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

Eric Chyn* Samantha Gold+ Justine Hastings§

August 6, 2020

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 probability of college enrollment, and a $66,997 decrease in social program expenditures by age 14. We explore potential mechanisms driving impacts.

Keywords:
JEL Codes: I11, I12, I18, I21, J13.

*Dartmouth College, RIPL, and NBER (email: eric.t.chyn@dartmouth.edu), +RIPL (email: sgold@ripl.org), §Brown University, RIPL, and NBER (email: justine_hastings@brown.edu). 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. Finally, we are grateful for financial support from the Smith Richardson Foundation and the Laura and John Arnold Foundation. Declarations of interests: none.

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 interventions delivered during early-life can counteract the negative effects of low weight and poor health at birth. For example, investments in additional medical care could generate substantial social benefits by enhancing later-life economic self-sufficiency and reducing reliance on social programs.

This paper provides new evidence on the effectiveness of early-life interventions for children born with birth weight below the 1,500-gram (3 pounds and 5 ounces) threshold for Very Low Birth Weight (VLBW) status. Prior research has studied this question using a regression design (RD) design that exploits changes in hospital treatment across the VLBW threshold. 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. Using data from Chile and Sweden, Bharadwaj et al. (2013) find similar results and show 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 previous 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 investment mechanisms, we use measures of parental behavior such as records on maternal labor market activity and school choice.[1]

We find significant and large positive impacts of crossing the 1,500-gram VLBW threshold from above 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

[1] As we discuss in Section 3, we study the reduced form impact of VLBW status. Prior studies have noted that crossing the VLBW threshold identifies the combined impact of changes in medical treatment provided in-hospital and changes in post-hospital inputs (Bharadwaj et al. 2013). Post-hospital inputs could change at the threshold because parents may also react to VLBW status and the additional medical care provided to their children.

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.[2]

Overall, our findings add to the literature on the effectiveness of early-life interventions and medical treatments for at-risk children. We extend Almond et al. (2010)

[2] 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. We present these results in table A.13. 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.

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).[3] 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).[4] Our results more generally contribute to the literature demonstrating 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 and Non-hospital 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

[3] Daysal et al. (2019) study data from Denmark and find that children born just below the VLBW threshold have higher test scores in ninth grade. They also document large positive spillovers on academic outcomes of siblings.

[4] 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.

using diagnostic ultrasounds.[5] 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 recommendation and created follow-up programs that specifically use VLBW status as a criterion for inclusion. In our context of Rhode Island, the Women and Infants Hospital routinely schedules follow-up appointments from birth to adolescence for children born less than 1,500 grams.[6] 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).[7]

3 Conceptual Framework and Empirical Strategy

Why might crossing the 1,500-gram threshold from above matter for long-run outcomes of children? This section provides a conceptual framework that highlights mechanisms that could drive an impact of crossing the threshold for VLBW status. Our discussion is based on the model introduced in Bharadwaj et al. (2013). We assume that initial health at birth (H) is based on birth weight (BW) and a component unobserved to

[5] 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.

[6] Note that we do not have data on follow-up appointments at the Women and Infants Hospital so we cannot examine participation in this program for our main analysis sample.

[7] Similarly, 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).

the econometrician (ϵ). Physicians and health care workers provide treatments (*D)* which are determined by a decreasing function of initial health and a random component (ν). The discussion from Section 2 also suggests that medical treatments increase discretely by the amount κ at the 1,500-gram threshold because health care providers are responsive to the VLBW designation. Formally, we assume the following two-equation model of health and medical treatments for individual *i*:

$$H_i = BW_i + \epsilon_i \tag{1}$$

$$D_i = g(H_i) + \kappa\, 1(BW_i < 1{,}500g) + \ v_i. \tag{2}$$

This framework makes clear that medical treatments D will be correlated with unobserved health status through the decreasing function $g(H)$. This implies that estimating the impact of early-life medical interventions on later-life outcomes can suffer from endogeneity even when conditioning on birth weight.

To address the identification challenge implied by Equations 1 and 2, we use the following regression discontinuity (RD) specification proposed by Almond et al. (2010) and Bharadwaj et al. (2013):

$$y_i = \alpha + \gamma\, d_i + \theta\, 1(BW_i < 1{,}500g) + \delta\, 1(BW_i < 1{,}500g)\, d_i + \varepsilon_i \tag{3}$$

where y is an outcome of interest, *d* is a running variable which is equal to the difference between an infant's BW and the 1,500-gram threshold for VLBW status, $1(BW_i < 1{,}500g)$ is an indicator for VLBW status, and ε is a regression error term. The key regression parameter of interest is θ which is the coefficient on the indicator for crossing the VLBW threshold from above. Estimates of this coefficient will reflect the two types of effects. First, the causal effects of medical care will be captured because treatments jump at the VLBW threshold by the amount $\Delta D_i = \kappa$ and this discrete shift is uncorrelated with unobserved determinants of initial health. Second, the effects of post-hospital investments will also be captured if parental behavior or other post-hospital inputs (e.g., follow-up care appointments for at-risk infants) respond to the VLBW designation.

To formalize this discussion of mechanisms, we assume there is a structural model of outcomes that features both medical treatments and post-hospital investments. For the latter, we assume investments up to period t are a function of post-hospital investments (ξ), initial child health, and medical treatments at birth: $I_t^{post}(\xi, H, D)$. The model for an outcome for individual i at period t can be written as:

$$y_{it} = \phi_t H_i + \psi_t D_i + \tau I_t^{post}(\xi_i, H_i, D_i) + \omega_{it}. \quad (4)$$

This framework makes clear that threshold crossing estimates from the RD model from Equation 3 will capture both the combined effects of local variation in D_i and $I_t^{post}(\xi, H, D)$. Formally, Equation 4 allows us to re-write the threshold crossing estimate as:

$$\hat{\theta} = \psi_t \kappa + \tau \Delta I_t^{post}, \quad (5)$$

where ΔI_t^{post} is the difference in the average post-hospital investments that children receive at the threshold for VLBW status. The estimated coefficient $\hat{\theta}$ should be thus interpreted as the total policy relevant effect of the increased medical care at the VLBW threshold, which can include potential changes (positive or negative) in accumulated investments. As discussed further in Section 6, we attempt to gauge the importance of post-hospital investments by studying threshold crossing impacts on several measures of parental behavior and non-hospital inputs.

An additional consideration for the RD analysis is that the composition of the sample for some outcomes could be affected if crossing the VLBW threshold reduces mortality. For example, we would not observe standardized test scores for a child who does not survive to school age. We can test whether concerns over composition bias are warranted by conducting an RD analysis of sample attrition. Specifically, Section 5.4 presents results where we estimate impacts of threshold crossing on indicators for whether we observe children enrolled in Rhode Island public schools.

Finally, note that we estimate the RD model from Equation 3 using OLS and cluster standard errors at the gram level to address concern about the discreteness of birth weight (Lee and Card 2008). Our main results are based on using a window of 1,300-1,700 grams for birth weight, which follows other recent studies of VLBW. We chose this bandwidth to align with other recent RD studies of VLBW such as Bharadwaj et al. (2013) and Daysal et al. 2019.[8] All our analysis excludes births where the infant's weight was within three grams of the 1,500-gram threshold (Barreca et al. 2011).

4 Data

We use two types of data for our analysis. First, the bulk of our analysis relies on 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). Second, we use hospital discharge data from the Healthcare Cost and Utilization Project (HCUP). Note that we cannot link the HCUP records to any other data sources since there are no identifiers in these records. In the following subsections, we provide further details on the samples and outcomes that we construct for our analysis.

4.1 HCUP Sample and Hospitalization Outcomes

To provide evidence on the early-life medical interventions delivered in our context, we rely on data from the Healthcare Cost and Utilization Project (HCUP). The HCUP

[8] Section 5.4 examines the sensitivity of results to changing this bandwidth. In addition, Section 5.4 studies robustness by reporting results based on specifications that feature quadratic functions for the running variable terms.

records for Rhode Island are available for the years 2002-2015. In these records, we focus on two samples of children. First, we study 1,724 children who weigh 1,300-1,700 grams at birth and for whom HCUP provides non-missing data on the length of hospital stay and total charges. Second, we study 1,245 children who weigh 1,300-1,700 grams at birth and for whom we have non-missing data on the number of days spent in the NICU.[9]

4.2 Main Birth Records Sample

For our main analysis of outcomes in childhood and early adulthood, we construct an analysis sample based on birth records from the state of Rhode Island. We construct this sample as follows. We start with the universe of birth records from 1984-2016 (N=407,697) and 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). Based on these restrictions, our main analysis sample contains 2,726 children who weigh 1,300-1,700 grams at birth.

4.3 Outcomes for the Birth Records Sample

We join main birth records sample to several administrative data sources to measure outcomes in childhood and early adulthood. As detailed in the following sections, each outcome is only defined for a subset of cohorts in the birth records. This is because data sources for outcomes do not perfectly overlap with the years covered by the birth records. As a result, the sample sizes vary for each outcome that we consider. Note that the samples

[9] We have fewer children in the analysis for days spent in the NICU because this variable is only available for the years 2006-2015.

for many of our main outcomes (i.e., test scores, achievement and social program expenditures) have a similar size to those used in prior studies examining the effects of VLBW.[10]

4.3.1 Supplemental Security Income (SSI) Program Participation

Records from the Rhode Island Department of Human Services (RIDHS) 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 is determined using birth weight or a "failure to thrive" benchmark (SSA 2018a; 2018b).[11] We measure whether a child enrolled in SSI at any point from birth up to age 3.

4.3.2 Grade Repetition, Special Education Services (IEP) and Test Scores

Records from the Rhode Island Department of Education (RIDE) provide information on several key educational outcomes. We use records from the years 2003-2016 to study whether children repeat a grade or have a written Individualized Education Program (IEP). We study grade repetition during grades 1-4.[12] Similarly, we study the number of school years that a child has an IEP during the same grade range.[13] We also use RIDE records from the years 2005-2014 to study performance on elementary and middle school tests. Specifically, we study test scores in grades 3, 5, and 8, in addition to conducting

---

[10] To better understand our prospective ability to detect impacts of crossing the 1,500-gram threshold, it is useful to compare our samples sizes to those from Bharadwaj et al. (2013). As we detail below, we pool standardized test scores resulting in a sample of 3,070 person-year observations. For their analysis of test scores in Chile and in Norway, Bharadwaj et al. (2013) use a sample of 2,877 children in Chile and 1,163 in Norway (see Column 2 of Table 3 of their paper).

[11] Specifically, the SSA uses two rules to define whether a child has sufficiently low birth weight to qualify for SSI (SSA 2018a). The first rule defines a child as being eligible based on whether a child has birth weight less than 1,200 grams. The second rule defines a child as being eligible based on birth weight and gestational age. Evidence presented in Guldi et al. (2018) suggests that low birth weight children who are eligible for SSI may not participate in the program.

[12] A child is indicated as repeating a grade if they are observed in the same grade in different school years.

[13] Note that an IEP can be given in preschool, and children are assessed each year.

a pooled analysis using all test scores for grades 3-8. Note that all RIDE outcomes are available only for children who enroll in public schools in Rhode Island. As a result, children in the birth records sample who never enroll in public schools are not included in the analysis of grade repetition, IEP outcomes, or academic achievement.[14]

4.3.3 High School Disciplinary Offenses and College Preparation

Records from RIDE provide information on disciplinary offenses and college preparation during high school. For the former, we use records from 2003-2016 to study the number of offenses during grades 9-12.[15] For the latter, we use SAT, PSAT and AP records from the years 2010-2015. We use this data to construct an index measure 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.[16] As with our other RIDE outcomes (e.g., test scores in grades 3-8), the disciplinary offense and college preparation outcomes are only defined for observed for children who enroll in public schools and are sufficiently old to be observed in high school.

4.3.4 Post-Secondary Enrollment

[14] One potential concern is that enrollment in public schools may change at the threshold for VLBW status. This would change the composition of the sample of children on either side of the cutoff. The analysis that we conduct in Section 5.4 addresses this concern by estimating threshold crossing effects on the likelihood of enrolling in public school. We find generally do not find any significant effects.

[15] There are 44 distinct types of offenses recorded in RIDE records. These include the following: alcohol, arson, assault/battery of another student, assault/battery of a teacher, skipping class, skipping detention, skipping school, leaving school grounds, tardiness, truant, making bomb threats, cheating/plagiarism, prohibited use of communication devices, use of controlled substance, disorderly conduct, fighting, fire regulation violation, forgery, gambling, gang activity, harassment-stalking, harassment-sexual, harassment-prejudice, hazing, insubordination, kidnapping, larceny, obscene language toward student, obscene language toward teacher, robbery, sexual assault, sexual misconduct, unauthorized use of computers, threat, tobacco, trespassing, vandalism, weapon possession.

[16] We define the index for children who are enrolled in a Rhode Island public school for at least one high school grade

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. Unlike our other measures of academic achievement, we can measure college-going outcomes for children regardless of whether they enroll in Rhode Island public schools.

4.3.5 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). Anonymized program enrollment and payment records are available from 1997-2016.[17] In addition, we include program costs for special education participation (i.e., having a written IEP), which we set to an average expenditure of $34,000.[18] We construct measures of expenditure received during childhood by adding up program payments received and IEP costs incurred by age 10, 12, or 14. Note that we use the per-capita payment value for family-level program payments.

4.3.6 Maternal Care and Stress

[17] 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 some payment records for the first years of their life. We address censoring issues by imputing program payments for children born before 2004 using the annual average payments during the years we observe them. 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.

[18] 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).

PRAMS survey data provide information on maternal care and stress outcomes.[19] This survey collects information on maternal attitudes and experiences before, during, and after pregnancy. All mothers with low birth weight children (less than 2,500 grams) 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 to separate questions on stressful events, postpartum depression, and degree of child difficulty. The stress index is available from 2004-2014. Note that each index is set to missing if a respondent fails to answer all the questions associated with the measure.

4.3.7 Maternal Labor Market Activity and Subsequent Fertility

We also study maternal labor market outcomes and subsequent fertility. For the former, we use records from the Rhode Island Department of Labor and Training covering the period 1991-2016. The outcomes that we measure are employment and average earnings during the first two years after birth. For the latter, we use Rhode Island birth records from 1984-2016 to measure subsequent maternal fertility. Specifically, we create a measure of whether a mother has had an additional child within three years following the birth of a child in our analysis sample.

4.3.8 Child Medicaid Enrollment and Expenditures by Age 2

Records from the RIDHS allow us to measure a child's 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

[19] Note that the birth records from Rhode Island contain information on mothers. As mentioned, researchers at RIPL created anonymous identifiers that can be used to link birth records with PRAMS respondents.

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.

4.3.9 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).[20] We use the school-level mean of the resulting residuals as the measure of value-added.

4.4 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.[21,22] Finally, children in the RD sample come from households with greater rates of pre-birth

[20] Details on the specifications are provided in the notes for Appendix Table A.16.

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

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

enrollment in social assistance programs. Household earnings (based on UI records in Rhode Island) are also lower in the RD sample.[23]

5. Results

5.1 Evidence on Medical Spending and Health 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.

Table 1 replicates the analysis of medical treatments for VLBW children using the HCUP samples described in Section 4. Specifically, 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 effects on total days spent in the hospital and charges are not statistically significant, but they are positive. While the point estimates are sometimes imprecisely estimated, our results are broadly consistent with results from prior studies that find evidence of an increase in the intensity of medical care increases for children born just under the 1,500-gram threshold (Almond et al. 2010; Bharadwaj et al. 2013; Daysal et al. 2019).

5.2 Evidence on Identifying Assumptions

Our research design relies on the assumption that all pre-birth factors vary continuously with birth weight across the 1,500-gram threshold for VLBW status. Appendix Table A.2 shows results from testing the validity of this assumption where we estimate

[23] 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.

threshold crossing effects for various demographic or pre-birth characteristics on the independent variables from Equation 3. Column 2 reports estimates for the coefficient of the VLBW indicator. The results generally 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. Panel B of 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).

5.3 Main Results for Child Outcomes

Tables 2-4 report regression results based on Equation 3. To supplement this analysis, Figure 1 and Appendix Figures A.2-A.4 plot means of outcome variables within 20-gram bins of birth weight. For all results, we use the sample of 1,300-1,700-gram births (excluding observations within three grams of the 1,500-gram threshold).

5.3.1 Development and Childhood Education

Table 2 reports results from Equation 3 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 1,500-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.[24]

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.

5.3.2 High School and Post-Secondary Education

Table 3 shows that crossing the threshold for VLBW status has significant and beneficial impacts on several outcomes measured in high school and early adulthood. For

[24] Appendix Table A.3 provides additional analysis of IEP outcomes. This analysis estimates impacts on an indicator for whether the child ever has an IEP for a learning disability or for a non-learning disability. We focus on IEPs associated with learning disabilities because this is the largest category of IEP that we observe in our sample of children who weigh 1,300-1,700 grams at birth. We find that crossing the 1,500-gram threshold from above reduces the likelihood that a child has an IEP due to a learning disability, and this result is statistically significant at the 10 percent level.

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 3 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.

5.3.3 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 5.1, 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 4, 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 4 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.4 provides corresponding visual results for the impact of threshold crossing for the age 10 measures. Panel A in Table 4 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 1,500-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.[25]

5.4 Robustness Tests and Attrition Analysis

We conduct several robustness tests to test whether our results are sensitive to different choices for the RD analysis. Appendix Tables A.5-A.7 report the results associated with our main test, which is focused on examining sensitivity for all outcomes by using quadratic RD specifications and changing the bandwidth for births included in analysis. We use the optimal bandwidth selection method from Calonico et al. (2014) for each outcome.

[25] Appendix Table A.4 reports threshold crossing impacts for participation and expenditures for each program used to construct the measure of total expenditures up to age 14. For each program, the estimated impacts of crossing the 1,500-gram threshold are generally not statistically significant, but the point estimates are consistently negative.

Our results are generally robust to these changes: coefficient signs and magnitudes remain similar, although estimates sometimes lose statistical significance.[26]

Next, we conduct three additional types of robustness tests. First, we investigate the sensitivity of our results to different choices in clustering in Appendix Table A.9. These results show that clustering at higher-levels (e.g., clustering all observations within 10-gram bins) generally does not affect the statistical significance of our results. Second, Appendix Table A.10 tests for discontinuities in outcomes at other intervals of 100 grams. A pattern of significant threshold crossing effects at 100-gram intervals of birth weight that exceed 1,500 grams would generate concern that heaping is driving our main results. The top row of Appendix Table A.10 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. Third, Appendix Table A.11 reports results after varying the window used to exclude observations around the 1,500-gram threshold and including additional fixed effects to further address concerns over non-random heaping. The second and third rows of Appendix Table A.11 show that we find similar results when excluding observations that are within 5 and 10 grams of the threshold.[27] The last row reports estimates based on an augmented version of Equation 3 that includes fixed effects for observations at each 10-gram interval of birth weight. The

[26] Appendix Table A.8 further tests for sensitivity based on the choice of bandwidth. Specifically, we provide additional results based using a sample of births from 1,200-1,800 grams and 1,100-1,900 grams. These results are similar to the results from our preferred bandwidth.

[27] One result in Appendix Table A.11 differs from our main estimates. In Column 5, we see that the threshold crossing estimate for total social expenditures by age 14 becomes slightly less negative and is no longer statistically significant when we exclude the observations that are within 10-grams of the 1,500-gram threshold.

point estimates remain statistically significant and are generally similar in magnitude when we include 10-gram fixed effects.[28,29]

Finally, Appendix Table A.13 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 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.

5.5 Heterogeneity

Appendix Table A.14 reports results for education and social expenditure outcomes after dividing the RD sample based on child sex (as reported in birth records). Panels A and B compare results for girls and boys, respectively. In most cases, the subgroup-specific effects are statistically significant, and the point estimates are similar to those from Table 2. We cannot reject the hypothesis that estimates for each outcome are equal for girls and boys.

---

[28] An additional robustness test assesses whether we obtain similar impacts when we focus only on the cohorts for the test score analysis that best correspond to the period for which we have HCUP data from Rhode Island. Appendix Table A.12 reports results from this analysis. The estimates are similar to those in Table 2.
[29] Another potential concern is that many of our findings could be an artifact of multiple hypothesis testing. To assess this, we have followed the recommended practice of adjusting per comparison p-values (Anderson, 2008). Specifically, we use the two-step procedure from Benjamini et al. (2006) to calculate "q-values" that control for the false discovery rate (FDR), which is the proportion of rejections that are false positives (Type I errors). When we compute the adjusted p-values, we find that our results for test scores, disciplinary offenses, college enrollment and total social expenditures are still significant at the 10-percent level after adjusting for the fact that we analyzed multiple outcomes in Tables 2, 3 and 4. Note that we compute adjusted p-values accounting for the fact that we estimate effects on the following 10 outcomes: SSI enrollment at age 0, SSI enrollment at ages 0-3, IEP participation in grades 1-4, repetition in grades 1-4, test scores in grades 3-8, disciplinary offenses, the college preparation index, enrollment in any college, enrollment in a 4-year college, and total social expenditures by age 14.

6. Mechanisms

As discussed in Section 3, 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. 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.[30]

Table 5 shows threshold crossing impacts on direct measures of parental responses and other factors that could plausibly mediate the impacts in our sample.[31] We first study indices for maternal care and stress based on PRAMS responses.[32] 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

[30] Note that the results from Datar et al. (2010) may be less relevant for understanding effects in our context since they focus on comparing children born with higher birth weight. In contrast, the analysis by Bharadwaj et la. (2013) may be more relevant since they also focus on VLBW children.

[31] Appendix Figure A.5 plot means of the early childhood investment outcomes within 20-gram bins of birth weight. Appendix Table A.15 provides results from robustness tests to test whether the findings from Table 5 are sensitive to different choices for the RD analysis.

[32] While all mothers with low birth weight (less than 2,500 grams) receive the PRAMS survey, only a fraction respond. The PRAMS response rate in our main analysis sample of children who weigh 1,300-1,700 grams at birth is 70 percent. Note that we find no evidence that crossing the 1,500-gram threshold has any impact on the likelihood of participating in the PRAMS survey.

anxiety. The results show that we find statistically insignificant impacts, which is perhaps not surprising given the relatively small sample sizes available for the PRAMS measures. That said, 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 5 shows no significant impacts on employment or mean annual earnings measured during the first two years postpartum. The standard errors are sufficiently large that the confidence intervals include substantially large positive and negative effects.

As a measure of early-life child health, we study Medicaid enrollment and expenditures from 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 care at birth should reduce post-birth medical needs. In Table 5, 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 5 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).[33]

[33] Appendix Table A.15 shows the point estimate is positive but smaller when we use the optimal bandwidth selected using the method from Calonico et al. (2014).

Finally, Appendix Table A.16 tests whether parents of children born just under 1,500 grams are more likely to send their children to higher-quality schools.[34] 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.

7 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 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.

[34] Appendix Table A.17 provides results from robustness tests to test whether the findings from Table A.16 are sensitive to different choices for the RD analysis.

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? Revisiting the Effect of Very Low Birth Weight Classification." *Quarterly Journal of Economics* 126 (4): 2117-2123.

Benjamini, Yoav, Abba M. Krieger, and Daniel Yekutieli. 2006. “Adaptive Linear Step-up Procedures That Control the False Discovery Rate.” *Biometrika* 93 (3): 491–507.

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.

Daysal, N. Meltem, Mircea Trandafir, Marianne Simonsen, and Sanni Breining. 2019. "Spillover Effects of Early-Life Medical Interventions." Manuscript.

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.

Guldi, Melanie, Amelia Hawkins, Jeffrey Hemmeter, and Lucie Schmidt. 2018. “Supplemental Security Income and Child Outcomes: Evidence from Birth Weight Eligibility Cutoffs. NBER Working Paper #24913. National Bureau of Economic Research.

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.

Anderson, Michael L. 2008. "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American Statistical Association* 103 (484): 1481–95.

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. 2013. “Going to a Better School: Effects and Behavioral Responses.” *American Economic Review* 103 (4): 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 on Hospitalization Outcomes (HCUP Analysis)

	Mean for BW > 1,500 grams (1)	RD Est. BW < 1,500 grams (2)	Observations (3)	Birth Cohort (4)
Days in the NICU	9.925	3.427** (1.508)	1,245	2006-2015
Length of Stay	24.25	0.192 (1.818)	1,724	2002-2015
Total Charges ($)	116,965	3,470.24 (10,793.38)	1,724	2002-2015

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 3. Standard errors clustered at the gram level are presented in parentheses. Colunns 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-gram threshold). The HCUP data for Rhode Island is available from 2002-2015. Statistical significance is denoted by ***p < 0.01; **p <0.05; *p <0.10.

Table 2: 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 (0.015)	1,790	1996-2015
SSI Ages 0 - 3	0.025	-0.024 (0.019)	1,577	1996-2012
Years on IEP (1-4)	1.068	-0.205 (0.259)	605	1996-2006
Grade Repetition (1-4)	0.105	-0.131*** (0.039)	608	1996-2006
Avg. Std. Test Score				
3rd Grade	-0.284	0.444*** (0.162)	544	1996-2005
5th Grade	-0.244	0.299* (0.159)	527	1994-2003
8th Grade	-0.176	0.314* (0.165)	485	1990-2000
All (3-8)	-0.220	0.338*** (0.128)	3,070	1990-2005

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 3. 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-gram threshold). Statistical significance is denoted by ***$p<0.01$; **$p<0.05$; *$p<0.10$.

Table 3: 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 3. 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-gram threshold). Statistical significance is denoted by ***p<0.01; **p<0.05; *p<0.10.

Table 4: 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 (0.071)	951	1997-2006
By Age 12	0.587	-0.077 (0.079)	740	1997-2004
By Age 14	0.588	-0.117 (0.089)	529	1997-2002
Panel B: Total Expenditures ($)				
By Age 10	91,296	-27,291* (15,736)	951	1997-2006
By Age 12	113,844	-44,067* (25,034)	740	1997-2004
By Age 14	145,486	-66,997** (33,622)	529	1997-2002
Panel C: Total Expenditures ($) | *Any Expenditures (=1)*				
By Age 10	150,334	-29,149 (22,832)	572	1997-2006
By Age 12	194,052	-53,444 (33,486)	434	1997-2004
By Age 14	247,515	-74,075* (37,596)	310	1997-2002

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 3. 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-gram threshold). Statistical significance is denoted by ***p<0.01; **p<0.05; *p<0.10.

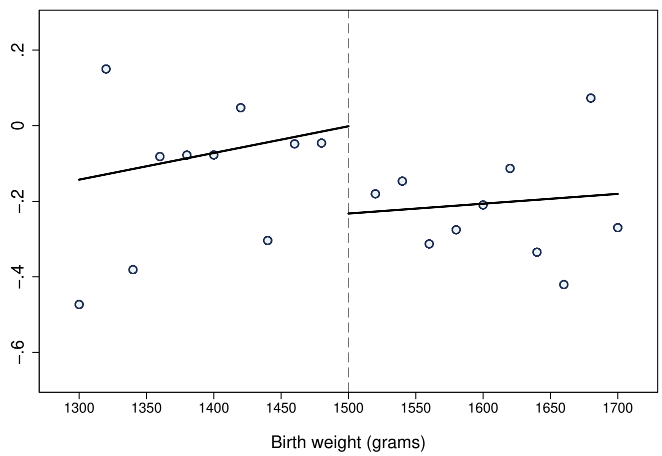
Table 5: Impacts for Early Childhood Investment Outcomes

	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)		

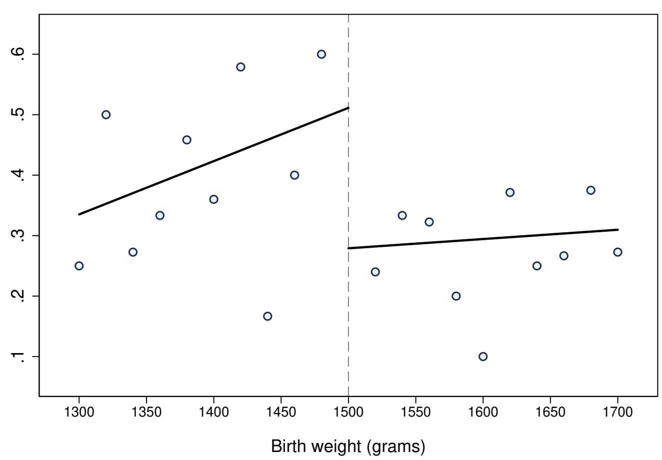
Notes: Column 1 reports the mean of the dependent variable for children born above 1,500 grams. Column 2 provides estimates of the impact of crossing the 1,500-gram threshold using Equation 3. Standard errors clustered at the gram level are presented in parentheses. Columns 3 and 4 report the total number 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-gram 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.

Online Appendix Materials

Table A.1: Descriptive Statistics for Births in Rhode Island

Variable		All Births (1)	Births Between 1,300 - 1,700 g (2)	p-value (3)
Child Demographics	White (=1)	0.715	0.665	0.000
		(0.451)	(0.472)	
	Black (=1)	0.090	0.129	0.000
		(0.286)	(0.335)	
	Hispanic (=1)	0.129	0.149	0.005
		(0.335)	(0.356)	
	Asian (=1)	0.024	0.024	0.958
		(0.153)	(0.152)	
	Other (=1)	0.042	0.043	0.040
		(0.201)	(0.180)	
	Gestational Age (Weeks)	38.696	31.819	0.000
		(3.145)	(3.682)	
Maternal Characteristics	Age	28.261	28.920	0.000
		(6.328)	(7.228)	
	Married (=1)	0.655	0.563	0.000
		(0.475)	(0.496)	
	Less than High School (=1)	0.117	0.140	0.000
		(0.322)	(0.347)	
	High School (=1)	0.216	0.199	0.036
		(0.412)	(0.399)	
	College or Higher (=1)	0.332	0.275	0.000
		(0.471)	(0.446)	
	Smokes (=1)	0.168	0.229	0.000
		(0.374)	(0.420)	
	Drinks (=1)	0.024	0.027	0.477
		(0.152)	(0.161)	
Economic Self-Sufficiency (Prior to Birth)	Medicaid (=1)	0.349	0.406	0.011
		(0.477)	(0.492)	
	SNAP (=1)	0.205	0.261	0.003
		(0.404)	(0.440)	
	TANF (=1)	0.073	0.101	0.019
		(0.260)	(0.301)	
	HH Earnings ($) in RI, Annual	36,784	34,288	0.014
		(46,802)	(47,698)	
	Observations	407,697	2,726	

Notes: All self-sufficiency measures (Medicaid, SNAP, and TANF) are defined for the mother using the four quarters prior to the child's birth. Annual household earnings are defined as the mother's and father's earnings. A father is only observed if the mother is married at the time of birth. The reported p-value in column (3) is based on the difference in means between births in the 1,300-1,700 gram range and births outside the 1,300-1,700 gram range.

Table A.2: Baseline Balance Test

Variable		Mean for BW > 1,500 grams (1)	RD Est. BW < 1,500 grams (2)	Observations (3)
Child Demographics	White (=1)	0.673	-0.022	2,237
			(0.052)	
	Black (=1)	0.137	0.024	2,237
			(0.028)	
	Hispanic (=1)	0.134	0.001	2,237
			(0.047)	
	Asian (=1)	0.023	0.001	2,327
			(0.017)	
	Other (=1)	0.033	-0.004	2,237
			(0.019)	
	Gestational Age (Weeks)	32.381	-0.131	2,263
			(0.471)	
Maternal Characteristics	Age	28.75	0.453	2,399
			(0.617)	
	Married (=1)	0.559	0.046	2,722
			(0.051)	
	Less than High School (=1)	0.037	-0.006	2,726
			(0.017)	
	High School (=1)	0.198	-0.001	2,726
			(0.065)	
	College or Higher (=1)	0.157	0.055	2,726
			(0.056)	
	Smokes (=1)	0.245	-0.090*	1,170
			(0.049)	
	Drinks (=1)	0.031	-0.007	1,166
			(0.019)	
Economic Self-Sufficiency (Pre-Birth)	Medicaid (=1)	0.381	0.000	448
			(0.101)	
	SNAP (=1)	0.230	0.042	448
			(0.101)	
	TANF (=1)	0.084	-0.013	477
			(0.064)	
	HH Earnings ($) in RI, Annual	34,043	6,658.696	2,077
			(4,329.335)	
Joint Test p-value			0.932	

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 3. Standard errors clustered at the gram level are presented in parentheses. Column 3 reports the total number of observations 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-gram threshold). All self-sufficiency measures (Medicaid, SNAP, and TANF) are defined for the mother using the four quarters prior to the child's birth. Annual household earnings are defined as the sum of the mother's and father's earnings. A father is only observed if the mother is married at the time of birth. Statistical significance is denoted by ***p<0.01; **p<0.05; *p<0.10.

Table A.3: Additional Results for IEP Categories

	Mean for BW > 1,500 grams (1)	RD Est. BW < 1,500 grams (2)	Observations (3)	Birth Cohort (4)
Panel A: Ever have IEP an IEP for...				
Learning Disability	0.126	-0.109*	468	1999-2006
		(0.058)		
Non-Learning Disability	0.142	0.014	468	1999-2006
		(0.081)		
Panel B: Ever have an IEP for... | Years on IEP 1-4 > 0				
Learning Disability	0.393	-0.218	154	1999-2006
		(0.133)		
Non-Learning Disability	0.440	0.127	154	1999-2006
		(0.162)		

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 3. 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-gram threshold). Information on primary disability categories for IEP come from Special Education records (2009-2015). The dependent variable "Ever IEP for Learning Disability" is an indicator equal to one if the child has a learning disability in grades 1-4. Non-learning disability categories include 13 additional classifcations of disability that include condtions such as autism, blindness, or speech impairment. Statistical significance is denoted by ***p<0.01; **p<0.05; *p<0.10.

Table A.4: Additional Results for Social Expenditure Categories

	Mean for BW > 1,500 grams (1)	RD Est. BW < 1,500 grams (2)	Observations (3)	Birth Cohort (4)
Panel A: Any Expenditures by Age 14 (= 1)				
SNAP	0.344	-0.114 (0.092)	529	1997-2002
TANF	0.244	-0.071 (0.082)	529	1997-2002
Medicaid Payments	0.469	-0.135* (0.080)	529	1997-2002
Medicaid Pharmacy Payments	0.439	-0.163* (0.095)	529	1997-2002
IEP Cost	0.359	-0.079 (0.096)	529	1997-2002
Panel B: Total Expenditures by Age 14 ($)				
SNAP	3,926	-2,000* (1,192)	529	1997-2002
TANF	2,276	-1,423 (899)	529	1997-2002
Medicaid Payments	12,638	-3,871 (5,744)	529	1997-2002
Medicaid Pharmacy Payments	1,439	-639 (633)	529	1997-2002
IEP Cost	123,850	-54,718 (33,415)	529	1997-2002

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 3. 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-gram threshold). Statistical significance is denoted by ***$p<0.01$; **$p<0.05$; *$p<0.10$.

Table A.5: Robustness Check - Impacts for Development and Education Outcomes

	Linear RD Model			Quadratic RD Model			
	RD Est. BW < 1,500 grams	Observations	Bandwidth	RD Est. BW < 1,500 grams	Observations	Bandwidth	Birth Cohort
Variable	(1)	(2)	(3)	(1)	(2)	(3)	
SSI, IEP, and Grade Rep.							
SSI Age 0	0.004	1,111	(1372-1628)	0.006	2,018	(1272-1728)	1996-2015
	(0.019)			(0.022)			
SSI Ages 0 - 3	-0.009	972	(1370-1630)	-0.006	1,683	(1285-1715)	1996-2012
	(0.025)			(0.029)			
Years on IEP (1-4)	-0.256	636	(1291-1709)	-0.298	1,169	(1144-1856)	1996-2006
	(0.248)			(0.294)			
Grade Repetition (1-4)	-0.228***	376	(1370-1630)	-0.214***	1,076	(1170-1830)	1996-2006
	(0.044)			(0.043)			
Avg. Std. Test Score							
3rd Grade	0.438***	677	(1262-1738)	0.266	689	(1252-1748)	1996-2005
	(0.139)			(0.209)			
5th Grade	0.342**	642	(1263-1737)	0.347**	1,218	(1093-1907)	1994-2003
	(0.139)			(0.164)			
8th Grade	0.358**	515	(1296-1704)	0.402**	830	(1170-1830)	1990-2000
	(0.160)			(0.199)			
All (3-8)	0.371***	2,818	(1316-1684)	0.391***	6,459	(1126-1874)	1990-2005
	(0.127)			(0.140)			

Notes: Table presents RD analysis using the optimal bandwidth selection procedure from Calonico et al. (2014). We also follow Calonico et al. (2014) to calculate robust standard errors. Statistical significance is denoted by ***$p<0.01$; **$p<0.05$; *$p<0.10$.

Table A.6: Robustness Check - Impacts for High School and Higher Education Outcomes

	Linear RD Model			Quadratic RD Model			
	RD Est. BW < 1,500 grams	Observations	Bandwidth	RD Est. BW < 1,500 grams	Observations	Bandwidth	Birth Cohort
Variable	(1)	(2)	(3)	(1)	(2)	(3)	
Disciplinary Off. (9-12)	-0.081	419	(1292-1708)	-0.133**	958	(1097-1903)	1988-1998
	(0.059)			(0.063)			
College Prep. Index	0.185	714	(1223-1777)	-0.102	947	(1149-1851)	1993-2001
	(0.309)			(0.419)			
Any College Enroll. by 22	0.172***	596	(1220-1780)	0.121	596	(1220-1780)	1984-1994
	(0.059)			(0.077)			
4-Year College Enroll. by 22	0.217***	655	(1204-1796)	0.235***	1,094	(1062-1938)	1984-1994
	(0.069)			(0.084)			

Notes: Table presents RD analysis using the optimal bandwidth selection procedure from Calonico et al. (2014). We also follow Calonico et al. (2014) to calculate robust standard errors. Statistical significance is denoted by ***$p<0.01$; **$p<0.05$; *$p<0.10$.

Table A.7: Robustness Check - Impacts for Social Program Expenditure by Age

	Linear RD Model			Quadratic RD Model			
	RD Est. BW < 1,500 grams	Observations	Bandwidth	RD Est. BW < 1,500 grams	Observations	Bandwidth	Birth Cohort
Variable	(1)	(2)	(3)	(1)	(2)	(3)	
Panel A: Any Expenditures (=1)							
By Age 10	-0.054	1,602	(1185-1815)	-0.018	1,811	(1136-1864)	1997-2006
	(0.054)			(0.073)			
By Age 12	(-0.049)	1,693	(1081-1919)	-0.017	1,280	(1176-1824)	1997-2004
	(0.052)			(0.085)			
By Age 14	-0.052	1,430	(1029-1971)	-0.147*	1,303	(1072-1928)	1997-2002
	(0.059)			(0.081)			
Panel B: Total Expenditures ($)							
By Age 10	-17,103	788	(1331-1669)	-22,086	1,168	(1262-1738)	1997-2006
	(17,517)			(23,125)			
By Age 12	-43,253*	692	(1318-1682)	-46,623	1,051	(1225-1775)	1997-2004
	(25,316)			(33,777)			
By Age 14	-56,634*	525	(1304-1696)	-52,871	912	(1181-1819)	1997-2002
	(32,320)			(41,355)			
Panel C: Total Expenditures ($) | Any Expenditures (=1)							
By Age 10	-41,476	434	(1345-1655)	-43,883	814	(1234-1766)	1997-2006
	(27,021)			(32,588)			
By Age 12	-71,118*	428	(1304-1696)	-76,593*	891	(1130-1870)	1997-2004
	(37,207)			(39,632)			
By Age 14	-61,533*	287	(1310-1690)	-48,905	460	(1218-1782)	1997-2002
	(34,886)			(47,689)			

Notes: Table presents RD analysis using the optimal bandwidth selection procedure from Calonico et al. (2014). We also follow Calonico et al. (2014) to calculate robust standard errors. Statistical significance is denoted by ***p<0.01; **p<0.05; *p<0.10.

Table A.8: Robustness Check - Alternative Bandwidth Results for Selected Outcomes

	1,200-1,800 Grams			1,100-1,900 Grams			
	Mean for BW > 1,500 grams	RD Est. BW < 1,500 grams	Observations	Mean for BW > 1,500 grams	RD Est. BW < 1,500 grams	Observations	Birth Cohort
Variable	(1)	(2)	(3)	(1)	(2)	(3)	
3rd Grade Avg. Std. Test Score	-0.233	0.443***	876	-0.227	0.273**	1,264	1996-2004
		(0.138)			(0.121)		
Grade Repetition (1-4)	0.094	-.070**	971	0.091	-0.047*	1,395	1996-2004
		(0.033)			(0.028)		
Disciplinary Offenses (9-12)	0.132	-0.130**	642	0.123	-0.067	957	1984-2001
		(0.055)			(0.049)		
Any College Enrollment by Age 22	0.594	0.153*	663	0.584	0.088*	1,005	1985-1994
		(0.067)			(0.052)		
Total Social Expenditures by Age 14	125,546	-97,092***	849	123,958	-49,662**	1,207	1997-2002
		(26,004)			(25,276)		

Notes: Table reports threshold crossing estimates for alternative gram bandwidths. 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. Column 3 and 4 report the total number of observations and describe the birth cohorts included in the analysis. Statistical significance is denoted by ***p<0.01; **p<0.05; *p<0.10.

Table A.9: Robustness Check - Alternative Clustering Results for Selected Outcomes

	3rd Grade Avg. Std. Test Score (1)	Grade Repetition (1-4) (2)	Disciplinary Offenses (9-12) (3)	Any College Enrollment by Age 22 (4)	Total Social Expend. by Age 14 (5)
RD Est. BW < 1,500 grams	0.444***	0.131***	-0.131**	0.171**	-66,697.35**
SE Clustering at gram level	(0.162)	(0.039)	(0.060)	(0.067)	(33,622.24)
SE Clustering at 10-gram bins	(0.182)	(0.032)	(0.060)	(0.071)	(34,469.59)
SE Clustering at 20-gram bins	(0.155)	(0.034)	(0.061)	(0.069)	(31,202.33)
SE Clustering at 50-gram bins	(0.177)	(0.041)	(0.046)	(0.065)	(29,047,83)
Observations	544	608	393	416	529

Notes: The first row reports the point estimate of the impact of crossing the 1,500-gram threshold. The remaining rows report standard error estimates that are based on different choices for clustering. 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-gram threshold). Statistical significance in the first row is based on clustering at the gram-level and is denoted by ***$p<0.01$; **$p<0.05$; *$p<0.10$.

Table A.10: Examining Impacts on Cutoffs for Selected Outcomes (Cutoffs Between 1,500 and 3,000 grams)

Gram Threshold	3rd Grade Avg. Std. Test Score (1)	Grade Repetition (1-4) (2)	Discplinary Offenses (9-12) (3)	Any College Enrollment by Age 22 (4)	Total Social Expend. by Age 14 (5)
1500	0.444***	-0.131***	-0.131**	0.171**	-66,997.351**
	(0.162)	(0.039)	(0.061)	(0.067)	(33,622.243)
1600	-0.210	0.065**	-0.039	-0.148	-34,816.011
	(0.143)	(0.031)	(0.066)	(0.089)	(44,008.394)
1700	0.142	-0.032	0.118**	-0.140	46,498.115
	(0.200)	(0.040)	(0.054)	(0.089)	(57,363.462)
1800	0.123	0.011	-0.068*	0.139**	-43,018.037*
	(0.100)	(0.029)	(0.041)	(0.061)	(21,751.957)
1900	-0.027	0.012	0.018	-0.076	-8,245.931
	(0.079)	(0.031)	(0.054)	(0.060)	(16,621.581)
2000	-0.055	0.075***	-0.088**	-0.009	28,473.172**
	(0.080)	(0.024)	(0.037)	(0.050)	(11,894.619)
2100	-0.081	-0.009	0.032	0.058	-17,838.113
	(0.074)	(0.022)	(0.033)	(0.044)	(12,909.573)
2200	-0.091	-0.002	-0.039	-0.009	2,403.226
	(0.056)	(0.020)	(0.027)	(0.038)	(12,483.468)
2300	0.026	0.005	-0.027	0.030	-3,919.248
	(0.056)	(0.016)	(0.021)	(0.030)	(14,430.155)
2400	-0.034	0.036***	-0.025	-0.051***	9,279.01
	(0.046)	(0.013)	(0.023)	(0.019)	(11,273.191)
2500	0.060	-0.014	0.015	0.008	-11,117.321
	(0.041)	(0.012)	(0.018)	(0.013)	(11,044.726)
2600	-0.065**	0.017	-0.001	0.020	-2,313.72
	(0.033)	(0.012)	(0.015)	(0.016)	(7,264.108)
2700	0.041	-0.002	-0.024	-0.003	3,328.503
	(0.035)	(0.012)	(0.017)	(0.013)	(7,507.523)
2800	-0.007	-0.006	0.018	-0.013	-3,759.513
	(0.039)	(0.008)	(0.015)	(0.022)	(6,346.463)
2900	-0.027	-0.003	-0.007	0.017	13,237.808***
	(0.026)	(0.008)	(0.014)	(0.014)	(3,882.636)
3000	0.021	-0.011	-0.010	-0.013	-4,185.7
	(0.027)	(0.011)	(0.013)	(0.019)	(5,273.798)

Notes: Each row provides estimates of the impact of crossing a birth weight threshold defined at 100 gram intervals between 1,500 and 3,000 grams using Equation 3. The first row replicates the main estimates of the impact of crossing the 1,500-gram threshold (as reported in Tables 2, 3, and 4). Standard errors clustered at the gram level are presented in parentheses. The sample for all results includes children born with birth weight within 200 grams of the specified 100 gram threshold (excluding children born within 3 grams of the specified gram threshold). Statistical significance is denoted by ***$p<0.01$; **$p<0.05$; *$p<0.10$.

Table A.11: Additional Robustness Tests for Selected Outcomes

	3rd Grade Avg. Std. Test Score (1)	Grade Repetition (1-4) (2)	Disciplinary Offenses (9-12) (3)	Any College Enrollment by Age 22 (4)	Total Social Expend. by Age 14 (5)
Main Specification	0.444***	0.131***	-0.131**	0.171**	-66,697.35**
	(0.162)	(0.039)	(0.060)	(0.067)	(33,622.24)
	[N=544]	[N=608]	[N=393]	[N=416]	[N=529]
Exclude+/- 5g Near 1500g	0.425***	-0.130***	-0.131**	0.171**	-64,558.99**
	(0.162)	(0.039)	(0.060)	(0.067)	(33,707.64)
	[N=543]	[N=606]	[N=393]	[N=416]	[N=527]
Exclude +/- 10g Near 1500g	0.448***	0.125***	-0.137**	0.184**	-56,806.97
	(0.172)	(0.042)	(0.062)	(0.071)	(34,890.94)
	[N=535]	[N=594]	[N=390]	[N=409]	[N=521]
Main Specification + 10g FEs	0.516**	-0.142***	-0.166***	0.199**	-93,890.94**
	(0.209)	(0.049)	(0.058)	(0.078)	(39,746.48)
	[N=544]	[N=608]	[N=393]	[N=416]	[N=529]

Notes: Each row presents estimates of the impact of crossing the 1,500-gram threshold using different samples and specifications. Results for selected outcomes are provided in each column. The first row replicates the main estimates of the impact of crossing the 1,500-gram threshold (as reported in Tables 2, 3, and 4). The sample for the main estimates excludes children born within 3 grams of the 1,500-gram threshold. The second and third rows report threshold crossing impacts by excluding children born within 5 and 10 grams of the 1,500-gram threshold. The fourth row reports threshold crossing impacts based on an augmented version of Equation 3 that includes fixed effects for observations at 10 gram intervals of birth weight.

Table A.12: Impacts for Test Score Outcomes Restricted to 2002-2015 Births

	Mean for BW > 1,500 grams (1)	RD Est. BW < 1,500 grams (2)	Observations (3)	Birth Cohort (4)
Avg. Std. Test Score				
3rd Grade	0.153	0.476*	222	2002-2005
		(0.280)		
5th Grade	-0.065	0.172	101	2002-2003
		(0.389)		
All (3-5)	-0.116	0.295	483	2002-2005
		(0.285)		

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 3. 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-gram threshold). Average test scores for the birth cohorts born in the HCUP data range (2002-2015) are only available for grades 3 through 5. Statistical significance is denoted by ***p<0.01; **p<0.05; *p<0.10.

Table A.13: Attrition Analysis

	Mean for BW > 1,500 grams (1)	RD Est. BW < 1,500 grams (2)	Observations (3)	Birth Cohort (4)
Has 3rd Grade Test Score (=1)	0.604	-0.080 (0.068)	940	1996-2005
Has 5th Grade Test Score (=1)	0.621	-0.098 (0.076)	878	1994-2003
Has 8th Grade Test Score (=1)	0.567	-0.104* (0.055)	894	1990-2000
Has Any Test Score (3-8) (=1)	0.634	-0.077 (0.048)	1,366	1990-2005
Enrolled in Public High School (=1)	0.698	0.026 (0.063)	676	1986-2001

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 3. 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-gram threshold). Statistical significance is denoted by ***$p<0.01$; **$p<0.05$; *$p<0.10$.

Table A.14: Impacts for Selected Outcomes for Girls and Boys

	Mean for BW > 1,500 grams (1)	RD Est. BW < 1,500 grams (2)	Observations (3)	Birth Cohort (4)
Panel A: Girls				
3rd Grade Avg. Std. Test Score	-0.242	0.461*	285	1996-2004
		(0.263)		
Grade Repetition (1-4)	0.063	-0.144**	308	1996-2004
		(0.055)		
Disciplinary Offenses (9-12)	0.165	-0.162**	201	1988-1998
		(0.077)		
Any College Enrollment by Age 22	0.538	0.227*	203	1985-1994
		(0.123)		
Total Social Expenditures by Age 14	98,089	-71,332*	276	1997-2002
		45,286		
Panel B: Boys				
3rd Grade Avg. Std. Test Score	-0.327	0.460**	259	1996-2004
		(0.179)		
Grade Repetition (1-4)	0.145	-0.102*	300	1996-2004
		(0.061)		
Disciplinary Offenses (9-12)	0.194	-0.101	192	1988-1998
		(0.092)		
Any College Enrollment by Age 22	0.535	0.141	213	1985-1994
		(0.097)		
Total Social Expenditure by Age 14	197,433	-70,397*	253	1997-2002
		(41,434)		

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 3. Standard errors 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-gram threshold). Statistical significance is denoted by ***p<0.01; **p<0.05; *p<0.10.

Table A.15: Robustness Check - Impacts for Early Childhood Investment Outcomes

	Linear RD Model			Quadratic RD Model			
Variable	(1)	(2)	(3)	(1)	(2)	(3)	
Maternal Care Index	0.111	550	(1247-1753)	-0.004	671	(1194-1806)	2002-2014
	(0.150)			(0.185)			
Maternal Stress Index	-0.120	534	(1254-1746)	-0.079	770	(1160-1840)	2004-2014
	(0.324)			(0.403)			
Maternal Employment Ages 0-2 (=1)	-0.011	3,473	(1135-1865)	-0.011	6,418	(915-2085)	1991-2014
	(0.035)			(0.041)			
Avg. Maternal Earn. Ages 0-2 ($)	715.5	2,031	(1274-1726)	-635.5	3,709	(1120-1881)	1991-2014
	(2,042)			(2,436)			
Medicaid Enrollment Ages 0 - 2 (=1)	-0.026	3,774	(1223-1777)	-0.030	5,616	(1119-1881)	1989-2014
	(0.057)			(0.078)			
Medicaid Exp. Ages 0 - 2 ($)	-2,520	2,143	(1193-1807)	-6,636	2,386	(1168-1832)	1984-2016
	(5,972)			(8,397)			
Birth Within 3 Years (=1)	0.037	4,131	(1146-1854)	0.073**	4,130	(1146-1854)	1989-2014
	(0.023)			(0.034)			

Notes: Table presents RD analysis using the optimal bandwidth selection procedure from Calonico et al. (2014). We also follow Calonico et al. (2014) to calculate robust standard errors. Statistical significance is denoted by ***$p<0.01$; **$p<0.05$; *$p<0.10$.

Table A.16: Impacts for School Based Mechanisms

	Mean for BW > 1,500 grams (1)	RD Est. BW < 1,500 grams (2)	Observations (3)	Birth Cohort (4)
Value Added (Elementary)	-0.029	0.014	1,637	1993-2006
		(0.009)		
Value Added (Middle)	-0.029	0.013	1,489	1990-2002
		(0.010)		
Value Added (College Enrollment)	-0.022	-0.007	799	1984-2001
		(0.012)		

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 3. 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-gram threshold). We construct value-added as follows. For test scores in elementary and middle school, we regress average standardized test scores on lagged test scores (including their square and cube), indicators for a student's race, gender, IEP status, Limited English Proficiency status, and Free and Reduced Price Lunch status. For college enrollment, we regress an indicator of whether a student enrolled in any college by the age of 22 on eighth grade standardized test scores (averaged over math and reading), indicators for a student's race, gender, IEP status, and Free and Reduced Price Lunch status. Based on these regressions, we use the school-level mean of the resulting residuals. Statistical significance is denoted by ***$p<0.01$; **$p<0.05$; *$p<0.10$.

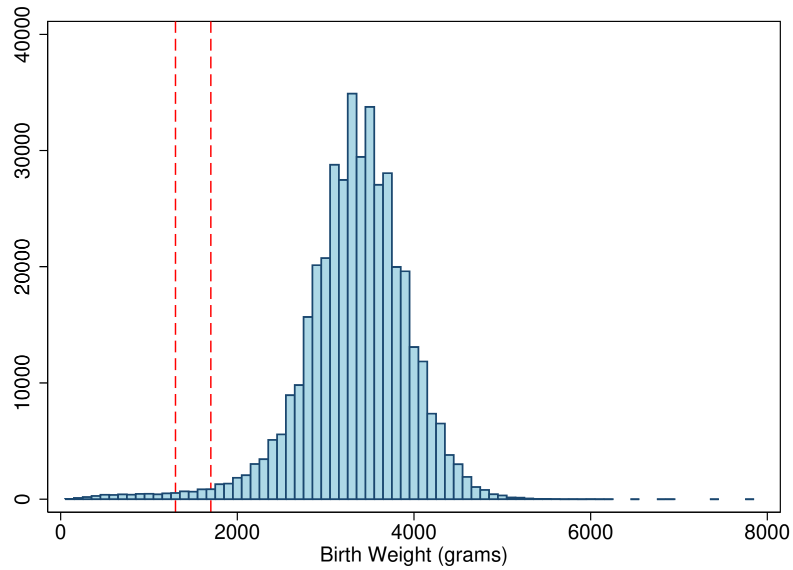
Table A.17: Robustness Check - Impacts for School Based Mechanisms

Variable	Linear RD Model: RD Est. BW < 1,500 grams (1)	Linear RD Model: Observations (2)	Linear RD Model: Bandwidth (3)	Quadratic RD Model: RD Est. BW < 1,500 grams (1)	Quadratic RD Model: Observations (2)	Quadratic RD Model: Bandwidth (3)	Birth Cohort
Value Added (Elementary)	0.013*	2,361	(1228-1772)	0.013	3,461	(1125-1875)	1993-2006
	(0.008)			(0.009)			
Value Added (Middle)	0.021*	1,380	(1314-1686)	0.026*	2,343	(1212-1788)	1990-2002
	(0.012)			(0.015)			
Value Added (College Enrollment)	0.000	1,140	(1225-1775)	-0.007	1,714	(1117-1883)	1984-2001
	(0.011)			(0.014)			

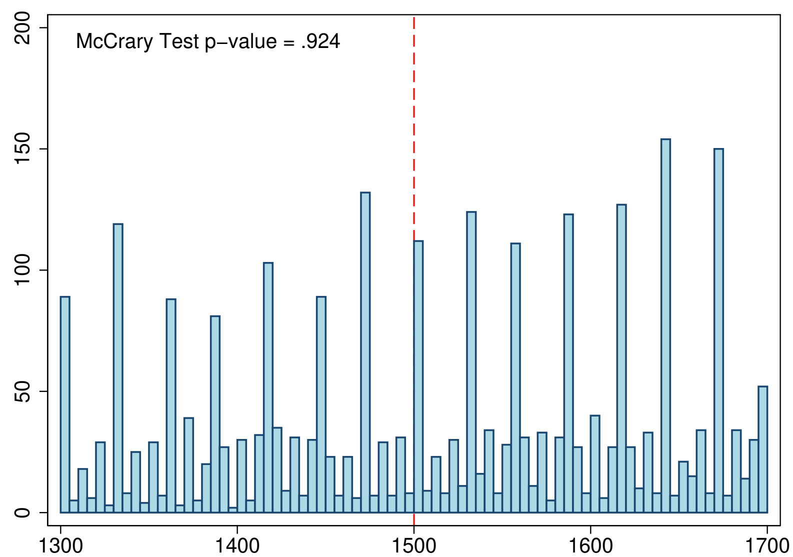
Notes: Table presents RD analysis using the optimal bandwidth selection procedure from Calonico et al. (2014). We also follow Calonico et al. (2014) to calculate robust standard errors. Statistical significance is denoted by ***p<0.01; **p<0.05; *p<0.10.

Figure A.1: Histograms of Birth Weight

Panel A: All Births

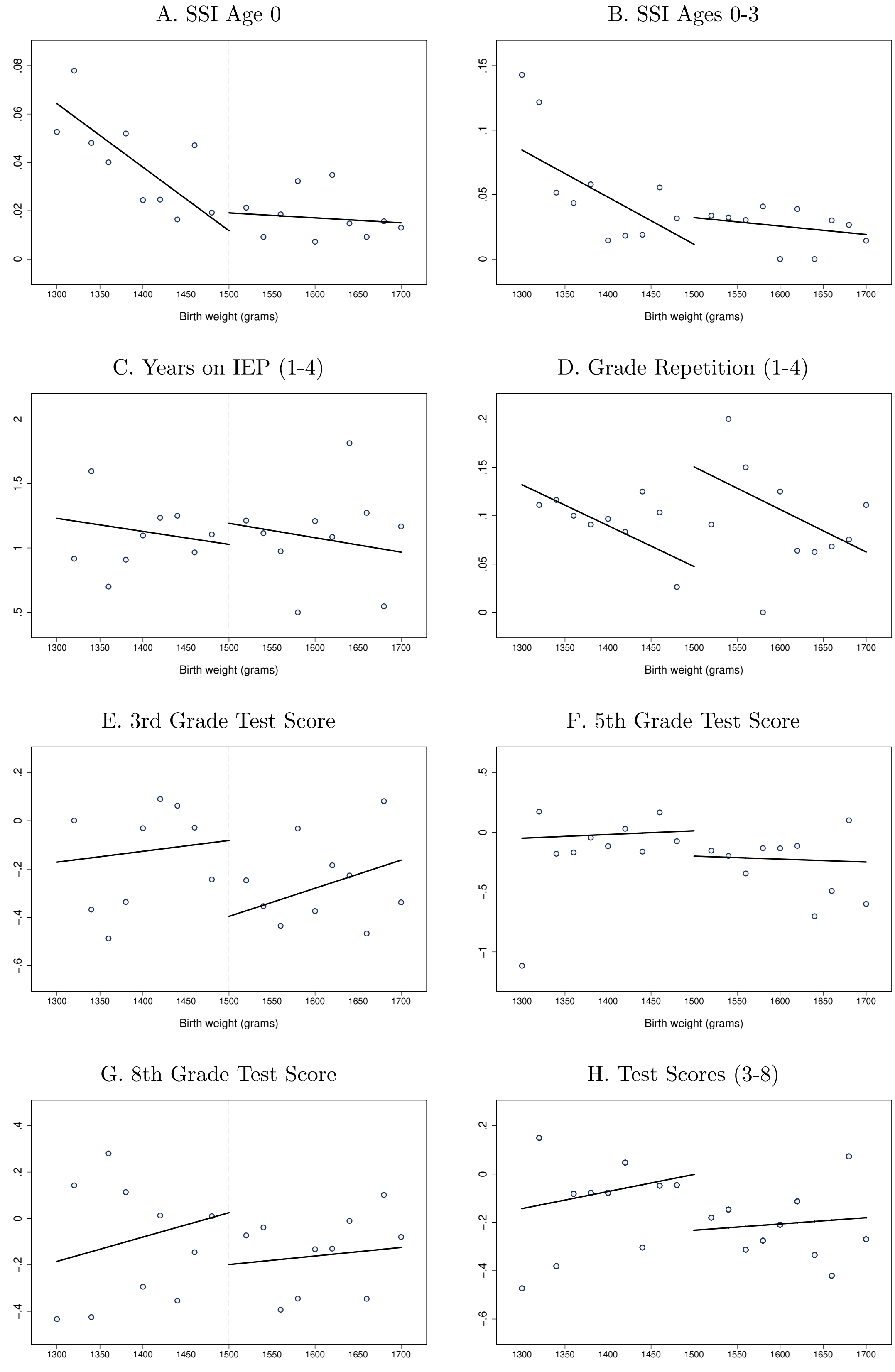


Panel B: Births 1,300-1,700 Grams



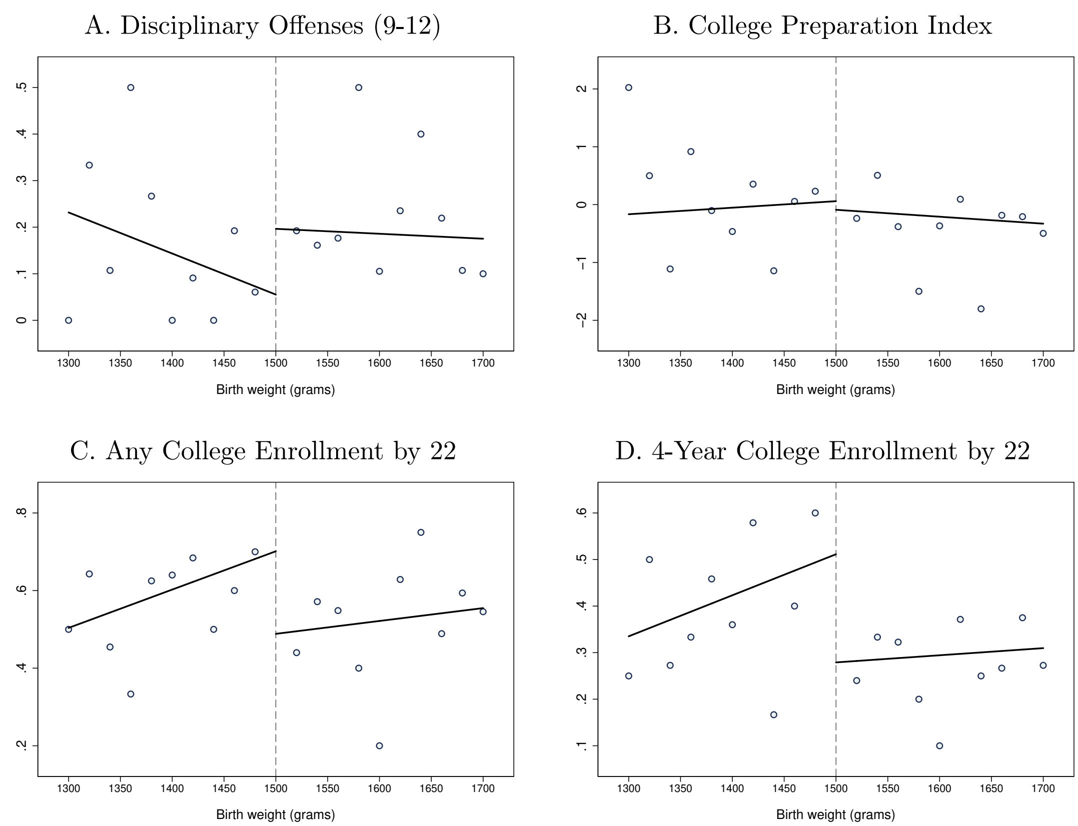
Notes: Panel A shows a histogram of the birth weight distribution for all births in Rhode Island (1984-2016). The bins have a width of 100 grams. Panel B shows a histogram of the birth weight distribution between 1,300 and 1,700 grams for all births in Rhode Island (1984-2016). The bins have a width of 5 grams. We conduct a McCrary (2008) test for a discontinuity in the density of the running variable at the 1,500 gram threshold. We fail to reject the null hypothesis of continuity with $p < 0.924$.

Figure A.2: Impacts for Development and Education Outcomes



Notes: Each panel shows the relationship between birth weight and various post-birth outcomes. Dots represents means within 20 gram bins of the running variable. The dark lines are predictions from a linear regression using the individual-level data.

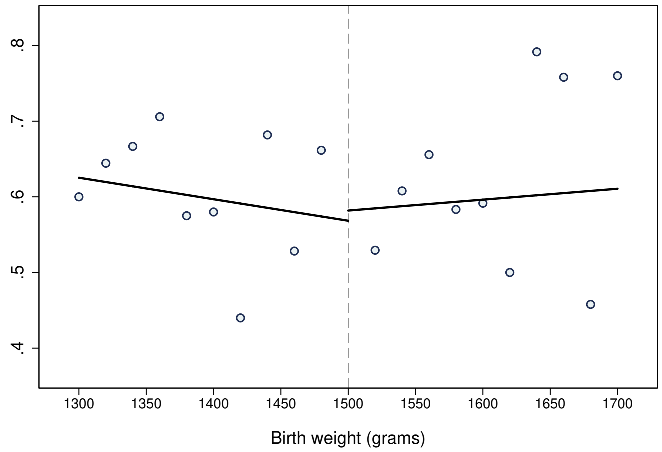
Figure A.3: Impacts for High School and Higher Education Outcomes



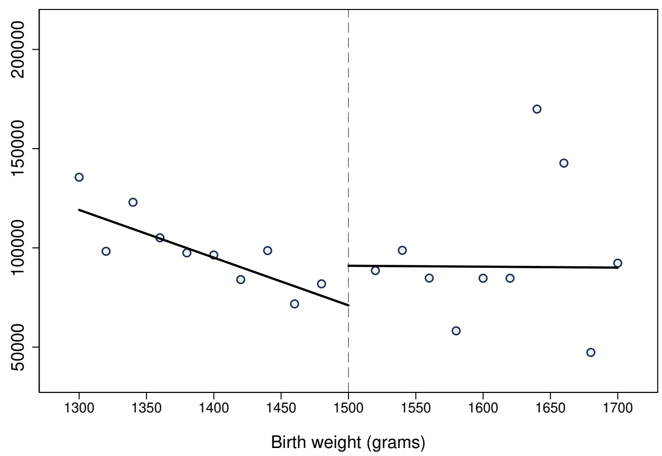
Notes: Each panel shows the relationship between birth weight and various post-birth outcomes. Dots represents means within 20 gram bins of the running variable. The dark lines are predictions from a linear regression using the individual-level data.

Figure A.4: Impacts for Social Program Expenditures by Age 10

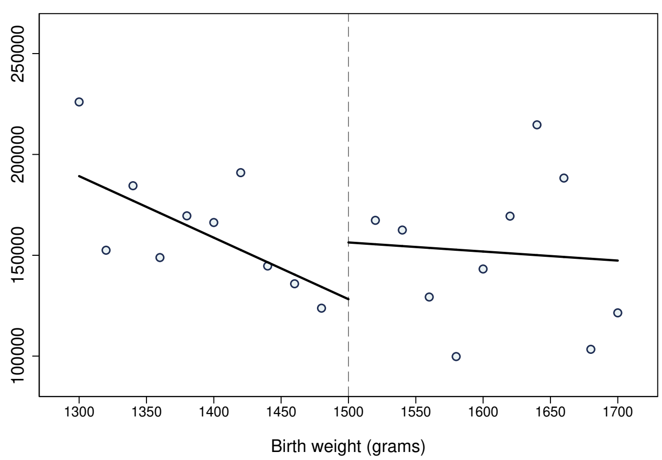
Panel A: Any Expenditures by Age 10 (=1)



Panel B: Total Expenditures by Age 10 ($)

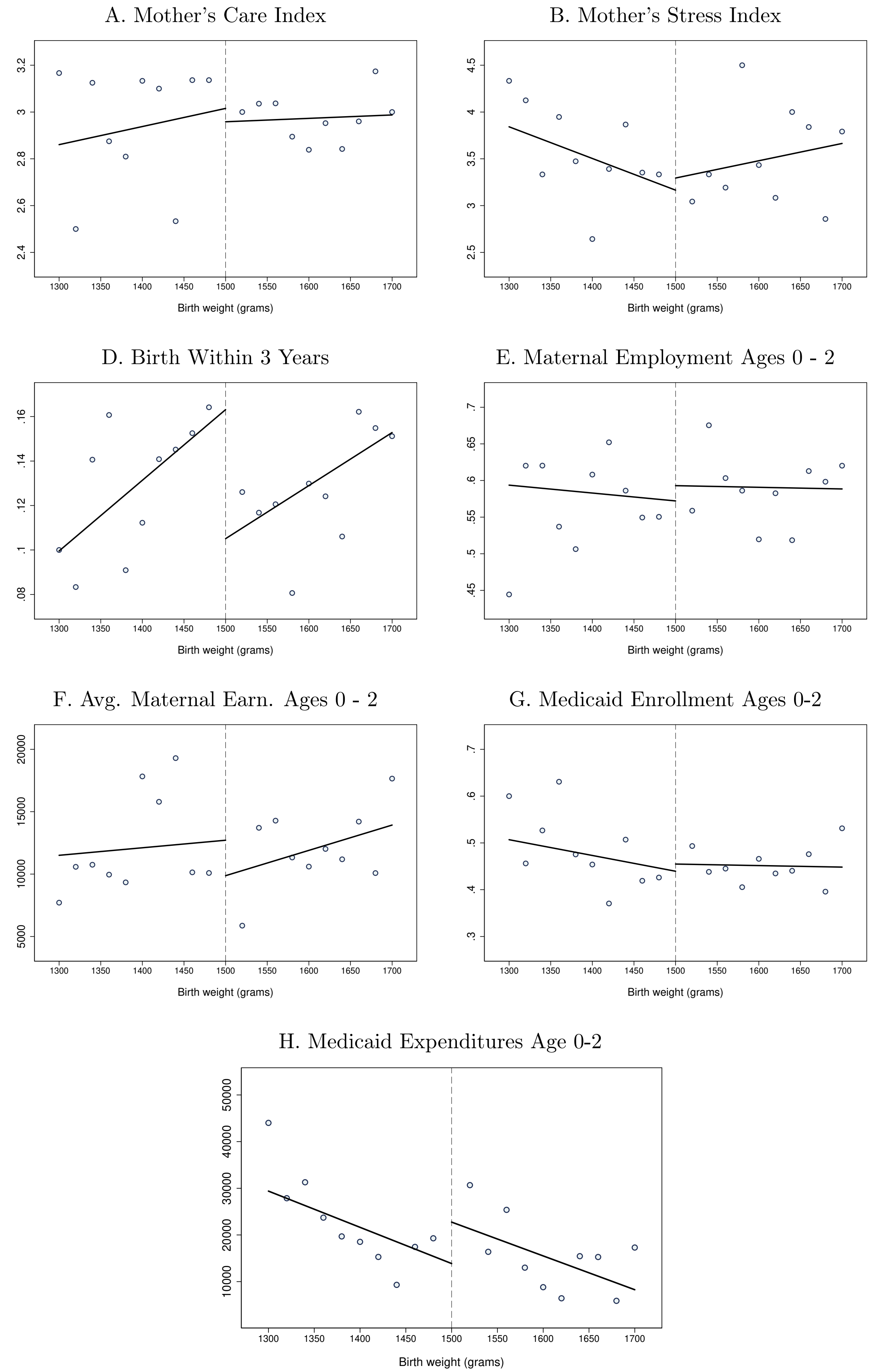


Panel C: Total Expenditures ($) | Any Expenditures (=1) by Age 10



Notes: Figure shows the relationship between birth weight and total social expenditures by age 10. Total social expenditures include SNAP payments, Medicaid expenditures, Medicaid pharmacy payments, TANF payments, and IEP costs. 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.

Figure A.5: Impacts for Early Childhood Investment Outcomes



Notes: Each panel shows the relationship between birth weight and various post-birth outcomes. Dots represents means within 20 gram bins of the running variable. The dark lines are predictions from a linear regression using the individual-level data.