Inequality in Lead Exposure and the Black-White Test Score Gap

Anna Aizer
Brown University and NBER

Janet Currie
Princeton University and NBER

Peter Simon
Brown University

Patrick Vivier
Brown University

August, 2015

We consider a new source of racial disparities in test scores: African American students' disproportionate exposure to environmental toxins, and, in particular, lead. Using a unique individual-level dataset of children's preschool lead levels linked with future educational outcomes for children in Rhode Island, we show that state policies aimed at reducing lead exposure led to significant declines in children's blood lead levels among cohorts born between 1997 and 2004. Exploiting the change in child lead levels as a result of the policy, we generate causal estimates of the impact of preschool lead levels on reading and math test scores from grade three through grade eight in an IV framework. We find that a 5 micrograms per deciliter increase in child lead levels (the threshold at which the CDC recommends intervention) reduces test scores by 6 points or 43 percent of a standard deviation. The effects are stronger in the lower tail of the test score distribution and do not fade over time. The decline in racial disparities in lead exposures can explain roughly half of the 32 percent decline in the racial gap in test scores in these cohorts.

JEL Codes: I24, J15, Q53, Q58

We thank Rebecca Lee and Alyssa Sylvaria of the Providence Plan for their generosity and help with the data, and the John D. and Catherine T. MacArthur Foundation for financial support. We also thank Michelle Marcus and Molly Schnell for outstanding research assistance as well as Pedro Dal Bó and seminar participants at the Harvard-BU-MIT joint seminar, the University of Wisconsin Institute on Poverty, Yale, LSE, PSE, the Rand Corporation, the New York Federal Reserve, and the NBER Summer Institute for helpful comments. The authors are solely responsible for any errors.

I. Introduction

Racial disparities in test scores have declined significantly over the past forty years, though a substantial gap remains. As measured by National Association of Educational Progress (NAEP) scores for 9 year olds, the black-white test score gap represented 1.25 of a standard deviation in 1971, falling to 0.71 by 2012, with the greatest gains made during the 1970s and 1980s (US DOE, 2013). In their review of the existing research on the factors responsible for the black-white test score gap, Jencks and Phillips (2011) conclude that family income, school segregation, family structure, parenting practices, and the quality of educational inputs explain only part of the gap, leaving a large share unexplained.

We consider an alternative explanation for the black-white test score gap: Black children's greater exposure to environmental toxins. A long literature documents racial disparities in environmental quality, exploring multiple sources of the disparities. These include differences in proximity to known hazards, exposure to air pollution, siting of landfills/other hazards, designation as Superfund sites, and other regulatory actions (Brown, 1995; Brulle and Pellow, 2006; Ringquist, 2005; Mohai and Saha, 2006). Currie (2011) demonstrated that even within zip codes, black mothers were more likely to reside near Superfund sites or plants with toxic emissions.

Because African American students are disproportionately burdened by exposure to harmful environmental toxins, we argue that recent efforts to reduce exposure are likely to have the greatest impact on them. If exposure to environmental pollutants negatively affects cognitive development, it follows that environmental regulations that lead to disproportionate reductions in exposure among African American children are also likely to reduce racial disparities in educational outcomes.

To examine this hypothesis, we focus on the environmental toxin lead, for which there is strong neurobiological and epidemiological evidence of a negative relationship between early exposure and future cognitive and non-cognitive outcomes. The major sources of lead in the environment today are residual lead in soil from past emissions and deteriorating lead-based paint remaining in old homes (Levin et al 2008), both of which are higher in urban areas. Because African American families are residentially segregated in older, urban areas, African American children suffer greater exposure to lead.

We construct a unique dataset for the state of Rhode Island starting with all children born between 1997 and 2004, and matching in information about children's blood lead levels measured from birth through age six, as well as test scores and educational outcomes for grades three through eight. Using these data, we first document the greater lead burden born by African American and poor children in our sample. African American RI children born in 1997, (the first birth cohort of our data), had lead levels of 6.1 micrograms per deciliter (ug/dl), on average, compared with 3.8 for white children; children on free lunch had lead levels of 5.1 on average, compared with 3.3 for those not on free lunch. By the end of the sample, only seven years later, lead levels had declined to 3.5 for black children and 2.4 for white children, and 3.2 and 2.1 for children on free lunch and not, respectively.

While there is an extensive literature documenting the negative relationship between elevated childhood lead levels and poor cognitive outcomes, identifying the causal impact of elevated lead levels on child outcomes is challenging for two reasons: First, the presence of multiple confounders can lead to omitted variable bias, and second, measurement error in child lead levels leads to attenuation bias. Confounding arises from the fact that children with higher lead levels are more disadvantaged: they are more likely to be African American, poor, live in

single parent homes, and have mothers with lower human capital – all of which are independently associated with lower test scores. Measurement error arises primarily from the imprecision in each blood test, but also from the fact that blood tests capture only recent exposure that may or may not have been persistent, whereas lead that persists in the body in organs such as the brain still causes cognitive impairment.

To identify the effect of elevated lead levels in early childhood on future cognitive achievement, we employ an instrumental variables strategy to address both omitted variable bias and measurement error. Our instrument is based on two policies implemented in RI in 1997 that required landlords to ensure that their rental homes were free of lead hazards. These policies followed new CDC statements issued in 1997 urging states to target testing and remediation of the highest risk children (Jones et al, 2009). The first state policy required the owners of any building in which a child had an elevated lead level to mitigate any lead hazard or face prosecution by the state Attorney General. The second state policy applied to all landlords, regardless of whether a child had tested positive for lead, and required them to obtain "lead-safe certificates" in order to rent their properties. For the latter, the state targeted the oldest urban centers of the state. We document that neighborhoods with a greater share of old housing (the primary source of lead paint), and with higher initial child lead levels (as measured in 1997) witnessed the biggest gains in the number of lead-safe certificates issued, and the biggest reductions in child lead levels.[1] These neighborhoods were also predominantly African American and poor. By including neighborhood (census tract) fixed effects in our analysis, we focus on changes in the number of lead safe certificates over time within a tract to identify our effects. We

[1] Previous work has documented that individual blood lead levels decline in a household after a certificate has been issued (Rogers et al, 2014).

assume that other underlying characteristics of the neighborhood did not change coincident with the growth in certificates, and provide empirical evidence to support our identifying assumption.

Using this instrumental variable strategy, we find that an average lead level of 5 ug/dl (the level at which the CDC currently recommends intervention) reduces average reading test scores by 6 points, or 43 percent of a standard deviation. This effect does not appear to fade over time, remaining strong by eighth grade, the last year for which we have data. Moreover, the effects are greatest at the lower tail of the test score distribution, particularly for math test scores. The IV estimates are larger than the OLS estimates, consistent with the considerable measurement error in lead scores which we document.[2]

Finally, we show that the policy change did disproportionately reduce lead levels in black children relative to white, with the black-white ratio in measured lead levels falling from an average of 1.61 for the 1997 birth cohort to 1.4 for the 2004 birth cohort. Based on our causal estimates, we calculate that the falling racial gap in child lead levels over this period corresponds to a 1.56 point reduction in the black-white test score gap (12 percent of a standard deviation), thereby explaining half of the total decline in the black-white test score gap witnessed over this period in RI.

These results do not simply reflect changes in lead exposure and scores among the disadvantaged more generally. When we compare outcomes for children on free lunch or not, we find that both the baseline (1997) levels of lead were lower and more similar to those of other children and that the declines in disparities in lead levels and test scores were smaller. The

[2] We explore whether African American children in addition to being more likely to be exposed, also suffer more harm from a given level of exposure. For example, good nutrition and cognitive stimulation may be protective against the negative effects of lead poisoning, and African American children may be less likely to enjoy these protective factors (Environmental Protection Agency, 2015). However, our data and identification do not allow such a breakdown.

greater effects for African American families are likely attributable to the fact that African American families are disproportionately located in urban areas of the state even conditional on poverty (89 percent of poor blacks relative to 60 percent of poor whites) and these areas are characterized by the highest (historical) lead-contaminated emissions, oldest housing and the highest baseline lead levels. These were also the areas targeted by the state in its outreach efforts.

Our results suggest that African American children's disproportionate exposure to environmental pollutants that affect child development can potentially explain a significant share of the black-white gap in educational outcomes. Policies aimed at reducing the environmental burden borne by low income and minority households have the potential to be an effective tool in reducing disparities in educational outcomes and, by extension, future income.

II. Background

A. Disparities in Educational Outcomes and Environmental Quality

Jencks and Phillips (2008) provide a comprehensive review of the existing research on various factors that have been considered to explain racial disparities in test scores and their decline over time. These include family income, family structure, school segregation, cultural influences, test bias, stereotype threat, parenting strategies, and the quality of school inputs.[3] They conclude that together these factors explain some, but not all of the racial gap in test scores and that alternative explanations should be considered. More recently, Chay, Guryan and Mazumder (2009) have done just that. They examined the impact of relative improvements in

[3] There is a long literature on this topic which includes Brooks-Gun et al (2003), Card and Rothstein (2007),Cook and Evans (2000), Dee (2005), Figlio (2005), Fryer and Levitt (2004), Hanushek and Rivkin (2006), Krueger and Whitmore (2001), and Reber (2010) among others.

infant health among African American children due to hospital desegregation in the American south on black cognitive achievement. They find that among cohorts born between 1963 and 1971, increasing access to health care among African Americans during infancy and the resulting improvements in health explain a significant share of the closing of the black-white gap in cognitive achievement in the American South during the 1980s.

This recent work underscores the importance of looking at factors beyond the educational system and even family characteristics to explain the racial gap in test scores – a strategy we pursue here. We focus on the role of racial disparities in environmental quality in explaining racial gaps in test scores.

Debate over the cause of environmental inequalities remains. While some argue that the pattern results from the deliberate placing of hazardous sites in minority communities (due to lack of organized opposition and/or low land prices), others argue that a more likely explanation is that housing prices reflect the quality of the environment (Chay and Greenstone, 2005; Banzhaf and Walsh, 2008; Currie et al., 2015).[4] Since African American families are more likely to be poor, it is not surprising then that they tend to live in homes of lower environmental quality.

Inequity in the regulatory response to pollution by race and/or class spawned new federal regulation in 1994 aimed at preventing discrimination in the "development, implementation, and enforcement of environmental laws, regulations and policies" (EPA), also referred to as environmental justice. Since then, work examining the probability of a hazardous site receiving priority for clean-up efforts (ie, Superfund status) has found that sites in black and low income areas are still less likely to receive a Superfund designation (O'Neal, 2007). But, conditional on

[4] For example, Woverton (2009) finds that race and income predict current locations of manufacturing plants but to not predict the original citing of plants.

Superfund designation, there no longer appears to be any disparity in the duration of clean-up (Burda and Harding, 2014). As part of our analysis, we examine both the burden of lead by race and income as well as the targeting of efforts to reduce lead exposure across groups.

B. Lead Exposure in American Children and the National Policy Response

In 1970, the US Surgeon General issued his first formal statement on lead poisoning, naming it a national health problem (US Dept of Health, Education and Welfare, 1970). Since then, American children's exposure to lead has declined dramatically, due largely to two major regulatory changes involving the elimination of lead in gasoline and in household paint.[5] These regulatory changes have been credited with reducing the lead concentrations in the air from 2.5 ug/dl in 1980 to less than 0.5 by 2012.[6] Coincident with this decline, the share of pre-school age American children with lead levels in excess of 10 ug/dl has likewise declined from 8.6% to 1.4% over the period 1988 -2004, with declines witnessed across all groups, but especially for historically high-risk groups of children (Jones et al, 2009).

Using laboratory data from the NHANES, we examine average BLLs by race and birth cohort for children less than seven years old. After 1978, BLLs of all children declined, but they declined more for African American children (Appendix Figure 1, first panel). For comparison, we also plot racial disparities in NAEP reading test scores for 9 year olds by birth cohort in the second panel. Just as the racial gap in lead levels has fallen after 1978, so too has the gap in test scores, though a substantial gap remains.

[5] The key regulations and legislation that reduced exposure to lead included the 1970 Lead Paint poisoning Prevention Act, the Clean Air Act and EPA rules regarding leaded gasoline.

[6] (http://www.epa.gov/air/airtrends/lead.html).

Despite the removal of lead from gasoline and household paint in the 1970s and 1980s, lead has remained a significant hazard in the environment, particularly in urban areas where African American families are disproportionately located (Bajari and Kahn, 2001).[7] This is due to the geographic concentration of three of the main sources of lead in urban areas: residual lead in soil near high traffic areas, lead in deteriorating paint found in old homes, and residual lead in soil at former industrial sites which were often located in central cities (Lanphear et al 1998; Levin et al, 2008). With respect to past emissions, not only are urban homes located closer to major roadways where there are high levels of residual lead found in soil, but conditional on distance to roadway, urban areas are characterized by higher lead levels due to greater (historical) traffic volume (Filipelli et al, 2005).[8] With respect to lead-based paint as another major source of exposure, urban areas are also more likely to be characterized by older housing that contains lead-based paint than non-urban areas. But even conditional on living in an old home, there may be important differences in exposure to lead by socio-economic status since the hazards associated with lead-based paint can be reduced by painting over old paint and ensuring that living areas are free of paint chips and dust. To the extent that more educated or wealthier families are more likely to take these precautions, similar exposures to old housing within an urban area are more likely to negatively affect the most disadvantaged.

While child lead levels have declined dramatically over the past 30 years, the CDC still considers exposure to lead an important public health issue. Current estimates suggest that 4.5 million households in the US are still exposed to high levels of lead and a half a million preschool aged children have elevated blood lead levels (BLLs). Moreover, disparities by race

[7] In 1990, among households living in metropolitan areas (75% of the US population), a black household was 31 percentage points more likely to live in a metropolitan area's center city than a white household (Bajari and Kahn, 2001), and this is true even conditional on poverty.

[8] High levels of residual lead found in soil near major roadways exhibit an exponential decay, declining by two thirds at a distance of 50 meters (Filippelli et al, 2005).

and income, while smaller, still remain, with African American and poor children two to three times as likely to have elevated lead levels, defined as above 5 (CDC, MMWR 2013).[9] Indeed, the U.S. Public Health Service recently included the elimination of elevated blood lead levels as a goal in Healthy People 2020, the ten-year national objectives for improving the health of all Americans. The CDC has also eliminated use of the term "level of concern" when referring to blood levels and lowered the threshold for which children should receive case-management services from 10 to 5 ug/dl in 2012 based on recent research that has found that even small amounts of lead in a child's system can have deleterious effects (CDC, 2012).

C. The Biology of Lead Poisoning

Lead has no biological value and is toxic to the human body, regardless of the pathway of exposure (ingestion or inhalation). Lead is toxic because of its ability to inhibit or mimic the actions of calcium, thereby affecting all calcium-dependent or related biological processes. Many systems including the renal, endocrine and cardiovascular systems have been found to be affected by lead exposure but the nervous system appears to be the most sensitive target. Within the brain, lead levels lead to damage in the prefrontal cerebral cortex, hippocampus and cerebellum (Finkelstein, Markowitz, and Rosen, 1998). Many neurodevelopmental studies have concluded that infants, children, and the developing fetus are at the greatest risk of toxicity from low-level exposure (National Research Council, 1993). This sensitivity is due to the fact that ingested lead is more likely to be absorbed from the gastrointestinal tract of children, and,

[9] Based on analysis of the 1999-2002 NHANES, the CDC reported that the average blood level for children aged 1-5 was 1.9 ug/dL, but for African American children this figure was 50% higher (2.8 ug/dL) and 30% higher for low income households (2.5 ug/dL) (CDC MMWR, 2005).

conditional on absorption, is more likely to affect the developing nervous system than the mature brain (Lidsky and Schneider, 2003).

D. Existing Evidence Regarding Elevated Lead and Child Outcomes

There is a long epidemiological literature documenting the significant negative relationship between exposure to lead and child outcomes. The negative health effects of lead ingestion have been known since at least Roman times. But until the mid-twentieth century, health officials considered only severe cases of lead poisoning to be harmful. This changed in the 1960s, when medical professionals recognized that less acute lead exposure also had harmful effects and began characterizing lead poisoning as epidemic (Berney, 1993).

Recent research has focused on the impact of relatively low levels of lead during childhood on cognitive and behavioral outcomes. Chandramouli et al. (2009), Canfield et al. (2003), Lamphear et al. (2005), McLaine et al (2013), Nigg et al. (2010) and Wasserman (1997) have all documented a significant relationship between relatively low levels of lead (<10 ug/ml) during childhood and cognitive and behavioral outcomes including ADHD and hyperactivity.[10]

In addition to the above studies based on individual level data linking BLLs in early childhood with later outcomes, other research has relied on cohort level analyses. Ferrie, Rolf and Troesken (2012) exploit variation in the use of lead water pipes across time and place in the early part of the 20th century to estimate a negative relationship between exposure to leaded water and later intelligence test scores among WWII army enlistees. Interestingly, the effects are

[10] McLaine et al. (2013) uses data on preschool BLLs to estimate the relationship between elevated BLLs and reading readiness at kindergarten, a marker for later school performance, among Providence, RI school children. They find an inverse relationship between school readiness test scores and BLLs: among children with BLLs below 5, 68% scored above the benchmark levels, among children with levels between 5 and 9, 62% scored above the benchmark, and among children with a BLL of at least 10, only 49% exceeded the benchmark.

greatest for children from low SES homes, consistent with work showing that good nutrition can reduce lead absorption. The focus of much of the cohort-level analyses has been to link declining lead levels in US children in the 1970s and 1980s to the dramatic declines in crime witnessed since the mid-1990s, exploiting cross-area and cross-cohort variation in lead exposure, typically from the staggered timing of de-leading gasoline during this period. These studies are based on aggregate level data that link average lead levels in a given state or city during a cohort's first years of life with crime levels 15-20 years later. Examples include Masters et al. (1998), Reyes (2015), Nevin (2000 and 2007), Mielke and Zahran (2012), and Grönqvist, Nilsson and Robling (2014). This work supports a strong relationship between declines in lead exposure early in life and declines in crime later in life, but like all cohort-level analysis, has been subject to the criticism of ecological fallacy.

E. Challenges to Estimation of the Impact of Elevated Lead on Child Outcomes

Despite the large body of evidence showing a strong inverse relationship between elevated BLLs in childhood and later cognitive and behavioral outcomes, there is still uncertainty regarding the magnitude of the effect as well as its duration. This uncertainty arises from two sources: the confounding of the relationship between lead exposure and child outcomes, and measurement error in child BLLs.

Regarding the former, disadvantaged children are more likely to be exposed to lead through older, substandard housing as well as proximity to high traffic roadways where residual lead in soil is found in high concentrations, and, as a result, are more likely to have elevated BLLs. While this is a concern we try to address, Bellinger (2008) argues that "Such confounding seems highly unlikely to account completely for the associations, given the wide range of

circumstances and settings in which they have been found. Evidence that so-called "subclinical" exposure to lead not only alters behavior but brain structure as well would make the argument of confounding even less tenable."

The second estimation challenge is measurement error in child lead levels obtained from blood samples. Measurement error can arise from contamination of the blood sample especially when the sample is a capillary sample (also known as a finger prick), which is a common measure due to its lower expense and minimal discomfort to the child. An alternative measure (venous) is more costly but less prone to sample contamination. Even without contamination, there is significant error in measurement, with greater error characterizing capillary measures. According to the CDC, the "ratio of imprecision to measurement value, particularly at BLLs <10ug/dL, is relatively high" (MMWR, 2002).[11] This can lead to considerable attenuation bias in the estimated effect of lead. Moreover, the half-life of lead in blood is relatively short (36 days). As such, BLLs only capture very recent exposure and will not necessarily capture the amount of lead that has settled in body organs and bone.[12] Because over the first 72 months of a child's life, a child is typically only sampled a small number of times, this problem introduces an additional source of measurement error, though with ambiguity with respect to the direction of any bias that may result.[13] Below we discuss our strategy for addressing these potential sources of bias in estimates of the impact of BLLs on later child cognitive outcomes.

[11] According to the CDC guidelines "Federal regulations allow laboratories that perform blood lead testing to operate with a total allowable error of ± 4 ug/dL or ±10% whichever is greater" (MMWR, 2007).
[12] "Deleterious health effects of lead resulting from long-term lead exposure will only be correlated with current blood-lead levels if lead exposure has been relatively constant over a long period of time, up to the time of sampling." Moreover, "Physiologically, the measurement of lead in blood is not a direct assessment of target organ dose, since the red cell is not a critical target for lead toxicity. Kinetically, blood is not a good analog for critical targets, such as soft tissue, because of the relatively short half-life of lead in blood as compared to target organs or bone." Mount Sinai School of Medicine. Accessed on 12/16/2014: http://research.mssm.edu/xrf/why.html
[13] If BLLs consistently underestimate the amount of lead in the child's system, this problem would result in an overestimate of the impact of lead on child outcomes. However, it is also possible that mismeasurement could result

III. Data and Empirical Strategy

A. Overview of Estimation

To explore the extent to which racial disparities in exposure to lead can explain the racial gap in test scores, we proceed in three stages. First we document disparities in BLLs by race and income, examining both levels at the beginning of our period (1997) and trends over time. We then examine the relationship between BLLs in early childhood and later cognitive achievement as measured by scores on standardized reading and math tests. As part of this effort we address issues related to confounding and measurement error in BLLs, relying on instrumental variable techniques to identify the causal effect of lead levels on future cognitive achievement. For identification, we exploit variation in lead safe housing over time induced by a policy change that required landlords to certify that their homes were lead safe. Neighborhoods with old housing stock and high lead burden in 1997 were targeted, and lead levels declined significantly faster in these area than in others. Importantly, neighborhoods that gained more lead-safe housing over this period (as measured by the number of certificates) do not appear to have changed over time in other dimensions that might influence child test scores.

We conclude by examining the role of lead in explaining large changes in the racial gap in test scores over our sample, conducting two exercises. We calculate how much of the decline in the gap in black-white test scores can be explained by declines in the gap in black-white lead levels over this period using our causal estimates.[14] We also directly examine the relationship

---

in overestimates of the amount of lead in the child's system if the measurement coincides with brief exposure that does not result in sustained BLLs.

[14] We find no evidence that the causal impact of lead exposure on test scores differs by race. Therefore we use a single causal estimate of the effect for this calculation.

between the falling gap in child lead levels and the falling black-white test scores in each county in Rhode Island. Below we describe the data we use for this analysis and our strategy for the IV estimates of the impact of lead on cognitive achievement in greater detail.

B. Data

Our sample consists of all children in the state of RI born between 1997 and 2004 with at least one BLL measure before the age of six, and who are enrolled in any RI public school. The data on BLLs comes from the RI Department of Health (RIDOH) and includes the age at each test, the test method (capillary or venous), the census tract in which the child lived at the time of the test, and the BLL. Educational outcomes come from the RI Department of Education (RIDE) and include NECAP[15] scaled test scores in grades three through eight, whether and when the child received an Individual Education Plan (IEP), and school lunch (free or reduced) status. The third source of data are birth certificate data for the child. These include: birth order, child gender, maternal race, maternal age and marital status at birth, maternal education, the child's birth weight and gestational age at birth as well as when the mother initiated prenatal care. The birth certificate data include a mother identifier so siblings can be identified in the dataset.

The final sample includes roughly 57,000 RI children born between 1997 and 2004 with linked BLL and educational test score data. Details on construction of the sample can be found in Appendix Table 1. Three things are worth noting. First, compared with a national lead screening rate of 25 percent, 80 percent of all children with a RI birth certificate are screened at

[15] NECAP stands for the New England Common Assessment Program. It is a series of exams developed collaboratively by the New Hampshire, Rhode Island and Vermont departments of education, with assistance from the National Center for the Improvement of Educational Assessments. The NECAP tests measure students' academic knowledge and skills relative to Grade Expectations which were created by teachers from the three states.

least once and the screening rate remained constant throughout the study period, suggesting that the types of children screened also likely remained constant. Second, using vital statistics data to compare the maternal characteristics (race, education, and marital status) of children with a lead level (80%) against those without a lead level (20%), we find that those with a lead screen are slightly more disadvantaged along most measures with the exception of birth weight which is the same for the two groups. Third, among those with a lead level, those matched with RIDE data are slightly more disadvantaged than those who we were not found in RIDE data (presumably because they attended private school). However, children with matched lead-RIDE data are representative of all children in the RIDE data.

Children have on average 4.7 BLL measures over the first 72 months of life (median of 4), after which they are typically not tested.[16] To construct a single measure we calculate the arithmetic mean of all samples, but also consider alternative measures: The maximum level, the geometric mean, the median BLL, a sum of all BLLs and an approximation of the "area under the curve" for which we assume a linear trend between two lead levels and calculate the mean over all 72 months.

While all children in our sample have a third grade reading test score, the number of children with 4^{th} – 8^{th} grade scores declines steadily so that only 18,000 children (those born in 1997 and 1998) have 8^{th} grade test scores. The decline in the availability of later test scores is due almost entirely to the age of the cohorts in 2010, the last year for which we have data. There does not appear to be any other significant source of attrition (eg – migration out of the state or absence of scores for other reasons such as IEP). For example, of those children born in 1997, we have 3^{rd} grade test scores for 7,084 students and of those, we have 8^{th} grade test scores for 6,756 of them, losing less than 5% of the sample.

[16] When there is a capillary and venous sample available for the same month, we drop the capillary measure.

To estimate the effect of early childhood lead levels on future child outcomes we exploit variation in BLLs induced by changes in policies and resources devoted to reducing exposure to lead among RI children over the period 1997-2010 in an instrumental variable (IV) framework. Below we discuss the nature of RI lead mitigation policies that underlie the exogenous variation we exploit for identification and the construction of our instrument.

C. Background on Lead Mitigation Policies in RI

Rhode Island, and particularly its urban areas, is characterized by old housing, with much of it (43 percent) built prior to WWII, and therefore containing the greatest concentration of leaded paint. As a result, many of the state's children are exposed to lead through deteriorating lead-based paint, with African American and low income children disproportionately exposed.[17] This disproportionate exposure is due largely to the residential segregation of the poor, and especially African Americans, in the core urban areas of the state located within the county of Providence. US Census data reveals that 81 percent of the homes in Providence county were built prior to 1978 and 49 percent before WWII (Appendix Table 2). Comparable numbers for the rest of the state are 68 and 27 percent. Moreover, low income and African American families are much more likely to live in Providence than elsewhere: 86 percent of African Americans live in Providence compared with 51 percent of whites, while 77 percent of the poor live in Providence compared with 55 percent of the non-poor. The disproportionate exposure of African American children does not just reflect poverty: Even conditional on poverty, African

[17] Previous work has established that within RI, children living in high (top quintile) poverty neighborhoods are nearly four times more likely to have elevated BLLs than those in low (bottom quintile) poverty neighborhoods (Vivier et al, 2011).

Americans are more likely to live in Providence (89 percent of poor African Americans compared to 60 percent of poor whites).

In an effort to reduce childhood exposure to lead through old deteriorating paint, RI state policy makers established two programs to encourage lead hazard mitigation in old homes. The first required all owners of homes in which an elevated lead level was reported to the DOH to mitigate such hazards. Once the home was mitigated, an inspection would be performed and the Department of Health would issue a certificate (DOH certificate). Non-compliant landlords were referred to the state Attorney General for presecution. The second was administered by the RI Housing Resources Commission (HRC) and required all landlords to mitigate lead hazards in the homes they rented – regardless of whether an elevated lead level was reported. In particular, the state provided landlords with training in the importance of lead hazard mitigation, information and training on how to reduce lead hazards in the homes in the least costly manner, and they required all landlords to get a lead-safe certificate (HRC certificate) for homes they planned to rent. Section 8 voucher rentals and family day care centers were prioritized. There were no resources dedicated to enforcement and though non-compliance was illegal, landlords were typically not penalized for non-compliance. However, low-cost loans were made available to landlords to help finance mitigation efforts. Moreover, landlords could be sued in civil court when children living in their homes were found to have elevated lead levels if the home did not have a lead-safe certificate.

Over the period 1997-2010, the total number of lead-safe certificates issued to landlords increased from 333 to over 41,000 (Appendix Table 3). The state contracted with community based non-profit organizations to provide outreach to families and landlords in the four "core cities" of RI where the housing stock is the oldest, the families most disadvantaged and the lead

burden the greatest. As a result, census tracts that witnessed the greatest increase in certificates issued per capita were characterized by older housing stock, a greater share of children with high lead levels in 1997, and a greater share of poor families (Figure 1).[18] This growth in the number of lead safe certificates within a neighborhood over time provides the exogenously occurring variation in exposure to lead that is the basis of our empirical strategy.

A potential concern over our ability to use certificates as an exogenous source of variation in area lead levels is that areas that saw greater growth in certificates may have been characterized by other changes in demographic composition, for example, that could also influence test scores. This does not appear to be the case. Tracts with the greatest increase in certificates do not appear to have experienced other demographic changes over the period 1997-2010. They are no different in terms of changes in the share black, the share poor, median family income or housing prices over this period than tracts that witnessed no or small changes in certificates per capita (Figure 2). Regression results bear this out: increases in the number of certificates are not associated with significant (statistically or economically) changes in race, maternal education, marital status or income of families who live in the neighborhoods (Appendix Table 4).[19]

Our preferred specification relies only on the certificates issued by the RI DOH as instruments as there is less discretion in these certificates since they are required after an elevated lead level and the state's Attorney General threatening prosecution for non-compliance.

[18] Certificates are scaled by 100

[19] We can examine a limited measure of mobility in the first five years of life directly with our data since the census tract of residence is recorded at each blood draw. Seventy two percent of the sample moves at some point during the first 72 months of life and movers are disproportionately disadvantaged with higher blood lead levels, 4.1 vs. 3.2 (Appendix Table 5). However, they appear to move to geographically proximate and very similar areas in terms of the underlying BLLs of children living there, the share poor in the neighborhood and the number of certificates issued (Appendix Table 5). That is, while there is a lot of mobility, movers seem to move to areas that are close and nearly the same.

We do however also use the RI HRC certificates and the amount of low interest rate loans made to landlords to help with remediation efforts as instruments in extensions shown below and the results are very similar.

D. Instruments and First Stage

We collected data on the date and address of all certificates issued in the state of RI from 1997-2010. For each tract and year we calculated: the number of certificates issued by the DOH in the past 5 years (recall that these are required if a child tests above a certain threshold), the number of certificates issued by the HRC in the past five years, the total number of certificates (DOH and HRC) that had been issued, and the loan amount. Based on this information, we constructed two measures to account for the fact that tracts differed in their lead burden and thus need for certificates. The first measure scales the number of certificates or the loan amount by the number of housing units in the tract built before 1978 (Certificates/Housing Units Built Pre 1978). The second multiplies the number of certificates by the average lead level of all tests of children in the tract in 1997 (Certificates*Average Lead Levels in 1997).[20] Because these measures are linked with RI children based on the tract and year of birth, they are not a function of any (potentially endogenous) migration of the family after the birth of the child. The first stage is:

$$(1) \qquad Lead_i = \alpha_0 + \alpha_1 Certificates_{nt-5} + \alpha_2 \mathbf{X}_i^c + \alpha_3 \mathbf{X}_i^m + \alpha_4 \mathbf{X}^n_{tn} + \gamma_n + \gamma_t + \mu_i.$$

[20] The average lead level is actually the leave-out average (ie, removing the focal child from the calculation of the average).

The vector $\mathbf{X}^c$ includes the child's birth weight, gestation at birth, birth order, gender, measure of the child's free/reduced lunch status, and the month in which the child's mother initiated prenatal care (a measure included to capture prenatal investments). The vector $\mathbf{X}^m$ includes maternal characteristics such as maternal age at birth, marital status at birth, race, and educational attainment (years of schooling). The vector $\mathbf{X}^n$ includes time-varying neighborhood (tract) characteristics from census data including share poor, median family income, and the share housing built post 1978. [21] Year of birth and tract fixed effects are also included.

The estimates reveal a strong negative relationship between the number of DOH certificates issued in the five years prior to birth over the number of old housing units and child lead levels (Table 1, column 1). Moreover, the effects are concave in nature (Table 1, column 2). The results are also strong when we interact the number of certificates with the baseline lead level in the tract in 1997 (column 3). We present additional first stage estimates based on the number of certificates issued by the HRC (columns 4-6) and total certificates (columns 7-10). Finally, we include the amount of loans provided for lead hazard mediation and find a negative relationship between the loan amount and average lead levels in the neighborhood (columns 11 - 12). The F statistics on the excluded instruments range from 41 to 133 depending on the specification. We focus on results using the first instrument (column 2) because it is the most arguably exogenous, but also present second stage estimates based on the other instruments for comparison in the robustness section.

To interpret the magnitude of the coefficients, we calculate the relationship between lead certificates and average lead levels in the tract based on the average increase in certificates from 1997 to 2004, and the increase in the lead certificates among the top 10% of tracts, because of the skewed distribution of certificates. An average increase in DOH certificates reduced the

[21] Values for intercensal years from linear interpolations between census years.

average lead levels by 0.36 ug/dl, while a large increase in DOH certificates reduced lead levels by 0.69 on average. For HRC certificates, the equivalent reductions are 0.30 and 0.85 respectively.

IV. Results

A. Lead Levels and Child, Family and Neighborhood Characteristics

In RI, disadvantaged children are characterized by higher lead levels, as they are nationally. Among those born between 1997 and 2004 in RI, African American children have average lead levels of 4.8 compared with 4.2 for Hispanic children and 3.1 for white children. Likewise, children of less educated mothers have lead levels of 4.6 on average, compared to children with more educated mothers whose lead levels are 3.2. The same patterns are true if we classify children by whether they are eligible for free lunch or not (4.1 vs. 2.7), their mothers are single or married (4.2 vs. 3.0) and whether they live in a poor neighborhood (Table 2, column 1). For comparison, we present differences in third grade reading test scores for these same groups. Groups with high lead levels are characterized by lower reading test scores across all types (Table 2, columns 3 and 4).[22] But these aggregate measures mask significant heterogeneity in declines over time, which we characterize in the next section.

B. Trends in Lead Levels Over Time

[22] The higher lead burden born by disadvantaged children is not due entirely to neighborhood segregation. Even within a neighborhood, disadvantaged children are more likely to have elevated lead levels.

Overall, average lead levels have fallen dramatically since the 1997 birth cohort (Figure 3A). The average blood lead level for those born in 1997 was 4.25, falling to 2.5 for those born in 2004, a remarkable decline in only seven years.[23] The declines in lead levels over this period were greatest for the most disadvantaged children: Black children as well as those eligible for free lunch witnessed greater declines over this period than other groups, reducing, but not eliminating, the gap in lead levels (Figure 3B). However, inspection of the figures show that not only were the initial BLLs of free lunch students considerably lower than they were for African American children, but that the decline over time was also less steep. This is likely due to the fact that both generally and among the disadvantaged population in RI, African American children are disproportionately located in the four core cities which is where the highest concentration of old housing is located and the state targeted its efforts.

We investigate the role of the intervention in explaining the decline in these disparities. Figure 1 shows that neighborhoods with a greater share of African American and poor families received more certificates, with a stronger relationship for the former than the latter. We also estimate whether the certificates appear to have had a greater impact on the lead levels of certain groups within a neighborhood or tract. To do so we repeat the first stage stratifying by race and then interacting certificates with the following child characteristics: Race, free-lunch status, maternal education and child gender. The results suggest that in addition to certificates being disproportionately issued in neighborhoods with a high share of black families (Figure 1), conditional on the neighborhood, the certificates did disproportionately reduce the lead levels of African American children, consistent with their living in housing with greater risk of lead exposure (Table 3, columns 1-3). Interactions with free lunch (column 5) suggest that poorer

[23] Changes in the distribution of average lead levels for the 1998 vs. 2004 birth cohort are in Appendix Figure 2.

children are also disproportionately affected, but interactions with gender and maternal education are smaller and imprecisely estimated (columns 6 and 7).

C. Lead Levels and Future Child Cognitive Achievement – Preliminary Evidence

Visual inspection reveals a strong linear relationship between child BLLs in the preschool years and reading test scores as averaged over grades 3-8, when the children were aged 9 to 14 (Figure 4). Children with an average BLL of 0 ug/dl score nearly 5 points higher (36% of a standard deviation) than those with BLLs of 5. Similar patterns are observed for math test scores, as well as for the probability of having an IEP which increases from 20 percent for those with 0 BLLs, to 27 percent for those with a BLL of 5 and over 30 percent for those with a BLL of 10.

For comparison we also graph average maternal years of schooling by child lead levels in the final panel of Figure 4. One also observes a clear linear and negative relationship between maternal schooling and elevated BLLs. This can be indicative of confounding, but could also potentially be another source of exposure if low-income mothers have higher lead levels that they pass on to their children in-utero (Dort, Hurlbut and Hoyer-Hassen, 2004; Bellinger, 2005).

D. Lead Levels and Future Child Cognitive Achievement – OLS and IV Results

We present estimates of the impact of preschool BLLs on future academic achievement as measured by NECAP scaled scores. The equation estimated is as follows:

$$(2) \qquad Y_i = \beta_0 + \beta_1 Lead + \beta_2 \mathbf{X}_i^c + \beta_3 \mathbf{X}_i^m + \beta_4 \mathbf{X}^n_{tn} + \boldsymbol{\tau}_n + \boldsymbol{\tau}_t + \varepsilon_i.$$

Where Y_i is a measure of test scores (in reading or math, in grades 3-8), $\mathbf{X}^c$, $\mathbf{X}^m$ and $\mathbf{X}^n$ are defined as above and year of birth and tract fixed effects are also included. In the IV regressions, child lead levels are instrumented based on the first stage described above. For the first set of results, the outcome is the average reading test score over all grades for which a score is available (grades 3 through 8). The OLS results with no controls except cohort fixed effects suggest that going from a lead level of 0 to 5 (the threshold at which the CDC recommends action), would decrease test scores by 5 points or 38% of a standard deviation (Table 4A, column 1). Adding a set of controls that includes race, maternal education, marital status, free lunch, gender, birth order, birth weight, gestation, and prenatal care reduce the estimated coefficient from -1.029 to -0.283 (column 2), adding neighborhood characteristics from census data (column 3) reduces the coefficient only slightly to -0.264 and the inclusion of tract FE does little to the estimated coefficient (column 4).

IV estimates of the impact of average lead levels on child reading test scores are larger than the tract FE estimates and suggest that going from a lead level of 0 to 5 would decrease average test scores by 6 points, or 46% of a standard deviation (Table 4A, column 5). This larger IV estimate is consistent with considerable measurement error in lead levels. We explore the potential attenuation bias inherent in a mis-measured lead level by constructing an additional measure of lead based on a single, randomly drawn lead BLL for each child. This measure should reflect greater classical measurement error than the one based on multiple measures and this should be reflected in attenuation of the OLS estimate. When we regress test scores on this measure (Table 4, column 6), the relationship between lead levels and test scores is significantly lower (-0.161) than the estimate based on a lead level constructed from an average of multiple

test scores (-0.275). These results are consistent with considerable attenuation bias from mismeasurement of lead levels, which we explore further in a later section.[24]

We consider two additional specifications that include school FE and mother FE. The school FE results (columns 7 and 8) are similar in magnitude to the tract FE. When we include mother FE, we limit our comparison to differences in lead levels across siblings which are considerably smaller than cross-sectional differences in lead levels. Including a mother FE significantly reduces the coefficient estimate to -0.105 (Table 4A, column 9). This lower estimate is consistent with both omitted variable bias in the estimates that include only tract FE, but also with previous work showing that the inclusion of a maternal FE can exacerbate the attenuation bias from measurement error (Griliches and Hausman, 1986). When we include the maternal FE and instrument for lead levels, the estimate is considerably larger than the maternal FE estimator, -2.3 (Table 4A, column 10). Given the potential for exacerbating measurement error in the estimates and the loss of precision with mother FE, we focus on IV with tract FE in what follows.

We also consider one falsification test: estimating whether childhood lead levels in the first 72 months of life affect birth weight in OLS and IV settings. The results (Table 4B) suggest a negative relationship between childhood BLLs and birth weight in an OLS regression with tract FE and a full set of controls. This is consistent with negative selection into high lead levels. However, once we instrument for child BLLs, the estimated effect is essentially zero and imprecisely estimated. We interpret this result as support for our identification strategy based since it suggests that the instrument meets the exclusion restriction.

[24] Though the average is itself a noisy measure, if we consider it to be the true measure and the randomly drawn measure to be the noisy measure, this would suggest a reliability ratio of 0.58 and downward attenuation bias of 42%.

E. Extensions: Achieving Proficiency, Math Scores and Changes Over Time

The NECAP tests are used primarily to assess the extent to which students have met certain benchmarks. A score of at least 40 indicates "proficiency" in the subject (71% of students for reading), while a score below 30 indicates "substantial deficiency" (10%) and a score of greater than 56 indicates "passing with distinction" (21%). We examine the extent to which lead levels affect a student's reaching each of these three benchmarks in both reading and math test scores as well as changes in the effects of lead over time.

We find the largest effects of elevated lead in the lower tail of the distribution of test scores. Elevated lead levels are associated with a greater likelihood of substantial deficiency in both Reading (Table 5, panel A, columns 1 and 2) and Math (Table 5, Panel B, columns 1 and 2). Specifically, students with a lead level of 5 are 39 percent more likely to be substantially deficient in reading and 20 percent more likely to be substantially deficient in math. They are 60 percent less likely to meet the standard of proficiency in reading (Table 5, Panel A, columns 3 and 4). There are no effects found at the high end of the distribution of test scores for either reading or math.

We also explore the extent to which these estimated effects increase or decrease over time by examining the impact of lead on test scores (Table 6 Panel A) as well as reaching proficiency in 3rd and 8th grade (Table 6 Panel B). When we reduce the sample size for this analysis to children with third and eighth grade test scores (birth cohorts 1997 and 1998), the estimates become less precise in the IV setting. But the point estimates show that the negative effects of lead either remain the same or increase slightly over this period.

Finally, we explore whether the effects are heterogeneous across children. Specifically we estimate whether the estimated effect of lead on reading scores differs by child race or school lunch status. The results, not presented here, are too imprecisely estimated (particularly for the IV estimates) to say definitively whether the results vary by race or school lunch status.

In the next section we consider additional specifications to assess the robustness of our results.

F. Robustness

Alternative Measures of Lead

First we consider alternative measures of lead. To rule out the possibility that the results are being driven by outliers, we trim observations with lead levels in excess of 25 and the results are unchanged (Table 7 columns 1 and 2). We try to account for the timing of the lead measures in a second alternative measure. To do so we assume that between two lead measures, lead decays (or increases) linearly, interpolate lead measures for non-tested months based on this assumption, and calculate a mean over imputed values for all 72 months (we refer to this measure as the "area under the curve"). These results (columns 3 and 4) are smaller but still large. We also consider the sum of all lead levels which is designed to capture persistence (column 5 and 6). A fourth measure is the single highest measure recorded for each child (Max lead level). This measure is negatively and significantly associated with children's test scores though the estimates are lower in magnitude in both the FE and IV FE specifications (columns 7 and 8). We also calculate the geometric mean of lead as is sometimes done in this literature, which yields somewhat larger point estimates in both the FE and IV specifications (columns 9 and 10). Finally we consider the median lead level (column 11 and 12) which reduces the influence of outlier values and the

results are similar to the average lead level. In sum, regardless of how we measure lead levels, the estimated relationship between lead levels and test scores is unchanged.

Alternative Instruments

We also present IV estimates based on alternative sets of instruments: certificates at birth*initial(1997)lead level (Table 8 column 1), the number of certificates/pre-war housing units (column 2), the number of HRC certificates/old housing (columns 3 and 4), HRC and DOH certificates (columns 5 and 6) and the total number of combined certificates (columns 7 and 8). Finally we include as instruments the number of certificates/all homes and the share of all homes that are old (column 9) as well as the amount of loans offered to landlords for remediation (column 10). Estimates across all specifications range from -1.32 to -1.623 and the precision also varies. However, even the smallest point estimate suggests a strong relationship.

Mean Reversion/Convergence

An alternative explanation for our finding that areas with the greatest number of certificates issued witnessed the largest declines in lead is one of mean reversion or convergence. More specifically, since the areas (or groups of children) that received the greatest number of certificates were also the areas or groups with the highest lead levels, one would expect that over enough time, the high lead levels would eventually converge to the average lead levels, regardless of the number of certificates. In other words, even in the absence of the certificates, one might expect the higher lead levels of African American and poor children to eventually converge to the lower levels characterizing other children.

To address this possibility, we consider a number of alternative specifications. The first addresses convergence by race by including race*year FE in the econometric specification to

account for different time trends by race (Table 9 column 1-2). This specification allows for black and white children to have different trends, which might be the case, for example, if there were other measures that affected black children's test scores differentially.[25] However, this change in specification has little effect.

To rule out that convergence of high lead levels to low lead level areas is driving the results, we consider another specification includes an interaction between 1997 lead level in the tract and year of birth. The results (Table 9, columns 3-4) show that the inclusion of different (linear) time trends for low vs. high lead areas does not alter the findings.

We also consider limiting our sample to those areas with the highest lead burden in 1997 (defined as the top quartile). The idea underlying this test is to take areas that have initially similar high lead levels and show that lead falls more where there are more certificates. In contrast, mean reversion would suggest that the places with the highest initial lead levels would experience large declines with or without the certificate program. The IV point estimates are actually greater in magnitude but not statistically distinguishable from those generated from the full sample (Table 9, columns 5 and 6). Finally, we predict the number of certificates that a tract should receive based on 1997 lead levels as well as demographic characteristics of the tract and then limit our sample to those tracts predicted to receive the greatest number of certificates. Again, the results are very similar (Table 9, columns 7 and 8).

Measurement Error in Lead Levels and Attenuation Bias

[25] For example, it has been suggested to us that black test scores might have been differentially affected by No Child Left Behind, which was signed into law in Jan. 2002. Dee and Jacob (2010) review the literature on NCLB and find little evidence that the effects were greater for inner city black children, or black children at all. Moreover, what improvements were seen were found throughout the test score distribution, not only towards the bottom of the distribution as we find. It is important to keep in mind that the 3rd grade test scores we examine were measured between 2007 and 2013, well after the implementation of NCLB in any case.

We assess measurement error in lead levels by comparing OLS estimates based on more and less noisy measures of underlying lead levels. As noted previously, lead levels from capillary tests are noisier than venous tests. For our subsample of children for which we have both capillary and venous test results (n=10,235), we present two sets of estimates: one based on an average of all capillary measures and another on the average of all venous measures. This exercise is similar to the previous results using the randomly drawn measure that resulted in a significant downward bias on the order of 50 percent (Table 4A column 6, reproduced in Table 10, column 1). We find that the "noisy" capillary measures yield an estimated effect of -0.137 (Table 10, column 2) whereas the less noisy venous measure yields an estimate of -0.328 (Table 10, column 3), nearly three times greater. These results are consistent with significant measurement error in lead levels leading to attenuation of OLS estimates.

Finally, we follow the work of Ashenfelter and Krueger (1994) in which they assess the extent to which measurement error in reports of schooling biases estimates of the returns to schooling using a sample of twins in an IV framework. To do so, they exploit multiple measures of educational attainment: One self-reported and the other reported by the twin. With these two measures, they calculate reliability ratios and use a twin's report of his sibling's educational attainment as an instrument for the self-report. Other similar work includes Chalfin and McCrary (forthcoming) in which the authors use multiple measures of police force size to estimate the impact of policy on crime in an IV framework.

Like Ashenfelter and Krueger (1994) and Chalfin and McCrary (forthcoming), we have multiple measures of our variable of interest – lead exposure in early childhood. Our exercise differs not only because we typically have more than 2 measures for each child, but also because these multiple measures were often taken at different times and therefore the underlying child

lead levels may have differed. However, each measure can still be interpreted as an (imperfect) measure of the underlying level of lead exposure during early childhood.

Following these two papers, we present estimates of the impact of child lead exposure on future academic achievement based on a single BLL (not the average) and instrument for that measure with other BLL measures for the same child. More specifically, we limit our sample to those with at least two BLLs (n=45,663), draw a single BLL at random, then instrument for it with either a second randomly drawn BLL or with an average of all other BLLs for that child.

The first finding is that the correlation between two randomly drawn lead levels is roughly 0.463 (the coefficient on the second randomly drawn BLL in the first stage regression). This can be interpreted as the reliability ratio and suggests that nearly 54 percent of the measured variance in a single BLL is error. Interestingly, when we repeat the first stage instrumenting for randomly drawn BLL with an average of all other BLLs, the coefficient increases to 0.65, consistent with the average of multiple measures being a more reliable (less noisy) measure.

As expected, the estimate of the impact of the randomly drawn BLL increases from -0.168 in the tract FE specification to -0.384 in the IV specification that relies on a second randomly drawn BLL as an instrument (Table 10, column 4), representing an increase of 139%. The estimated coefficient increases to -0.395 when the average of the remaining BLLs is used as an instrument (column 5), though statistically these are indistinguishable. These estimates, while highly statistically significant, are considerably smaller than those obtained using instruments derived from the certificate programs. One reason is that these IV estimates represent the effect of lead on any child with multiple lead measures, whereas the IV estimates shown earlier represent the "treatment on the treated" effects of lead mitigation on children living in the

targeted areas. Thus, these different estimates may suggest heterogeneity in the effects of lead on test scores and/or greater ex poste mitigation by families of more advantaged children.

Recall that areas targeted for the certificates were overwhelmingly urban, poor, and African American. Some previous studies have suggested that disadvantaged children may be more negatively affected by lead exposure due to the fact that nutrition and cognitive stimulation, which are both weaker among disadvantaged children, may each help to counter the negative effects of lead on cognitive functions. For example, Ferrie, Rolf and Troeskin (2012, 2014) find that the introduction of leaded piping has negative effects on the IQ of sons born to low SES families, but not high SES families, which may reflect these children's better nutrition (especially calcium and iron) or greater cognitive stimulation. In animal studies, Guilarte et al (2002) has shown that an "enhanced environment" can reverse the negative effects of lead on cognitive functioning in rat pups. Recent work by Billings and Schnepel (2015) found that an intervention in North Carolina targeting children with very high levels of lead with a treatment that consisted of a bundle of medical and nutritional assistance as well as lead remediation reversed some of the negative effects of lead on cognition and behavior. We take this as evidence that the negative effects of lead on cognitive development may well be higher among the most disadvantaged children, thus potentially explaining why the IV estimates based on the targeted treatment are greater.

V. Implications for Disparities in Child Educational Outcomes

Finally we explore whether and to what extent African American children's greater lead burden can explain racial disparities in test scores. We previously documented how RI state

efforts to reduce lead levels among RI children disproportionately targeted urban neighborhoods with a greater initial lead burden and thus a greater share of African American and low income families (Figure 1). Even within neighborhoods, African American children's lead levels appear to have been disproportionately affected by state efforts to ensure that homes were lead safe (Table 3) and as a result of these efforts, the lead levels of African American, and to a lesser extent, low income children, have fallen disproportionately faster (Figure 3B).

The decline in racial disparities in lead levels over this period was accompanied by a decline in racial disparities in test scores as well. We present the distribution of third grade reading scores by race and year of birth (1997 vs. 2004) in order to examine how the gap has changed over time (Figure 5). These scores were measured over approximately 2006 to 2013, when the children would have been about 9 years old. We are not aware of other measures in Rhode Island that would have been expected to have a differentially large effect on the test scores of black children in inner city neighborhoods over this seven year time period.

The test scores of white children have improved slightly, but black test scores have improved considerably more, thereby sharply reducing the racial gap in test scores. Calculating disparities based on average grade 3 test scores by race, we find that the racial gap in test scores fell from 9.2 points for those born in 1997 to 6.3 for those born in 2004, while the gap in lead scores fell from 2.2 to 0.9 over this period. Based on our causal estimates of the impact of lead on test scores, we calculate that the falling disparity in lead levels would correspond to a 1.56 point decline in the racial gap in test scores, explaining 54 percent decline in the racial test score gap witnessed over this short interval. While we do not have the data to directly examine the effect of lead mitigation on the racial gap in test scores, our work would appear to confirm speculation by others (c.f. Lang, 2006) that lead abatement could be responsible for a third or

more of the narrowing of the black-white test score gap over the 80s and 90s. It has not previously been possible to examine this hypothesis rigorously.

This narrowing does not simply represent trends in lead levels and test scores by family income. When we examine trends in disparities in lead and test scores by free lunch status, the declines in both are considerably smaller: over this same period, the income gap (linear difference) in lead levels fell from 1.83 to 0.99, while the test score gap fell from 9.3 to 8.4 (67% of a standard deviation to 60%). That the decline in the disparity in lead levels by race was so much larger than the decline by income is likely due to two factors: that African American children bore a much greater lead burden in 1997 than poor children more generally, and that African American children were more likely to reside in areas of the state that were targeted for lead hazard mitigation – the four core urban centers.[26]

We examine the relationship directly, relating racial gaps in lead levels to racial gaps in test scores at the municipality level. To do so, we calculate for each of the 40 municipalities in RI and birth cohort, the racial disparities in lead levels and test scores, as measured by the ratio of average levels by race.[27] We then regress disparities in test scores (the ratio of white/black test scores) on disparities in lead levels (the black/white ratio) including municipality and year of birth fixed effects, as well as controls for some time varying characteristics of mothers residing in the municipality for each birth cohort (average maternal age, education level, marital status and free lunch status). These estimates show that as the racial gap in lead levels increases, the gap in test scores likewise increases. Over this period, the ratio of black/white lead levels

[26] These differences in trends by income and race are comparable to national trends. According to the CDC, the share of BLLs above 5 ug/ul fell more quickly for African Americans than for poor children over the period 1999-2010. In particular the share of African American children with BLLs above 5 fell from 18.5 to 5.6 of the national population, while for children below 130% of the FPL, the share fell from 12.9 to 4.4. (MMWR, 2013)

[27] The total number of observations, 187, is much lower than the 320 municipality*birth cohorts (40 municipalities x 8 birth cohorts) because of the many municipality*birth cohorts with no black children with lead levels. Moreover, because of the large number of cells with very few blacks, we weight the regressions by the number of black children with lead levels in our data.

declined from 1.6 to 1.4, which would, according to the OLS estimate in Table 11, column 1, correspond to a decline in the white/black test score ratio of 0.015.[28] Over this period, the white/black test score ratio declined from 1.25 to 1.15 (a difference of 0.10), and the decline in racial disparities in lead explains 15% of this decline. When we instrument for that black-white gap in lead levels with the number of certificates (Table 11, column 2), the resulting estimates are larger and suggest that the decline in lead levels witnessed over this period resulted in a decline in the ratio of white/black test scores of 0.06, or 63% of the observed decline, which is similar to our estimate that was based on coefficients obtained from the individual-level data.

Lead remediation in Rhode Island was implemented in a way that was relatively inexpensive. The annual spending on the programs we analyze here was approximately a half a million dollars annually, or $4 million over an eight year period. Thus, the cost per African-American child in our sample was about $679. This expenditure resulted in an improvement of .2 standard deviations in African American test scores relative to white test scores. Relative to other education expenditures that have yielded similar or smaller improvements in test scores, this is a very inexpensive intervention. For example, the Tennessee STAR experiment, which reduced class sizes from about 22 to 15 in the treatment group, cost more than $8000 per child. This intervention also increased test scores by .2 standard deviations in the short run, though the longer run effects were smaller (unlike the longer run effects of lead remediation) (Krueger, 2003).

[28] We do not present the first stage estimates, but the instrument does significantly predict reductions in lead levels within tract and county over time.

VI. Conclusions

The racial gap in family income has, depending on the measure considered, either held steady or increased over time.[29] As a result, one would not necessarily have expected the racial gap in child test scores to decline, but it did. We consider a new explanation for both racial disparities in test scores and their recent decline – African American children's disproportionate exposure to environmental toxins that can affect cognitive achievement. We find that since 1997, when the state of RI instituted measures to reduce lead hazards in the homes of RI families, lead levels fell across the state, but significantly more so for African American children. This differential effect is likely because their lead levels were considerably higher than other children in the state in 1997, including other low income children, and African American families were disproportionately located in the areas where lead abatement outreach efforts were focused. We find that this policy led to reductions in lead exposure that can explain a large fraction of the reductions in the black-white test score gap in RI witnessed over this period.

Jencks and Phillips (2008) have argued that eliminating black-white disparities in test scores would dramatically reduce not only educational inequality but economic inequality as well. Moreover, they argue that traditional sources of the gap (eg, inequalities in school inputs or family income), likely explain only a small share of the racial gap in test scores, and alternative sources of the gap should be considered. Our results suggest that environmental regulation, when targeted to those children at greatest risk, is effective in reducing environmental exposure to toxins and that declines in racial disparities in exposure, test scores, and, ultimately, economic outcomes, could follow.

[29] The share of African Americans in poverty increased from 23 to 28% between 2000 and 2011, after years of steady declines and for whites it increased from 7 to 10% (Pew Research Center). The ratio of White to Black income increased from 1.8 to 1.9 over the period 1983-2010; however, the wealth ratio has increased from 5.3 to 5.7 (Urban Institute, Interactive Race Graphic).

References

Ashenfelter, Orley and Alan Krueger (1994) “Estimates of the Economic Return to Schooling from a New Sample of Twins” The American Economic Review, 84(5): 1157-1173.

Bellinger, DC (2005). "Teratogen update: lead and pregnancy".*Birth defects research. Part A, Clinical and molecular teratology* 73 (6): 409–20.

Bellinger, D. C. (2004) “Lead.” *Pediatrics*, **113**(4 Suppl):1016–22.

Bellinger, D. C. (2008) “Very Low Lead Exposures and Children's Neurodevelopment.” *Current Opinion in Pediatrics*,**20**(2):172–77.

Bellinger DC (2008) “Neurological and behavioral consequences of childhood lead exposure” PlosMedicine.5(5):e115-e117.

Berney, B. (1993) “Round and round it goes: The epidemiology of childhood lead poisoning, 1950–1990.” *The Milbank Quarterly*. 71:3–39.

Billings, Stephen and Kevin Schnepel (2015) “Life Unleaded: Effects of Early Interventions for Children Exposed to Lead” unpublished mimeo.

Brooks-Gunn, J., Klebanov, P. K., Smith, J., Duncan, G. J., & Lee, K. (2003). The Black-White test score gap in young children: Contributions of test and family characteristics. *Applied Developmental Science*, *7*(4), 239-252.

Brown, Phil (1995) “Race, Class, and Environmental Health: A Review and Systematization of the Literature.” Environmental Research,69(1): 15–30

Brulle, Robert J. and David N. Pellow (2007) Environmental Justice: Human Health and Environmental Inequalities Annu. Rev. Public Health. 27:103–24

Burda, Martin and Matthew Harding (2014) “Environmental Justice: Evidence from Superfund Cleanup durations” Journal of economic behavior and organization 107: 380-401.

Canfield, RL, Henderson CR, Cory-slechta DA, Cox C, Jusko TA and BP Lanphear (2003) “Intellectual impairment in children with blood lead concentrations below 10 microg per deciliter” new England Journal of Medicine, 248(16):1517-26.

Chandramouli, K, Steer CD, Eillis M, and AM Emond (2009) “Effects of learly childhood lead exposure on acadmici performance and behavior of school aged children” Arch Dis Child 94:844-848.

Chay, Kenneth, Guryan, Jonathan and Bhashkar Mazumder (2009) "Birth Cohort and the Black-White Achievement Gap: The Roles of Access and Health Soon After Birth" NBER WP 15078.

Chay, Kenneth and Michael Greenstone (2005) "Does air quality matter? Evidence from the housing market" Journal of Political Economy 113(2):376-424.

Cook, M. D., & Evans, W. N. (2000). Families or schools? Explaining the convergence in white and black academic performance. *Journal of labor Economics*, *18*(4), 729-754.

Chalfin, Aaron and Justin McCrary (forthcoming) "Are US Cities Underpoliced? Theory and Evidence" *Review of Economics and Statistics*

Card, David and Jesse Rothstein (2007) "Racial Segregation and the Black-White Test Score Gap." Journal of Public Economics 91, 11-12: 2158-2184

Currie, Janet. (2011) "Inequality at Birth: Some Causes and Consequences," American Economic Review, 101(3): 1-22.

Currie, Janet, Davis, Lucas, Greenstone, Michael and Reed Walker (2015) "Environmental Health Risks and Housing Values: Evidence from 1600 Toxic Plant Openings and Closings," American Economic Review, 105(2):678-709.

CDC MMR Weekly "Blood Lead Levels in Children Aged 1–5 Years — United States, 1999–2010 April 5, 2013 / 62(13);245-248

Dart, R.C., Hurlbut, K.M., Boyer-Hassen, L.V. (2004). "Lead". In Dart, RC. Medical Toxicology (3rd ed.). Lippincott Williams & Wilkins.

Dee, T. S. (2005). A teacher like me: Does race, ethnicity, or gender matter?. *American Economic Review*, 158-165.

Dee, Thomas and Brian Jacob. "The Impact of NCLB on Students, Teachers, and Schools," *Brookings Papers on Economic Activity*, Fall 2010, 2007-2013.

Ewart, Stephanie. "The Decline in Private School Enrollment" SEHSD Working Paper Number FY12- 117 January, 2013, U.S. Census Bureau Social, Economic, and Housing Statistics Division

Ferrie, Joseph, Rolf, Karen and Werner Troesken (2014) "Attainment, Income and Longevity: Differential Impact of Bhildhood Household Socioeconomic Status and the Perpetual Disadvantage Across Generations, 1890-2010"

Ferrie, Joseph, Rolf, Karen and Werner Troesken (2012) "Cognitive disparities, lead plumbing, and water chemistry: prior exposure to water-borne lead and intelligence test scores among World War Two U.S. Army enlistees." Econ Hum Biol. 10(1):98-111.

Figlio, D. N. (2005). *Names, expectations and the black-white test score gap* (No. w11195). National Bureau of Economic Research.

Filippelli, GM, Laidlaw, MAS, Latimer, JC and R Raftis (2005) Geological Society of American Today, 15(1):4-11.

Finkelstein, Y, Markowtiz, ME and JF Rosen (1998) "Low level lead-induced neurotoxicity in children: an update on central nervous system effects' Brain Research Review 27(2):168-176.

Fryer Jr, R. G., & Levitt, S. D. (2004). Understanding the black-white test score gap in the first two years of school. *Review of Economics and Statistics*, *86*(2), 447-464.

Griliches, Zvi and Jerry Hausmn. (1986). "Errors in Variables in Panel Data." Journal of Econometrics. 31:93-118.

Grönqvist, Hans, Nilsson, Peter and Per-Olof Robling, (2014) "Early-childhood Lead Exposure and Criminal Behavior: Lessons from the Swedish Phase-out of Leaded Gasoline" SOFI working paper 2014/9, Stockholm University.

Guilarte, Tomás R., Toscano, Christopher D., McGlothan, Jennifer L. and Shelley Weaver (2002) "Environmental enrichment reverses cognitive and molecular deficits induced by developmental lead exposure." Annals of Neurology, 53(1):50-56.

Hanushek, E. A., & Rivkin, S. G. (2006). *School quality and the black-white achievement gap* (No. w12651). National Bureau of Economic Research.

Hanushek, E. A. (2001). Black-white achievement differences and governmental interventions. *American Economic Review*, 24-28.

Jencks, Christopher and Meredith Phillips, eds. (2008) The Black-White Test Score Gap. Washington DC: Brookings Institution Press.

Johnson, William R., and Derek Neal (1998). Basic skills and the black-white earnings gap. *The Black-White test score gap*, 480-497.

Jones, Robert, Homa, David, Meyers, Pamela, Brody, Debra, Caldwell, Kathleen, Pirkle, James and Mary Jean Brown (2009) "Trends in Blood Lead Levels and Blood Lead Testing Among US Children aged 1 to 5 Year, 1988-2004" Pediatrics 123(3): E376-85.

Krueger, Alan. "Economic Considerations and Class Size," The Economic Journal, 113 (February), F34–F63.

Krueger, A. B., & Whitmore, D. M. (2001). *Would smaller classes help close the black-white achievement gap?* (No. 451). Industrial Relations Section, Princeton University.

Lang, Kevin. *Poverty and Discrimination*, Princeton University Press: Princeton NJ, 2006.

Lanphear, Bruce P., Richard Hornung, Jane Khoury, Kimberly Yolton, Peter Baghurst, David C. Bellinger, Richard L. Canfield, et al. (2005). "Low-level environmental lead exposure and children's intellectual function: an international pooled analysis." Environmental Health Perspectives 113(7): 894-899.

Levin, R., Brown, M. J., Kashtock, M. E., Jacobs, D. E., Whelan, E. A., Rodman, J., … Sinks, T. (2008). Lead Exposures in U.S. Children, 2008: Implications for Prevention. *Environmental Health Perspectives*, *116*(10), 1285–1293

Lidsky, Theodore and Jay Schneider (2003) "Lead neurotoxicity in children: basic mechanisms and clinical correlates" Brain: A Journal of Neurology. 126:5-19.

McLaine, Pat, Navas-Acien, Ana, Lee, Rebecca, Simon, Peter, Diener-West, Marie and Jacqueline Agnew (2013) "Elevated Blood Lead Levels and Reading Readiness at the Start of Kindergarten" Pediatrics, 131(6):1081-1089.

Magnuson, Katherine and Jane Waldfogel, eds. (2008) Steady Gains and Stalled Progress: Inequality and the Black-White Test Score Gap. New York: Russell Sage Press.

Masters, Roger D, Hone, Brian and Anil Dosh (1998) "Environmental pollution, Neurotoxicity, and Criminal Violence" in J. Rose., ed., *Environmental Toxicology: Current Developments* (London: Taylor and Francis, 1998), pp. 13-48.

Mielke, H. W., and S. Zahran. (2012) "The Urban Rise and Fall of Air Lead (Pb) and the Latent Surge and Retreat of Societal Violence." *Environment International*, **43**: 48–55.

Mohai, Paul and Robin Saha (2006) "Reassessing racial and socioeconomic disparities in environmental justice research" Demography 43(2):383-399.

Nevin, R..(2007) "Understanding International Crime Trends: The Legacy of Preschool Lead Exposure." *Environmental Research*, **104**(3): 315–36.

Nigg, Joel, Nikolas, Molly, Knottnerus, Mark, Cavanagh, Kevin and Karen Friderici (2010) "Confirmation and extension of association of blood lead with attention definit/hyperactivity disorder (ADHD) and ADHS symptom domains at population-typical exposure levels" Journal Child Psychology and Psychiatria, 51(1):58-65.

O'Neil SG (2007) "Superfund: evaluating the impact of executive order 12898**."** *Environ Health Perspec***,** 115:1087-1093.

Reber, Sarah (2010) School desegregation and the educational attainment for blacks. Journal of Human Resources, 45 (4):893-914.

Reyes, Jessica Wolpaw (2015) "Lead Exposure and Behavior: Effects on Aggression and Risky Behavior among Children and Adolescents," *Economic Inquiry,* 53:3.

Reyes, Jessica Wolpaw (2015) "Lead Policy and Academic Performance: Insights from Massachusetts," *Harvard Educational Review,* 85(1).

Ringquist, Evan (2005) "Assessing evidence of environmental inequities: a meta-analysis" Journal of Policy Analysis and Management 24(2): 223-247.

Rogers, Michelle, Lucht, James A., Sylvaria, Alyssa, Cigna, Jessica, Vanderslice, Robert and Patrick M. Vivier. (2014) "Primary Prevention of Lead Poisoning: Protecting Children From Unsafe Housing." American Journal of Public Health. 104(8): e119-e124.

RI Dept of Health (2002) "Lead Poisoning in Rhode Island: The Numbers (2002). Providence, RI: Rhode Island.

US Dept of Education NAEP 2012 Trends in Academic Progress NCES 2013-456.

U.S. Environmental Protection Agency, "Lead and a Healthy Diet: What You Can Do to Protect Your Child," 2015, www2.epa.gov/sites/production/files/documents/nutrition.pdf.

Vivier, Patrick, Hauptman, M, Weitzen, SH, Bell, S, Quilliam, DN and JR Logan (2011) "The important health impact of where a child lives: neighborhood characteristics and the burden of lead poisoning" Journal of Maternal and Child Health, 15(8):1195-1202.

Fig 1: Certificates and 1997 Tract Characteristics

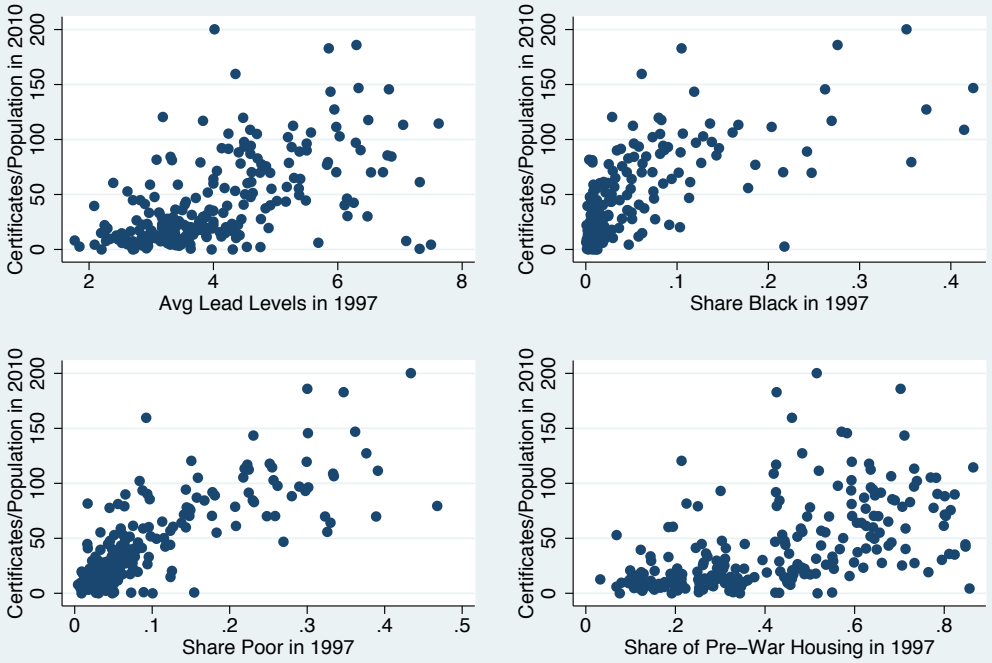
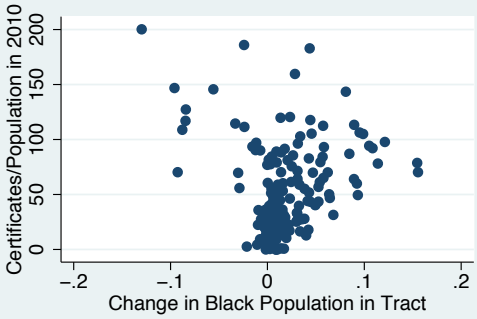
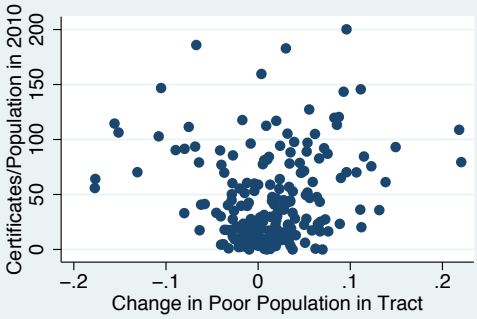


Fig 2: Certificates and Changing Tract Characteristics





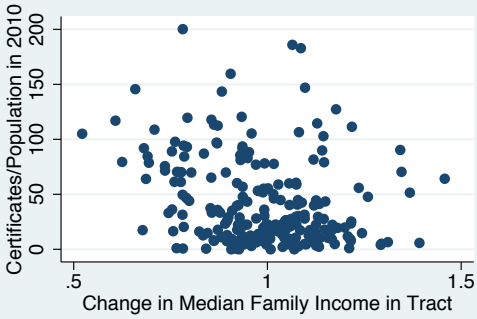


Figure 3A: Average Preschool Lead Level by Birth Cohort

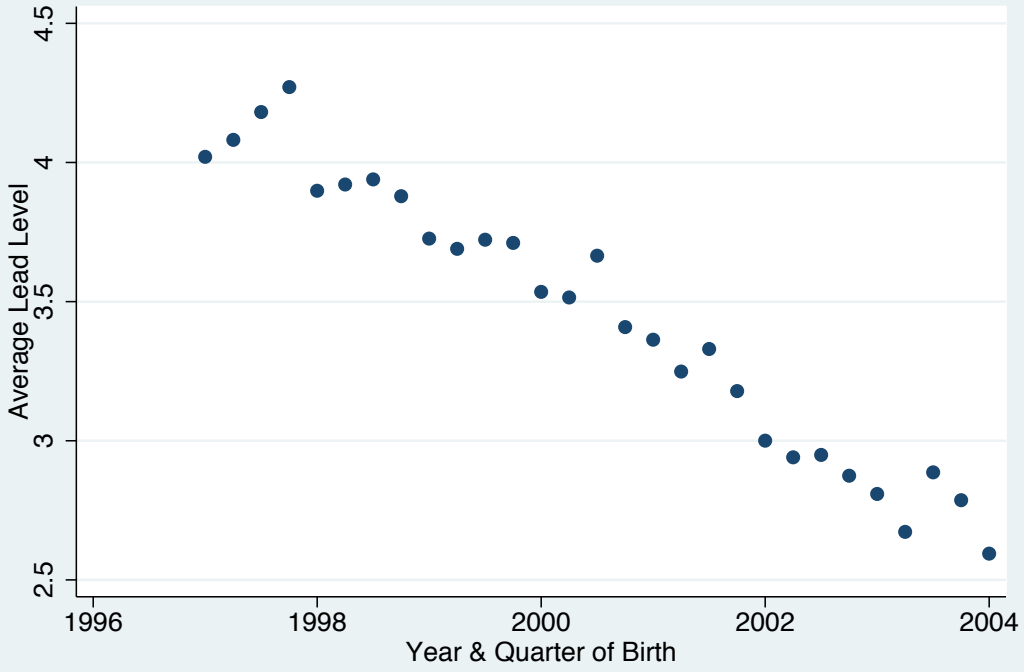


Figure 3B: Trends in Lead by Child Characteristic

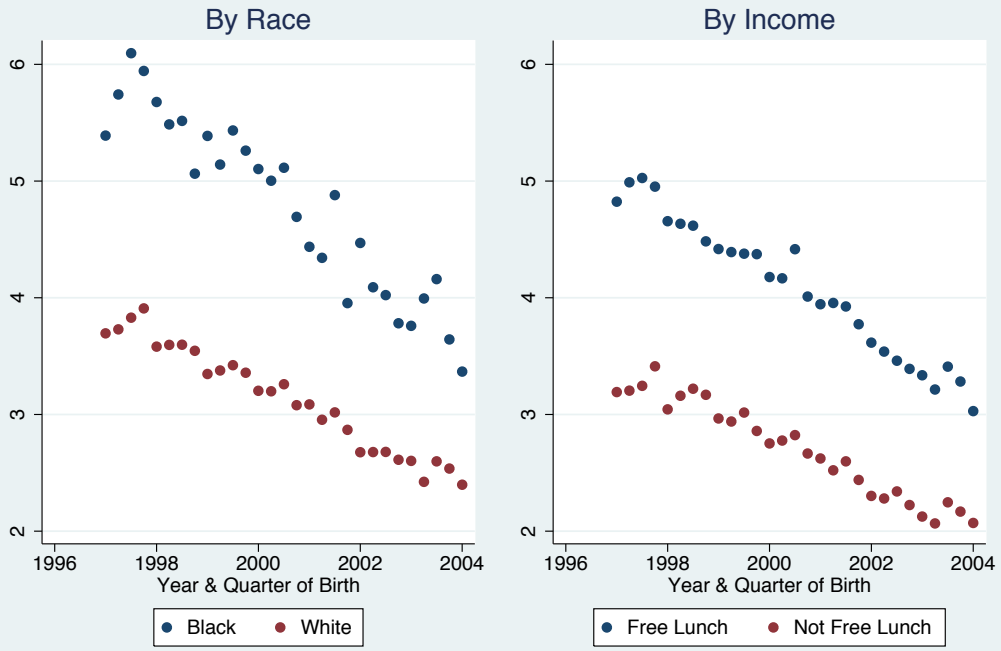


Figure 4: Lead Levels, Test Scores and Maternal Education

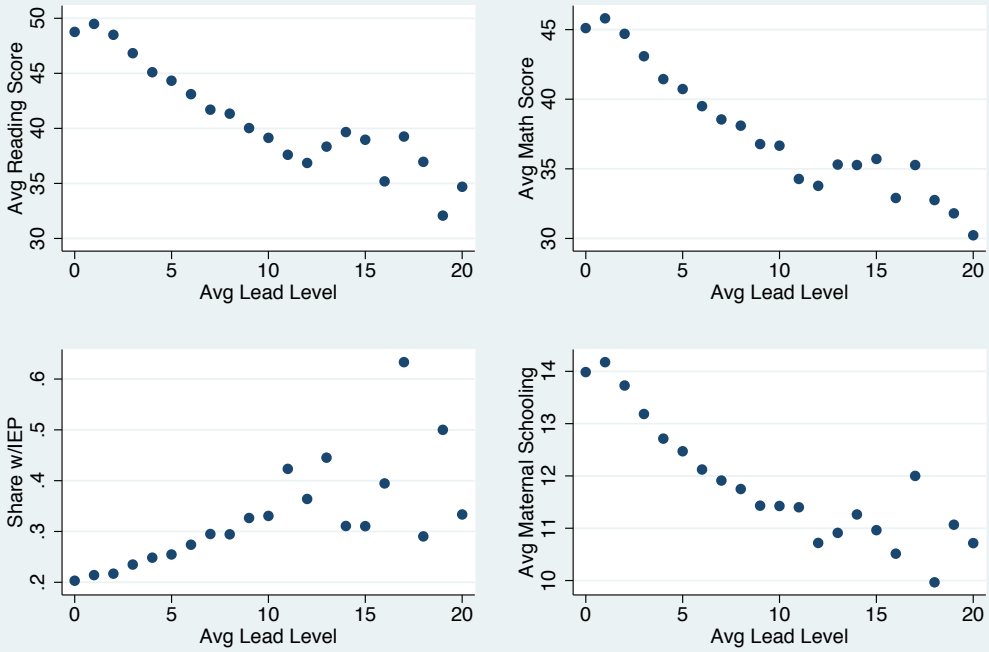


Figure 5A: Reading Scores by Race

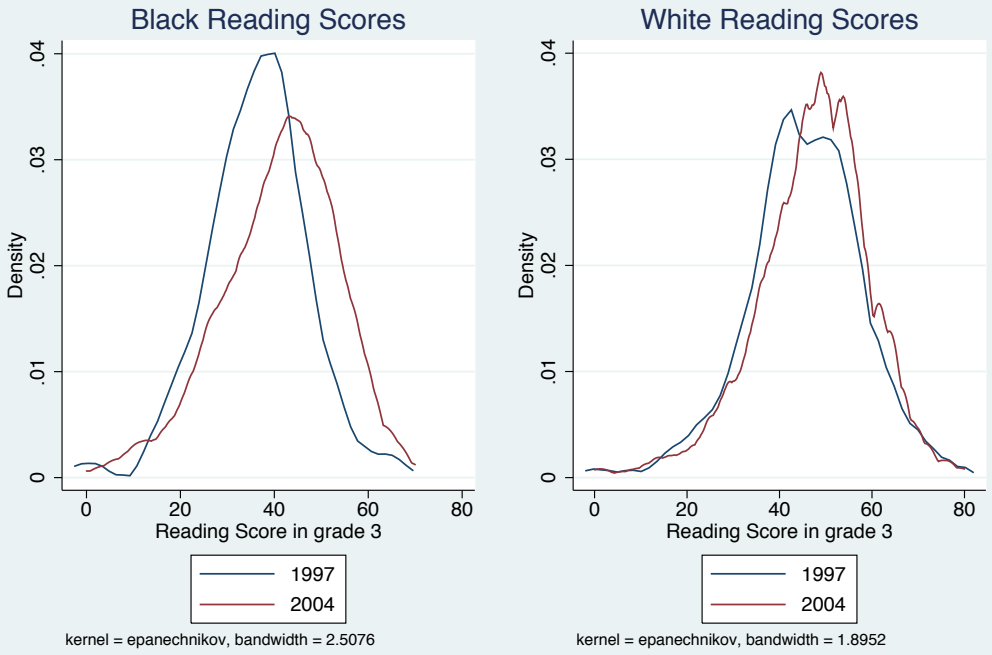
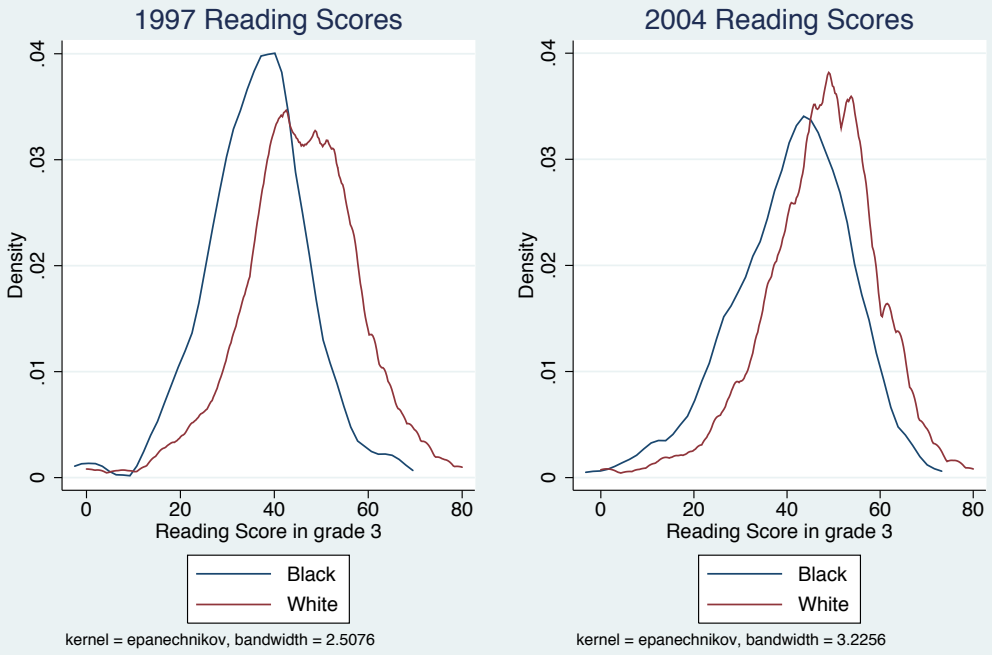
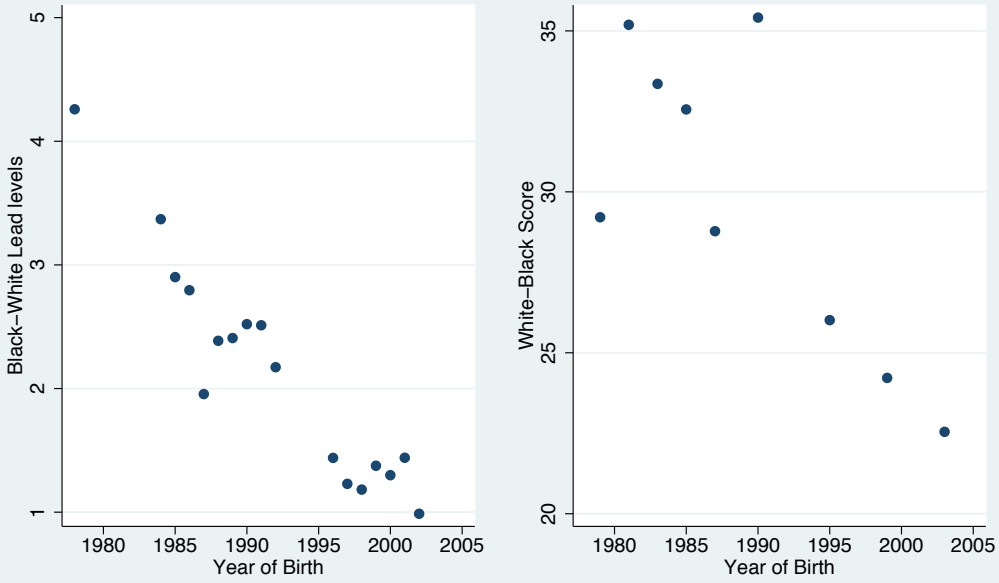


Figure 5B: Reading Scores by Year



App Figure 1: National Black–White Differences in Lead and Test Scores by Birth Cohort



Note: Lead data come from combined NHANES data and test scores from 9 Year old NAEP English tests

Appendix Figure 2: Average Lead Levels Over Time