Online Appendix

Background on Data Sources and Sample Construction

Our primary source of data is the blood lead surveillance data from the state registry maintained by the NC Childhood Lead Poisoning Prevention Program of the Children's Environmental Health Branch. This dataset includes a child's name, gender, birth date, test date, blood lead level (BLL) and home address. The North Carolina State Laboratory for Public Health (Raleigh, NC) conducted 90 percent of the lead analyses of the blood samples and all BLL values are stored as integers with a value of 1μg/dL (micrograms per deciliter) given to children without any detectable lead.

Our analysis focuses only on children living in Mecklenburg County and includes all BLL tests for a child between 1993 and 2008. North Carolina requires all children participating in Medicaid or the Special Nutrition Program for Women, Infants and Children (WIC) to be screened for lead at 1 or 2 years of age. Other children are screened if a parent responds "yes" or "don't know" to any of the questions on a CDC Lead Risk Assessment Questionnaire. The North Carolina Blood Lead Surveillance Group estimates that it screened between 21.9 and 30.4% of children one and two years of age from 1995 through 1998 and we expect screening rates were similar during our analysis period (Miranda et al. (2007)). This dataset provides multiple blood lead level tests per child which allows us to determine which children received various lead policy interventions due to two tests with BLL of 10μg/dL or above.

We subsequently match individual children to two additional databases in order to examine the impact of interventions on educational and behavioral outcomes. All matches are conducted using first and last name as well as date of birth and we incorporate fuzzy matches for names in some cases. Our first database is the administrative records from Charlotte-Mecklenburg Schools (CMS) that span kindergarten through 12th grade and the school years 1998-1999 through 2010-2011. This dataset includes each student that attended a public school in the City of Charlotte for at least one semester and provides annual data for each year of matriculation. Specifically, we incorporate student demographics on race and home address, yearly end-of-grade (EOG) test scores for grades 3 through 8 in math and reading, number of days absent, days suspended from school as well as the number of incidents of school crime.[56] We are able to match 65 percent of lead tests to a student record in CMS. This match rate improves to 74 percent for our policy sample of individuals with two tests and one test>10μg/dL.

[56] According to NC State Statute 115C – 288(g), any incident at school involving any violent or threats of violent behavior, property damage, theft or drug possession must officially be reported to the NC school crimes division. This statute ensures that this measure of school crime is consistently reported across schools and cannot be treated differently based on school administrators.

In order to examine adult criminal outcomes we match our lead database to a registry of all adult (defined in North Carolina as age 16 and above) arrests in Mecklenburg County from 2006 to 2013. We use first name, last name and date of birth to link individuals across the two data sources. While over 90 percent of the matches are exact, we recover additional matches using an algorithm for partial matches that has been used and validated in Deming (2011). The Mecklenburg County Sheriff (MCS) tracks arrests and incarcerations across individuals using a unique identifier that is established with fingerprinting. The arrest data include information on the number and nature of charges as well as the date of arrest. This data allows us to observe adult criminality regardless of whether a child later transferred or dropped out of CMS schools with the main limitation being that it only includes crimes committed within Mecklenburg County. The quality of matching between the lead and arrests databases is not directly measurable since one cannot distinguish between those lead tested individuals never arrested versus individuals who do not match due to clerical errors in names or moving out of the county. We can speak to the quality of matches using the arrest database by the fact that we are able to match approximately 94 percent of arrest records for a given cohort to our CMS education database.

In order to provide some basic controls for parental and housing factors, we draw on two additional databases. The first database is the universe of birth certificate records from the state of North Carolina from 1990-2002. As with previous databases, we are able to match our lead database to the birth records database using name and date of birth. In the case of birth records we are primarily interested in two variables, birth weight and mother's years of education.[57] We are able to match approximately 54 percent of birth records to our lead database. Even though this match rate is somewhat lower than our other databases, the variables from this database are simply used as control variables and we later show that this match rate is unrelated to our analysis of lead policy interventions. The second database is county assessor's data for all parcels on an annual basis from 2002-2012 in Mecklenburg County, NC. For this database, we match our lead data to parcel records based on home address given for an individual's first lead test. We augment this parcel data with building permits for all home renovations from 1995-2012. This database on parcels allows us to generate variables for prior home renovations, age and type of housing structure. The lead database is matched to parcels records 86 percent of the time with differences primarily a result of incomplete homes address information.

In some of our analysis, we merge into our dataset two additional data elements. First, we merge data from the LeadSafe Charlotte program which contains detailed data on the addresses of approximately 2,500 homes (single-family and multi-family) which have been lead inspected or

[57] We focus on maternal education because paternal information is not consistently reported on birth records in NC during this time period.

lead remediated and certified lead safe since 1998. We match LeadSafe addresses to our county parcel data based on parcel addresses with 20 LeadSafe homes unable to be successfully matched to parcel records. Second, we construct a measure of siblings using birth records data. In order to be characterized as a sibling, two individuals must share a mother's first name, last name and date of birth based on Mecklenburg County birth records.

Summary Index Construction

We follow the methodology in Anderson (2008) to create two summary index outcome measures: educational performance and adolescent antisocial behavior. The antisocial behavior index is created to include measures of number of days suspended and unapproved absences (6th through 10th grade), school reported crimes, and criminal arrests between the ages of 16 and 18. The education index includes 3rd through 8th grade math and reading test score results and grade retention between 1st and 9th grade.

A summary of the steps to create an index are listed below. See Anderson (2008) for additional detail in calculation of a summary index.

1. Switch signs where necessary so the positive direction indicates a larger outcome effect.
2. Demean outcomes and convert to effect sizes by dividing by its control group standard deviation.
3. Define groupings of outcomes.
4. Create a new variable that is a weighted average of the outcomes in each grouping. When constructing the weighted average, weight each element by the inverse of the covariance matrix of the standardized outcomes in each group.
5. Regress the new weighted average for each group on intervention status to estimate treatment effects.

Table A1: Variable Definitions

Variable Name	Description	Source
Outcomes		
Education Index	An index with mean zero and standard deviation one based on the academic outcomes listed below	CMS pupil records 1999-2011
Reading Test Score	The average reading test scores for given grades All test scores are normalized to mean zero and standard deviation one based on all NC tests in the same grade and testing year	
Math Test Score	The average math test scores for given grades	
Repeat a Grade	Indicator if a student ever repeats a given school grade.	
Antisocial Behavior Index	An index with mean zero and standard deviation one based on the behavioral outcomes listed below	
Days Suspended	Total days suspended (in + out of school) (6th-10th grade)	CMS pupil records 1999-2011
Days Absent	Total days absent (6th-10th grade)	CMS pupil records 1999-2011
School Crimes	Total crimes reported for a given student (6th-10th grade)	CMS pupil records 1999-2011
Ever Arrested (16-18 years old)	Indicator if ever arrested for a crime in Mecklenburg County/Charlotte NC	Mecklenburg County Sheriff 1998-2014
Ever Arrested - Violent Crime	Indicator if Ever Arrested for a violent crime (assault, manslaughter, robbery)	Mecklenburg County Sheriff 1998-2014
Ever Arrested - Property Crime	Indicator if Ever Arrested for a property crime (burglary, larceny, forgery/fraud, auto theft)	Mecklenburg County Sheriff 1998-2014
Background Characteristics		
Male	Indicator for gender	CMS pupil records 1999-2011
Minority	Indicator for race equal to black or hispanic	CMS pupil records 1999-2011
Stand Alone Residence*	Indicator for single unit housing structure (single-family residence)	Mecklenburg County Assessor 1999-2015
Home Built pre-1978*	Indicator for home built before 1978	Mecklenburg County Assessor 1999-2015
Past Lead Tests at a Home (mean BLL)*	Average of all prior BLL tests with the same residential address as a given observation	NC Blood Surveillance 1993-2008
Age at BLL Test	Age in months for initial BLL test	NC Blood Surveillance 1993-2008
Mother's Education*	Years of education reported for a child's mother	NC Birth Records 1990-2008
Birth Weight*	An observation's birth weight (ozs)	NC Birth Records 1990-2008
Index of Neighborhood Attributes	Average of standardized values (mean zero, standard deviation one) for four Census Block Group neighborhood variables: Percent in Poverty; Percent of adults with no high school diploma; Percent of households with single female-headed households; Population density	US Census 2000

* Some values of these variables are missing for our main estimation sample. For estimation, any observations with missing values are assigned a zero value and we create an indicator for missing value for that variable.

Table A2: Means of demographic, housing, and neighborhood characteristics

	All Students	Lead Tested	BLL 5-9	BLL ⩾10
Background Characteristics				
Male	0.51	0.51	0.52	0.55
	(0.50)	(0.50)	(0.50)	(0.50)
Minority	0.49	0.60	0.69	0.70
	(0.50)	(0.49)	(0.46)	(0.46)
Stand Alone Residence	0.67	0.65	0.63	0.66
	(0.47)	(0.48)	(0.48)	(0.48)
Home Built pre 1978	0.43	0.65	0.72	0.74
	(0.49)	(0.48)	(0.45)	(0.44)
Past Lead Tests at a Home (mean μg/dL)	3.91	4.09	4.20	4.43
	(1.21)	(1.16)	(1.18)	(1.52)
Age at Blood Lead Test	2.12	2.20	2.15	1.89
	(1.50)	(1.53)	(1.42)	(1.26)
Mother Education (years)	13.28	12.69	12.33	12.08
	(2.48)	(2.52)	(2.44)	(2.40)
Birth Weight (ozs)	115.81	113.52	112.54	111.22
	(21.86)	(21.95)	(21.39)	(20.56)
Index of Neighborhood Attributes	0.08	−0.28	−0.42	−0.44
	(0.76)	(0.85)	(0.86)	(0.87)
Observations	153,039	19,731	5,857	935

This table reports means and standard deviations for variables given in Table 1 for the full population of public students as well as by different lead testing BLL values for individuals born between 1990 and 1997.

Table A3: Means of education and behavior outcomes

	All Students	Lead Tested	BLL 5-9	BLL ⩾10
Blood lead level (μg/dL)	4.144	4.220	6.169	13.129
	(3.115)	(3.236)	(1.245)	(7.900)
Education Outcomes				
Reading Test Score (avg 3-5th grade)	−0.030	−0.204	−0.364	−0.474
	(0.965)	(0.956)	(0.934)	(0.916)
Math Test Score (avg 3-5th grade)	−0.033	−0.205	−0.366	−0.427
	(0.973)	(0.953)	(0.921)	(0.918)
Repeat a grade (grades 1-5)	0.046	0.102	0.133	0.140
	(0.210)	(0.303)	(0.339)	(0.347)
Reading Test Score (avg 6-8th grade)	−0.033	−0.174	−0.335	−0.409
	(0.967)	(0.952)	(0.932)	(0.920)
Math Test Score (avg 6-8th grade)	−0.038	−0.175	−0.324	−0.378
	(0.969)	(0.935)	(0.888)	(0.888)
Repeat a grade (grades 6-9)	0.101	0.142	0.193	0.197
	(0.302)	(0.349)	(0.395)	(0.398)
Antisocial Behavior Outcomes				
Days Suspended (6-10th grade)	4.34	8.49	11.29	14.35
	(13.39)	(19.85)	(22.88)	(26.75)
Days Absent (6-10th grade)	20.78	30.64	37.23	41.31
	(31.00)	(39.30)	(45.74)	(47.65)
School Reported Crimes (6-10th grade)	0.93	1.96	2.44	2.77
	(3.02)	(4.63)	(5.09)	(5.40)
Ever Arrested	0.05	0.08	0.11	0.12
	(0.21)	(0.27)	(0.31)	(0.33)
Ever Arrested - Violent	0.02	0.04	0.05	0.06
	(0.13)	(0.18)	(0.21)	(0.24)
Ever Arrested - Property	0.02	0.04	0.05	0.06
	(0.14)	(0.19)	(0.22)	(0.24)
Observations	153,039	19,731	5,857	935

This table reports means and standard deviations for variables given in Table 2 for the full population of public students as well as by different lead testing BLL values for individuals born between 1990 and 1997.*Note:* The mean blood lead level for All Students does not equal the mean blood lead level for the Lead Tested individuals since some students are not matchable to lead testing data.

Table A4: Balancing test for observable characteristics

	(1) Intervention (2 tests 10+)
Male	−0.008
	(0.058)
Minority	0.083
	(0.076)
Home Built pre 1978	0.144**
	(0.068)
Past Lead Tests at a Home (mean μg/dL)	−0.028
	(0.031)
Stand Alone Residence	−0.029
	(0.064)
Age at Blood Lead Test (months)	0.002
	(0.002)
Birth Weight (ozs)	0.001
	(0.002)
Mother Education (years)	0.023*
	(0.013)
Index of Neighborhood Attributes	0.047
	(0.037)
F-Stat (p-value)	0.237
Observations	301

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors robust to arbitrary within-CBG correlation in parentheses.

This table presents results from a balance test in which the dependent variable is an indicator equal to one if individual received two tests ⩾10μg/dL. The reported p-value at the bottom of each panel is the result of an F-test of joint-significance of all of the reported variables above. The regression also includes year-of-birth indicators and indicator variables for no match to a birth record or a parcel record.

Table A5: Does the elevated BLL intervention affect matching to data sets or residential mobility?

	Intervention	Control	Difference
Matched to CMS record (incl. in est. sample)	0.76	0.78	−0.022
	(0.43)	(0.41)	(0.043)
Observations	156	232	388
Did Not Enroll in CMS Secondary School	0.11	0.09	0.019
	(0.32)	(0.29)	(0.037)
Observations	108	174	280
Parcel Info Missing	0.17	0.21	−0.046
	(0.38)	(0.41)	(0.047)
Birth Record Missing	0.36	0.28	0.081
	(0.48)	(0.45)	(0.055)
Change in Residence btw Test and School	0.64	0.72	−0.086
	(0.48)	(0.45)	(0.058)
Observations	119	182	301

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors robust to arbitrary within-CBG correlation in parentheses.

This table reports means and standard deviations of data matching and mobility indicators for the group eligible for intervention (two tests $\geqslant$ 10μg/dL) and our control group (first test $\geqslant$ 10μg/dL, second test $\geqslant$ 5 but $<$ 10μg/dL) as well as the mean difference and the standard error of the difference. Any statistically significant differences are noted with * p < 0.1, ** p < 0.05, *** p < 0.01. The first row presents a comparison across the two groups of the probability of an individual meeting the BLL testing criteria being matched to a CMS school record (which is required for inclusion in our estimation sample). The second row compares rates (conditional on primary school enrollment) of not enrolling in a CMS secondary school to test for differential attrition rates for adolescent outcomes. The third and fourth rows compare rates of matching records to the birth record and parcel characteristic databases for those in our estimation sample. Finally, the fifth row compares rates of residential mobility between the first blood lead test and CMS school enrollment for our estimation sample. An indicator for a change in residence is created through a comparison of the residential address listed in each database.

Table A6: Effects of eligibility for the elevated BLL intervention adding sets of controls

	(1) Education Index	(2) Adolescent Antisocial Behavior Index
1. Uncontrolled mean differences		
Intervention	0.129*	−0.247***
	(0.068)	(0.081)
2. Add controls for bith year		
Intervention	0.142**	−0.244***
	(0.070)	(0.080)
3. Add controls for individual characteristics		
Intervention	0.138**	−0.221***
	(0.066)	(0.075)
4. Add controls for age at BLL test		
Intervention	0.136**	−0.219***
	(0.066)	(0.075)
5. Add controls for maternal educ		
Intervention	0.116*	−0.199**
	(0.066)	(0.077)
6. Add controls for parcel characteristics		
Intervention	0.125*	−0.194**
	(0.068)	(0.081)
7. Add controls for nbhd characteristics		
Intervention	0.117*	−0.184**
	(0.067)	(0.082)
P-value (model 1 = model 7)	0.47	0.11
Observations	301	301

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors robust to arbitrary within-CBG correlation in parentheses.

Notes: This table reports the estimated coefficient on an indicator for eligibility for the BLL intervention for specifications which incrementally add sets of control variables as indicated by the row title. All regressions are based on the primary estimation sample with the control and intervention groups defined as in Table 4. The first row reports the mean uncontrolled differences as previously reported in Table 2. The following controls are added in subsequent rows: (2) indicators for birth year and an indicator for whether there are any missing outcomes in the creation of the index variable; (3) birthweight and indicators for missing birth record, gender, and minority (black or Hispanic) status; (4) age at the time of the initial BLL test; (5) years of maternal education ; (6.) single family home indicator, pre-1978 indicator, building age, blood lead lead levels for previous children tested in home ; (7) Index of Neighborhood Attributes (an unweighted z-score sum of the percent of households without a high school graduate, the CBG poverty rate, the fraction of single female headed households, and the CBG population density).

Table A7: Effects of an elevated BLL intervention on summary index outcomes: *Robustness Check - Different Control Variables*

	(1) Education Index	(2) Adolescent Antisocial Behavior Index
1. Main Results		
Intervention	0.117*	−0.184**
	(0.067)	(0.082)
2. Add control for test type		
Intervention	0.115	−0.185**
	(0.070)	(0.079)
3. Control for initial BLL fixed effects		
Intervention	0.055	−0.150
	(0.077)	(0.110)
4. Control for days between tests		
Intervention	0.096	−0.170**
	(0.078)	(0.081)
Observations	301	301

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors robust to arbitrary within-CBG correlation in parentheses.

Notes: This table presents results from specifications which add various controls to test for robustness. All regressions include the full set of control variables listed in the table notes of Table 4 and the primary estimation sample with the control and intervention groups defined as in Table 4. The first panel reports our primary estimation results for comparison purposes. The second panel includes an indicator for the type of BLL test used for a second (confirmatory) test. The majority of tests are capillary specimens typically obtained by the a finger prick while some follow up tests are a venous blood draw. The third panel of this table presents results from a specification which includes indicator variables for each initial BLL value to assess whether there may be substantial selection concerns arising from parents responding differently to initial results. Finally, the fourth panel presents results controlling for the days between the initial and follow-up confirmatory test. This variable is correlated with treatment assignment through a mechanical mechanism: due to the 30 day half-life of lead in the blood, a quicker follow-up is more likely to yield a higher confirmatory test result.

Table A8: Regression Discontinuity Results

	(1) Education Index	(2) Adolescent Antisocial Behavior Index
1. Full Sample RDD, BW=5-14		
Intervention	0.071	−0.228**
	(0.090)	(0.099)
Observations	6,575	6,575
2. Estimation Sample RDD, 2nd BLL Test BW=5-14		
Intervention	−0.163	−0.163
	(0.153)	(0.188)
Observations	248	248
3. Estimation Sample RDD, any 2nd BLL Test		
Intervention	−0.062	−0.228
	(0.200)	(0.336)
Observations	301	301

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors robust to arbitrary within-CBG correlation in parentheses.

The top panel presents results from a local linear regression discontinuity design (RDD) using the initial BLL test result as a running variable and thus utilizing all students with an initial BLL from 5μg/dL through 15μg/dL. The second and third panels present results from a local linear RDD using the second (confirmatory) BLL test result as the running variable. The second panel restricts the bandwidth again to those with a 2nd test result of 5 through 14μg/dL. The third panel allows data from the entire estimation sample and estimates an RDD again using the confirmatory test as the running variable but without any restriction on the bandwidth. All regressions include the full set of control variables listed in the table notes of Table 4.

Table A9: Effects of an elevated BLL intervention on summary index outcomes for siblings

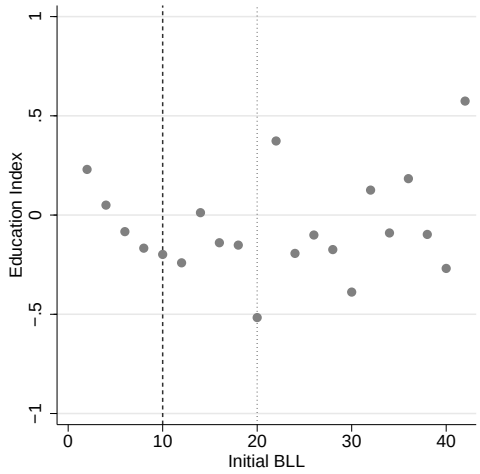
	(1) Education Index	(2) Adolescent Antisocial Behavior Index
Younger Siblings		
Younger Sibling of Child (>10 , >10)	0.091	−0.216
	(0.170)	(0.241)
Observations	120	120
Older Siblings		
Older Sibling of Child (>10 , >10)	−0.009	−0.123
	(0.278)	(0.262)
Observations	88	88

* p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors robust to arbitrary within-CBG correlation in parentheses.

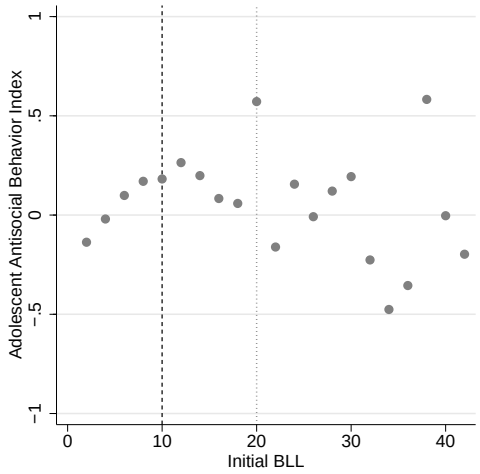
Notes: This table presents results for a sample of siblings of our estimation sample. Due to small sample sizes, we compare siblings between the intervention group and the broader control group defined by individuals whose first BLL test result was $\geqslant$10µg/dL. We limit our analysis to siblings within 3 years of age. Siblings are defined based on being born to the same mother (identified by first name, last name and date of birth). All regressions include the full set of control variables listed in the table notes of Table 4.

Figure A7: Outcomes by BLL

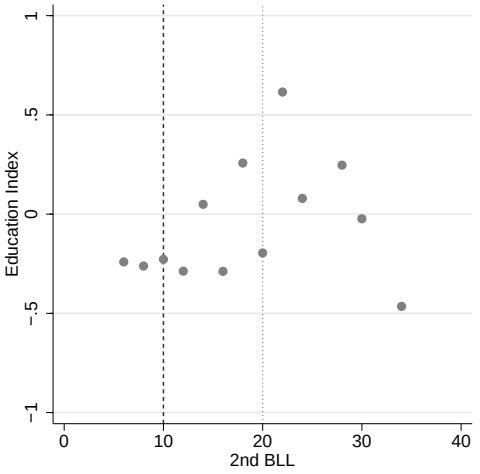
A1. Education Index, Initial BLL



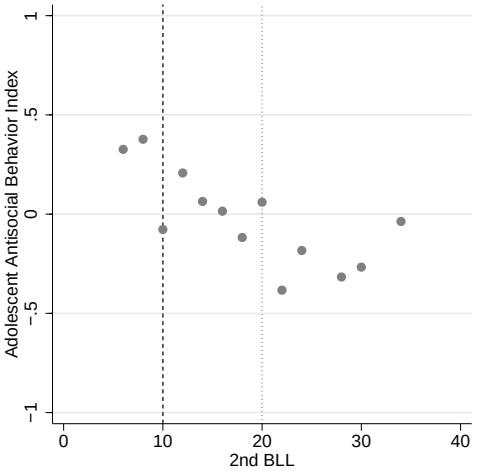
B1. Antisocial Beh. Index, Initial BLL



A2. Education Index, Estimation Sample, 2nd BLL test



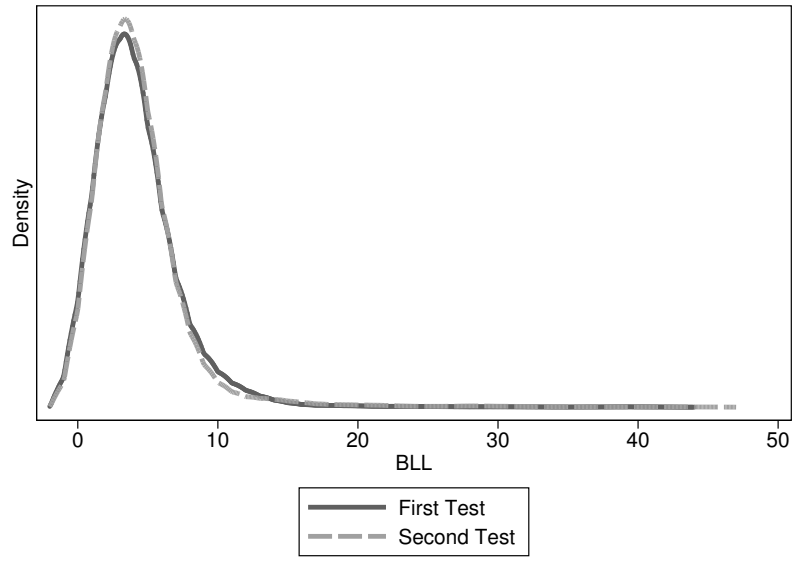
B2. Antisocial Beh. Index, Estimation Sample, 2nd BLL test



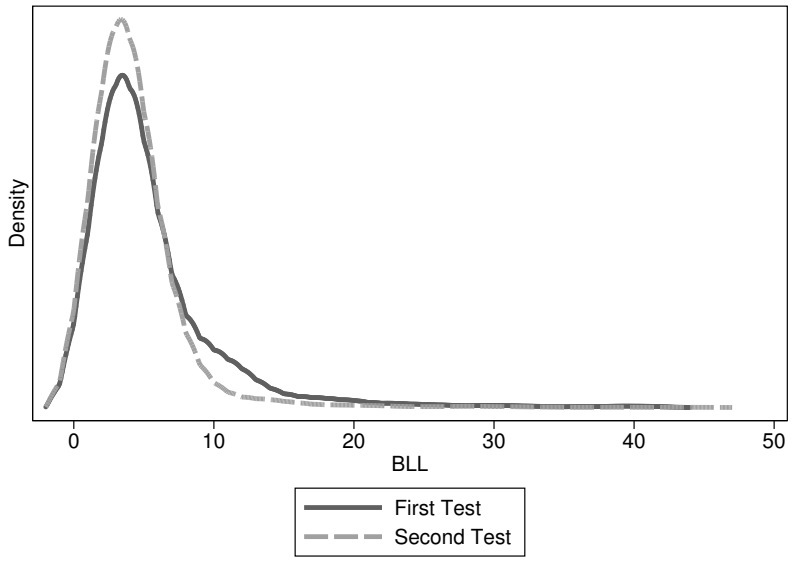
Notes: This figure provides plots average index outcomes by BLL values in 2μg/dL bins. The top panel (A1 and B1) plots outcomes for all children with BLL test results based on the value of their first test result. The second panel (A2 and B2) includes only those in our estimation sample and plots outcomes by the second (confirmatory) BLL test result horizontal variable.

Figure A8: Blood Lead Testing Variation

A: Distribution of all BLL test results



B: Distribution of BLL test results - only children with at least 2 BLL tests



Notes: Panel A of the figure provides the distribution of all first BLL tests in comparison to those individuals that ever had a second BLL test for the full blood surveillance dataset. Panel B further restricts this comparison such that both distributions only contain individuals with two BLL tests.