizygotic. For same-sex twin pairs, the ECLS-B asked both parents and interviewers questions drawn from existing research that attempt to reliably distinguish identical from fraternal twins^{26–28}. Specifically, both parents and observers were asked whether the twins differed in their hair texture, shades of hair colors, eye color, complexion, facial appearance, and shape of ear lobes. Parents were additionally asked about the similarity of the twins when their first teeth erupted, whether the twins were ever mistaken for each other by either parent, other family members, or strangers and, if yes, how frequently, and whether there were any physical differences between the twins. Parents were also asked about the twins' blood types and RH factors. Using these questions, approximately 16 percent of twins were classified as monozygotic.

Birth Weight Measures—The ECLS-B obtained birth weight information from birth certificates. We created binary variables for VLBW, defined as birth weight less than 1500 grams, and MLBW, defined as birth weight greater than or equal to 1500 grams and less than 2500 grams, to include as the primary explanatory variables of interest. The omitted category is NBW defined as 2500 grams and above. Categorical variables were chosen instead of a continuous birth weight variable since these cutoffs have been traditionally used in the medical literature and this specification allows comparison of our findings with other studies. In alternate models, we used small-for-gestational age (SGA), defined as birth weight less than the 10th percentile for gestational age²⁹, as an alternate measure of low birth weight. Although not accurate, SGA has traditionally been used to capture intrauterine growth restriction (IUGR)³⁰.

Mental and Motor Development Scores—Data on children's mental and motor development are collected through direct child assessments using the Bayley Short Form — Research Edition (BSF-R) in both waves of data. This instrument was specifically designed for the ECLS-B and is derived from the Bayley Scales of Infant Development, Second Edition (BSID-II)³¹, which

is a standardized assessment of developmental status for children from birth to 42 months of age (More details about the assessments are available at <http://nces.ed.gov/pubs2005/2005100.pdf>). The BSF-R administered at 9 months consisted of a 31-item mental scale (33 items at 2 years) and a 35-item motor scale (32 items at 2 years). Each scale included a set of core items that all children were administered as well as a set of basal and ceiling items that were administered depending upon the child's responses on the core items. The mental scale assesses early cognitive and language ability as manifested in memory, habituation, preverbal communication, problem solving, and concept attainment. The motor scale captures gross and fine motor abilities (e.g., picking up objects, crawling, walking), perceptual-motor integration, and problem solving.

Responses on the BSF-R scale were first weighted to adjust for oversampling and nonresponse and subsequently equated to the full BSID-II scale using Item Response Theory (IRT)³¹. This makes it possible to compare the BSF-R scores to those obtained using the full set of BSID-II items. There are a total of 178 mental items and 111 motor items on the full BSID-II, which means that the ECLS-B mental scale scores could range from 0–178 and the motor scale scores could range from 0–111, although in reality all scores lie well within these intervals since the items are designed for children from birth through 42 months. At 9-months the mental scale scores range from 32 to 131 and the motor scale scores range from 21 to 87 in the full sample. At 2 years the mental scale scores range from 92 to 174; while the motor scale scores range from 56 to 109. These scale scores are not age-normed and simply measure the number of items a child would have answered correctly were she administered the full set of BSID-II mental and motor development scale items. However, this is not a concern in our within-twins analyses since both children within the twin pair are assessed at the same age. The reliability of the BSF-R scale scores in the ECLS-B is high; for example, in the 9-month wave the IRT reliability coefficient is 0.80 for the mental scale and 0.93 for the motor scale²⁴. The corresponding numbers in the 2-year wave are 0.88 and 0.73, respectively.

Growth—Children's weight and length/height were measured by the ECLS-B interviewers in both waves of data. Standard procedures were followed to obtain the physical measurements to ensure that all measurements were obtained in the same way for all children. A Seca digital bathroom scale (accuracy = 0.1 kilogram) was used to measure the child's weight in both waves. The mother first stood on the scale and her weight was recorded. The interviewer then handed the child to the mother and the scale's tare feature (in which the child's weight is automatically subtracted from the combined weight of the mother and child) then displayed the child's weight alone. The child's recumbent length was measured at the 9-month interview using a Seca pediatric measure mat. At the 2-year interview, a stadiometer (Seca Model 214 Road Rod) was used to measure the child's height. Assessors were trained to understand the purpose of the measurements and were made familiar with the equipment prior to the study by practicing with training staff as well as infants and 2-year old children specifically recruited for staff training. The reliabilities of the physical measurements were quite high; the correlation between the first and second recordings was $r = .81$ for both weight and height measurements.

Each of these physical measurements was obtained twice to minimize measurement error. If the two measures were more than 5 percent apart a third measurement was taken. Composite values for height and weight were created by using the two measurements that were closest together: the value of the composite was the average of these two values. If the three measurements were equidistant from one another, the value that was furthest from the weighted average of the smallest and largest value was discarded and the composite took the value of the average of the remaining two measures. If the child had two height or weight values that were less than 5 percent apart, the average of the two values was computed and used as the composite value.

We examine weight-for-age (9-month and 2-year data), weight-for-length (9-month data), weight-for-height (2-year data), and length-for-age (9-month data) z-scores. The parameters necessary to calculate these z-scores are from the Centers for Disease Control (CDC) website at <http://www.cdc.gov/nchs/about/major/nhanes/growthcharts/datafiles.htm>. Physical anthropometry protocols call for the measurement of recumbent length up to 2 years of age, and standing height from this age onwards. However, only height data and not recumbent length data were collected at the 2-year wave regardless of the toddler's age at the time of assessment. At the 2-year survey, variation in assessment age across children resulted in roughly half the assessed sample being below 24 months. Hence, the corresponding values (length/height) and z-scores were used depending on the availability of data and the range of the indices (e.g., gender, age) necessary to merge in the information from the CDC. For example, weight-for-height is indexed by height and gender; therefore, we can create weight-for-height z-scores for all births with non-missing data regardless of age at assessment. In comparison, height-for-age is indexed by age and gender starting at 24 months, thus prohibiting the calculation of z-scores for toddlers less than 24 months.

Other Explanatory Variables—Data on child, maternal, household, and pregnancy-related characteristics collected from birth certificates and the nine-month and the two-year surveys are included as explanatory variables. Note that these controls are used only in the cross-sectional models since they compare outcomes across children who differ in maternal, environmental, and genetic factors. Child-level controls include indicators for female, race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, other), first birth, and whether the child was a twin or higher-order birth. In addition, age at the time of assessment in weeks is included because some of our outcome measures such as mental and motor scale scores are not age-normed. The age of assessment for wave 1 is included when examining an outcome measured using the 9-month data and the age of assessment for wave 2 is measured when investigating an outcome measured using the 2-year data. We also include a continuous variable for gestational age, measured in weeks. Controlling for age at assessment and gestational age is critical since mental and motor outcomes are measured at different post conceptual ages. We do not include variables capturing an infant's later health conditions as they may be the result of birth weight status and would prevent us from measuring the total effect of birth weight status on our outcomes.

Maternal characteristics include height in inches and indicator variables for age (less than 20 years, between 20 and 24.99 years, between 25 and 29.99 years, between 30 and 34.99 years, at least 35 years), education (less than high school diploma, high school diploma or equivalent, some college or vocational program, bachelors degree or higher), and married at birth. Household characteristics include indicator variables for income (less than or equal to \$20000, between \$20001 and \$40000, between \$40001 and \$75000, between \$75001 and \$100000, \$100001 or higher), region (Northeast, Midwest, South, West), and urbanicity (population less than 2,500, population of 2,500 to 49,999, population at least 50,000). When examining outcomes at 2 years, variables capturing changes in income and marital status are included.

Pregnancy characteristics include gestation in weeks at pregnancy recognition and as mentioned earlier, gestation in weeks. Other pregnancy-related variables include indicators for whether mother smoked in the last trimester, whether mother drank alcohol in the last trimester, adequacy of prenatal care utilization (inadequate, intermediate, adequate, adequate plus), whether the pregnancy was at-risk, and presence of labor and/or delivery complications. Adequacy of prenatal care is measured by the Adequacy of Prenatal Care Utilization Index³³, which is a function of the month prenatal care began (i.e., adequacy of initiation of prenatal care) and the proportion of the number of visits recommended by the American College of Obstetricians and Gynecologists that a mother received from the time prenatal care began until the time of delivery (i.e., adequacy of received services). Table 1 reports means or proportions and standard errors for the outcome variables, birth weight measures, and selected explanatory variables.

Statistical Analysis

We estimate 3 linear regression models for each of the nine outcome variables discussed above. The first model is a linear regression of the outcome variable on indicators of VLBW and MLBW using the full sample of singletons and twins. This model controls for the child, maternal, household, and pregnancy-related characteristics described in the previous section and mimics the cross-sectional study design typically seen in studies of birth weight effects. The remaining 2 models estimate “within-group” estimates of birth weight effects using samples of fraternal and identical twin pairs, respectively. These within-twins models essentially use one twin as the control for the other twin by regressing within twin pair differences in outcomes on within twin pair differences in birth weight (model (3) described in Carlin et al³⁴). When relating outcomes differences to birth weight differences within fraternal twin pairs, we include a control for child’s gender. Controls for gestational age are not needed in within-twin models since both children in the twin pair have the same gestation. Comparing outcomes of fraternal twins allowed us to control for all maternal and environmental factors shared by the twins. The final model regresses outcome differences within identical twin pairs on differences in their birth weights. Comparing outcomes of identical twins allowed us to additionally control for genetic differences between the twins. The effects of SGA on all the outcomes are also estimated using the 3 models described above.

The exact sample size used to estimate each model varies across different outcomes due to different number of missing values for the dependent variable. For each outcome we include only those births in the sample for whom data on the outcome are available at 9 months and 2 years to allow testing for any catch up between LBW and NBW births at age 2. To ensure that the results from the 2 twins-based models are comparable for each outcome, we only include twin-pairs that have non-missing values for a given outcome in the final 2 models.

All analyses are conducted using the statistical program STATA version 9.2 (StataCorp, College Station, TX). All estimates are weighted using the ECLS-B provided weight, W2C0. For the cross-sectional model, we adjust the standard errors from the regression to account for within-twin pair correlations in the outcomes using the *cluster* option in STATA.

RESULTS

Descriptive Statistics

Table 2 reports descriptive statistics for the outcome variables, birth weight measures, and selected explanatory variables in our analysis for singleton births, twin births, identical twins, and fraternal twins. Two tailed t-tests for differences in means for continuous variables and proportions for binary variables between singletons and all twins and between fraternal and identical twins are conducted, and those with $p < 0.01$ are discussed below. Twins are more likely to be VLBW and MLBW than singleton births; although they are not more likely to be SGA. Twins tend to be born to mothers who are more advantaged. Mothers of twins are more likely to be non-Hispanic White, be older at birth, have a bachelor's degree or higher, have higher income, be married at the time of birth, and receive adequate plus prenatal care compared to mothers of singleton births. As expected, pregnancies with twins tend to be more complicated than those that lead to singleton births, with twin pregnancies more likely to be at risk, to experience labor and delivery complications, and to have shorter mean gestation in weeks. Upon comparing fraternal and identical twins, we see that fraternal twins are born to more advantaged mothers, with mothers of fraternal twins more likely to be non-Hispanic White, married at the time of birth, older at the time of birth, more educated, have higher income, and receive adequate plus prenatal care.

Bivariate Results

Table 2 also reports the means for all the outcomes for singleton births, all twins, fraternal twins, and identical twins. Two-tailed t-tests for mean differences between singletons and twins, and between fraternal and identical twins are conducted. Singletons score significantly higher on all outcomes compared to twins ($p < 0.01$) with the exception of weight-for-length at 9 months ($p = 0.02$). However, there is no statistically significant difference in outcomes between fraternal and identical twins using an $\alpha = 0.01$.

Regression Results

Table 3 reports regression coefficients on VLBW and MLBW for mental and motor development scores. Table 4 reports corresponding estimates for growth outcomes.

Mental Development—Columns 1 and 2 in Table 3 report estimates of VLBW and MLBW on children's mental development scale scores at 9 months and 2 years, respectively, from the 3 models described in the previous section. Cross-sectional estimates using the full sample (Panel A) show that VLBW and MLBW confer a significant disadvantage at 9 months in the form of lower mental development scores compared to NBW, with VLBW having a larger negative effect compared to MLBW. VLBW infants score 9.1 points lower ($p<0.01$), i.e., about 9 fewer correct items, compared to NBW infants. MLBW infants score 3.3 points lower ($p<0.01$), i.e., about 3 fewer correct items compared to NBW infants. At 2 years, the effects of VLBW and MLBW on mental scores remain statistically significant at -7.6 points ($p<0.01$) and -2.6 points, respectively, but the change in the estimates from 9-months to 2-years is not statistically significant at $\alpha=0.01$.

Within-twin estimates from the sample of fraternal twins (Panel B) show much smaller, yet statistically significant, effects of VLBW and MLBW on mental scores compared to the cross-sectional estimates, suggesting that unobserved maternal and environmental factors not controlled for in the cross-sectional regressions are important in determining birth weight and children's mental development. VLBW infants score 2.6 points lower ($p=0.021$) and MLBW infants score 2.0 points lower ($p<0.01$) on the mental development scale at 9 months compared to their NBW twin. At 2 years, the effects of VLBW and MLBW appear to have increased, however, the change in coefficients is not statistically significant at $\alpha=0.01$.

Within-twin estimates from the sample of identical twins are reported in Table 3, Panel C. The effect of VLBW and MLBW on mental development scores at 9 months is about -3.1 points ($p=0.38$) and -2.51 points ($p=0.05$), respectively. A difference of 2 or 3 correct items on the BSID-II test between low birth weight and NBW infants is a very small effect in comparison to the mean number of items answered correctly (77 items). At 2 years, the point estimates for VLBW (0.54 points; $p=0.84$) and MLBW (1.01 points; $p=0.65$) become small and statistically insignificant. This suggests that when maternal, environmental, and genetic factors are controlled for, differences in mental development of VLBW, MLBW, and NBW children at 2 years do not differ significantly.

Motor Development—Columns 3 and 4 in Table 3 report estimates of VLBW and MLBW on children's motor development scale scores at 9 months and 2 years, respectively. The cross-sectional results in Panel A are similar to those seen for mental development. At 9 months, the effects of VLBW (-8.8 points; $p<0.01$) and MLBW (-2.9 points; $p<0.01$) are large and statistically significant. At 2 years, however, the cross-sectional estimates of VLBW and MLBW become considerably smaller and this change is statistically significant at $\alpha=0.01$. Therefore, the cross-sectional estimates suggest some catch-up taking place between LBW and NBW children by age 2.

Within-twins estimates using the sample of fraternal twins (Panel B) show that only the effect of VLBW at 9 months is statistically significant (−3.8 points; $p<0.01$). Within-twins estimates from the sample of identical twins (Panel C) show that the effect of VLBW on motor development scores at 9 months is −2.5 points ($p=0.06$) but becomes smaller and statistically insignificant at 2 years (−0.52 points; $p=0.78$). In contrast, the effect of MLBW on motor development scores is small and statistically insignificant at 9 months (−0.93 points; $p=0.22$), increases slightly at 2 years (−1.64 points; $p=0.08$). However, the differences between the 9-month and 2-year estimates are not statistically significant. Again, these estimates suggest that when maternal, environmental, and genetic factors are controlled for, differences in motor development of VLBW, MLBW, and NBW children at 9 months and 2 years are of small magnitude and do not differ statistically.

Weight-for-Age—Columns 1 and 2 in Table 4 report estimates of VLBW and MLBW on children's weight-for-age z-scores at 9 months and 2 years, respectively. In all the three models, VLBW and MLBW children have significantly lower weight-for-age compared to NBW children. The effects are larger in the cross-sectional models compared to the within-twins models.

Estimates from cross-sectional regressions on the full sample (Panel A) show that VLBW infants have 1.64 SD ($p<0.01$) lower weight-for-age and MLBW infants have 0.77 SD ($p<0.01$) lower weight-for-age than NBW infants at 9 months. Estimates from within-twins analysis on fraternal twins (Panel B) show that VLBW infants have 1.24 SD ($p<0.01$) lower weight-for-age and MLBW infants have 0.33 SD ($p<0.01$) lower weight-for-age than NBW infants at 9 months. Estimates from within-twins analysis on identical twins (Panel C) show that VLBW infants have 0.96 SD ($p<0.01$) lower weight-for-age and MLBW infants have 0.40 SD ($p=0.02$) lower weight-for-age than NBW infants at 9 months. There is no statistically significant change in the effects of VLBW and MLBW at 2 years in any of the three models.

Weight-for-Length/Height—Column 3 in Table 4 reports estimates of the effect of VLBW and MLBW on children's weight-for-length z-scores at 9 months. Estimates from the cross-sectional model (Panel A) show that weight-for-length at 9 months is significantly lower for VLBW (−0.53 SD; $p<0.01$) and MLBW (−0.31 SD; $p<0.01$) infants compared to NBW infants.

Within-twins estimates from the sample of fraternal twins show that the effect of VLBW is strong and negative (−0.63 SD; $p=0.014$) but the effect of MLBW becomes smaller (−0.15 SD; $p=0.16$). Within-twins estimates from the sample of identical twins show that the effects of VLBW (−0.13; $p=0.74$) and MLBW (−0.25; $p=0.22$) are small and statistically insignificant at 9 months. At 2 years, the effects of VLBW and MLBW increase across all the three models. In fact, the within-twins estimates of VLBW (−0.53; $p=0.03$) and MLBW (−0.46; $p=0.03$) from the identical twins sample are significant at the 5% level, but not at the 1% level.

Length-for-Age—Column 5 in Table 4 reports estimates of the effect of VLBW and MLBW on length-for-age z-scores at 9 months. In all 3 of the models, VLBW and MLBW are associated with significantly lower length-for-age than NBW at 9 months ($p=0.06$ for MLBW effect in identical twins).

model; $p < 0.01$ for all other estimates). The estimates are largest in the cross-sectional models, followed by the within-twins estimates from the fraternal twins sample. The estimates from the identical twins sample are the smallest, although statistically significant in the case of VLBW. The cross-sectional estimates from the full sample (Panel A) indicate that VLBW infants have 1.75 SD ($p < 0.01$) lower length-for-age and MLBW infants have 0.75 SD ($p < 0.01$) lower length-for-age compared to NBW infants at 9 months. Estimates from the within-twins analysis on the sample of identical twins (Panel C) show that VLBW infants have about 1.1 SD ($p < 0.01$) lower length-for-age compared to their NBW twin. MLBW infants also have a lower length-for-age compared to their NBW twin (-0.26 SD; $p = 0.06$).

To examine catch-up for this outcome, we also estimated similar models using height-for-age at 2 years as the outcome. For reasons described earlier, we estimated models for length-for-age (at 9 months) and height-for-age (at 2 years) for the sample of children for whom both measures could be constructed. These results are reported in the Appendix. There is a substantial drop in the coefficients between 9 months and 2 years, however, this change is only statistically significant for the VLBW estimates in the cross-sectional models and not in the within-twins models.

Effects of small-for-gestational age—Estimates of the effect of SGA on developmental and growth outcomes at 9 months and 2 years from each of the three models are reported in Table 5. Cross-sectional estimates suggest that SGA infants answer about 1.6 ($p < 0.01$) and 1.3 ($p < 0.01$) fewer correct items on the mental and motor developmental scales, respectively, compared to non-SGA infants at 9 months. There is no statistically significant change at a 1 percent level in these estimates at 2 years. However, within-twins estimates of the effects of SGA on mental and motor development are not significant at either 9 months or 2 years at a 1 percent level.

The effects of SGA on weight-for-age in all 3 models are large and negative at 9 months and remain large even at 2 years. Within-twins estimates from the identical twins sample are smaller compared to estimates from the cross-sectional and fraternal twins models. Estimates from identical twin models show that SGA infants have 0.59 SD ($p < 0.01$) lower weight-for-age at 9 months and 0.41 SD ($p = 0.02$) lower weight-for-age at 2 years compared to their non-SGA twins. The effects of SGA on weight-for-stature/height at 9 months and 2 years are statistically significant ($p < 0.01$) and sizeable in the cross-sectional and within-fraternal twins models. However, the effects of SGA are insignificant in the within-identical twins models (9 months: -0.34 SD; $p = 0.10$; 2 years: -0.31 SD; $p = 0.09$). The effect of SGA on length-for-age z-scores at 9 months is large and statistically significant ($p < 0.01$) in all the three models. The within-identical twins estimate, which controls for maternal, environmental, and genetic factors shows that SGA infants have 0.47 SD ($p < 0.01$) lower length-for-age compared to their non-SGA identical twin.

DISCUSSION

This study examined whether VLBW and MLBW confer a sizeable disadvantage in terms of mental and motor development and physical growth during the first two years of life based on a within-twins analysis

of a nationally representative data set. Conventional cross-sectional models that only controlled for sociodemographic and observable maternal risk factors, including gestational age, showed large detrimental effects of VLBW and MLBW. However, as discussed earlier, when comparing the outcomes of LBW to NBW children, one can only control for a limited set of potential confounders. Other unobserved maternal, environmental, and genetic factors that are correlated with LBW and child development could lead to biased estimates of LBW effects if they are not controlled for. We addressed this concern by using the sample of twins in our data to control for these unobserved influences. We used a “within-group” estimator to examine whether within twin pair differences in mental, motor, and physical growth are related to differences in their birth weight. As such, the within-twins analyses offered a powerful strategy for estimating birth weight effects that are independent of maternal, environmental, and genetic influences.

Similar to findings from twins-based studies that have examined the effects of birth weight on adult health outcomes^{35–42} results from this study suggest that maternal, environmental, and genetic factors play an important role in determining LBW and early life developmental outcomes and that birth weight effects tend to be overstated when these factors are not controlled for. We found that estimates from our within-twins analyses were generally much smaller than what the conventional cross-sectional analysis on the full sample suggested. Moreover, estimates from the within-fraternal twins analyses tended to be generally larger relative to estimates from the within-identical twins analyses. This is not surprising since within-identical twins analyses also control for genetic influences on birth weight and infant outcomes. Within-identical twins analyses did not find any statistically significant effects on children’s mental and motor development in the first two years. However, there was a large and statistically significant effect of VLBW and MLBW on physical growth at 9 months that remained even at the end of 2 years. Moreover, the effects of VLBW tended to be larger than the effects of MLBW.

We also examined whether SGA infants had worse developmental and physical growth outcomes compared to non-SGA infants. Although not perfect, SGA (i.e., birth weight below the 10th percentile for gestation) has traditionally been used to measure low birth weight that results from IUGR. Our results for SGA were qualitatively similar to those for VLBW and MBLW. After controlling for maternal, environmental, and genetic factors, SGA did not have a statistically significant effect on mental and motor development outcomes in the first two years. However, SGA did have large and statistically significant negative effects on children’s growth at 9 months without any evidence of a catch-up at 2 years.

Case-control studies based on longer follow-up of non-twin samples have found significant differences in height, weight, motor development, and IQ between VLBW and NBW children even in adolescence^{43–48}, suggesting slow catch-up between VLBW and NBW children. However, as mentioned before, these studies are limited by their inability to control for unobserved maternal, environmental, and genetic confounders. A recent study looking at within-twin differences in birth weight and IQ found birth weight was positively correlated with IQ scores in a sample of 5 year old twins in the

UK²². As more waves of the ECLS-B data become available, we will be able to examine whether the effects on physical growth persist beyond age 2 and whether the effects on mental and motor development become salient over time for a sample of U.S. born twins.

Despite the advantages that twins-based analyses possess, there are some concerns regarding the generalizeability and interpretation of twins-based estimates of birth weight effects. First, twins have been known to be different from singleton births on a variety of dimensions, which may make it difficult to generalize these results to singleton births. Twins are more likely to have congenital anomalies, adult cancers, and some childhood conditions such as autism⁴⁹. Moreover, most twin births are preterm and experience compromised growth in the last trimester of pregnancy compared to singletons. Another concern as illustrated in Table 2, is that twinning may be the result of fertility treatments that are more common among higher socioeconomic populations. Fertility treatments might also increase the incidence of fraternal twinning relative to identical twinning, which may affect comparability between fraternal twins-based estimates and identical twins-based estimates. Second, a twin may not always be a perfect control for the other twin. Recent studies suggest that even monozygotic twins may be discordant for complex genetic traits and disorders⁵⁰. The ECLS-B collected information on the presence of congenital anomalies and we dropped children with them from our sample.

There are other limitations of our study that should be noted. First, even though comparison of the sample that attrited between the two waves with the sample that remained in the study showed small differences based on socioeconomic status, it is possible that differential attrition of lower socioeconomic status mothers from the sample may have contributed to the attenuation of low birth weight effects. Families of higher socioeconomic status may be able to provide increased stimulation along with a good environment (e.g., diet) that may prevent malnourished or underweight infants from being cognitively delayed. Second, the ECLS-B is limited because the data come from a national survey and therefore lack information on details that may be available from clinical samples of low birth weight children. For example, there is no information in these data about chorionicity of the twins which may determine the specific mechanisms that lead to birth weight discordance within twin pairs such as twin-to-twin transfusion and cord compromise. These different mechanisms may have different implications for children's developmental outcomes. Larger samples of twins will be needed to separately examine the effects of each of these pathways. The ECLS-B also lacks comprehensive information about fertility treatments used, which may in turn influence the pathways leading to birth weight discordance within twins and consequently have implications for whether twins based estimates can be generalized to singletons.

Notwithstanding concerns about generalizeability to singleton births, our findings help to understand the effects of LBW in the rapidly growing population of twins in the U.S. Between 1980 and 2004, the twin birth rate increased 70% from 18.9 to 32.2 per 1000 births⁵¹, and this rise is likely to continue if current trends in use of Assisted Reproductive Technologies and maternal age of child bearing persist. Consequently, estimates based on

twins data will become increasingly important. The prevalence of LBW has also risen steadily during the same period, from 6.8 percent in 1980 to 8.2 percent in 2005, capturing both the increased prevalence of multiple births in the population as well as increased survival among LBW singleton births. Understanding the consequences of LBW and the role of maternal, environmental, and genetic factors in determining LBW and subsequent developmental outcomes is, therefore, important for designing effective public health strategies.

CONCLUSION

Maternal, environmental, and genetic factors play an important role in determining birth weight as well as children's mental, motor, and physical development in the first two years of life. Consequently, the effects of VLBW and MLBW are significantly reduced when these factors are controlled for. Nevertheless, VLBW and MLBW remain major risk factors for children's physical growth in the early years. Future studies need to examine the longer term developmental effects of VLBW and MLBW using twins data.

List of Abbreviations

BSF-R	Bayley Short Form — Research Edition
CDC	Centers for Disease Control
DZ	Dizygotic
ECLS-B	Early Childhood Longitudinal Study — Birth Cohort
IRT	Item Response Theory
IUGR	Intrauterine growth restriction
LBW	Low birth weight
MLBW	Moderately low birth weight
MZ	Monozygotic
NBW	Normal birth weight
NCES	National Center for Education Statistics
NCHS	National Center for Health Statistics
SGA	Small-for-gestational age
US	United States
VLBW	Very low birth weight

Acknowledgments

The authors thank Dr. Homero Martinez, Dr. Jennifer Park, three anonymous referees, and participants at the Early Childhood Longitudinal Study — Birth Cohort (ECLS-B) First Release Conference: Development from Birth through Age Two for helpful comments. This research was funded by a National Institute for Child Health and Human Development grant (R03HD051578). All opinions expressed are those of the authors and not of the funding agency.

REFERENCES

1. National Center for Health Statistics (NCHS). Health, United States, 2006 with Chartbook on Trends in the Health of Americans. U.S. Department of Health and Human Services; 2006. DHHS 2006-1232
2. Barker DJ. Fetal origins of coronary heart disease. *British Medical Journal* 1995;311:171–174. [PubMed: 7613432]
3. Barker, D. Mothers, Babies and Health Later in Life. Edinburgh: Churchill Livingstone; 1998.
4. Barker DJ. The malnourished baby and infant. *British Medical Bulletin* 2001;60:69–88. [PubMed: 11809619]
5. Aylward GP, Pfeiffer SI, Wright A, Verhulst SJ. Outcome Studies of Low Birth Weight Infants Published in the Last Decade: A Meta-analysis. *Journal of Pediatrics* 1989;115:515–520. [PubMed: 2795341]
6. Brooks AM, Byrd RS, Weitzman M, Auinger P, McBride JT. Impact of Low Birth Weight on Early Childhood Asthma in The United States. *Archives of Pediatrics and Adolescent Medicine* 2001;155(3):401–406. [PubMed: 11231809]
7. Lee H, Barratt M. Cognitive Development of Preterm Low Birth Weight Children at 5 to 8 Years Olds. *Developmental Behavioral Pediatrics* 1993;14:242–248.
8. McCormick MC. The Contribution of Low Birth Weight to Infant Mortality and Childhood Morbidity. *The New England Journal of Medicine* 1985;312:82–90. [PubMed: 3880598]
9. McCormick MC, Gortmaker SL, Sobol AM. Very Low Birth Weight Children: Behavior Problems and School Difficulty in a National Sample. *Journal of Pediatrics* 1990;117:687–693. [PubMed: 2231200]
10. Paneth NS. The Problem of Low Birth Weight. *The Future of Children* 1995;5(1): 19–34. [PubMed: 7633862]
11. Richards M, Hardy RJ, Kuh D, Wadsworth MEJ. Birth Weight and Cognitive Function in the British 1946 Birth Cohort: Longitudinal Population Based Study. *British Medical Journal* 2001;322:199–203. [PubMed: 11159613]
12. Wu YW, Croen LA, Shah SJ, Newman TB, Najjar DV. Cerebral palsy in a term population: Risk factors and neuroimaging findings. *Pediatrics* 2006;118:690–697. [PubMed: 16882824]
13. Korhonen P, Tammela O, Koivisto AM, Laippala P, Ikonen S. Frequency and risk factors in bronchopulmonary dysplasia in a cohort of very low birth weight infants. *Early Human Development* 1999;54(3):245–258. [PubMed: 10321791]
14. Hack M, Taylor HG, Drotar D, Schluchter M, Cartar L, Andreias L, Wilson-Costello D, Klein N. Chronic Conditions, Functional Limitations, and Special Health Care Needs of School-aged Children Born With Extremely Low-Birth-Weight in the 1990s. *Journal of the American Medical Association* 2005;294:318–325. [PubMed: 16030276]
15. Iliadou A, Cnattingius S, Lichtenstein P. Low birthweight and Type 2 diabetes: A study on 11,162 Swedish twins. *International Journal of Epidemiology* 2004;33:948–953. [PubMed: 15166205]
16. Hubinette A, Cnattingius S, Ekbom A, de Faire U, Kramer M, Lichtenstein P. Birthweight, early environment, and genetics: a study of twins discordant for acute myocardial infarction. *Lancet* 2001;357:1997–2001. [PubMed: 11438130]
17. Hubinette A, Cnattingius S, Johansson AL, Henriksson C, Lichtenstein P. Birth weight and risk of angina pectoris: analysis in Swedish twins. *European Journal of Epidemiology* 2003;18:539–544. [PubMed: 12908719]
18. Babson SC, Kangas J, Young N, Bramhall JL. Growth and development of twins of dissimilar size at birth. *Pediatrics* 1964;33:327. [PubMed: 14129079]
19. Churchill JA. The relationship between intelligence and birth weight in twins. *Neurology* 1965;4:341. [PubMed: 14280597]

20. Scarr S. Effects of birthweight on later intelligence. *Soc. Biol* 1969;16:249. [PubMed: 5392814]
21. Willerman L, Churchill JA. Intelligence and birth weight in identical twins. *Child Develop* 1967;38:623. [PubMed: 6049626]
22. Newcombe R, Milne BJ, Caspi A, Poulton R, Moffitt TE. Birthweight Predicts IQ: Fact or Artefact? *Twin Research and Human Genetics* 2007;Volume 10(Number 4): 581–586. [PubMed: 17708699]
23. Stein REK, Seigel MJ, Bauman LJ. Are Children of Moderately Low Birth Weight at Increased Risk for Poor Health? A New Look at an Old Question. *Pediatrics* 2006;118:217–223. [PubMed: 16818568]
24. National Center for Education Statistics (NCES). User's Manual for the ECLS-B Nine Month Restricted-Use Data File and Electronic Codebook. U.S. Department of Education. Institute of Education Sciences; 2004. Early Childhood Longitudinal Study, Birth Cohort (ECLS-B). NCES 2004-092
25. National Center for Education Statistics. Early Childhood Longitudinal Study, Birth Cohort (ECLS-B), User's Manual for the ECLS-B Longitudinal 9-Month–2-Year Data File and Electronic Codebook. U.S. Department of Education. Institute of Education Sciences; 2006. NCES 2006-046
26. Goldsmith HH. A zygosity questionnaire for young twins: A research note. *Behav Genet* 1991;21:257–269. [PubMed: 1863259]
27. Nichols RC, Bilbro WC. The diagnosis of twin zygosity. *Acta Genet Stat Med* 1966;16:265–275. [PubMed: 4959152]
28. Cohen DJ, Dibble E, Grawe JM, Pollin W. Reliably separating identical from fraternal twins. *Arch Gen Psychiatry* 1975;32:1371–1375. [PubMed: 1239251]
29. Alexander GH, Himes JH, Kaufman RB, Mor J, Kogan M. A United States national reference for fetal growth. *Obstetrics and Gynecology* 1996;87(2):163–168. [PubMed: 8559516]
30. Resnick R. Intrauterine growth restriction. *Obstetrics and Gynecology* 2002;99(3):490–496. [PubMed: 11864679]
31. Bayley, N. The Psychological Corporation. 2nd Edition. San Antonio, TX: 1993. Bayley Scales of Infant Development. Manual
32. Hambleton, RK.; Swaminathan, H.; Rogers, HJ. Fundamentals of Item Response Theory. Newbury Park, CA: Sage Press; 2000.
33. Kotelchuck M. The Adequacy of Prenatal Care Utilization Index: Its US Distribution and Association with Low Birthweight. *American Journal of Public Health* 1994;84(9):1486–1489. [PubMed: 8092377]
34. Carlin JB, Gurrin LC, Sterne JAC, Morley R, Dwyer T. Regression models for twin studies: a critical review. *International Journal of Epidemiology* 2005;34:1089–1099. [PubMed: 16087687]
35. Baird J, Osmond C, MacGregor A, Snieder H, Hales CN, Phillips DI. Testing the fetal origins hypothesis in twins: the Birmingham twin study. *Diabetologia* 2001;44:33–39. [PubMed: 11206409]
36. Christensen K, Stovring H, McGue M. Do genetic factors contribute to the association between birth weight and blood pressure? *Journal of Epidemiology and Community Health* 2001;55:583–587. [PubMed: 11449017]
37. Dwyer T, Blizzard L, Morley R, Ponsonby AL. Within pair association between birth weight and blood pressure at age 8 in twins from a cohort study. *British Medical Journal* 1999;319:1325–1329. [PubMed: 10567134]
38. Ijzerman R, Stehouwer CD, Boomsma DI. Evidence for genetic factors explaining the birth weight-blood pressure relation. Analysis in twins. *Hypertension* 2000;36:1008–1012. [PubMed: 11116116]
39. Johansson-Kark M, Rasmussen F, De Stavola B, Leon DA. Fetal growth and systolic blood pressure in young adulthood: the Swedish Young Male Twins Study. *Paediatric and Perinatology Epidemiology* 2002;16:200–209.

40. Loos RJ, Fagard R, Beunen G, Derom C, Vlietinck R. Birth weight and blood pressure in young adults: a prospective twin study. *Circulation* 2001;104:1633–1638. [PubMed: 11581141]
41. Nowson CA, MacInnis RJ, Hopper JL, et al. Association of birth weight and current body size to blood pressure in female twins. *Twin Research* 2001;4:378–384. [PubMed: 11869492]
42. Poulter NR, Chang CL, MacGregor AJ, Snieder H, Spector TD. Association between birth weight and adult blood pressure in twins: historical cohort study. *British Medical Journal* 1999;319:1330–1333. [PubMed: 10567135]
43. Zhang J, Brenner RA, Klebanoff MA. Differences in birth weight and blood pressure at age 7 years among twins. *American Journal of Epidemiology* 2001;153:779–782. [PubMed: 11296150]
44. Ford GW, Doyle LW, Davis NM, Callanan C. Very low birth weight and growth into adolescence. *Arch Pediatr Adolesc Med* 2000 Aug;154(8):778–784. [PubMed: 10922273]
45. Hack M, Schluchter M, Cartar L, Rahman M, Cuttler L, Borawski E. Growth of very low birth weight infants to age 20 years. *Pediatrics* 2003 Jul;112(1 Pt 1):e30–e38. [PubMed: 12837903]
46. Evensen KA, Vik T, Helbostad J, Indredavik MS, Kulseng S, Brubakk AM. Motor skills in adolescents with low birth weight. *Arch Dis Child Fetal Neonatal Ed* 2004 Sep;89(5):F451–F455. [PubMed: 15321969]
47. Whitaker AH, Feldman JF, Lorenz JM, Shen S, McNicholas F, Nieto M, McCulloch D, Pinto-Martin JA, Paneth N. Motor and cognitive outcomes in nondisabled low-birth-weight adolescents: early determinants. *Arch Pediatr Adolesc Med* 2006 Oct;160(10):1040–1046. [PubMed: 17018463]
48. Martinez-Cruz CF, Poblano A, Fernandez-Carrocer LA, Jimenez-Quiroz R, Tuyu-Torres N. Association between intelligence quotient scores and extremely low birth weight in school-age children. *Arch Med Res* 2006 Jul;37(5):639–645. [PubMed: 16740436]
49. Hall JG. Twinning. *The Lancet* 2003;362(9385):735–743.
50. Hall JG. Twinning: mechanisms and genetic implications. *Current Opinions in Genetic Development* 1996 Jun;6(3):343–347.
51. Martin, JA.; Hamilton, BE.; Sutton, PD., et al. National vital statistics reports. Vol. vol 55. Hyattsville, MD: National Center for Health Statistics; 2006. Births: Final data for 2004.

Table 1

Means and Proportions for Key Variables

	Mean or proportion	Standard error
<i>Outcomes</i>		
Mental development scale scores at 9-months	76.724	0.151
Mental development scale scores at 2-years	127.280	0.160
Motor development scale scores at 9-months	55.962	0.144
Motor development scale scores at 2-years	81.497	0.076
Weight-for-age z-scores at 9-months	0.046	0.021
Weight-for-age z-scores at 2-years	0.199	0.020
Weight-for-length z-scores at 9-months	0.293	0.028
Weight-for-height z-scores at 2-years	0.508	0.023
Length-for-age z-scores at 9-months	0.117	0.016
<i>Birth weight measures</i>		
Normal birth weight	0.928	0.002
Moderately low birth weight	0.060	0.002
Very low birth weight	0.012	0.000
Small-for-gestational age	0.102	0.004
<i>Child characteristics</i>		
Female	0.490	0.008
Race/ethnicity:		
Non-Hispanic White	0.546	0.007
Non-Hispanic Black	0.141	0.005
Hispanic	0.242	0.007
Other race/ethnicity	0.071	0.003
First birth	0.406	0.008
<i>Maternal characteristics</i>		
Height (inches)	64.520	0.047
Age (years)	27.338	0.091
Education:		
Less than high school diploma	0.262	0.007
High school diploma or equivalent	0.220	0.006
Some college or voc/tech program	0.271	0.007
Bachelors degree or higher	0.247	0.006
Married at birth	0.680	0.007
<i>Household characteristics</i>		
Income:		
Less than or equal to \$20000	0.251	0.006
Between \$20001 and \$40000	0.296	0.007
Between \$40001 and \$75000	0.247	0.007
Between \$75001 and \$100000	0.101	0.005
\$100001 or higher	0.104	0.005
<i>Pregnancy characteristics</i>		

	Mean or proportion	Standard error
Gestation at pregnancy recognition (weeks)	5.393	0.054
Smoked in last trimester of pregnancy	0.112	0.005
Drank alcohol in last trimester of pregnancy	0.033	0.003
Adequacy of prenatal care utilization:		
Inadequate	0.099	0.005
Intermediate	0.142	0.005
Adequate	0.442	0.008
Adequate plus	0.317	0.007
Presence of pregnancy risk factors	0.306	0.007
Presence of labor and delivery complications	0.324	0.007
Gestation (weeks)	38.720	0.029

Notes: Means are calculated for continuous variables and proportions are shown for binary variables. Continuous variables are the following: the outcome variables, height, age, gestation at pregnancy recognition, and gestation. The remainder of the variables are binary. Values are weighted. Sample sizes are rounded to the nearest 50 per NCES regulations. The sample sizes for the outcome variables vary. The sample sizes for the independent variables is 8,050.

Table 2

Means and Proportions for Key Variables by Twin Status

	Singleton births	All twins	p-value	Fraternal twins	Identical twins	p-value
Outcomes						
Mental development scale scores (9 mos)	76.863	72.082	0.000	72.029	72.334	0.685
Mental development scale scores (2 yrs)	127.398	123.412	0.000	123.587	122.585	0.197
Motor development scale scores (9 mos)	56.091	51.739	0.000	51.918	50.902	0.213
Motor development scale scores (2 yrs)	81.542	80.147	0.000	80.167	80.054	0.777
Weight-for-age z-scores (9 mos)	0.059	-0.337	0.000	-0.378	-0.153	0.044
Weight-for-age z-scores (2 yrs)	0.208	-0.095	0.000	-0.132	0.074	0.101
Weight-for-length z-scores (9 mos)	0.298	0.161	0.022	0.115	0.369	0.024
Weight-for-height z-scores (2 yrs)	0.514	0.309	0.000	0.270	0.486	0.107
Length-for-age z-scores at (9 mos)	0.131	-0.303	0.000	-0.321	-0.217	0.316
Birth weight measures						
Normal birth weight	0.942	0.476	0.000	0.480	0.460	0.647
Moderately low birth weight	0.048	0.454	0.000	0.448	0.481	0.449
Very low birth weight	0.010	0.070	0.000	0.072	0.059	0.371
Small-for-gestational age	0.102	0.107	0.598	0.103	0.128	0.401
Child characteristics						
Female	0.489	0.510	0.232	0.505	0.533	0.512
Race/ethnicity:						
Non-Hispanic White	0.543	0.633	0.000	0.661	0.499	0.000
Non-Hispanic Black	0.141	0.136	0.602	0.134	0.145	0.694
Hispanic	0.244	0.172	0.000	0.149	0.283	0.001
Other race/ethnicity	0.071	0.059	0.100	0.056	0.073	0.411
First birth	0.406	0.419	0.467	0.444	0.298	0.001
Maternal characteristics						
Height (inches)	64.507	64.989	0.000	65.161	64.189	0.000
Age (years)	27.280	29.299	0.000	29.619	27.805	0.000
Education:						
Less than high school diploma	0.264	0.201	0.000	0.165	0.367	0.000

	Singleton births	All twins	p-value	Fraternal twins	Identical twins	p-value
High school diploma or equivalent	0.221	0.178	0.002	0.184	0.150	0.235
Some college or voc/tech program	0.271	0.276	0.723	0.287	0.228	0.091
Bachelors degree or higher	0.245	0.344	0.000	0.363	0.254	0.002
Married at birth	0.678	0.744	0.000	0.776	0.596	0.000
Household characteristics						
Income:						
Less than or equal to \$20000	0.253	0.203	0.001	0.185	0.286	0.014
Between \$20001 and \$40000	0.297	0.255	0.005	0.241	0.317	0.049
Between \$40001 and \$75000	0.246	0.254	0.600	0.262	0.221	0.230
Between \$75001 and \$100000	0.101	0.116	0.160	0.124	0.078	0.047
\$100001 or higher	0.103	0.172	0.000	0.188	0.099	0.000
Pregnancy characteristics						
Gestation at pregnancy recognition (weeks)	5.403	5.086	0.007	5.016	5.411	0.172
Smoked in last trimester of pregnancy	0.112	0.110	0.874	0.100	0.162	0.040
Drank alcohol in last trimester of pregnancy	0.034	0.024	0.073	0.029	0.000	0.000
Adequacy of prenatal care utilization:						
Inadequate	0.099	0.095	0.765	0.076	0.183	0.003
Intermediate	0.145	0.045	0.000	0.043	0.054	0.510
Adequate	0.451	0.169	0.000	0.161	0.207	0.205
Adequate plus	0.306	0.691	0.000	0.720	0.555	0.000
Presence of pregnancy risk factors	0.301	0.441	0.000	0.447	0.410	0.383
Presence of labor and delivery complications	0.318	0.519	0.000	0.528	0.478	0.246
Gestation (weeks)	38.813	35.850	0.000	35.834	35.926	0.660
Sample size for independent variables	6750	1250		1050	200	

Notes: Means are calculated for continuous variables and proportions are shown for binary variables. Continuous variables are the following: the outcome variables, height, age, gestation at pregnancy recognition, and gestation. The remainder of the variables are binary. Values are weighted. Sample sizes are rounded to the nearest 50 per NCES regulations. The sample sizes for the outcome variables varies. P-values are reported for two-tailed t-tests of differences in means or proportions.

Table 3

Multivariate Linear Regression Estimates of Birth Weight Effects on Mental and Motor Development

		Mental development scale scores		Motor development scale scores	
		9-months	2-years	9-months	2-years
		(1)	(2)	(3)	(4)
A). Cross-sectional regression using full sample					
VLBW	Coefficient	-9.069	-7.627	-8.764	-4.123
	P-value	[0.000]	[0.000]	[0.000]	[0.000]
MLBW	Coefficient	-3.252	-2.550	-2.901	-1.383
	P-value	[0.000]	[0.000]	[0.000]	[0.000]
	n	8000	8000	7950	7950
B). Within-twins regression using fraternal twins					
VLBW	Coefficient	-2.581	-5.101	-3.784	-1.340
	P-value	[0.021]	[0.000]	[0.003]	[0.166]
MLBW	Coefficient	-2.007	-2.367	-0.917	0.015
	P-value	[0.000]	[0.002]	[0.093]	[0.970]
	n	1050	1050	1000	1000
C). Within-twins regression using identical twins					
VLBW	Coefficient	-3.130	0.544	-2.503	-0.519
	P-value	[0.378]	[0.841]	[0.062]	[0.783]
MLBW	Coefficient	-2.506	1.008	-0.931	-1.641
	P-value	[0.051]	[0.648]	[0.216]	[0.075]
	n	200	200	200	200

Notes: Notes: VLBW=very low birth weight; MLBW=moderately low birth weight. Coefficient estimates capture the number of additional correct items on the BSID-II mental or motor development scale that VLBW and MLBW children obtain relative to NBW children. Results are weighted. P-values from Model A are adjusted to account for within twins correlations. Sample sizes are rounded to the nearest 50 per NCES regulations. Control variables in model A include: female, assessment age in months, race/ethnicity, first birth, twin or higher-order birth, height in inches, age at birth, education, married at birth, household income, region of residence, urbanicity, gestation at pregnancy recognition in weeks, smoked in last trimester of pregnancy, drank alcohol in last trimester of pregnancy, adequacy of prenatal care utilization, presence of pregnancy risk factors, presence of labor and delivery complications, and gestation in weeks. Indicators for missing values are also included in these regressions for the following variables: first birth, adequacy of prenatal care utilization, presence of pregnancy risk factors, and presence of labor and delivery complications. Variables measured at 2-year wave and those that represent changes between waves are only included in regressions examining outcomes measured two years after birth. Female is the control variable in Panel B. Bold indicates that the 9-month coefficient is statistically different from the 2-year coefficient at a 1% level.

Table 4
Multivariate Linear Regression Estimates of Birth Weight Effects on Physical Growth

		Weight-for-age z-scores		Weight-for-length/height z-score		Length-for-age z-score at 9-month	
		9-months	2-years	9-months	2-years	9-months	2-years
		(1)	(2)	(3)	(4)	(5)	(6)
<i>A). Cross-sectional regression using full sample</i>							
VLBW	Coefficient	-1.643	-1.392	-0.534	-1.02	-1.749	
	P-value	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	
MLBW	Coefficient	-0.772	-0.676	-0.309	-0.475	-0.747	
	P-value	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	
	n	7400	7400	7150	7150	8000	
<i>B). Within-twins regression using fraternal twins</i>							
VLBW	Coefficient	-1.236	-1.288	-0.631	-1.209	-1.149	
	P-value	[0.000]	[0.000]	[0.014]	[0.000]	[0.000]	
MLBW	Coefficient	-0.332	-0.393	-0.152	-0.385	-0.341	
	P-value	[0.000]	[0.002]	[0.159]	[0.017]	[0.000]	
	n	900	900	850	850	1000	
<i>C). Within-twins regression using identical twins</i>							
VLBW	Coefficient	-0.962	-0.726	-0.132	-0.532	-1.052	
	P-value	[0.000]	[0.000]	[0.738]	[0.030]	[0.000]	
MLBW	Coefficient	-0.395	-0.462	-0.253	-0.459	-0.262	
	P-value	[0.019]	[0.016]	[0.223]	[0.030]	[0.056]	
	n	200	200	150	150	200	

Notes: VLBW=very low birth weight; MLBW=moderately low birth weight. Coefficient estimates capture the difference (in standard deviation units) in the value of the outcome variable for VLBW and MLBW children relative to NBW children. Results are weighted. P-values in Panel A are adjusted to account for within twins correlations. Sample sizes are rounded to the nearest 50 per NCES regulations. Control variables in Panel A include: female, assessment age in weeks, race/ethnicity, first birth, twin or higher-order birth, height in inches, age at birth, education, married at birth, household income, region of residence, urbanicity, gestation at pregnancy recognition in weeks, smoked in last trimester of pregnancy, drank alcohol in last trimester of pregnancy, adequacy of prenatal care utilization, presence of pregnancy risk factors, presence of labor and delivery complications, and gestation in weeks. Indicators for missing values are also included in these regressions for the following variables: first birth, adequacy of prenatal care utilization, presence of pregnancy risk factors, and presence of labor and delivery complications. Variables measured at 2-year wave and those that represent changes between waves are only included in regressions examining outcomes measured two years after birth. Female is the control variable in Panel B. None of the 9-month coefficients are statistically different from their counterpart 2-year coefficients at a 1% level.

Multivariate Linear Regression Estimates of Small-for-Gestational Age on Mental and Motor Development and Physical Growth

Table 5

	Mental development scale scores		Motor development scale scores		Weight-for-age z- scores		Weight-for-length/height z- score		Length-for-age z- score at 9- month	
	9-months	2-years	9-months	2-years	9-months	2-years	9-months	2-years	9-months	2-years
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
A). Cross-sectional regression using full sample										
Coefficient	-1.613	-1.665	-1.309	-0.751	-0.699	-0.701	-0.328	-0.556	-0.654	-0.654
P-value	[0.000]	[0.000]	[0.000]	[0.002]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]
n	8000	8000	7950	7950	7400	7400	7150	7150	8000	8000
B). Within-twins regression using fraternal twins										
Coefficient	-0.020	-2.531	-0.315	-0.270	-0.855	-0.861	-0.645	-0.910	-0.555	-0.555
P-value	[0.982]	[0.044]	[0.704]	[0.667]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]
n	1050	1050	1000	1000	900	900	850	850	1000	1000
C). Within-twins regression using identical twins										
Coefficient	-2.568	0.887	-0.870	-0.667	-0.585	-0.409	-0.336	-0.312	-0.472	-0.472
P-value	[0.078]	[0.667]	[0.401]	[0.464]	[0.000]	[0.015]	[0.103]	[0.090]	[0.000]	[0.000]
n	200	200	200	200	200	200	150	150	200	200

Notes: For mental and motor development scores, coefficient estimates capture the number of additional correct items on the BSID-II mental or motor development scale that small for gestational age children obtain relative to children who are not small for gestational age. For the physical growth outcomes, coefficient estimates capture the difference (in standard deviation units) in the value of the outcome variable for small for gestational age children relative to children who are not small for gestational age. Results are weighted. P-values in Panel A are adjusted to account for within twins correlations. Sample sizes are rounded to the nearest 50 per NCES regulations. Control variables in Panel A include: female, assessment age in weeks, race/ethnicity, first birth, twin or higher-order birth, height in inches, age at birth, education, married at birth, household income, region of residence, urbanicity, gestation at pregnancy recognition in weeks, smoked in last trimester of pregnancy, drank alcohol in last trimester of pregnancy, adequacy of prenatal care utilization, presence of pregnancy risk factors, presence of labor and delivery complications, and gestation in weeks. Indicators for missing values are also included in these regressions for the following variables: first birth, adequacy of prenatal care utilization, presence of pregnancy risk factors, and presence of labor and delivery complications. Variables measured at 2-year wave and those that represent changes between waves are only included in regressions examining outcomes measured two years after birth. Female is the control variable in Panel B. None of the 9-month coefficients are statistically different from their counterpart 2-year coefficients at a 4% level.

Appendix

Comparison of Length-for-Age Z-scores at 9-months and Height-for-Age Z-scores at 2-years

		Length-for-age z-score at 9-months	Height-for-age z-scores at 2-years
		(1)	(2)
<i>A). Cross-sectional regression using full sample</i>			
VLBW	Coefficient	-1.691	-1.013
	P-value	[0.000]	[0.000]
MLBW	Coefficient	-0.712	-0.475
	P-value	[0.000]	[0.000]
	n	4700	4700
<i>B). Within-twins regression using fraternal twins</i>			
VLBW	Coefficient	-0.924	-0.531
	P-value	[0.000]	[0.003]
MLBW	Coefficient	-0.242	-0.061
	P-value	[0.032]	[0.443]
	n	550	550
<i>C). Within-twins regression using identical twins</i>			
VLBW	Coefficient	-1.344	-0.652
	P-value	[0.000]	[0.001]
MLBW	Coefficient	-0.214	-0.295
	P-value	[0.270]	[0.016]
	n	100	100

Notes: VLBW=very low birth weight; MLBW=moderately low birth weight. Coefficient estimates capture the difference (in standard deviation units) in the value of the outcome variable for VLBW and MLBW children relative to NBW children. Sample sizes are rounded to the nearest 50 per NCES regulations. Results are weighted. P-values in Panel A are adjusted to account for within twins correlations. Control variables in Panel A include: female, assessment age in weeks, race/ethnicity, first birth, twin or higher-order birth, height in inches, age at birth, education, married at birth, household income, region of residence, urbanicity, gestation at pregnancy recognition in weeks, smoked in last trimester of pregnancy, drank alcohol in last trimester of pregnancy, adequacy of prenatal care utilization, presence of pregnancy risk factors, presence of labor and delivery complications, and gestation in weeks. Indicators for missing values are also included in these regressions for the following variables: first birth, adequacy of prenatal care utilization, presence of pregnancy risk factors, and presence of labor and delivery complications. Variables measured at 2-year wave and those that represent changes between waves are only included in regressions examining outcomes measured two years after birth. Female is the control variable in Panel B. Bold indicates that the 9-month coefficient is statistically different from the 2-year coefficient at a 1% level.