NBER WORKING PAPER SERIES

THE RETURNS TO EARLY-LIFE INTERVENTIONS FOR VERY LOW BIRTH WEIGHT CHILDREN

Eric Chyn
Samantha Gold
Justine S. Hastings

Working Paper 25753
http://www.nber.org/papers/w25753

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2019

We thank the team at Research for Improving People's Lives (RIPL) and the State of Rhode Island for making this project possible. We are also thankful for helpful comments that we received from Prashant Bharadwaj and Amanda Kowalski, and research teams at the Rhode Island Departments of Education and Health. We are grateful for financial support from the Smith Richardson Foundation and the Laura and John Arnold Foundation. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Returns to Early-life Interventions for Very Low Birth Weight Children
Eric Chyn, Samantha Gold, and Justine S. Hastings
NBER Working Paper No. 25753
April 2019
JEL No. H53,I1,I12,I14,I18,I21,I24,J13

ABSTRACT

We use comprehensive administrative data from Rhode Island to measure the impact of early-life interventions for low birth weight newborns on later-life outcomes. We use a regression discontinuity design based on the 1,500-gram threshold for Very Low Birth Weight (VLBW) status. We show that threshold crossing causes more intense in-hospital care, in line with prior studies. Threshold crossing also causes a 0.34 standard deviation increase in test scores in elementary and middle school, a 17.1 percentage point increase in the probability of college enrollment, and $66,997 decrease in social program expenditures by age 14. We explore potential mechanisms driving these impacts.

Eric Chyn
Department of Economics
University of Virginia
P.O. Box 400182
Charlottesville, VA 22904
ericchyn@virginia.edu

Samantha Gold
RIPL
1 Park Row
Providence, RI 02906
sgold@ripl.org

Justine S. Hastings
Brown University
Department of Economics
64 Waterman Street
Providence, RI 02912
and NBER
justine_hastings@brown.edu

A data appendix is available at http://www.nber.org/data-appendix/w25753

1 Introduction

Children born with low birth weight tend to have more health difficulties and worse later-life outcomes relative to their peers born at normal birth weight (Almond and Currie 2011; Currie 2011). A key question for public policy is whether medical interventions delivered during early-life can counteract the negative effects of low weight and poor health at birth. Investment in child health may generate substantial private and social benefits by enhancing later-life economic self-sufficiency and reducing reliance on social programs.

This paper provides new evidence on the impacts of additional medical care for children born with birth weight below the 1,500-gram threshold for Very Low Birth Weight (VLBW) status. We identify causal effects using a regression discontinuity (RD) design that exploits changes in medical treatment across the VLBW threshold. Using this approach, Almond et al. (2010) find that children in the U.S. born just under 1,500 grams have higher hospital costs and lower one-year mortality rates. Bharadwaj et al. (2013) find similar results using data from Chile and Sweden. They also find that children with birth weight just below 1,500 grams have 0.13 to 0.22 standard deviations higher test scores measured during elementary and high school grades.

We build on these results from prior studies using comprehensive administrative data from Rhode Island. The data contain over three decades of birth records that we join to a range of outcomes for education and social program participation. We study academic achievement throughout primary and secondary public education, as well as college attendance. We also examine enrollment in social programs and total benefits received during childhood. To explore post-hospital investments, we use measures of parental investment such as records on maternal labor market activity and school choice.

We find significant and large positive impacts of crossing the 1,500-gram threshold for VLBW status on educational outcomes at all levels. Children born just under 1,500 grams perform 0.34 standard deviations better on test scores measured during grades 3-8 relative to those born just above the threshold. For each grade throughout childhood, we find similar magnitude impacts, suggesting that factors influencing higher academic performance accrued prior to enrollment in school. These gains in test scores are matched by significant and positive effects on post-secondary educational enrollment. Threshold crossing is associated with a 17.1 percentage points (32 percent) increase in the likelihood of enrolling in any college by age 22.

There are also significant threshold crossing effects on the amount of social program expenditures that a child received. We measure spending using payment records for SNAP and TANF, claims files from Medicaid, and school data on enrollment in special education programs (IEP). By age 14, children born just under the 1,500-gram threshold receive $66,997 *less* in social program expenditures. This is a large reduction relative to $145,486 spent on children born above the threshold. As another comparison, Almond et al. (2010) estimate that hospital costs increase by an additional $4,000 for children just under the 1,500-gram threshold.

Given the size of estimated impacts relative to the in-hospital expenditures, we analyze two types of potential post-hospital mechanisms. First, we examine proxies for parental behavior using maternal employment records and responses to the Center for Disease Control's Pregnancy Risk Assessment Monitoring Program System (PRAMS) survey. We find no significant impacts on employment or earnings for mothers within the first two years after birth. Similarly, we find no significant impacts on measures for maternal care and stress based on the PRAMS survey. Although imprecise, the point estimates suggest that, if anything, threshold crossing increases care and reduces stress for mothers.

Second, we study whether threshold crossing affects school choice by creating value-added measures for public schools in Rhode Island. For elementary and middle school grades, we study school value-added for standardized test scores. For high school grades, we study school value-added in terms of enrollment in college. We find small, insignificant impacts of crossing the 1,500-gram threshold on measures of school quality. The results suggest impacts on academic achievement are not driven by parents selecting substantially better schools.

The results are robust to standard regression discontinuity checks for balance in baseline characteristics across the 1,500-gram threshold and tests of the sensitivity of our estimates to higher-order polynomial RD specifications. In addition, we address concerns about non-random heaping at the VLBW threshold by excluding observations within a small window around the 1,500-gram threshold (Barreca et al. 2011). Finally, we show that our results do not appear to be affected by differential survival rates (or other forms of attrition) around the VLBW threshold. In our sample, we find little evidence of significant threshold crossing effects on the likelihood of having a test score or enrolling in high school.[1]

Overall, our findings add to the literature on the effectiveness of medical interventions for at-risk children. We extend Almond et al. (2010) by showing positive impacts on long-run measures of human capital and reductions in social program expenditures throughout childhood. The impacts on childhood test scores are consistent with findings for Chile and Norway reported in Bharadwaj et al. (2013). The effect sizes that we detect are also similar in magnitude to experimental and quasi-experimental evaluations of the test-score impact of early childhood education programs (Duncan and Magnuson 2013).[2] Our results more generally contribute to the literature demonstrating

---

[1] Using the census of U.S. births from 1983-2002, Almond et al. (2010) estimate that crossing the VLBW threshold decreases one-year mortality by one percentage point. In our sample of births in Rhode Island, there is no statistically significant discontinuity in the likelihood of having a test score in grades 3 through 8. At the same time, the standard error for our estimate does not rule out threshold crossing effects on mortality (and survivorship) that are as small as one percentage point.

[2] Duncan and Magnuson (2013) review of 84 evaluations of early childhood education programs and find that the average effect size is 0.35 standard deviations.

that interventions delivered in early-life can generate notable gains for children (Currie and Thomas 1995; Chetty et al. 2011; Havnes and Mogstad 2011; Heckman et al. 2013; Chetty et al. 2016; Hoynes et al. 2016; Chyn 2018; Bald et al. 2019).

2 Background on Medical Care for VLBW Children

The American Academy of Pediatrics (AAP) publishes guidelines for U.S. neonatal care that recommend admitting all VLBW children into a Neonatal Intensive Care Unit (NICU) (AAP, 2012). At the NICU, a child receives continuous, specialized care from a broad range of specialists including neonatologists, pediatric nurses, and respiratory therapists. The NICU also has specialized equipment (e.g., mechanical ventilators) to provide life support for as long as necessary. In 2012, the admission rate to the NICU for VLBW children was 844.1 per 1,000 live births (Harrison and Goodman 2015).

The VLBW designation also may affect receipt of specific medical treatments and procedures. For example, the 1,500-gram threshold is commonly cited as a threshold for using diagnostic ultrasounds.[3] As noted in Almond et al. (2010), it is likely that VLBW infants also receive differential treatment in terms of health inputs that are often difficult to measure in administrative data. For example, Angert and Adam (2009) recommend that VLBW infants should be "handled gently and not placed in a head down or Trendelenburg position" in order to minimize risk of intraventricular hemorrhage or brain injury.

Finally, children born under 1,500 grams often receive additional post-hospital treatment. The AAP guidelines recommend that high-risk infants should receive follow-up services to conduct neurological assessment. Many hospitals have adopted this

[3] Almond et al. (2010) note that the neonatal manual at the Longwood Medical Area (in Boston, MA) stipulates the use of diagnostic ultrasounds for at-risk newborns.

recommendation and created follow-up programs that specifically use VLBW status as a criterion for inclusion.[4,5]

3 Data

We use data from anonymized administrative records from a relational database created by researchers at Research Improving People's Lives (RIPL) in partnership with the state of Rhode Island. All personally identifiable information has been removed and replaced with anonymous identifiers. These identifiers allow researchers with approved access to join records across administrative sources (Hastings et al. 2019, Hastings 2019).

3.1 Birth Sample

We construct our analysis sample using birth records from the state of Rhode Island. We focus on all births from 1984-2016 (N=407,697). Using these data, we impose two restrictions. First, because we identify impacts using an RD design in a window around the 1,500-gram cutoff, we remove births where the infant's weight was less than 1,300 or greater than 1,700 grams. Second, following Barreca et al. (2011), we remove births where the infant's weight was within a "doughnut hole" of three grams around the 1,500-gram threshold to exclude observations that could be subject to technological constraints in measurement precision and rounding tendencies by attending physicians (which cause bunching in birthweight around the cutoff).

Our final sample contains 2,726 children who weigh 1,300-1,700 grams at birth. We join this sample to administrative data sources to measure outcomes in childhood and early

[4] In Rhode Island, the Women and Infants Hospital routinely schedules follow-up appointments from birth to adolescence for children born less than 1,500 grams. This follow-up program provides medical management for infants with respiratory problems and developmental assessments for cognitive, language, motor skills, behavior, executive functioning, memory, and phonological processing (Vohr et al. 2010).

[5] For example, 90 percent of NICUs in California are funded by the California Children's Services (CCS) program. All CCS-funded NICUs are required to provide a neonatal follow-up service for high-risk infants (California Perinatal Quality Care Collaborative 2018).

adulthood. Each outcome is available for a subset of cohorts since the data sources have partial overlap with the birth records.

3.2 Development and Childhood Education Outcomes

3.2.1 Supplemental Security Income (SSI)

Records from the Rhode Island Department of Human Services provide data on enrollment in the SSI program during the period 1996-2015. The SSI program is a federal program that provides cash assistance to a variety of recipients, including disabled children. For children, eligibility for SSI payments is determined using birth weight or a "failure to thrive" benchmark (SSA 2018a; 2018b). We measure SSI enrollment at birth and at any point from birth to age 3.

3.2.2 Grade Repetition (Grades 1-4)

Records from the Rhode Island Department of Education (RIDE) provide information on grade repetition in public school during 2003-2016. A child is indicated as repeating a grade if they are observed in the same grade in different school years. We study repetition during grades 1-4.

3.2.3 Special Education Services (IEP)

Records from RIDE indicate enrollment in special education services when a child has a written Individualized Education Program (IEP) during 2003-2016. An IEP can be given in preschool, and children are assessed each year. We study the number of school years that a child has an IEP in grades 1-4.

3.2.4 Elementary and Middle School Test Scores

Records from RIDE provide information on standardized test scores in grades 3-8 from 2005-2014. We study individual test scores observed in grades 3, 5 and 8. We also conduct analysis pooling all test scores for grades 3-8.

3.3 High School and Higher Education Outcomes

3.3.1 High School Disciplinary Offenses

Records from RIDE provide information on disciplinary offenses during high school for children enrolled in public school during 2003-2016. We study the number of offenses during grades 9-12. This measure is not defined for children who are too young or too old to be observed in all high school years.

3.3.2 College Preparation

RIDE data contain SAT, PSAT and AP records from 2010-2015. We use this data to construct an index for college preparation. For the PSAT and SAT components, we create indicators for whether a child took either exam in grades 9-12. For the AP component, we create an indicator for whether a child took at least one AP class in grades 9-12. We define the index for children who are enrolled in a Rhode Island public school for at least one high school grade.

3.3.3 Post-Secondary Enrollment

Records from the National Student Clearinghouse (NSC) provide information on post-secondary enrollment during 2003-2016. We study two enrollment outcomes: whether a child ever enrolls in a two- or four-year college by age 22 and whether a child ever enrolls in a four-year college by age 22.

3.4 Social Program Expenditures

Records from the Rhode Island Department of Human Services allow us to construct social program expenditure measures. We include payments from the following social programs: Supplemental Nutrition Assistance Program (SNAP), Medicaid, and Temporary Assistance for Needy Families (TANF). We also include program costs for having an IEP, which we set to an average expenditure of $34,000.[6]

[6] We calculate IEP per pupil costs using expenditure statistics from the Annual Per Pupil Expenditure Report from RIDE for the years 2012-2013 and IEP pupil data from the U.S. Department of Education (Rhode Island Department of Education 2013; U.S. Department of Education 2015).

We construct the social program expenditure measures by adding up all payments received and IEP costs by age 10, 12, or 14. For family-level payments, we use the per-capita payment value. Anonymized enrollment and payment records are available from 1997-2016. Children who are not old enough by the end of the data source for a given age range are not included in this measure. Note that the anonymized payment records do not have uniform coverage during the period 1997-2016. For example, SNAP payment records do not begin until 2004, while Medicaid claim records begin in 2000. Thus, children born prior to the latest source start date (2004) will be missing payment records for the first years of their life. We address this censoring issue by imputing program payments for children born before 2004 using the annual average payments during the years we observe them.[7]

3.5 Parental Behavior and School-Quality Measures

3.5.1 Maternal Care and Stress

PRAMS survey data provide information on maternal care and stress outcomes. This survey collects information on maternal attitudes and experiences before, during, and after pregnancy. All mothers with low birth weight (less than 2,500 grams) children receive the PRAMS survey, though only a fraction respond. We construct two indices from these data. The first is an index of maternal care based on survey questions on childcare practices such as breastfeeding or knowledge of proper sleeping position. This care index is available from 2002-2014. The second is an index of maternal stress based on responses about stressful events, postpartum depression, and degree of child difficulty. The stress index is available from 2004-2014.

[7] For example, suppose a child was born in 2003. This child will be missing one year of social program participation since complete coverage for all children starts in 2004. To address this, we impute social program spending in 2003 using the average spending level observed in the years we do observe them (2004 - X), where X is the year in which the child turns 10, 12, or 14.

3.5.2 Maternal Labor Market Activity

Records from the Rhode Island Department of Labor and Training provide information on employment and earnings during 1991-2016. We study maternal employment during the first two years after birth and average maternal earnings over the same two-year period.

3.5.3 Medicaid Enrollment and Expenditures by Age 2

Records from the Rhode Island Department of Human Services allow us to measure enrollment and total expenditures from Medicaid. The data on enrollment and expenditures are available for the periods 1984-2016 and 1999-2015, respectively. The measure of total expenditures is based on claims for emergency, institutional, physician, and pharmacy services. In our analysis, we focus on enrollment and total expenditures during the first two years of the child's life.

3.5.4 Subsequent Maternal Fertility

Birth records from the state of Rhode Island allow us to measure subsequent maternal fertility. The subsequent maternal fertility measure indicates whether a mother has had an additional child within three years following the birth of a child in our analysis sample. Birth records in RI are available through 2016.

3.5.5 School Value-Added Measures

Records from RIDE and NSC allow us to measure school-level value-added for the public elementary, middle, and high schools that children attend in Rhode Island. For elementary and middle schools, we measure school-level value-added in terms of standardized math and reading test scores in RIDE test score records. For high schools, we measure school-level value-added in terms of college enrollment observed in the NSC data. Value-added for each school is estimated using all years available, excluding the students in our RD sample. To construct value-added, we estimate models which regress individual test scores or college enrollment on prior achievement or background characteristics (Kane et al. 2008; Chetty et al. 2014).[8] We use the school-level mean of the resulting residuals as the measure of value-added.

3.6 Descriptive Statistics

Appendix Table A.1 presents summary statistics for all births in Rhode Island during our sample period (Column 1) and the RD sample of births (Column 2). Column 3 reports the p-value from a test of the equality of each statistic between the RD sample and all other births. The RD sample contains children who differ in terms of demographic and economic background characteristics. For example, mothers of children in the RD sample are less likely to be married or have a post-secondary degree (college or higher). Based on PRAMS survey measures, these mothers are also about 7 percentage points more likely to smoke.[9,10] Finally, children in the RD sample come from households with greater rates of pre-birth enrollment in social assistance programs. Household earnings (based on UI records in Rhode Island) are also lower in the RD sample.[11]

4 Empirical Strategy

We follow Almond et al. (2010) and Bharadwaj et al. (2013) by estimating the impact of classification as a VLBW newborn using the following regression-discontinuity model:

$$y_i = \alpha + \gamma d_i + \theta 1(d_i < 0) + \delta\ 1(d_i < 0) \times d_i + \varepsilon_i \tag{1}$$

where y_i is an outcome of interest, X_i is a set of child characteristics measured prior to birth, d_i is a running variable which is equal to the difference between an infant's birth weight and the 1,500-gram threshold for VLBW status, $1(d_i < 0)$ is an indicator for VLBW status,

---

[8] Details on the specifications are provided in the notes for Appendix Table A.13.

[9] The PRAMS survey response data is available only for a subsample of mothers.

[10] Prior research suggests maternal smoking is the leading cause of low birth weight in the U.S. (Kramer 1987).

[11] We define household income as earnings for mothers and fathers who work in Rhode Island in the four quarters prior to birth. Earnings outside of Rhode Island are not captured in this measure. If a woman is not married, the father's earnings will also not be captured in our definition of household income.

and ε_i is an error term. We estimate Equation 1 using OLS and cluster standard errors at the gram level to address concern about the discreteness of birth weight (Lee and Card 2008). In Section 5.4, we examine the sensitivity of results to a changing the bandwidth for births and the polynomial for either side of the 1,500-gram threshold. All robustness tests exclude births where the infant's weight was within three grams of the 1,500-gram threshold (Barreca et al. 2011).

4.1 Evidence on Medical Spending and Interventions

As discussed in Section 2, medical practitioners may use the VLBW classification for considering whether an infant receives additional medical care. Almond et al. (2010) examine the impact of crossing the 1,500-gram threshold on measures of hospital care from discharge records. Their analysis finds that hospital costs increase by $4,000 and the length of hospital stay is extended by 2 days for children just under the VLBW threshold.

Appendix Table A.2 replicates the analysis of medical treatments for VLBW children using data for Rhode Island from the Healthcare Cost and Utilization Project (HCUP).[12] In Panel B, we estimate the impact of crossing the 1,500-gram threshold on the number of days spent in the NICU, length of stay in the hospital, and total charges. We find a statistically significant 3.4 day increase in the number of days spent in the NICU for children born just under the 1,500-gram threshold. The effect on total charges is not statistically significant, and the point estimate of $3,470 is smaller than the $9,450 estimate from Almond et al. (2010).[13]

4.2 Evidence on Identifying Assumptions

[12] The HCUP data for Rhode Island is available for the years 2002-2015. Our sample contains 1,724 children born during this time period for which we have non-missing data on the length of stay and total charges. The variable recording days spent in the NICU is available for the years 2006-2015.

[13] Almond et al. (2010) use cost-to-charge ratios from the Centers for Medicare and Medicaid Services to adjust the $9,450 estimated threshold crossing impact into a $4,000 estimated effect on total hospital costs.

Our research design relies on the assumption that factors other than receipt of additional initial medical care vary continuously with birth weight across the 1,500-gram threshold for VLBW status. Appendix Table A.3 shows results from tests of the validity of this assumption where we regress various demographic or pre-birth characteristics on the independent variables from Equation 1. Column 2 reports estimates for the coefficient of the VLBW indicator. The results show no detectable discontinuities in baseline measures.

We also test our identification assumption by examining the density of births around the threshold for VLBW status. Appendix Figure A.1 is a histogram of births between 1,300 and 1,700 grams. Note that this figure includes births from the RD sample and births where the infant's weight was within three grams of the 1,500-gram threshold (which we exclude from our main analysis). As in prior studies, there is pronounced heaping at the "round" gram numbers (such as multiples of 100) and at the gram equivalent of ounce intervals (Almond et al. 2010; Bharadwaj et al. 2013). There is no apparent irregular pattern of heaping near the 1,500-gram threshold. We conduct a McCrary (2008) test and fail to reject the null hypothesis of continuity at the threshold (p-value=0.924).

5. Main Results

Tables 1-4 report regression results based on Equation 1. To supplement this analysis, Figure 1 and Appendix Figures A.2-A.5 plot means of outcome variables within 20-gram bins of birth weight. For all results, we use the sample of 1,300-1700 gram births (excluding observations within three grams of the 1,500 threshold).

5.1 Development and Childhood Education

Table 1 reports results from Equation 1 for development and childhood education outcomes. Children below the 1,500-gram threshold have 0.444 higher average (math and reading) standardized test scores measured in third grade. There are similar impacts on standardized test scores in fifth and eighth grade, suggesting that the permanent increases in child ability accrued by third grade persists through middle school. To give a clearer sense of our results, Figure 1 (Panel A) shows means for test scores in the pooled sample (all grades 3-8 scores) in 20-gram bins of birth weight. This figure shows a clear discontinuity at the 1500-gram threshold.

While we find no significant effects of threshold crossing for enrollment in SSI or special education services (IEP), the point estimates are consistent with children just under the threshold having improved outcomes. For example, the point estimates suggest that children who marginally qualify for VLBW status are 2.4 percentage points (96 percent) less likely to participate in SSI from ages 0-3. Similarly, the point estimate suggests that there is a 20.5 percentage point reduction in the years that a child has an IEP in grades 1-4.

Overall, these results are similar to previous findings of significant threshold crossing impacts on standardized test scores in Chile and Norway. Bharadwaj et al. (2013) find that children in Chile who are born just under the 1,500-gram threshold have 0.13 standard deviations higher test scores measured in first to eighth grade in Chile. Similarly, they find that threshold crossing has a positive impact of 0.22 standard deviations on test scores in tenth grade for children in Norway. An additional comparison for these results comes from studies of early childhood education programs such as Head Start. Duncan and Magnuson (2013) review 84 experimental and quasi-experimental evaluations of early childhood education programs and report that the average impact is 0.35 standard deviations.

5.2 High School and Post-Secondary Education

Table 2 shows that crossing the threshold for VLBW status has significant and beneficial impacts on several outcomes measured in high school and early adulthood. For high school outcomes, we find benefits in terms of non-cognitive outcomes. Threshold crossing is associated with a 13-percentage point reduction in the likelihood of having a disciplinary offense in high school. This represents a 72-percent decrease relative to the mean offense rate of 17.9 percent. We find no significant impacts on a high school index of readiness for college, although the point estimate is positive.

After high school, we find notable improvements in post-secondary education enrollment. Table 2 and Panel B of Figure 1 shows that children just under the 1,500-gram threshold are 20 percentage points more likely to enroll in a 4-year college by age 22. This is a large impact given the mean enrollment rate of 30.5 percent for children born above the birth weight threshold. When we examine enrollment specifically at four-year institutions, we find similarly significant and positive impacts.

Our estimates suggest that infant health investments at birth can reverse previously documented negative impacts of the low birth weight on long-run human capital outcomes. For example, Currie and Hyson (1999) find that low birth weight children are more than 25 percent less likely to pass high school exit exams in Britain even after controlling for a range of background characteristics. Using data on twins from Norway and a fixed effects approach, Black et al. (2007) find that lower birth weight significantly reduces the likelihood of graduating from high school. Similarly, Oreopoulos (2008) and Royer (2009) provide estimates based on within family variation that show reduced weight has negative impacts on schooling.

5.3 Social Program Expenditures

We provide direct evidence on the social impact of VLBW threshold crossing by studying social program spending for children in our sample. Any realized impact on social spending can be compared to the costs associated with additional hospital care provided to children born just below the 1,500-gram threshold. As noted in Section 4.1, we find a statistically insignificant (positive) threshold crossing effect on hospital expenditures using data for Rhode Island. Given the imprecision in this estimate, it may be helpful to refer to estimates from Almond et al. (2010). In a larger sample of states, they find a significant \$4,000 increase in hospital costs for children born just below 1,500 grams.

As outlined in Section 3, we calculate total social program expenditures (in 2015 dollars) that accrue to the child by the time they reach age 10, 12, and 14. Table 3 presents estimates of the impact of crossing the VLBW threshold on the extensive (Panel A) and intensive (Panel B) margin of these measures of social expenditures. Appendix Figure A.5 provides corresponding visual results for the impact of threshold crossing for the age 10 measures. Panel A shows no significant impacts on the likelihood of having any social program expenditures, although the point estimates are consistently negative. At the same time, Panel B shows that there are marginally significant reductions on total spending for all measures. For example, we find that social safety net expenditures by age 10 are \$27,291 *lower* (p-value<0.10) for children who are born just to the left of the 1500-gram threshold, and \$66,997 *lower* (p-value<0.05) by the age of 14. This suggests a sizeable direct public return relative to the \$4,000 spent on additional hospital costs.

5.4 Robustness

Appendix Tables A.4-A.8 conduct sensitivity tests by using quadratic RD specifications and changing the bandwidth for births included in analysis.[14] We use the optimal bandwidth selection method from Calonico et al. (2014) for each outcome. Our results are generally robust to these changes: coefficient signs and magnitudes remain similar, although estimates sometimes lose statistical significance.

Appendix Table A.10 examines attrition from the sample of test scores and high school enrollment (recall that once a child reaches high school, we will see them in the NSC data regardless of whether the child moves from Rhode Island). The only statistically

[14] Appendix Table A.9 provides test scores results for the sample of births that overlaps best with the period for which we have HCUP data from Rhode Island. The point estimates from this analysis are similar to the results in Table 1.

significant results is a negative point estimate for eighth grade (p-value<0.10). The fact that we generally find insignificant impacts on attrition in high school suggests differential attrition is not the primary factor driving the results on educational attainment or social program expenditures.

Finally, Table A.11 tests for discontinuities in outcomes at other intervals of 100 grams. A pattern of significant threshold crossing effects the 100-gram intervals of birth weight that exceed 1,500 grams would generate concern that heaping is driving our main results. The first row of Table A.11 reproduces the 1,500-gram threshold crossing estimates on selected outcomes. The remaining rows provide estimates for every 100-gram cutoff between 1,600 and 3,000 grams. There is no consistent pattern of threshold crossing effects on test scores or other selected outcomes for any cutoff above 1,500 grams.

5.5 Heterogeneity

Appendix Table A.12 reports results for education and social expenditure outcomes after dividing the RD sample based on the mother's education (as reported in birth records). We compare threshold crossing estimates for mothers who have at least some college or higher (Panel A) and mothers who have a high school education or lower (Panel B). The subgroup-specific estimates are generally not statistically significant. In most cases, the point estimates for both subgroups are similar to the results reported in Table 1. We cannot reject the hypothesis that estimates for each outcome are equal across education subgroups.

6. Mechanisms

The potential for parental responses to child health complicates the interpretation of the reduced form estimates of the impact of crossing the 1,500-gram threshold. Parents, teachers, and the broader community may respond to a child's health status, thereby amplifying or reducing the direct impact of increased early-life medical interventions for VLBW children. Existing empirical studies provide mixed evidence as to whether we should expect to find evidence of parental responses in our setting.[15] Datar et al. (2010) use a family fixed effects approach and find that parents report investing more resources into children that have heavier birth weight. In contrast to these results, Bharadwaj et al. (2013) examine data from Chile and find that time spent reading with a child does not change as a function of crossing the VLBW threshold. They also study whether parents respond to child health through school choice, finding no evidence that children just under the 1,500-gram threshold attended elementary schools with higher or lower average test scores.

Table 4 shows threshold crossing impacts on direct measures of parental responses and other factors that could plausibly mediate the impacts in our sample. We first study indices for maternal care and stress based on PRAMS responses. The care index is the number of correct responses to questions about infant care such as sleeping position. The stress index is the number of affirmative responses to questions about depression and anxiety. Although we find statistically insignificant impacts, the point estimates for the care and stress indices are positive (indicating better care) and negative (indicating less stress), respectively.

Next, we study maternal labor market outcomes using earnings records. The effect of crossing the 1,500-gram threshold could be positive if mothers work more due to the impact of additional in-hospital care on child health. Yet, there could be negative impacts if mothers respond to health improvements by investing more time with their child. Table 4 shows no significant impacts on employment or mean annual earnings measured during the first two years postpartum.

As a measure of early-life child health, we study Medicaid enrollment and expenditures after birth up to age 2. Impacts on either outcome could signal increased parental involvement in health. At the same time, we also expect that increased medical

[15] For a review of studies of parental responses to child health, see Almond and Mazumder (2013).

care at birth should reduce post-birth medical needs. In Table 4, we find no statistically significant effect on either measure of Medicaid program use. Although imprecise, the point estimate for expenditures is negative and represents a reduction of about 50 percent relative to the mean for children just above 1,500 grams at birth.

Another parental response that could affect child outcomes is subsequent maternal fertility. Table 4 shows that mothers of children born just under 1,500 grams are about 6.4 percentage points more likely to have a child relative to children born just above the cutoff. This estimate represents a relatively large effect size (compared to the mean fertility rate above the threshold).[16]

Finally, Appendix Table A.13 tests whether parents of children born just under 1,500 grams are more likely to send their children to higher-quality schools. Prior research suggests that there is a strong link between school quality and child achievement (Hastings and Weinstein 2008; Hoxby and Murarka 2009; Abdulkadiroglu et al. 2011; Pop-Eleches and Urquiola 2013; Deming et al. 2014). We find statistically insignificant and small point estimates for school quality measured in elementary, middle and high school.

8 Conclusion

This paper uses comprehensive administrative data to provide new evidence on the impacts of providing additional medical care for children with low birth weight. We identify causal effects using an RD approach based on the 1,500-gram threshold for VLBW status. Prior studies by Almond et al. (2010) and Bharadwaj et al. (2013) document that physicians and hospitals make decisions about medical care using this threshold.

We find significant, large and positive impacts of crossing the 1,500-gram threshold on academic achievement, college enrollment and reliance on social programs. Children just

[16] Appendix Table A.7 shows the point estimate is positive but smaller when we use the optimal bandwidth selected using the method from Calonico et al. (2014).

under the threshold for VLBW status have 0.34 standard deviations higher test scores in grades 3-8. They are also 32 percent more likely to enroll in college and receive $66,997 *less* in social program expenditures by age 14.

When we analyze potential mediators of these impacts, the findings suggest there is a large role for direct effects of additional medical care. We find no statistically significant impacts of threshold crossing on measures of parental responses such as maternal stress or labor market outcomes. At the 1,500-gram threshold, there are also no detectable changes in the quality of schools that children attend during childhood.

Overall, this paper suggests there are private and public benefits to investing in health for children near the VLBW threshold. There are long-run private gains given that increased schooling should translate into higher lifetime earnings and better health. In terms of public benefits, the analysis suggests more immediate returns in the form of reduced participation in social programs during childhood. The magnitude of these gains is large relative to the initial hospital costs of providing VLBW infants with additional care.

References

Abdulkadiroğlu, Atila, Joshua D. Angrist, Susan M. Dynarski, Thomas J. Kane, and Parag A. Pathak. “Accountability and Flexibility in Public Schools: Evidence from Boston’s Charter and Pilots.” *The Quarterly Journal of Economics* 126 (2): 699-748.

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

Almond, Douglas, and Janet Currie. 2017. "Childhood Circumstances and Adult Outcomes: Act II. ." NBER Working Paper #23017. National Bureau of Economic Research.

Almond, Douglas, and Mazumder Bhashkar. 2013. "Fetal Origins and Parental Responses." *Annual Review of Economics* 5 (1): 37-56.

Almond, Douglas, Joseph J. Doyle, Amanda E. Kowalski, and Heidi Williams. 2010. "Estimating Marginal Returns to Medical Care: Evidence from At-Risk Newborn." *Quarterly Journal of Economics* 125 (2): 591-634.

American Academic of Pediatrics, Committee on Fetus and Newborn. 2008. "Hospital Discharge of the High-Risk Neonate." *Pediatrics* 122 (5): 119-26.

American Academic of Pediatrics, Committee on Fetus and Newborn. 2012. "Levels of Neonatal Care." *Pediatrics* 130 (3): 587-97.

Angert, Robert, and Henry Adam. 2009. “Care of the Very Low-birthweight Infant.” *Pediatrics in Review 30,* no. 1 (January 2009): 32-35.

Bald, Anthony, Eric Chyn, Justine Hastings and Margarita Machelett. 2019. “The Causal Impact of Removing Children from Abusive and Neglectful Homes.” NBER Working Paper #25419. National Bureau of Economic Research.

Barreca, Alan I, Melanie Guldi, Jason M. Lindo, and Glen R. Waddell. 2011. "Saving Babies? Revisitng the Effect of Very Low Birth Weight Classification." *Quarterly Journal of Economics* 126 (4): 2117-2123.

Bharadwaj, Phrashant, Katrine Veelsen Loken, and Christopher Neilson. 2013. "Early Life Health Interventions and Academic Achievement." *American Economic Review* 103 (5): 1862-91.

Black, Sandra E., Paul J. Devereux, and Kjell G. Salvanes. 2007. "From the Cradle to the Labor Market? The Effect of Birth Weight on Adult Outcomes." *Quarterly Journal of Economics* 122 (1): 409-439.

California Perinatal Quality Care Collaborative. 2018. High Risk Infant Follow-up. Accessed August 24, 2018. https://www.cpqcc.org/perinatal-programs/high-risk-infant-follow.

Calonico, Sebastian, Matias D. Cattaneo, Rocio Titiunik, and others. 2014. “Robust Data-Driven Inference in the Regression-Discontinuity Design.” Stata Journal 14 (4): 909–946.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104 (9): 2593-2632.

Chetty, Raj, John N. Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How does your kindergarten classroom affect your earnings? Evidence from Project STAR." *Quarterly Journal of Economics* 126 (4): 1593-1660.

Chetty, Raj, Nathaniel Hendren, and Lawrence Katz. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Project." *American Economic Review* 106 (4): 855-902.

Chyn, Eric. 2018. "Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children.*" American Economic Review* 108 (10): 3028-3056.

Currie, Janet. 2011. “Inequality at Birth: Some Causes and Consequences.” *American Economic Review* 101 (3): 1–22.

Currie, Janet, and Douglas Almond. 2011. "Human Capital Development Before Age Five." Chap. 15 in Handbook of Labor Economics, edited by David Card and Orley Ashenfelter, 1315-1486. Elsevier.

Currie, Janet, and Duncan Thomas. 1995. "Does Head Start Make a Difference?" *The American Economic Review* 85 (3): 341–64.

Currie, Janet, and Rosemary Hyson. 1999. "Is the Impact of Health Shocks Cushioned by Socioeconomic Status? The Case of Low Birthweight." *American Economic Review* 89 (2): 245-50.

Datar, Ashlesha, Rebecca M. Kilburn, and David S. Loghran. 2010. "Endowments and Parental Investments in Infancy and Early Childhood." *Demography* 47 (1): 145-162.

Deming, David J., Justine S. Hastings, Thomas J. Kane, and Douglas O. Staiger. 2014. "School Choice, School Quality, and Postsecondary Attainment." *American Economic Review* 104 (3): 991-1013.

Duncan, Greg J., and Katherine Magnuson. 2013. "Investing in Preschool Programs." *Journal of Economic Perspectives* 27 (2): 109-32.

Harrison, Wade, and David Goodman. 2015. "Epidemiologic Trends in Neonatal Intensive Care." *JAMA Pediatrics* 169 (9) 855-62.

Hastings, Justine S., and Jeffrey M. Weinstein. 2008. "Information, School Choice, and Academic Achievement: Evidence from Two Experiments." The Quarterly Journal of Economics 123 (4): 1373–1414.

Hastings, Justine S. 2019. *Fact-Based Policy: How Do State and Local Governments Accomplish It?* Research Improving People's Lives, Brookings Institute.

Hastings, Justine S., Mark Howison, Ted Lawless, John Ucles, and Preston White. 2019. "Unlocking Data to Improve Public Policy." Communications of the ACM. Forthcoming.

Havnes, Tarjei, and Magne Mogstad. 2011. "No Child Left Behind: Subsidized Child Care and Children's Long-Run Outcomes." *American Economic Journal: Economic Policy* 3 (2): 97-129.

Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review* 103 (6): 2052-2086.

Hoxby, Caroline M., and Sonali Murarka. 2009. "Charter Schools in New York City: Who Enrolls and How They Affect Their Students' Achievement." NBER Working Paper #14852. National Bureau of Economic Research.

Hoynes, Hilary, Diane Whitemore Schanzenbach, and Douglas Almond. 2016. "Long-Run Impacts of Childhood Access to the Safety Net." *American Economic Review* 106 (4): 903-934.

Kramer, Michael S. 1987. "Intrauterine Growth and Gestational Duration Determinants." *Pediatrics* 80 (4): 502–11.

Kane, Thomas J., and Douglas O. Staiger. 2008. "Estimating Teacher Impacts on Student Achievement: An Experimental Evaluation." NBER Working Paper #14607. National Bureau of Economic Research.

Lee, David S., and David Card. 2008. "Regression Discontinuity Inference with Specification Error." *Journal of Econometrics* 142 (2): 655-74.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics*, 142 (2): 698–714.

Oreopoulos, Philip, Mark Stabile, Randy Wald, and Leslie L. Roos. 2008. "Short-, Medium-, and Long-Term Consequences of Poor Infant Health An Analysis Using Siblings and Twins." *Journal of Human Resources* 43 (1): 88-138.

Pop-Eleches, Cristian, and Miguel Urquiola. "Going to a Better School: Effects and Behavioral Responses." *American Economic Review* 103, no.4 (June 2013): 1289-1324.

Rhode Island Department of Education. 2013. Annual Per Pupil Expenditure Report.

Royer, Heather. 2009. "Separated at Girth: US Twin Estimates of the Effects of Birth Weight." *American Economic Journal: Applied Economics* 1 (1): 49-85.

SSA. 2018a. *100.00 Low Birth Weight And Failure To Thrive - Childhood.* Retrieved March 2019 from https://www.ssa.gov/disability/professionals/bluebook/100.00-GrowthImpairment-Childhood.htm#100_04B.

SSA. 2018b. *Understanding Supplemental Security Income SSI for Children--2018 Edition.* Retrieved March 2009 from https://www.ssa.gov/ssi/text-child-ussi.htm.

U.S. Department of Education. 2015. Local Education Agency Universe Survey.

Vohr, Betty, Bonnie Stephens, and Richard Tucker. 2010. "35 Years of Neonatal Follow-up in Rhode Island." *Medicine & Health Rhode Island* 93 (5): 151-253.

Table 1: Impacts for Development and Education Outcomes

	Mean for BW > 1,500 grams (1)	RD Est. BW < 1,500 grams (2)	Observations (3)	Birth Cohort (4)
SSI, IEP, and Grade Rep.				
SSI Age 0	0.017	-0.008	1,790	1996-2015
		(0.015)		
SSI Ages 0 - 3	0.025	-0.024	1,577	1996-2012
		(0.019)		
Years on IEP (1-4)	1.068	-0.205	605	1996-2006
		(0.259)		
Grade Repetition (1-4)	0.105	-0.131***	608	1996-2006
		(0.039)		
Avg. Std. Test Score				
3rd Grade	-0.284	0.444***	544	1996-2005
		(0.162)		
5th Grade	-0.244	0.299*	527	1994-2003
		(0.159)		
8th Grade	-0.176	0.314*	485	1990-2000
		(0.165)		
All (3-8)	-0.220	0.338***	3,070	1990-2005
		(0.128)		

Notes: Column 1 reports the mean of the dependent variable for children born above 1,500 grams. Column 2 provides estimates of the impact of crossing the 1,500 gram threshold using Equation 1. Standard errors clustered at the gram level are presented in parentheses. Columns 3 and 4 report the total number of observations and describe the birth cohorts included in the analysis. The sample for all results includes children born with birth weight between 1,300 and 1,700 grams (excluding children born within 3 grams of the 1,500-threshold.) Statistical significance is denoted by *** $p<0.01$; ** $p<0.05$; * $p<0.10$.

Table 2: Impacts for High School and Higher Education Outcomes

	Mean for BW > 1,500 grams (1)	RD Est. BW < 1,500 grams (2)	Observations (3)	Birth Cohort (4)
Disciplinary Offenses (9-12)	0.179	-0.131**	393	1988-1998
		(0.061)		
College Preparation Index	-0.207	0.059	493	1993-2001
		(0.379)		
Any College Enrollment by Age 22	0.536	0.171**	416	1984-1994
		(0.067)		
4-Year College Enrollment by Age 22	0.305	0.200**	416	1984-1994
		(0.077)		

Notes: Column 1 reports the mean of the dependent variable for children born above 1,500 grams. Column 2 provides estimates of the impact of crossing the 1,500 gram threshold using Equation 1. Standard errors clustered at the gram level are presented in parentheses. Columns 3 and 4 report the total number of observations and describe the birth cohorts included in the analysis. The sample for all results includes children born with birth weight between 1,300 and 1,700 grams (excluding children born within 3 grams of the 1,500-threshold.) Statistical significance is denoted by *** $p<0.01$; **$p<0.05$; *$p<0.10$.

Table 3: Impacts for Social Program Expenditure by Age

	Mean for BW > 1,500 grams (1)	RD Est. BW < 1,500 grams (2)	Observations (3)	Birth Cohort (4)
Panel A: Any Expenditures (= 1)				
By Age 10	0.607	-0.057	951	1997-2006
		(0.071)		
By Age 12	0.587	-0.077	740	1997-2004
		(0.079)		
By Age 14	0.588	-0.117	529	1997-2002
		(0.089)		
Panel B: Total Expenditures ($)				
By Age 10	91,296	-27,291*	951	1997-2006
		(15,736)		
By Age 12	113,844	-44,067*	740	1997-2004
		(25,034)		
By Age 14	145,486	-66,997**	529	1997-2002
		(33,622)		
Panel C: Total Expenditures ($) | *Any Expenditures (=1)*				
By Age 10	150,334	-29,149	572	1997-2006
		(22,832)		
By Age 12	194,052	-53,444	434	1997-2004
		(33,486)		
By Age 14	247,515	-74,075*	310	1997-2002
		(37,596)		

Notes: Column 1 reports the mean of the dependent variable for children born above 1,500 grams. Column 2 provides estimates of the impact of crossing the 1,500 gram threshold using Equation 1. Standard errors clustered at the gram level are presented in parentheses. Columns 3 and 4 report the total number of observations and describe the birth cohorts included in the analysis. The sample for all results includes children born with birth weight between 1,300 and 1,700 grams (excluding children born within 3 grams of the 1,500-threshold.) Statistical significance is denoted by *** $p<0.01$; ** $p<0.05$; * $p<0.10$.

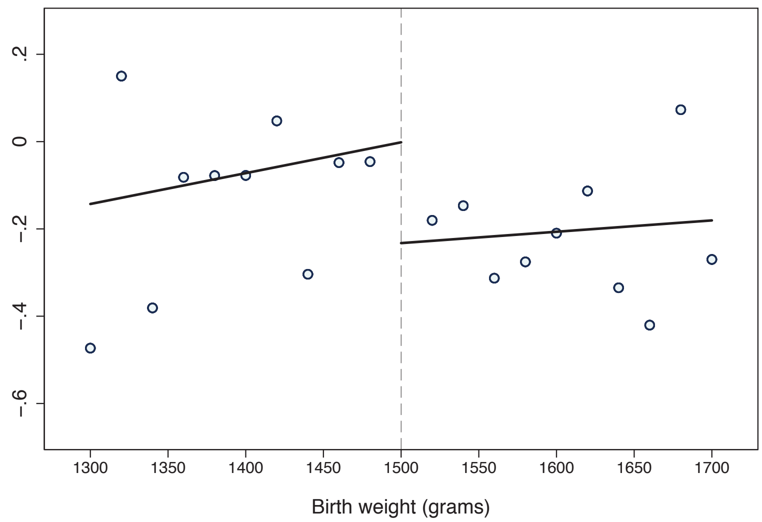
Table 4: Impacts for Early Childhood Investment Outcomes

	Mean for BW > 1,500 grams (1)	RD Est. BW < 1,500 grams (2)	Observations (3)	Birth Cohort (4)
Maternal Care Index	2.965	0.121	429	2002-2014
		(0.160)		
Maternal Stress Index	3.498	-0.156	429	2004-2014
		(0.338)		
Maternal Employment Ages 0-2 (=1)	0.588	-0.013	1,796	1991-2014
		(0.044)		
Avg. Maternal Earn. Ages 0-2 ($)	12,355	2,053	1,796	1991-2014
		(2,265)		
Medicaid Enrollment Ages 0-2 (=1)	0.455	-0.035	2,627	1989-2014
		(0.056)		
Medicaid Exp. Ages 0-2 ($)	14,171	-7,528	1,360	1984-2016
		(6,518)		
Birth Within 3 Years (=1)	0.132	0.064**	2,153	1989-2014
		(0.028)		

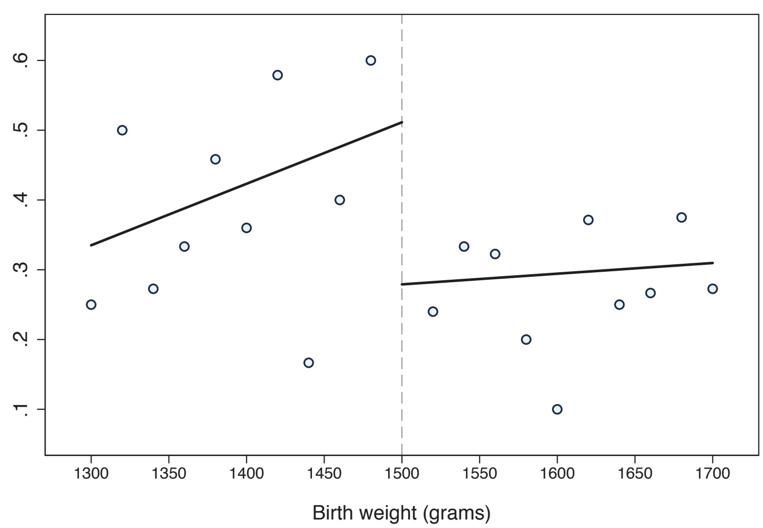
Notes: Column 1 reports the mean of the dependent variable for children born above 1,500 grams. Column 2 provides estimates of the impact of crossing the 1,500 gram threshold using Equation 1. Standard errors clustered at the gram level are presented in parentheses. Columns 3 and 4 report the total number of observations and describe the birth cohorts included in the analysis. The sample for all results includes children born with birth weight between 1,300 and 1,700 grams (excluding children born within 3 grams of the 1,500-threshold.) Statistical significance is denoted by *** $p<0.01$; **$p<0.05$; *$p<0.10$.

Figure 1: Impacts for Selected Education Outcomes

A. Test Scores (3-8)



B. 4-Year College Enrollment by 22



Notes: Each panel shows the relationship between birth weight and a selected education outcome. Dots represent means within 20 gram bins of the running variable. The dark lines are predictions from a linear model using the individual-level data.