The Long-Run Spillover Effects of Pollution: How Exposure to Lead Affects Everyone in the Classroom*

Ludovica Gazze
Department of Economics, University of Warwick
Ludovica.Gazze@warwick.ac.uk

Claudia Persico,
Department of Public Administration and Policy, School of Public Affairs, American University and IZA
cpersico@american.edu

Sandra Spirovska
Department of Economics, University of Wisconsin-Madison
spirovska@wisc.edu

November 20, 2020

Abstract

Children exposed to pollutants like lead are more disruptive and have lower achievement. However, little is known about whether lead-exposed children affect the long-run outcomes of their peers. We estimate these spillover effects using new data on preschool blood lead levels (BLLs) matched to education data for all students in North Carolina public schools. We compare siblings whose school-grade cohorts differ in the proportion of children with elevated BLLs, holding constant school and peers' demographics. Having more lead-exposed peers is associated with lower high-school graduation and SAT-taking rates and increased suspensions and absences. Peer effects are larger for same-gendered students.

* We are grateful to the North Carolina Division of Public Health and the North Carolina Education Research Data Center for providing the merged blood lead level data linked to education records data used in this analysis. The opinions expressed herein do not represent the views of the North Carolina Departments of Public Instruction or Health. We are thankful to Manuel Bagues, Marco Bertoni, James Fenske, Fanny Landaud, Victor Lavy, Lauren Russell, Viviane Sanfelice, and seminar and conference participants at UChicago Urban Labs, University of Pennsylvania, York University, APPAM, Temple University, University of Warwick, OSWEET, and WE_ARE for helpful comments and suggestions. We are also grateful to the Institute for Research on Poverty at the University of Wisconsin-Madison, the Becker Friedman Institute at the University of Chicago, and American University's School of Public Affairs and Provost for their initial financial support for data construction. All errors are our own.

I. Introduction

A growing literature shows that early life exposure to pollution hinders health and human capital accumulation (Persico, Figlio, and Roth 2020; Alexander and Currie 2017). For example, lead-poisoned children are more likely to be suspended and commit crimes (Aizer and Currie 2019; Reyes 2015) and have worse academic achievement (Ferrie, Rolf, and Troesken 2012; Grönqvist, Nilsson, and Robling 2020), consistent with lead's impacts on children's neurological development. These associations manifest at blood lead levels (BLLs) as low as 1–2 micrograms per deciliter (μg/dL) of blood (Aizer et al. 2018; Feigenbaum and Mueller 2016). Lead may also affect children's disability status (Gazze 2016). These negative effects of lead exposure are costly to children, families, and society in terms of reduced tax revenues and increased expenditure on special education, crime, and health care (Reyes 2014). Recent estimates suggest that at least half a million young children are still poisoned by lead each year in the US (Aizer et al. 2018). Low-income children are up to 12 times more likely to have elevated BLLs (CDC 2005), and Black children are more than twice as likely to be lead-poisoned than their White peers (CDC 2005).

So far, the literature has focused on estimating the effects of pollution and lead poisoning[1] on directly exposed children. However, these children interact daily with peers, which could lead to long-run spillover effects. In this paper, we ask whether there are spillover effects of lead poisoning on children who are not directly exposed to lead, but are exposed to school peers with high blood lead levels. Because children exposed to lead are more disruptive, have lower achievement, and engage in risky behavior, the effects of lead exposure might spill over to affect everyone in the classroom. These spillovers also may have long term consequences for students, but less is known about the long-run impacts of

[1] In this paper, we use the words lead poisoning and lead exposure interchangeably.

childhood peers. Thus, the spillover effects of lead exposure are an unexplored mechanism through which social context, pollution, and built environment could affect schools and children's outcomes.

Rigorously estimating peer effects is challenging because peers influence each other simultaneously, so it is unclear whether a disruptive child causes their classmates to misbehave, or whether the classmates cause them to be disruptive (i.e., the reflection problem). In addition, peer groups are not randomly assigned; they are selected based in part on unobserved characteristics (i.e., the selection problem). Using a novel identification strategy and data set, we plausibly estimate how early health shocks (i.e., pollution exposure) spill over within school contexts. We use rich education data from public schools in North Carolina linked to data on children's BLLs measured by age six and compare siblings whose cohorts happen to randomly differ in the proportion of children with elevated preschool BLLs in their grade-cohort in the same school. Our preferred specification includes family, school, grade, birth month, birth order, and year fixed effects, and controls for a broad set of time varying child and cohort demographic characteristics, as well as school quality. Since elevated BLLs have been linked to behavioral incidents, criminality, and lower test scores, we use them as a proxy for peers with potentially disruptive behavior and lower academic achievement. This methodology avoids the reflection problem because a child cannot affect the BLLs of their peers, but elevated BLLs might affect children negatively, which in turn might affect peers. Including family fixed effects mitigates the selection problem by controlling for unobserved family characteristics that could be correlated with both peers' quality and child's outcomes, such as the propensity to attend schools with fewer lead-poisoned children. Controlling for peers' race and economic status suggests that our estimated effects are due to lead poisoning and not peer demographics.

We find that a ten percent increase in the share of peers in a child's cohort that are exposed to lead is associated with a 0.3 percentage point increase in the

likelihood of suspension from school, a 2.8 percent increase. A ten percent increase in the share of peers in a cohort exposed to lead is also associated with a 1.7 percentage point decrease in the likelihood that a child graduates high school, a 2 percent decrease in the graduation rate. We also find that having more lead-exposed children in a child's cohort is associated with a higher likelihood of chronic absenteeism and dropping out of school, and a decrease in the likelihood of taking the SAT. Disruptive peers disproportionally affect the outcomes of economically disadvantaged and Black students, suggesting that the spillover effects of pollution could be contributing to persistent inequality in human capital accumulation. These findings generally hold in samples limited to locations where we measure lead exposure and sibling matches less noisily, and where students are less likely to switch schools in response to disruptive peers. Our results are also largely robust to further testing for school-switching directly and by including sibling-by-school fixed effects.

We find that disruptive peers disproportionally affect same-gendered students, and especially those students who also share the same race. Furthermore, we find that students going to school with a higher share of lead-exposed peers are more likely to be involved in incidents leading to suspensions with these disruptive students. Finally, exposure to disruptive peers in middle school, rather than elementary school, appears to drive long-run outcomes. We interpret our results as suggestive that homophily in network formation might drive the spillover effects of lead poisoning through peers influencing each other to engage in similar disruptive behavior.

This paper makes three main contributions. First, this is the first study to investigate the spillover effects of lead exposure on peers' academic achievement, behavior, and long-run outcomes. By exploiting rich individual-level data, we assess the costs of the spillover effects of lead exposure. Furthermore, our findings have implications for more than just lead: our estimates imply that the true costs of

pollution are likely higher than the direct costs alone, especially for pollutants that are known to affect behavior and suspensions from school.

Second, this is among the first studies to examine the long-run impacts of disruptive peers, as well as the channels through which these effects manifest. Current evidence on the long-run effects of peers is mixed. While Carrell, Hoekstra, and Kuka (2018) show that peers exposed to domestic violence lower wages and educational attainment, Bietenbeck (2020) finds positive long-run peer effects from peers who repeat kindergarten. We show that exposure to lead-poisoned peers in middle school can have long term consequences, including dropping out of high school, even for those children who were not themselves exposed to lead. We also find suggestive evidence on the mechanisms – that homophily within groups and exposure to disruptive peers in middle school might drive some of these effects. Understanding how the organization of schools mitigates these negative effects is crucial to design policies that curb the negative consequences of lead poisoning and pollution exposure.

Third, we contribute to a growing literature documenting the importance of neighborhood effects for health, education, and behavior outcomes. Our findings on the long-term effects of exposure to lead-poisoned children might help explain why high-poverty and high-pollution neighborhoods have persistent effects (Chetty, Hendren, and Katz 2016). Low-income children are more likely to live near sources of toxic waste (Persico, Figlio, and Roth 2020; Banzhaf, Ma, and Timmins 2019), and neighborhood characteristics contribute significantly to health disparities, for example in asthma rates (Alexander and Currie 2017). Our paper presents another channel through which inequalities in prevalence of pollutants at the neighborhood level contribute to the persistence of inequality in the US.

II. Background

Lead Exposure

Ingestion or inhalation of lead causes lead poisoning, which, if severe, can induce widespread brain damage (Meyer, McGeehin, and Falk 2003). Small children are especially exposed to lead-contaminated soil and dust from paint due to normal hand-to-mouth activity. Moreover, lead is most damaging to small children: they absorb and retain more lead than adults and their neurological development is particularly susceptible to neurotoxins (Meyer, McGeehin, and Falk 2003). Specifically, lead causes the axons of nerve cells to degenerate and lose their myelin coats (Meyer, McGreehin and Falk, 2003). Early life exposure to lead has been shown to cause cognitive disabilities, lower test scores, increase suspensions from school, and even affect crime and wages in adulthood (Persico, Figlio, and Roth 2020; Gazze 2016; Gronqvist, Nilsson and Robling 2020). Lead has also been associated with problems in cognition, executive functioning, abnormal social behavior (including aggression), and fine motor control (Cecil et al, 2008). Reyes (2014) estimates that lead costs $200 billion for a single birth-year cohort.

Peer Effects in the Classroom

Peer effects can work through different channels, both positively and negatively. Children teaching each other is an example of a positive peer effect, while disruptive behavior can negatively affect the learning of all children in a classroom (Carrell and Hoekstra 2010; Figlio 2007; Hoxby 2000). Using the random assignment of roommates in college, Sacerdote (2001) finds that roommates can influence college grade point averages positively or negatively. A variety of mechanisms link peer composition and academic outcomes, including differential curricular offerings and instructional practices in classes with higher average ability (Jackson 2013); social dynamics in a student's reference group (Hoxby 2000); and low performing students not keeping up with higher-achieving

peers (Imberman, Kugler and Sacerdote 2009).[2] Peers might also draw disproportionately on a teacher's time and influence class culture and standards.

One strand of the literature examines how low-performing and disruptive children affect peers. Having more low achieving peers or peers with learning disabilities is associated with lower achievement (Hoxby 2000; Fletcher 2010). Moreover, boys with feminine sounding names and children exposed to domestic violence are both more likely to be disruptive and negatively affect peers' achievement and behavior (Figlio 2007; Carrell and Hoekstra 2010).

Less is known about the long-term impacts of disruptive peers or the mechanisms through which disruptive peers affect long-run outcomes. Carrell, Hoekstra, and Kuka (2018) find that having more disruptive peers in elementary school leads to lower earnings in adulthood and lower college attendance. Bifulco, Fletcher, and Ross (2011) find that a higher percentage of high school classmates with college-educated mothers increases school completion and college attendance. Bobonis and Finan (2009) find that the PROGRESA program in Mexico increased college attendance of non-eligible peers. Black, Devereux, and Salvanes (2013) find that a higher share of girls in ninth grade reduces educational attainment and the likelihood of selecting an academic track for college but lowers teen birth rates.

In this paper, we show that children who were exposed to lead are associated with both short- and long-run negative outcomes for their peers. We provide fresh evidence on the spillover effects of lead, the long-run effects of having disruptive peers, and the mechanisms through which peers affect long-run outcomes.

III. Data Description

Education Data

[2] See Epple and Romano (2011) and Sacerdote (2011) for overviews of the literature on peer effects.

We use 1997-2017 population-level data on every child attending public school in North Carolina, including charter schools. The data provide individual-level education outcomes, rich demographic information, and blood lead test records when available. These unique data include home address identifiers that enable us to match siblings. To our knowledge, this is the first state-level data set linking individual BLLs to schooling records that allow the matching of siblings and students to classrooms. The data also include detailed information on students' race and economic disadvantage status in a given year, annual standardized test scores administered by the state, suspensions, absences, high school dropout and completion information, college intentions, as well as teacher characteristics.

While we use the entire sample to calculate the number of children per school-grade-year cohort who have elevated BLLs (as well as all of our cohort controls), for our main analysis we drop children who do not have siblings, as well as children who live in large buildings since we cannot reliably identify families in those buildings. We test for whether error in sibling matches affects the results in Section VD. Our main analysis also drops students who themselves have an elevated BLL and estimates the spillover effects of lead exposure on children without known lead poisoning. The Data Appendix provides more information on the linkage performed by the North Carolina Education Research Data Center (NCERDC), our sibling identification algorithm, and other details of our variables.

For our contemporaneous outcomes, we use the average of standardized mathematics and reading test scores and construct indicators for being absent for more than 22 days and receiving an in-school or out-of-school suspension, as well the number of days the child was suspended out-of-school each year. For our long-term outcomes, we use indicators for high school graduation, dropping out,[3] college intentions in 12th grade, and whether the student took the SAT in high school.

[3] The data include separate variables for dropping out and graduating. Dropping out of school is distinct from school switching, death, moving, promotion, graduation, and other confounding

Blood Lead Levels Data

We obtained individual blood lead test records for children up to age six from the North Carolina Department of Health and Human Services for the years 1992-2016. Test records include the date of blood draw, test result in µg/dL, and the child's identifier and address. We define a child as having an elevated BLL (EBLL) if their highest BLL is ≥5 µg/dL, the upper reference interval value per the 2012 guidelines by the Centers of Disease Control and Prevention (CDC 2013).[4]

Childhood lead screening is not mandatory in North Carolina. However, federal guidelines mandate that all children on Medicaid are screened for lead poisoning at ages one and two. Thus, we expect screening to be higher among low-income children, who have a higher likelihood of lead exposure. We construct indicators for children missing blood lead tests and include these children in our analysis. We compute the share of a child's peers with EBLLs using all children in the cohort or classroom as the denominator, independently of whether they have a blood lead test. Figure 1 plots the share of children with blood tests and the share of children with EBLLs by birth cohort in our sample, showing that as lead screening increases over time, the incidence of lead poisoning decreases.

Sample Description

Table 1 presents summary statistics for our original dataset (2.75 million children, Column 1) and our analysis sample of siblings (1.3 million children,

factors, and specific reason codes are given for dropping out. If a student is ever observed as graduating, we count them as graduating and not dropping out.

[4] This value is the 97.5th percentile of BLLs in U.S. children aged 1–5 years from the combined 2007–2008 and 2009–2010 cycles of the National Health and Nutrition Examination Survey. Starting in 1991 and prior to 2012, CDC defined BLLs ≥10 µg/dL as the "level of concern" for children aged 1–5 years. In robustness checks, we define a child as having an elevated BLL if alternatively the mean of their BLLs is ≥5 µg/dL or their highest BLL is ≥10 µg/dL.

Column 2). The Data Appendix details our sample selection criteria. 39.6 percent of children in our analysis sample have a blood lead test, and 10.9 percent have at least one test greater or equal than 5 μg/dL, slightly higher shares than in the original data. Children in our sample are also marginally more likely to be economically disadvantaged, less likely to be Black, and have slightly better outcomes. Children with EBLLs are more likely to be Black, be economically disadvantaged (ED) as measured by an indicator for having ever received free or reduced-price lunch, and have teachers without Master's degrees (Columns 4 and 5). The average cohort in our sample includes 225 children. Children who spend at least one elementary school year in a cohort with above median share of lead-exposed children (or >10.1 percent of cohort peers) have lower test scores, higher suspension rates, lower graduation and SAT taking rates, and have a lower probability of intending to attend a four-year college (Columns 6 and 7). These children are also more likely to be Black, be economically disadvantaged, have teachers without Master's degrees, and have a blood lead test themselves. Our identification strategy controls for family background with family fixed effects, assuaging concerns of omitted variable bias due to these differences.

IV. Identification Strategy

Rigorously estimating peer effects has proven difficult methodologically and due to limitations of existing data. First, peers influence each other simultaneously, so it is unclear whether a disruptive child causes their classmates to misbehave, or whether the classmates cause them to be disruptive. This is called the reflection problem (Manski 1993). Second, peer groups are not randomly assigned; they are selected based in part on unobserved characteristics. Children in the same classroom often share similar backgrounds. Moreover, attentive parents might remove their children from classrooms with more disruptive peers. Because of this self-selection into groups, it is impossible to determine whether the outcome

is a causal effect of the peers or the reason the individuals joined the peer group (Carrell and Hoekstra 2010; Hoxby 2000). Third, unobserved factors might simultaneously cause students and their peers to perform poorly.

We solve the reflection problem by finding a predetermined proxy for peer ability: lead exposure. Consistent with the literature on lead exposure and academic outcomes, being exposed to lead is strongly associated with worse academic achievement, a higher likelihood of suspension, and a lower probability of graduating or intending to attend a four-year college in our sample (Figure 2).[5] Previous research has proxied for peer ability and behavior using preexisting measures such as peers' race and gender (Hoxby and Weingarth 2006; Hoxby 2000), feminine-sounding names of male peers (Figlio 2007), peers' retention status (Lavy, Paserman, and Schlosser 2012), peers' disability (Fletcher 2010), or peers' exposure to domestic violence (Carrell and Hoekstra 2010). Our approach is similar in that we use the presence of peers with elevated blood lead levels to estimate how early health shocks (i.e., lead exposure) spill over within school contexts to exacerbate inequality through peer effects. This is a valid approach as a student cannot affect their peers' elevated blood lead levels.

Yet, a child's lead exposure could be correlated with their socioeconomic status, which in turn has been associated with peers' learning disruptions (Hoxby and Weingarth 2006; Hoxby 2000). Thus, to causally identify the spillover effect of a child's lead exposure on their peers we further control for the share of cohort peers who are non-White or economically disadvantaged. We also control for the share of the student's peers who have been tested for lead exposure. Because screening rates are higher among low-socioeconomic status students, additionally controlling for screening rates mitigates concerns about selection into testing.

[5] Our companion paper (Gazze, Persico, and Spirovska 2020) estimates the effects of lead exposure on short- and long-run outcomes in our sample using a family fixed effects model. That paper finds that having an EBLL is associated with worse educational and behavioral outcomes.

Finally, we exclude children with known EBLLs so that we can isolate the spillover effects of lead poisoning on peers who are not lead-poisoned.

We compare siblings whose grade cohorts randomly happen to have different proportions of children with EBLLs. Including family fixed effects mitigates the selection problem by controlling for unobserved family characteristics that could be correlated with both peer quality and child's outcomes, such as parents' propensity to move their children to schools with fewer lead-poisoned children. Including school fixed effects further controls for students' characteristics that are common to the school's catchment area. Remaining idiosyncratic variation in the BLLs of siblings' cohorts offers plausibly exogenous variation to identify the spillover effects of lead and the effects of peer quality more broadly.

We first examine how lead exposure affects contemporaneous outcomes, that is test scores, suspensions from school, and absences of peers without known EBLLs. Our main estimation equation is given by:

(1) $Y_{ijsgt} = \beta_1 \frac{\sum k \neq i \; PeersEBLLs_{ksgt}}{n_{sgt}-1} + \pi X_{it} + \omega S_{sgt} + \theta_j + \delta_s + \tau_g + \sigma_t + \gamma_e + \varepsilon_{ijsgt}$

where Y_{ijsgt} is some outcome for child *i* who either has not been screened for lead exposure or has always tested below 5 µg/dL, born to family j, attending school *s*, in grade g and in year *t*. $\frac{\sum k \neq i \; PeersEBLLs_{ksgt}}{n_{sgt}-1}$ is the share of students in a child's school-grade-year cohort (or school-classroom-grade-year cohort) with *known* EBLLs not including the student themselves. The coefficient β_1 on $\frac{\sum k \neq i \; PeersEBLLs_{ksgt}}{n_{sgt}-1}$ captures the effect of having 100 percent of a child's peers in a given year with *known* EBLLs. X_{it} is a vector of child-specific control variables, including gender, race, birth month fixed effects, birth order fixed effects, economically disadvantaged (ED) status in each year, and an indicator for whether a child was tested for lead. The vector S_{sgt} controls for time-varying school-grade

characteristics: the percent non-White students by school-grade-year, the percent economically disadvantaged by school-grade-year, and the share of students who have been tested for lead exposure by school-grade-year. We also control for school time-varying characteristics: annual school size, the share of teachers with Master's degrees and the school-level stability rate.[6] θ_j is a family fixed effect. τ_g is a grade fixed effect to account for statewide grade-specific shocks. δ_s is a school fixed effect to account for constant school characteristics over time, and we adjust for secular trends using a year fixed effect σ_t. γ_e is an exam type fixed effect that restricts our comparison to children who took the same exam.[7] We cluster standard errors at the school level to account for arbitrary correlation in the error terms.

There are three main threats to the internal validity of our estimates. First, our estimates would be biased if a child's peers' BLLs were correlated to the child's own BLLs, or their ability, other than through classroom interactions. To address this issue, we measure lead exposure prior to school entrance. Moreover, family fixed effects account for omitted variables such as unobserved lead exposure or parental characteristics that could confound the effects of peer quality. School fixed effects help us account for selection into schools. Second, our estimates could be biased in the presence of common shocks that are systematically correlated with the proportion of peers with BLLs in a school-grade-year. Time-varying school and teacher controls help assuage concerns that these channels drive our results. Third, bias could arise if high-quality students systematically select out of schools when there are more students with EBLLs. We test for school switching across siblings

[6] The stability rate is defined as the percentage of students from the October membership count who are still present in the second semester (90 days later).

[7] Exams and exam scales in North Carolina changed multiple times over this time period. The scale for the math EOG exam changed in 2001-2002, 2005-2006, and in 2012-2013, while the reading EOG exam scale changed in 2002-2003, 2007-2008, and again in 2012-2013. Finally, in 2014-2015 a new 5-level scale replaced the long-standing 4-level scale. See the Data Appendix for more details.

in Section VD, where we also show that our results are largely robust to including school-by-family fixed effects.

Finally, it is important that we have enough variation in our regressor of interest, the share of children with EBLLs within school-grade-year. Figure A1 shows the distribution of the within-school interquartile range of this variable for 3rd grade. Panel A shows this distribution for the share of children with EBLLs over all children in the cohort, which we use in the analysis, while Panel B shows the distribution for the share over the number of children with lead tests in the cohort. Despite the distribution being right-skewed, it has significant mass above 0.1, meaning the difference between the 25th and the 75th percentile is 10 percentage points or more.

V. Results

A. The Contemporaneous Effects of Peers Exposed to Lead on Child Outcomes

We begin by showing the effects of peers with elevated BLLs on contemporaneous standardized test scores, suspensions from school, and absences. Figure 3 shows that the share of a child's peers with EBLLs is negatively correlated with the child's test scores, and positively correlated with their likelihood of receiving a suspension in the raw data. Table 2 confirms these patterns are causal. Panel A presents the results for the effect of additional cohort peers who are lead-poisoned on a child's outcomes using our primary specification in equation (1). We find that a ten percent increase in the proportion of cohort-level peers with elevated BLLs in a given year leads to a 0.3 percentage point increase in the likelihood of suspension, compared to siblings in the same school. In other words, attending school with 10 percent more lead-poisoned peers increases the suspension rate by 2.8 percent above the mean of 10.4 percent, and increases out-of-school suspensions by one hour based on a 7-hour school day. Moreover, these increased suspensions appear to be driven by suspensions on the same day as suspensions for

lead-poisoned children. Using incident indicators to link children to the same incident resulting in suspensions, we find in fact an increase in suspensions linked to incidents including lead-poisoned children. Finally, we note that the effect of lead-poisoned peers on suspensions is similar to the effect of economically disadvantaged peers, while the coefficient on non-White peers is negative.

One concern, however, is that increased suspensions for peers of lead-poisoned children could be due to more punitive policies at the cohort-level. For example, teachers might be more prone to suspending students for minor misbehavior in cohorts with more disruptive students. To disentangle peers' behavior from school policies, we look at the effects of lead-poisoned peers on absences, which should not be driven by school policies. We find that a 10 percent increase in the proportion of cohort-level peers with elevated BLLs increases the likelihood of chronic absenteeism by 0.2 percentage points, or 5 percent on a base of 3.9 percent, suggesting that our results are driven by students misbehaving more when they have more lead-poisoned peers and not blanket-style school policies. Finally, some specifications show a decrease in test scores among students with more lead-poisoned peers, although this result is not statistically significant at conventional levels in our primary specification.

While we use cohort-level variation in our primary specification to avoid the issue of selection into classrooms by students, Panel B of Table 2 presents the estimates of the effect of having more lead-poisoned peers in the same *classroom* using family, school, grade and year fixed effects and all controls specified in equation (1). We define peer exposure at the classroom level by averaging the number of peers with EBLLs across all classes a child takes in that year. Thus, if students switch classrooms, they will have more peers overall.[8] We find that peers

[8] We only have classroom-level data for a subset of the children in the sample from 2006 to 2017, whereas the cohort level variation is available from 1997-2017. Children in grades 3-5 do not typically switch classrooms for different subjects, so they are counted as in the same classroom

in the same classroom have an impact on a child's test score that is almost ten times as large, and statistically significant at the 0.1 percent level, as peers in the same cohort but potentially different classrooms. Yet, classroom peers have a slightly smaller effect on suspensions and absences than cohort peers. These results suggest that while peers outside the classroom but in the same cohort do not directly affect a child's ability to learn, they can still influence each other's behavior, for example by spending time together at lunch or recess. Both classroom and cohort peers matter.

In Panel C, we estimate how average exposure to lead-poisoned peers in elementary school classrooms (in grades 3-5) affects outcomes in eighth grade. We find that elementary school peers have strong impacts on 8th-grade test scores: an additional student in a class of 25 decreases average eighth grade test scores by 0.97 percent of a standard deviation. However, elementary school peers do not appear to affect suspensions, incidents with lead-poisoned peers, or absences in 8th grade, suggesting that it is contemporaneous exposure to misbehaving peers that increases suspensions and absences.

B. Long Term Effects of Peers Exposed to Lead

We next examine whether a child's lead-poisoned peers in elementary and middle school affect that child's long-run outcomes. Table 3 presents estimates of these long-run effects by estimating the following regression:

$$(2)\ Y_{ijsgt} = \beta_1 \frac{\overline{\sum k \neq \iota\ PeersEBLLs_{ksgt}}}{n_{sgt}-1} + \pi \overline{X_\iota} + \omega \overline{S_{st}} + \theta_j + \delta_s + \tau_g + \sigma_t + \varepsilon_{ijst}$$

where $\overline{X_\iota}$ and $\overline{S_{st}}$ include all of the individual-level controls from our primary specification, as well as the average share of non-White peers, the average share of economically disadvantaged peers, the average share of peers tested for lead, and

based on their mathematics classroom peers. Children in grades 6 and up usually do switch classrooms, so they are counted as many times as the number of classes they take with each student.

the average school size, school stability rate and share of teachers with masters degrees over elementary and middle school. δ_s, τ_g and σ_t are school, grade and year fixed effects for the child's last observation. The coefficient β_1 captures the effect of having 100 percent of peers with *known* EBLLs in elementary and middle school.

Panel A of Table 3 shows that a child whose average cohort in elementary and middle school has 10 percent more lead-poisoned peers has a 1.7 percentage point lower likelihood of graduating high school, representing a 2 percent decrease on the mean graduation rate of 89 percent. We also find that having 10 percent more lead-poisoned peers increases the likelihood of dropping out by 0.47 percentage points and decreases the likelihood of taking the SAT while in high school by 2.3 percentage points, or a 4.3 percent decrease on the mean rate of 53.2 percent. Finally, a higher share of lead-poisoned peers decreases the likelihood that a student intends to attend a four-year college in some specifications, although this result is not statistically significant at conventional levels in our primary specification.

Panel B of Table 3 estimates the effect of lead-poisoned peers at the *classroom* level on long-run outcomes. While we do not find an effect of lead-poisoned peers in the classroom on graduation, we find larger effects on SAT taking. We also find evidence that lead-poisoned peers in the classroom might lead students to substitute college intentions from a four-year college to a community college. While we cannot measure whether students followed through on their college plans because we lack college enrollment data, these results suggest that prolonged exposure to lead-exposed peers could worsen long-run outcomes.

Panel C of Table 3 estimates the long-run effect of lead-poisoned peers in elementary and middle school cohorts *separately*. We find that long-run outcomes are largely driven by middle school peers. This result is in line with our interpretation of the findings in Panel C of Table 2—that behavior in 8th grade is driven by exposure to disruptive peers in middle school, which in turn could set students on a path to lower graduation and college attendance rates. Peers in middle

school also could be especially impactful for long-run outcomes if middle school is a time when some students are deciding whether to remain in school. Finally, student learning and behavior in middle school might be especially important for college readiness (Naven 2019).

While the peer effects we find on test scores are smaller than those obtained by Carrell, Hoekstra, and Kuka (2018), the effects we find on college going are similar in magnitude. Those authors find that one male peer exposed to domestic violence decreases four-year college going by 1.4 percentage points. Using our cohort results and assuming that there are 25 students in a class, we calculate that adding an additional lead-poisoned peer to each class, a 0.04% increase in the share of lead-poisoned peers, would lead to a 0.92 percentage point reduction in the likelihood of taking the SATs, a proxy for college intentions, and a 0.68 percentage point reduction in graduating high school.

C. Mechanisms and Heterogeneity of Estimated Effects

Given our findings that cohort peers shape children's contemporaneous behavior just as much as classroom peers (Table 2), we hypothesize that children friends' groups outside the classroom might drive peer effects. As we lack data on friendship networks, we exploit the fact that children likely sort into groups with similar characteristics (Jackson 2010). Table 4 presents both the effect of exposure to a higher share of lead-poisoned peers and the additional effect of exposure to a higher share of lead-poisoned peers of the same gender (Panel A), race (Panel B), and same gender and same race (Panel C), and male same-race peers (Panel D). We find that same gender peers with EBLLs have larger effects on suspensions and high school graduation, while same-race peers have larger effects on test scores. Peers who are of both the same race and same gender have larger effects on both suspensions and test scores, and male same-race peers increase suspensions among male students, suggesting that there are larger effects for same-race peers once we

take gender into account. While exposure to similar lead-poisoned peers appears to have an additional effect on contemporaneous outcomes, we find limited effects on long-run outcomes, suggesting the effects of homophily in networks might diminish over time.

Because exposure to lead-poisoned peers could interact with a child's background to shape their outcomes, we next study heterogeneity in peer effects by demographic subgroups. For example, students of different socioeconomic status might have differential access to resources, such as academic help outside of school, that could ameliorate the effects of peers with EBLLs. Table 5 presents our preferred estimates by race/ethnicity (White, non-Hispanic in Panel A, Black students in Panel B, and Hispanic students in Panel C), by economically disadvantaged status (never economically disadvantaged in Panel D, sometimes economically disadvantaged in Panel E, and always economically disadvantaged in Panel F), and by gender (girls in Panel G and boys in panel H).

We find some evidence of heterogeneous effects of lead-poisoned peers by race. Black students see the largest increase in suspensions and the largest decrease in high school graduation from lead-poisoned peers. Black students have higher suspension rates, so our results suggest that Black students might disproportionally get in trouble when there are disruptive peers.

Students who are economically disadvantaged in all grades have the greatest test score losses in the presence of more lead-poisoned peers. However, students who are economically disadvantaged only in some grades see larger increases in suspensions and decreases in graduation rates than students who are either never or always economically disadvantaged. Appendix Table A1 presents estimates of the effects of lead-poisoned peers for children in schools with different levels of poverty. We find stronger negative peer effects on suspensions in schools with the highest share of economically disadvantaged students. This finding suggests that poverty might exacerbate the effects of having lead-poisoned peers.

Finally, we find that in cohorts with a higher share of lead-poisoned peers, boys have lower test scores and graduation rates than girls. Panel A of Appendix Table A2 shows that lead-poisoned boys have larger negative effects on their peers than lead-poisoned girls. Together, these results support the hypothesis that peer effects are mediated by assortative matching of peer groups, as shown in Table 4.

D. Additional Threats to Internal Validity

This section discusses and tests for threats to internal validity, including spurious correlation, selection into lead testing, measurement error, and endogenous sorting.

If our results are driven by increases in peers' blood lead levels, we would expect students exposed to a higher percentage of cohort peers with elevated BLLs to do worse. Figure 4 plots estimates from equations (1) and (2) using bins for different percentages of cohort peers with elevated BLLs (0-5%, 5%-10%, 10%-15%, 15%-20%, 20%-100%).[9] We find a stronger effect of lead-poisoned peers on test scores, suspensions, and graduation rates as the percentage of peers with elevated BLLs increases. We do not find a statistically significant gradient for college intentions. Moreover, Appendix Figure A2 shows that our estimates are unlikely to be due to random chance. This figure plots the results from estimating 500 placebo specifications in which we assign a random share of lead-poisoned peers to each school-grade-year cohort drawn from a distribution with the same mean and standard deviation as the empirically observed peers' distribution. Our true estimates for the effects of lead-poisoned peers on suspensions and graduation rates fall well outside the distribution of estimates from the placebo specifications.

However, there may be selection in who is tested for lead. As we do not observe lead exposure for all children, we measure the share of lead-poisoned children in each cohort with error. Since we compute the share of lead-poisoned

[9] We omit the indicator for having 0-5% of cohort peers with elevated BLLs from the regression.

peers over all students in a cohort, irrespective of whether they have a blood lead test, *unknown* lead-poisoned peers would attenuate our results. North Carolina requires screening for all children living in zip codes where at least one block group within the zip code has 27 percent or more homes built prior to 1950.[10] Panel A of Table 6 shows the effects of lead-poisoned peers on children in these high-risk zip codes, where screening rates are 16 percent higher than average. We find a larger effect on test scores than in the full sample suggesting attenuation bias due to measurement error could be a concern. However, the coefficients on graduation and taking the SAT are comparable to our main results in Table 3.

Furthermore, we identify siblings based on home addresses, which could lead to error, particularly in multi-family homes. Thus, Panel B of Table 6 shows results on the sample of Census tracts where the majority of homes are single family homes. The pattern of results largely holds in this sample, suggesting that error in matching siblings does not bias the findings.

In addition, if parents of high-achieving students pull their children out of a cohort with particularly disadvantaged or lead-poisoned students, such nonrandom selection could lead us to misattribute poor peers' performance to the larger presence of lead-poisoned students. Importantly, most of North Carolina did not offer school choice options for public schools over our sample period: with one exception, up until the 2014-2015 school year, students could only switch schools if they switched into a charter or magnet school, which we observe in our data.[11] Panel C of Table 6 shows that our long-run results hold for children in zip codes with no charter schools or other school choice options (at the time), which are

[10] The designation also adjusts for prevalence of elevated BLLs (Hanchette 1999).

[11] In the 2014-2015 school year, North Carolina implemented the Opportunity Scholarships program, a voucher program for low income children. Children whose families make less than 133 percent of the qualifying amount for the federal free or reduced-price lunch program qualify for the voucher, which can be used for any school. In addition, the Charlotte Mecklenburg Public School district has had a school choice program from 2002 so we exclude that district.

effectively no-choice zip codes. Moreover, Table 7 formally investigates the association between a student's share of lead-poisoned peers and school switching. We find no evidence of increased switching to public or charter schools of students with higher shares of lead-poisoned peers or of their siblings. Thus, differential sorting does not appear to drive our results. To further test whether differential school switching biases our results, Panel A of Appendix Table A3 controls for siblings-by-school fixed effects, effectively comparing siblings only in grades during which they attend the same school.[12] We find spillover effects of lead-poisoned peers on long-run outcomes that are half to two-thirds the size of our main results.

Our estimates could also be biased if the share of peers with EBLLs in a school-grade-year is systematically correlated with students' or peers' characteristics other than those included in equation (1). Panel B of Table A3 adds Census block group fixed effects. The results are similar to those in our main specification despite the sample size being smaller due to missing block group information, suggesting that neighborhoods, including contemporaneous pollution exposure not captured by BLLs by age 6, do not drive the results. Panel C of Appendix Table A3 further shows that estimates using school-grade fixed effects are virtually indistinguishable from our main results that include school and grade fixed effects. Finally, because the incidence of lead poisoning has decreased over time (Figure 1), our primary estimates might capture similarly occurring trends in outcomes despite controlling for grade and year fixed effects. To assuage this concern, in Panel D of Appendix Table A3, we control for grade-year fixed effects and find peer effects that are, relative to our main specification, larger on test scores, similar on SAT taking, but smaller on suspension and graduation rates.

[12] Bertoni, Brunello, and Cappellari (2020) use this design to study the effects of privileged peers.

In addition, peer characteristics could be correlated with a child's own characteristics. We investigate this possibility in Table 8 by regressing child characteristics on the proportion of peers with EBLLs to see if peer EBLLs predict these characteristics. Generally, the fraction of peers with EBLLs shows only a very small correlation with a student's characteristics, and we control for these characteristics in our primary specification. Moreover, some characteristics correlated with a higher share of lead-poisoned peers would predict better outcomes, such as being female. A 10 percent increase in peers with EBLLs is associated with a 0.27 percentage point increase in the likelihood that a child is female and an increase in cohort size of 0.64 students. Finally, we find that a higher share of peers with EBLLs is associated with a slight increase in the school stability rate, suggesting that differential school-switching does not account for our findings.

To address the concern that blood lead levels are measured with some error, in Panels B and C of Appendix Table A2, we show that our results are largely robust to using different measures of lead-exposed peers, although when we define EBLLs as BLLs≥10 μg/dL we are left with smaller residual variation in this rarer condition after controlling for family, school, year, and grade fixed effects. In Panel D we include all students, even those who are exposed to lead, and control for one's own lead exposure. The estimates are largely similar to our main results.

Finally, Appendix Table A4 shows the robustness of our specification to different sets of controls. Panel A shows that when omitting all controls other than family, school, grade and year fixed effects, we would find larger contemporaneous but smaller long-run peer effects, suggesting that spurious correlations might arise even with our conservative specification. Reassuringly, Panel B shows that once we add individual and school-level controls, omitting the share of students in a school-grade-year who are non-White and the share of students who are economically disadvantaged does not affect our estimates compared to our main results. In other words, peers' characteristics other than lead poisoning do not

appear to explain much of the variation in students' outcomes after controlling for the set of fixed effects that provides our identification. This finding suggests that the share of lead-poisoned peers does not just capture the effect of non-White or poor peers. Panel C shows that excluding school fixed effects yields slightly larger contemporaneous peer effects but similar long-run effects to our main results. These results suggest that our more conservative primary specification controls for unobserved time invariant school characteristics.

VI. Conclusion

This is the first study documenting the spillover effects of lead onto school peers. By comparing siblings who attend the same school, we find that a child's own lead exposure spills over to affect other children's behavior and long-run outcomes. A ten percent increase in peers with elevated BLLs in a given year leads to a 2.8 percent increase above the mean in the likelihood of being suspended and a 5 percent increase in chronic absenteeism. A ten percent increase in peers with elevated BLLs over a student's elementary and middle school career causes a 2 percent decrease in the likelihood of graduating high school, and a 4.3 percent decrease in the likelihood of taking the SAT. The magnitude of these effects is substantively important, suggesting that the social cost of lead exposure has been underestimated so far. Our results suggest that environmental hazards are an important factor contributing to human capital accumulation even for children who are not themselves exposed to these hazards. In addition, we show that peers can have long term consequences on human capital formation and reveal some mechanisms through which peer effects manifest, namely homophily in network formation and behavior shaping while in middle school. Furthermore, our findings have implications for other types of pollution that are known to cause suspensions from school, such as traffic pollution and pollution from TRI or Superfund sites

(Persico and Venator 2020; Heissel, Persico and Simon 2020), suggesting that the true cost of pollution has been underestimated.

We likely estimate a lower bound of the effect of lead-poisoned peers. We find strong evidence of worse outcomes for children exposed to more lead-poisoned peers despite their siblings are likely exposed to disruptive peers as well and despite potential spillovers within siblings, too. Moreover, missing BLLs for some lead-poisoned children would attenuate our findings.

While external validity issues make it difficult to extrapolate how lead exposure might affect labor market outcomes, we attempt a back of the envelope calculation for the effect of one lead-poisoned peer in a cohort of 225. We find that being exposed to one additional lead-poisoned peer is associated with $71 in lost earnings per student from lower graduation rates alone.[13] This estimate does not include the additional costs of behavioral issues and potentially lower test scores (although our estimates are sensitive to specifications). Yet, this figure implies a spillover effect of a lead-poisoned child of $15,976 on their 224 school peers through elementary and middle school. As half a million young children appear to still be poisoned by lead each year (Aizer et al. 2018), these spillovers total almost $8 billion per birth-year cohort. Reyes (2014) estimates the direct social cost of lead poisoning at $200 billion per birth-year cohort. Thus, our estimates suggest that the social cost of lead has been underestimated by at least 4 percent by not including these spillover effects.

[13] Following Heckman, Lochner, and Todd (2006), we estimate the net present value of graduating high school to be $93,188. We estimate a schooling-experience-earnings profile non-parametrically in the 2018 March Current Population Survey data and predict earnings conditional on years of schooling at each age between 18 and 65, assuming a growth rate of real labor productivity growth of 1.9 percent and a discount rate of 3.38 (i.e., the 30-year Treasury bond rate). As 1 in 225 students is a 0.4% increase in the share of peers with elevated BLLs, we multiply that by our estimate of the effect of 100% of peers with elevated BLLs on graduation (-17.22 percentage points) to obtain the impact of one child with EBLLs through elementary and middle school on graduation rates: -0.077 percentage points, or a decrease in the probability of 0.00077. Thus, one child with EBLLs in a cohort would decrease the net present value of lifetime earnings by 0.00077*$93,188=$71.

Our results imply some important lessons for policy. Remediating lead hazards is likely to be more cost effective than previously supposed since lead exposure affects everyone in the classroom. Lead remediation efforts have shown positive impacts on children's blood lead levels and test scores (Sorensen et al. 2019). In addition, Billings and Schnepel (2018) show that offering early interventions for lead-poisoned children improves their outcomes. Thus, early interventions might help both lead-poisoned children and their peers.

Finally, school segregation by race and socioeconomic status likely exacerbates these peer effects, suggesting that additional efforts to desegregate students might be beneficial. Low income schools have some of the largest achievement gaps (e.g., see Reardon 2015). Our results suggest that peer effects and lead exposure contribute to low performance in high-poverty schools, as well as to some of the negative long-run outcomes associated with poverty. Lead exposure and exposure to lead-poisoned peers are both mechanisms through which poverty produces worse human capital outcomes.

References

Aizer, A., & Currie, J. 2019. "Lead and Juvenile Delinquency: New Evidence from Linked Birth, School and Juvenile Detention Records." *Review of Economics and Statistics,* 101(4): 575–587.

Aizer, A., Currie, J., Simon, P., & Vivier, P. 2018. "Do Low Levels of Blood Lead Reduce Children's Future Test Scores?" *American Economic Journal: Applied Economics,* 10(1): 307–341.

Alexander, D., & Currie, J. 2017. "Is It Who You Are or Where You Live? Residential Segregation and Racial Gaps in Childhood Asthma." *Journal of Health Economics,* 55: 186–200.

Banzhaf, S., Ma, L., & Timmins, C. (2019). "EJ: The economics of race, place, and pollution." *Journal of Economic Perspectives,* 33(1): 185-208.

Bertoni, M., Brunello, G., & Cappellari, L. 2020. "Who Benefits from Privileged Peers? Evidence from Siblings in Schools*." Journal of Applied Econometrics.*

Bietenbeck, J. (2020). "The Long-Term Impacts of Low-Achieving Childhood Peers: Evidence from Project STAR." *Journal of the European Economic Association*, 18(1): 392-426.

Bifulco, R., Fletcher, J. M., & Ross, S. L. 2011. "The Effect of Classmate Characteristics on Post-secondary Outcomes: Evidence from the Add Health." *American Economic Journal: Economic Policy,* 3(1): 25–53.

Billings, S. B., & Schnepel, K. T. 2018. "Life after Lead: Effects of Early Interventions for Children Exposed to Lead." *American Economic Journal: Applied Microeconomics,* 10(3): 315–344.

Black, S., Devereux, P., & Salvanes, K. 2013. "Under Pressure? The Effect of Peers on Outcomes of Young Adults." *Journal of Labor Economics,* 31(1): 119–153.

Bobonis, G. J., & Finan, F. 2009. "Neighborhood Peer Effect in Secondary School Enrollment Decisions." *Review of Economics and Statistics,* 91(4): 695–716.

Carrell, S. E., & Hoekstra, L. H. 2010. "Externalities in the Classroom: How Children Exposed to Domestic Violence Affect Everyone's Kids." *American Economic Journal: Applied Economics*, 2(1), 211–228.

Carrell, S. E., Hoekstra, L. H., & Kuka, E. 2018. "The Long-Run Effects of Disruptive Peers." *American Economic Review,* 108(11): 3377–3415.

CDC (Centers for Disease Control and Prevention). 2005. "Preventing Lead Poisoning in Young Children."

CDC (Centers for Disease Control and Prevention). 2013. "Blood Lead Levels in Children Aged 1–5 Years — United States, 1999–2010." *Morbidity and Mortality Weekly Report*.

Chetty, R., Hendren, N., & Katz, L. F. 2016. "The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment." *American Economic Review,* 106(4): 855–902.

Currie, J., Greenstone, M., & Moretti, E. 2011. "Superfund Cleanups and Infant Health." *American Economic Review,* 101(3): 435–441.

Epple, D., & Romano, R. E. 2011. "Peer Effects in Education: A Survey of the Theory and Evidence." *Handbook of Social Economics,* 1B: 1053–1163.

Feigenbaum, J. J., & Muller, C. 2016. "Lead Exposure and Violent Crime in the Early Twentieth Century." *Explorations in Economic History,* 62: 51–86.

Ferrie, J. P., Rolf, K., & Troesken, W. 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." *Economics & Human Biology,* 10(1): 98–111.

Figlio, D. N. 2007. "Boys Named Sue: Disruptive Children and Their Peers." *Education Finance and Policy*, 2(4): 376–394.

Fletcher, J. 2010. "Spillover Effects of Inclusion of Classmates with Emotional Problems on Test Scores in Early Elementary School." *Journal of Policy Analysis and Management*, 29(1): 69–83.

Gazze, L. 2016. "Lead Policies, Lead Poisoning, and Government Spending." Working Paper.

Gazze, L., Persico, C., & Spirovska, S. 2020. "The Long-Term Effects of Early Childhood Lead Exposure." Working Paper.

Grönqvist, H., Nilsson, J. P., & Robling, P. O. 2020. "Understanding How Low Levels of Early Lead Exposure Affect Children's Life Trajectories." *Journal of Political Economy*, 128(9): 3376–3343

Hanchette, C. L. (1999). "GIS and decision making for public health agencies: childhood lead poisoning and welfare reform." *Journal of Public Health Management and Practice*, 5(4), 41-46.

Heckman, J. J., Lochner, L. J., & Todd, P. E. (2006). "Earnings functions, rates of return and treatment effects: The Mincer equation and beyond." *Handbook of the Economics of Education*, 1, 307-458.

Heissel, Jennifer, Persico, Claudia and Simon, David. (2020). "Does Pollution Drive Achievement? The Effect of Traffic Pollution on Academic Performance." Journal of Human Resources. In press

Hoxby, C. 2000. "Peer Effects in the Classroom: Learning from Gender and Race Variation." Working Paper 7867, National Bureau of Economic Research.

Hoxby, C., & Weingarth, G. 2006. "School Reassignment and the Structure of Peer Effects." NBER Conference Paper.

Imberman, S., Kugler, A., & Sacerdote, B. 2012. "Evidence on the Structure of Peer Effects from Hurricane Evacuees." *American Economic Review*, 102(5): 2048–2082.

Isen, A., Rossin-Slater, M., & Walker, R. 2017. "Relationship between Season of Birth, Temperature Exposure, and Later Life Wellbeing." *PNAS,* 114(51): 13447–13452

Jackson, C. K. 2013. "Can Higher-Achieving Peers Explain the Benefits to Attending Selective Schools? Evidence from Trinidad and Tobago." *Journal of Public Economics,* 108: 63–77.

Lavy, V., Paserman, M. D., & Schlosser, A. 2012. "Inside the Black Box of Ability Peer Effects: Evidence from Variation in the Proportion of Low Achievers in the Classroom." *Economic Journal,* 122(559): 208–237.

Manski, C. F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *The Review of Economic Studies,* 60: 531–542.

Meyer, P. A., McGeehin, M. A., & Falk, H. 2003. "A Global Approach to Childhood Lead Poisoning Prevention." *International Journal of Hygiene and Environmental Health*, 206(4–5): 363–69.

Persico, C., Figlio, D., and Roth, J. 2020. "The Developmental Consequences of Superfund Sites." *Journal of Labor Economics.* 38, no. 4.

Persico, Claudia and Venator, Joanna. (2020). "The Effects of Local Industrial Pollution on Students and Schools." Journal of Human Resources. In press.

Reardon, S. F. 2015. “School Segregation and Racial Academic Achievement Gaps.” *Russel Sage Foundation Journal of the Social Sciences*, 2(5): 34–57.

Reyes, J. W. 2014. “The Social Costs of Lead” in *Lead: The Global Poison – Humans, Animals, and the Environment*, 1–4.

Reyes, J. W. 2015. “Lead Exposure and Behavior: Effects on Antisocial and Risky Behavior among Children and Adolescents.” *Economic Inquiry, 53*(3): 1580–1605.

Sacerdote, B. 2001. “Peer Effects with Random Assignment: Results for Dartmouth Roommates.” *Quarterly Journal of Economics*, 116(2): 681–704.

Sacerdote, B. 2011. “Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far?” *Handbook of the Economics of Education,* 3: 249–277.

Sorensen, L. C., Fox, A. M., Jung, H., & Martin, E. G. 2019. “Lead Exposure and Academic Achievement: Evidence from Childhood Lead Poisoning Prevention Efforts.” *Journal of Population Economics,* 32(1): 179–218.

Table 1: Characteristics of children and schools

Sample:	(1) All children attending school in North Carolina	(2) Children in sample	(3) Children with BLL test	(4) Children with EBLLs	(5) Children without EBLLs	(6) Children with above-median share of EBLL peers in at least one elementary grade	(7) Children with below-median share of EBLL peers in all elementary grades
Average test score	0.040	0.066	-0.097	-0.268	0.114	-0.104	0.287
Any suspension	0.277	0.322	0.369	0.494	0.301	0.382	0.255
Ever graduated	0.842	0.872	0.866	0.816	0.881	0.850	0.895
4-year college intentions	0.436	0.454	0.398	0.346	0.471	0.389	0.519
Has taken the SAT	0.452	0.466	0.411	0.366	0.482	0.407	0.525
Cohort size	224	225	199	203	229	194	263
Share of teachers with an MA degree	0.361	0.362	0.353	0.340	0.366	0.343	0.386
Share economically disadvantaged	0.534	0.586	0.712	0.822	0.557	0.712	0.445
Stability rate	0.955	0.957	0.957	0.953	0.958	0.955	0.960
Share Black	0.261	0.249	0.308	0.432	0.226	0.305	0.186
Share Hispanic	0.125	0.124	0.144	0.107	0.126	0.136	0.111
Share with a BLL test	0.314	0.396	1	1	0.322	0.530	0.249
Share with EBLL	0.086	0.109	0.276	1	0	0.162	0.050
N Students	2,749,324	1,326,622	525,535	144,957	1,181,665	696,924	629,698

Notes: The table presents summary statistics for selected variables in our sample. Observations are at the student-year level. Cohort is defined as student-grade-year. Column 1 shows the means for all children in North Carolina. Column 2 shows means for children with siblings, that is our main sample. Column 3 shows means for children that have a blood lead level test. Column 4 shows means for children with elevated blood lead levels (EBLLs), and Column 5 shows means for children without elevated blood lead levels. Column 6 shows means for children whose share of elementary school peers with elevated BLLs was above the median share at the grade-year level in at least one grade, while Column 7 shows means for children whose share was below the median in all elementary grades. Test scores are standardized at the grade-year level. The stability rate is defined as the percentage of students from the October membership count who are still present in the second semester (90 days later).

Table 2: Contemporaneous Effects of Attending School with an Increased Share of Children with Elevated BLLs

Dependent Variable:	(1) Average Test Score	(2) Any suspension	(3) Days Suspended	(4) Suspended Same Day as Lead-Exposed Child	(5) Incident with Lead-Exposed Child	(6) Absent 22 or More Days
Panel A: Cohort Peers with Family and School Fixed Effects						
Share of peers with BLLs over 5 µg/dL	-0.0203	0.0295***	0.1526+	0.0930***	0.0429***	0.0215***
	(0.0193)	(0.0083)	(0.0835)	(0.0058)	(0.0037)	(0.0040)
Share of Non-White Children in School-Grade-Year	-0.0403**	-0.0298***	-0.1997*	-0.0101+	-0.0027	-0.0010
	(0.0143)	(0.0081)	(0.0888)	(0.0060)	(0.0039)	(0.0032)
Share of Economically Disadvantaged Children in School-Grade-Year	0.0098	0.0261***	0.2753***	0.0292***	0.0072**	0.0069***
	(0.0094)	(0.0048)	(0.0486)	(0.0037)	(0.0025)	(0.0021)
Observations	5,452,009	6,940,254	6,940,254	6,136,248	5,459,509	7,611,487
N Students	1,135,915	1,161,968	1,161,968	1,155,334	1,123,086	1,158,135
Mean of outcome	0.1175	0.1037	0.4552	0.0241	0.0108	0.0394
Panel B: Peers in the Same Classroom with Family and School Fixed Effects						
Share of peers with BLLs over 5 µg/dL	-0.1943***	0.0160*	0.0166	0.1286***	0.0718***	0.0229***
	(0.0202)	(0.0079)	(0.0794)	(0.0067)	(0.0039)	(0.0040)
Observations	3,063,248	4,906,795	4,906,795	4,201,754	4,201,754	4,775,787
N Students	878,866	1,073,872	1,073,872	965,999	965,999	1,063,034
Mean of outcome	0.1169	0.1238	0.5041	0.0199	0.0073	0.0424
Panel C: Average Exposure in Classrooms Grades 3-5 on 8th Grade Outcomes						
Share of peers with BLLs over 5 µg/dL	-0.2428**	0.0069	0.4684	0.0056	0.0257	-0.0075
	(0.0905)	(0.0438)	(0.4208)	(0.0307)	(0.0178)	(0.0231)
Observations	119,363	120,258	120,258	120,258	120,258	116,336
N Students	118,591	119,405	119,405	119,405	119,405	115,542
Mean of outcome	0.1602	0.1668	0.6007	0.0334	0.0155	0.0365

Notes: The table reports the effect of a child's share of peers with EBLLs on the child's school outcomes. Panel A uses the share of peers with maximum BLLs over 5 µg/dL *at the school-grade-year level* as the main explanatory variable, while panels B and C use the share of peers with maximum BLLs over 5 µg/dL *at the classroom level*. All regressions include cohort and individual controls, as well as family, birth month, birth order, grade, school, and year fixed effects. In column 1 we take the average of math and reading test scores and additionally control for subject-by-type test fixed effects. Individual controls include indicators for gender, race, economically disadvantaged status, and whether the student has a blood lead level test. Cohort controls include the share of non-White peers, share of children with a lead test, and the share of peers who are economically disadvantaged at the school-grade-year level. We also control for school size, the stability rate, and the percent of teachers with an MA degree. Panel C additionally includes our cohort and school controls averaged over elementary grades. Standard errors are in parentheses and clustered at the school level. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 3: Long-Run Outcomes of Exposure to Peers with Elevated BLLs by Timing of Exposure

Dependent Variable:	(1) Ever graduated	(2) Ever dropped out	(3) Intention to Attend a 4-Year College	(4) Intention to Attend a Community College	(5) Took SAT
	Panel A: Share of All Peers with EBLLs Over Elementary and Middle School (Cohort Variation)				
Share of peers with BLLs over 5 µg/dL	-0.1722***	0.0474+	-0.1017	0.0388	-0.2302**
	(0.0357)	(0.0244)	(0.0731)	(0.0800)	(0.0745)
Mean of outcome	0.8902	0.0530	0.5066	0.3291	0.5316
N Students	281,098	412,514	204,213	204,141	200,186
	Panel B: Share of All Peers with EBLLs Over Elementary and Middle School (Classroom Variation)				
Share of peers with BLLs over 5 µg/dL	-0.0458	0.0089	-0.3405***	0.2589*	-0.3990***
	(0.0453)	(0.0347)	(0.0960)	(0.1008)	(0.0853)
Mean of outcome	0.9382	0.0346	0.5225	0.3180	0.5331
N Students	145,518	182,153	148,776	148,015	147,608
	Panel C: Share of Elementary Versus Middle School Peers with EBLLs (Cohort Variation)				
Share of peers with BLLs over 5 µg/dL in Elementary School	-0.0547+	-0.0068	-0.0157	0.0188	0.0076
	(0.0295)	(0.0218)	(0.0674)	(0.0689)	(0.0677)
Share of peers with BLLs over 5 µg/dL in Middle School	-0.1270**	0.0749*	-0.0393	-0.0021	-0.2406**
	(0.0426)	(0.0311)	(0.0942)	(0.1018)	(0.0903)
Mean of outcome	0.8940	0.0523	0.5098	0.3305	0.5372
N Students	243,118	347,820	178,049	177,992	174,634

Notes: The table reports the effect of a child's share of peers with elevated blood lead levels on the child's long-run outcomes. We restrict the sample to the highest grade a student is observed in. Column 1 reports the effects on the likelihood a student ever graduates from high school, and column 2 shows the effects on the likelihood of ever dropping out of school. Columns 3 and 4 show the effects on self-reported intention of enrolling in a four-year college and community college, respectively. Column 5 shows the effects on the likelihood of taking the SAT test by grade 12. All regressions include individual controls, as well as family, birth month, birth order, school, grade and year fixed effects. Individual controls include indicators for gender, race, economically disadvantaged status, and whether the student has a blood lead level test. We also control for the average share of elementary and middle school peers that are non-White or economically disadvantaged, average share of children with a lead test, school size, the stability rate, and the percent of teachers with an MA degree averaged over elementary and middle school. Standard errors are in parentheses and clustered at the school level. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 4: Heterogeneity by Peer Gender, Race and Neighborhood

	Short Run Outcomes		*Long-run Outcomes*			
	(1)	(2)	(3)	(5)	(6)	(6)
Dependent Variable:	Average Test score	Any Suspension	Ever graduated	Intention to Attend a 4-Year College	Intention to Attend a Community College	Took the SAT
Panel A: By Same-Gender Lead Poisoned Peers						
Share of same-gender peers with BLLs ≥5 µg/dL	-0.0328	0.0774***	-0.1031+	0.1679	-0.2161	0.0403
	(0.0271)	(0.0116)	(0.0534)	(0.1223)	(0.1332)	(0.1305)
Share of peers with BLLs ≥5 µg/dL	-0.0036	-0.0101	-0.1204**	-0.1861*	0.1472	-0.2507*
	(0.0236)	(0.0103)	(0.0429)	(0.0940)	(0.1055)	(0.0982)
Panel B: By Same-Race Lead Poisoned Peers (White)						
Share of same-race peers with BLLs ≥5 µg/dL	-0.2525***	-0.0160	0.0284	-0.0691	0.3516**	-0.1616
	(0.0271)	(0.0117)	(0.0586)	(0.1160)	(0.1271)	(0.1220)
Share of peers with BLLs ≥5 µg/dL	0.1480***	0.0416***	-0.1913***	-0.0539	-0.2000+	-0.1199
	(0.0260)	(0.0106)	(0.0485)	(0.1074)	(0.1206)	(0.1124)
Panel C: By Same Gender-Race Lead Poisoned Peers (White)						
Share of same gender-race peers with BLLs ≥5 µg/dL	-0.2212***	0.0413***	-0.0343	0.0735	0.0852	-0.0796
	(0.0292)	(0.0121)	(0.0610)	(0.1356)	(0.1505)	(0.1432)
Share of peers with BLLs ≥5 µg/dL	0.0544*	0.0168+	-0.1604***	-0.1262	0.0097	-0.2028*
	(0.0214)	(0.0089)	(0.0376)	(0.0834)	(0.0949)	(0.0887)
N Students	1,135,912	1,161,968	281,098	204,213	204,141	200,186
Mean of outcome	0.1176	0.1037	0.8902	0.5066	0.3291	0.5316
Panel D: By Same-Race Lead Poisoned Peers (White), Male Sample						
Share of male same-race peers with BLLs ≥5 µg/dL	-0.1511***	0.0494**	-0.0486	-0.0116	-0.0984	-0.0860
	(0.0399)	(0.0189)	(0.1149)	(0.2547)	(0.2906)	(0.2681)
Share of peers with BLLs ≥5 µg/dL	0.0052	0.0025	-0.2491***	-0.0520	0.0724	-0.1741
	(0.0272)	(0.0130)	(0.0736)	(0.1631)	(0.1817)	(0.1573)
N Students	558,528	575,934	78,090	56,282	56,264	54,952
Mean of outcome	0.0803	0.1396	0.8771	0.4607	0.3285	0.4878

Notes: The table reports the effect of a child's share of peers with elevated blood lead levels on the child's school outcomes. Panel A reports the effect of a child's share of same-gendered peers with elevated blood lead levels on the child's school outcomes, Panel B reports the reports the effect of a child's share of same-race peers with elevated blood lead levels, and Panel C reports the effect of a child's same-race and same-gender share of peers with elevated blood lead levels. Panel D shows the effect of male same-race peers on male students. All regressions include the cohort, school-level and individual controls listed in equation (1), as well as family, birth month, birth order, school, grade, and year fixed effects. Cohort and school controls are averaged over elementary and middle school in Columns 3-6. Standard errors are in parentheses and clustered at the school level. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 5: Heterogeneity by Demographic Subgroups

	Short Run Outcomes		*Long-run Outcomes*			
Dependent Variable:	(1) Average Test Score	(2) Any Suspension	(3) Ever Graduated	(5) Intention to Attend a 4-Year College	(6) Intention to Attend a Community College	(6) Took the SAT
Panel A: White, non-Hispanic students						
Share of peers w/ BLLs ≥5µg/dL	0.0340	0.0176*	-0.1538***	-0.0864	0.0058	-0.1220
	(0.0247)	(0.0077)	(0.0409)	(0.0901)	(0.0995)	(0.0875)
Panel B: Black non-Hispanic students						
Share of peers w/ BLLs ≥5µg/dL	0.0363	0.0364*	-0.3181***	0.0837	0.0258	-0.3591*
	(0.0300)	(0.0165)	(0.0797)	(0.1777)	(0.1618)	(0.1724)
p-val. =White	0.95	0.30	0.07	0.39	0.92	0.22
Panel C: Hispanic students						
Share of peers w/ BLLs ≥5µg/dL	-0.0206	0.0117	-0.1335	-0.1228	-0.1692	-0.4146+
	(0.0467)	(0.0168)	(0.1259)	(0.2264)	(0.2484)	(0.2431)
p-val. =White	0.30	0.75	0.88	0.88	0.51	0.26
Panel D: Never Economically Disadvantaged students						
Share of peers w/ BLLs ≥5µg/dL	0.1251***	0.0134+	-0.0978*	-0.0274	-0.0654	-0.0308
	(0.0308)	(0.0070)	(0.0409)	(0.1044)	(0.1144)	(0.1116)
Panel E: Sometimes Economically Disadvantaged students						
Share of peers w/ BLLs ≥5µg/dL	-0.0117	0.0261*	-0.1901**	-0.1265	0.1213	-0.2706*
	(0.0272)	(0.0120)	(0.0688)	(0.1417)	(0.1410)	(0.1371)
p-val. =Never	0.00	0.36	0.25	0.57	0.30	0.17
Panel F: Always Economically Disadvantaged students						
Share of peers w/ BLLs ≥5µg/dL	-0.0535+	-0.0169	0.0218	-0.1510	-0.0132	-0.1347
	(0.0297)	(0.0135)	(0.1037)	(0.1826)	(0.2026)	(0.2043)
p-val. =Never	0.00	0.05	0.28	0.56	0.82	0.66
Panel G: Girls						
Share of peers w/ BLLs ≥5µg/dL	0.0032	0.0334***	-0.1136+	-0.0424	0.0380	-0.2491+
	(0.0234)	(0.0080)	(0.0611)	(0.1316)	(0.1446)	(0.1318)
Panel H: Boys						
Share of peers w/ BLLs ≥5µg/dL	-0.0483*	0.0194+	-0.2648***	-0.0563	0.0385	-0.2037
	(0.0231)	(0.0114)	(0.0676)	(0.1316)	(0.1486)	(0.1370)
p-val. =Girls	0.12	0.32	0.10	0.94	1.00	0.81

Notes: The table reports the effect of a child's share of peers with elevated blood lead levels on the child's school outcomes for children with different observable characteristics in each panel. For each outcome, results are from three regressions, one for each characteristic (race, economic status, gender). All regressions include cohort and individual controls, as well as family, birth month, birth order, school, grade, and year fixed effects. Individual controls include indicators for whether the student has a blood lead level test, gender, race, and economically disadvantaged status. Cohort controls include the share of non-White peers, share of children with a lead test, and the share of peers who are economically disadvantaged at the school-grade-year level. We also control for school size, the stability rate, and the percent of teachers with an MA degree. Cohort and school controls are averaged over elementary and middle school in Columns 3-6. Standard errors are clustered at the school level. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 6: Results for Zip Codes with Universal Screening, No School Choice, and Tracts with Majority Single Family Homes

	Short Run Outcomes		*Long-run Outcomes*			
	(1) Average Test Score	(2) Any Suspension	(3) Ever Graduated	(5) Intention to Attend a 4-Year College	(6) Intention to Attend a Community College	(6) Took the SAT
Panel A: Zip Codes with Universal Screening Based on >27% of Housing Being Built Before 1950						
Share of peers with BLLs over 5 µg/dL	-0.0533*	0.0176	-0.1895***	-0.0320	0.0103	-0.1829+
	(0.0251)	(0.0110)	(0.0460)	(0.1087)	(0.1131)	(0.0982)
N Students	459,255	465,438	145,787	102,116	102,099	100,217
Mean of outcome	0.0395	0.1173	0.8798	0.4833	0.3453	0.5140
Panel B: >50% of Homes in Census Tract are Single Family						
Share of peers with BLLs over 5 µg/dL	-0.1161***	0.0145	-0.1515*	-0.1508	0.0374	-0.1351
	(0.0333)	(0.0130)	(0.0703)	(0.1174)	(0.1226)	(0.1350)
N Students	321,303	355,820	84,612	61,308	61,300	60,214
Mean of outcome	0.0626	0.1061	0.8831	0.4739	0.3505	0.5106
Panel C: Zip Codes with No School Choice Options (No Charter Schools or Voucher Programs)						
Share of peers with BLLs over 5 µg/dL	-0.0316	0.0150	-0.2685***	-0.0577	-0.0106	-0.2185*
	(0.0258)	(0.0108)	(0.0706)	(0.0906)	(0.1042)	(0.0888)
N Students	717,120	705,750	175,103	125,337	125,282	123,114
Mean of outcome	0.1225	0.0957	0.8932	0.4846	0.3498	0.5128

Notes: The table reports the effect of a child's share of peers with elevated blood lead levels on the child's school outcomes. Each cell reports results from a separate regression. Panel A restricts the sample to students who live in zip codes that are subject to universal lead screening. Panel B restricts the sample to Census tracts where more than half of homes are single family homes. Panel C restricts the sample to zip codes without charter schools or voucher programs. All regressions include cohort and individual controls, as well as family, birth month, birth order, grade, school, and year fixed effects. In column 1 we additionally control for subject-by-type test fixed effects. Individual controls include indicators for gender, race, economically disadvantaged status, and whether the student has a blood lead level test. Cohort controls include the share of non-White peers, share of children with a lead test, and the share of peers who are economically disadvantaged at the school-grade-year level. We also control for school size, the stability rate, and the percent of teachers with an MA degree. Cohort and school controls are averaged over elementary and middle school in Columns 3-6. Standard errors are in parentheses and clustered at the school level. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table 7: The Effects of Peers on Switching Schools

	(1) Changed Schools	(2) Changed to a Charter School	(3) Child's Sibling Changed Schools	(4) Both Siblings Changed Schools	(5) Consecutive Younger Sibling is in a Different School for the Same Grade
Share of peers with BLLs over 5 µg/dL	0.0217	0.0017	0.0082	-0.0022	-0.0237
	(0.0142)	(0.0019)	(0.0058)	(0.0046)	(0.0158)
Observations	6,372,937	6,372,937	6,900,757	6,928,249	4,172,510
N Students	1,011,814	1,011,814	1,045,756	1,049,982	575,582
Mean of outcome	0.3082	0.0046	0.1113	0.0507	0.1878

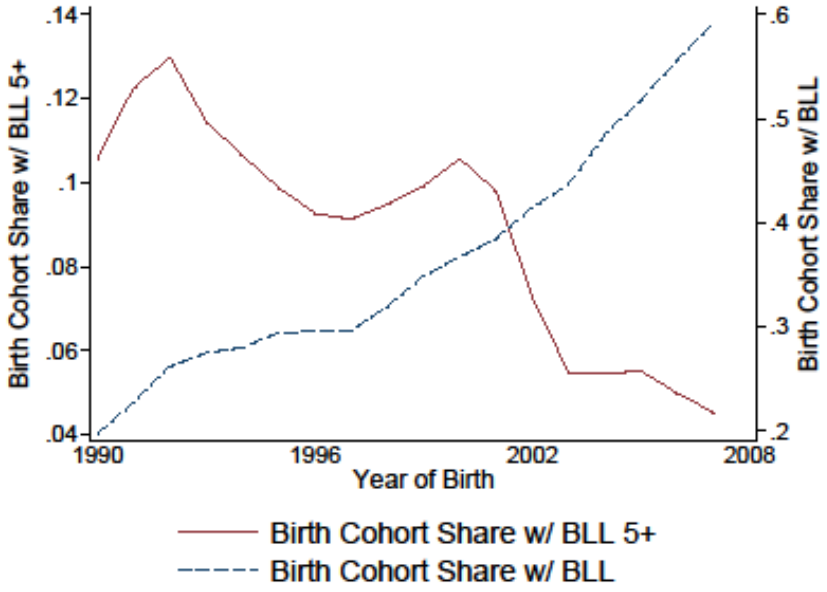
Notes: The table reports the association of a child's share of peers with elevated blood lead levels with the child's own likelihood of switching schools (columns 1 and 2), the child's sibling's likelihood of switching schools (column 3), both children switching schools (column 4), and the likelihood that a consecutive younger sibling attends a different school than the child's school for the same grade (column 5). All regressions include cohort and individual controls, as well as family, birth month, birth order, school, grade, and year fixed effects. Individual controls include indicators for gender, race, economically disadvantaged status, and whether the student has a blood lead level test. Cohort controls include the share of non-White peers, share of children with a lead test, and the share of peers who are economically disadvantaged at the school-grade-year level. We also control for school size, the stability rate, and the percent of teachers with an MA degree. Standard errors are clustered at the school level. + $p<0.10$, * $p<0.05$, ** $p<0.01$, *** $p<0.001$.

Table 8: Correlation Between Share of Peers with Blood Lead Levels at or above 5µg/dL and a Child's Own Characteristics

Dependent Variable:	(1) Child is Female	(2) Cohort size in School-Grade-Year	(3) Child is Economically Disadvantaged	(4) Child is missing test scores	(5) Cohort (School-Grade-Year) Stability Rate	(6) Share of teachers with Masters or higher
Share of peers with BLLs over 5 µg/dL	0.0270**	6.3857+	0.0057	-0.1330***	0.0052**	-0.0067
	(0.0086)	(3.2806)	(0.0069)	(0.0114)	(0.0019)	(0.0076)
Observations	1,171,475	1,171,475	1,171,475	1,171,475	1,171,475	1,171,475
Mean of Outcome	0.4935	231.4184	0.4032	0.2816	0.9572	0.3657

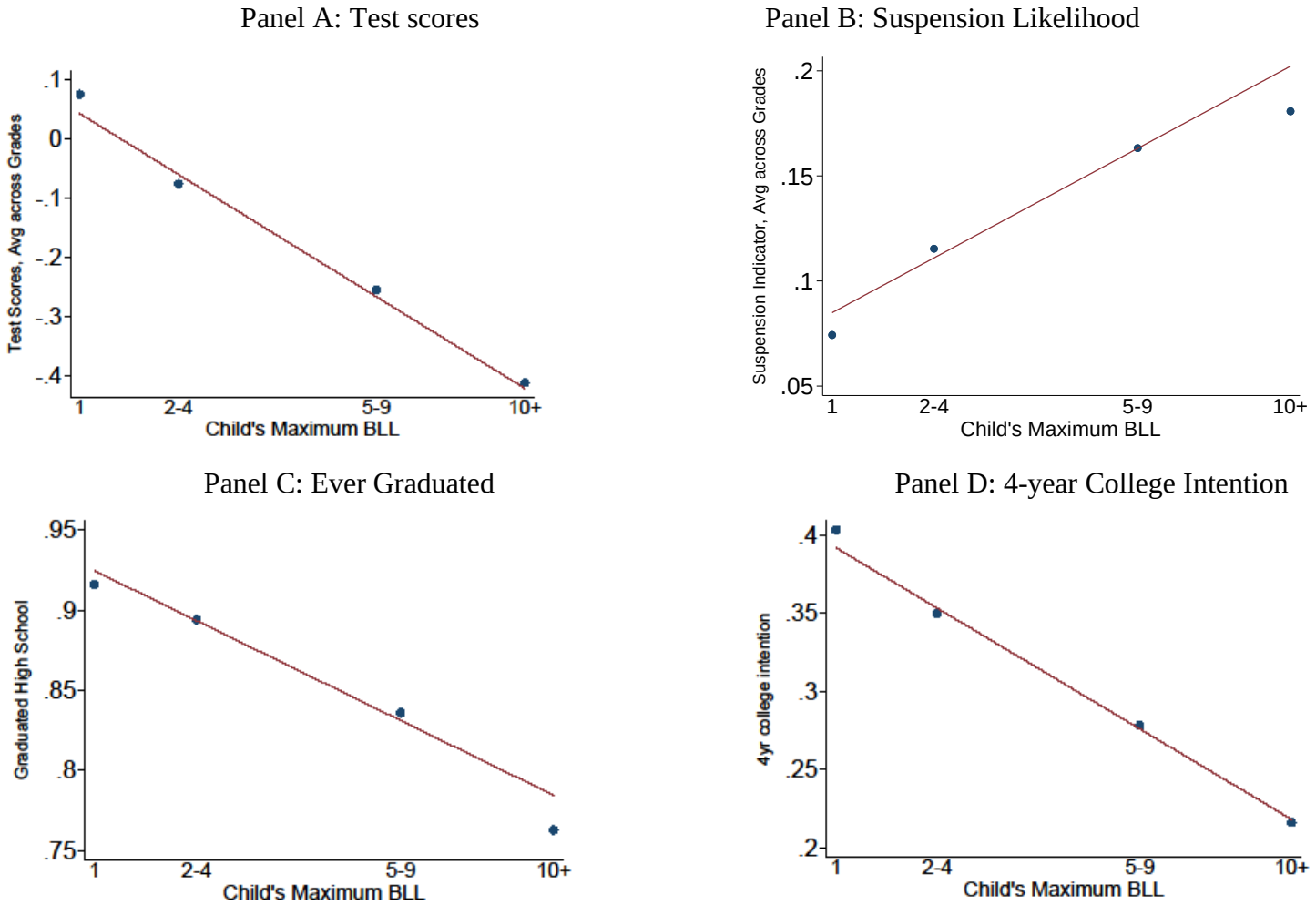
Notes: The table reports the correlation between a child's share of peers with elevated blood lead levels and the child's characteristics indicated in each column. Each cell reports results from a separate regression. All regressions include controls for gender, race, economically disadvantaged status, whether the student has a blood lead level test, share of non-White peers, share of children with a lead test, and the share of peers who are economically disadvantaged at the school-grade-year level. All regressions include family, birth month, birth order, school, grade, and year fixed effects. Standard errors are in parentheses and clustered at the school level. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Figure 1: Share of Children with Blood Lead Levels at or above 5µg/dL by Birth Cohort and Share of Children with Blood Lead Tests by Cohort



Notes: The figure plots the share of children in a school-grade-year cohort with at least one blood lead test (blue dashed line) and with a blood lead level of at least 5µg/dL (red solid line)

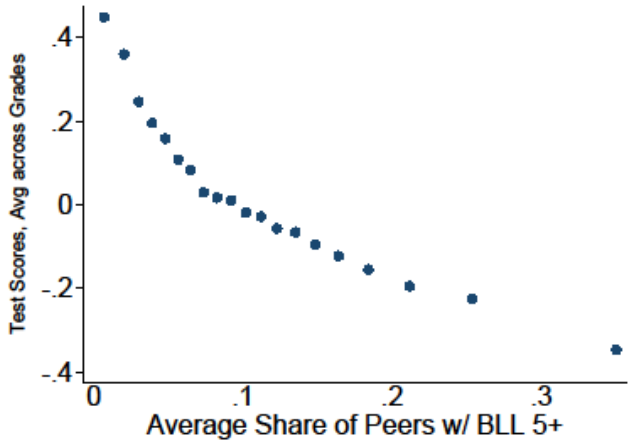
Figure 2: The Relationship Between a Child's Own Blood Lead Levels and, Test Scores, Suspensions, and Dropping out of School



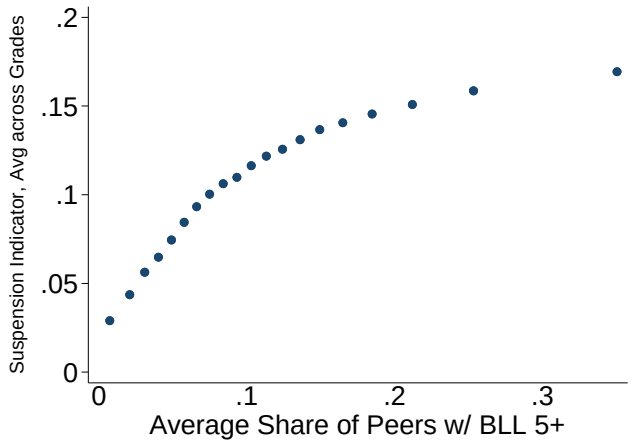
Notes: The figure plots average test scores (Panel A), suspension rates (Panel B), graduation rates (Panel C), college and community college intention rates (Panels D and E), and SAT taking rates (Panel F) by students' blood lead levels and adds the line of best fit.

Figure 3: The Relationship Between Peers' Blood Lead Levels and Test Scores, Suspensions, and Dropping out of School

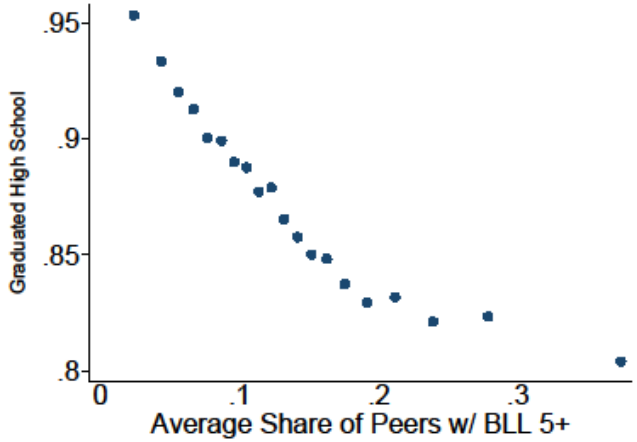
Panel A: Test scores



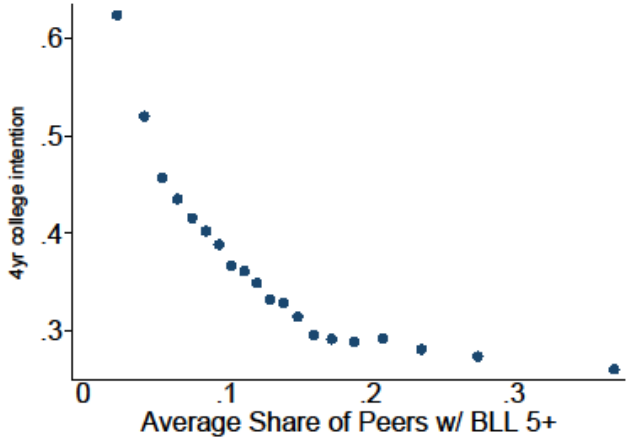
Panel B: Suspension Likelihood



Panel C: Ever Graduated



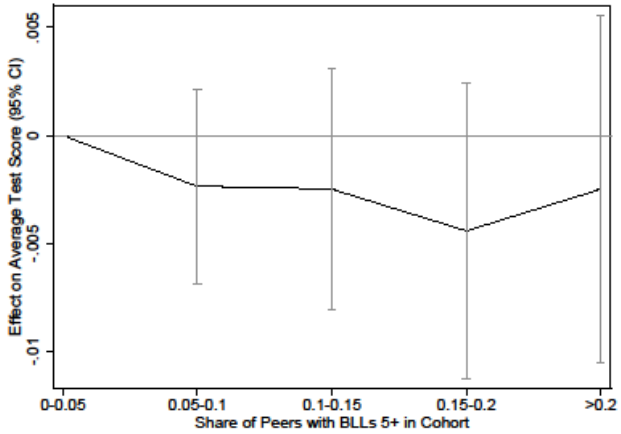
Panel D: 4-year College Intention



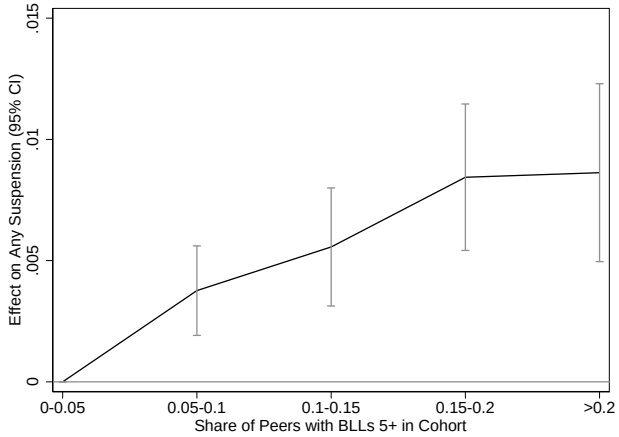
Notes: The figure plots average test scores (Panel A), suspension rates (Panel B), graduation rates (Panel C), college and community college intention rates (Panels D and E), and SAT taking rates (Panel F) by quintiles of students' share of peers with blood lead levels at or above 5µg/dL.

Figure 4: Binned Effects of Share of Peers with Blood Lead Levels above 5 µg/dL

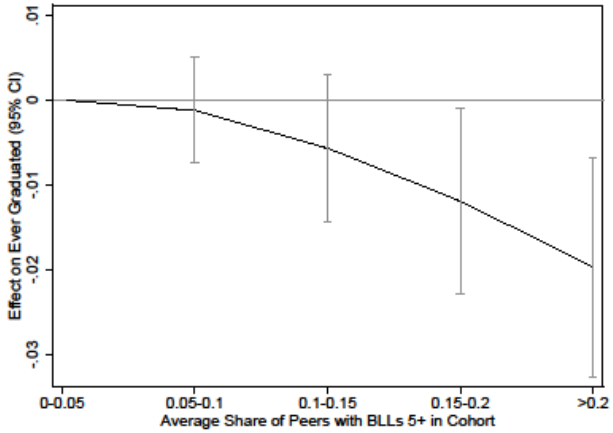
Panel A: Test scores



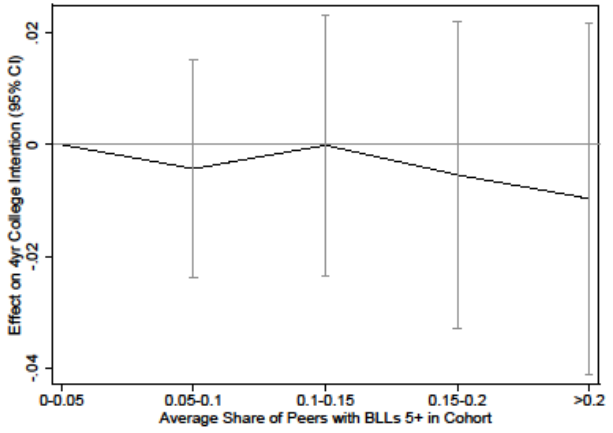
Panel B: Suspension Likelihood



Panel C: Ever Graduated



Panel D: 4-year College Intention



Notes: Each figure plots non-parametric estimates of the effect of having different proportions (binned) of peers with BLLs 5+ in a child's cohort on average test scores (Panel A), suspension rates (Panel B), graduation rates (Panel C), college and community college intention rates (Panels D and E), and SAT taking rates (Panel F). The omitted category is an indicator for share of peers with BLLs 5+ that is lower than 0.05. We control for all fixed effects and controls in our primary specification (which includes family, school, year, and grade fixed effects, and individual and demographic controls by cohort, averaged over elementary and middle school in Panels C and D.). Vertical bars represent 95% confidence intervals based on standard errors clustered at the school level.

APPENDIX: FOR ONLINE PUBLICATION

A. Additional Tables and Figures

Table A1: Heterogeneity by School-Level Demographics

	Short Run Outcomes		*Long-run Outcomes*			
	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable:	Average Test Score	Any Suspension	Ever Graduated	Intention to Attend a 4-Year College	Intention to Attend a Community College	Took the SAT
Panel A: Schools in Lowest Tercile of Share Students who are Economically Disadvantaged						
Share of peers with BLLs over 5 µg/dL	0.1714**	0.0113	-0.1739+	-0.0665	-0.2822	-0.4692*
	(0.0576)	(0.0173)	(0.1046)	(0.1638)	(0.2027)	(0.2205)
Panel B: Schools in Middle Tercile of Share Students who are Economically Disadvantaged						
Share of peers with BLLs over 5 µg/dL	0.0560	0.0334*	-0.1210*	-0.0902	0.0593	-0.0841
	(0.0378)	(0.0147)	(0.0583)	(0.1330)	(0.1515)	(0.1361)
p-val = First Tercile	0.09	0.33	0.66	0.91	0.18	0.14
Panel C: Schools in Highest Tercile of Share Students who are Economically Disadvantaged						
Share of peers with BLLs over 5 µg/dL	-0.0126	0.0520***	-0.1814*	-0.2763*	0.2439+	-0.3690**
	(0.0268)	(0.0128)	(0.0775)	(0.1359)	(0.1415)	(0.1384)
p-val = First Tercile	0.00	0.06	0.95	0.32	0.03	0.70
N Students	1,124,003	1,154,452	199,394	153,846	153,798	151,399
Mean of outcome	0.1213	0.1034	0.9038	0.5177	0.3233	0.5396

Notes: The table reports the effect of a child's share of peers with elevated blood lead levels on the child's school outcomes for children in schools with different shares of children who are economically disadvantaged in each panel. For each outcome, results are from a single regression. All regressions include cohort and individual controls, as well as family, birth month, birth order, school, grade, and year fixed effects. Individual controls include indicators for gender, race, economically disadvantaged status, and whether the student has a blood lead level test. Cohort controls include the share of non-White peers, share of children with a lead test, and the share of peers who are economically disadvantaged at the school-grade-year level. We also control for school size, the stability rate, and the percent of teachers with an MA degree. Cohort and school controls are averaged over elementary and middle school in Columns 3-6. Standard errors are clustered at the school level. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table A2: Alternative measures of BLL

	Short Run Outcomes		*Long-run Outcomes*			
	(1) Average Test Score	(2) Any Suspension	(3) Ever Graduated	(4) Intention to Attend a 4-Year College	(5) Intention to Attend a Community College	(6) Took the SAT
Panel A: Share of Male and Female Peers with Max BLL over 5 µg/dL						
Share of male peers with BLLs over 5 µg/dL	-0.0231	0.0791***	-0.1561***	-0.1148	0.0494	-0.2448*
	(0.0246)	(0.0110)	(0.0462)	(0.0961)	(0.1075)	(0.1034)
Share of female peers with BLLs over 5 µg/dL	-0.0137	-0.0260*	-0.1902***	-0.0851	0.0303	-0.2066+
	(0.0270)	(0.0115)	(0.0478)	(0.1067)	(0.1165)	(0.1078)
Panel B: Mean BLL						
Share of peers with BLLs over 5 µg/dL	-0.0398+	0.0164	-0.2094***	-0.0820	-0.0015	-0.2357**
	(0.0209)	(0.0101)	(0.0356)	(0.0769)	(0.0825)	(0.0741)
Panel C: Max BLL is over 10 µg/dL						
Share of peers with BLLs over 10 µg/dL	-0.0052	-0.1233***	-0.3955***	-0.2008	0.0515	-0.4545**
	(0.0449)	(0.0250)	(0.0858)	(0.1518)	(0.1701)	(0.1520)
N Students	1,135,915	1161968	281,098	204,213	204,141	200,186
Mean of outcome	0.1175	0.1037	0.8902	0.5066	0.3291	0.5316
Panel D: Including Students with Elevated Blood Lead Levels						
Share of peers with BLLs over 5 µg/dL	-0.0241	0.0219*	-0.1810***	-0.1045+	0.0495	-0.2042***
	(0.0185)	(0.0087)	(0.0331)	(0.0616)	(0.0687)	(0.0600)
N Students	1,284,429	1,310,320	359,422	260,673	260,581	255,215
Mean of outcome	0.0684	0.1125	0.8768	0.4775	0.3435	0.5031

Notes: The table reports the effect of a child's share of peers with elevated blood lead levels on the child's school outcomes using different measures of peer exposure based on blood lead levels. Panel A uses the share of male and share of female peers with maximum BLL over 5 µg/dL. Panel B uses the share of peers with average BLL above 5 µg/dL. Panel C uses the share of peers with maximum BLL over 10 µg/dL. Panel D includes children who have maximum BLL over 5 µg/dL. All regressions include cohort and individual controls, as well as family, birth month, birth order, school, grade, and year fixed effects. In column 1 we additionally control for subject-by-type test fixed effects. Individual controls include indicators for gender, race, economically disadvantaged status, and whether the student has a blood lead level test. Cohort controls include the share of non-White peers, share of children with a lead test, and the share of peers who are economically disadvantaged at the school-grade-year level, school size, the stability rate, and the percent of teachers with an MA degree. Cohort and school controls are averaged over elementary and middle school in Columns 3-6. Standard errors are clustered at the school level. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Table A3: Results with More Stringent Sets of Fixed effects

	Short Run Outcomes		*Long-run Outcomes*			
Dependent Variable:	(1) Average Test Score	(2) Any Suspension	(3) Ever Graduated	(4) Intention to Attend a 4-Year College	(5) Intention to Attend a Community College	(6) Took the SAT
Panel A: Sibling-School Fixed Effects						
Share of peers with BLLs over 5 µg/dL	-0.0162	0.0046	-0.1096***	-0.0570	-0.0016	-0.2160***
	(0.0232)	(0.0089)	(0.0272)	(0.0608)	(0.0691)	(0.0635)
Observations	4,928,691	6,500,548	226,517	165,983	165,943	163,507
N Students	1,091,665	1,129,993	226,517	165,983	165,943	163,507
Panel B: With Family, School, and Block Group Fixed Effects						
Share of peers with BLLs over 5 µg/dL	-0.0637**	0.0097	-0.2211***	-0.1030	0.1010	-0.2120+
	(0.0247)	(0.0102)	(0.0601)	(0.1236)	(0.1424)	(0.1243)
Observations	2,771,076	3,638,450	118,027	70,128	70,116	68,148
N Students	733,456	799,803	118,027	70,128	70,116	68,148
Panel C: Family and School-Grade Fixed Effects						
Share of peers with BLLs over 5 µg/dL	-0.0180	0.0227**	-0.1530***	-0.0864	0.0261	-0.2227**
	(0.0195)	(0.0084)	(0.0314)	(0.0737)	(0.0811)	(0.0749)
Observations	5,451,919	6,940,142	280,651	203,737	203,696	200,062
N Students	1,135,909	1,161,966	280,651	203,737	203,696	200,062
Panel D: Family, School, and Grade-Year Fixed Effects						
Share of peers with BLLs over 5 µg/dL	-0.0532**	-0.0060	-0.0848**	-0.0796	0.0633	-0.2256**
	(0.0197)	(0.0086)	(0.0326)	(0.0733)	(0.0807)	(0.0744)
Observations	5,452,009	6,940,254	281,090	204,203	204,135	200,184
N Students	1,135,915	1,161,968	281,090	204,203	204,135	200,184
Mean of outcome	0.1175	0.1037	0.8902	0.5066	0.3291	0.5316

Notes: The table reports the effect of a child's share of peers with elevated blood lead levels on the child's school outcomes. Each cell reports results from a separate regression. All regressions include controls for gender, race, economically disadvantaged status, whether the student has a blood lead level test, share of non-White peers, share of children with a lead test, and the share of peers who are economically disadvantaged at the school-grade-year level. We also control for school size, the stability rate, and the percent of teachers with an MA degree. Cohort and school controls are averaged over elementary and middle school in Columns 3-6. All regressions include birth month and birth order fixed effects. Panel A includes sibling-school, grade and year fixed effects. Panel B includes family, school, year, grade and block group fixed effects. Panel C includes sibling, year and school-grade fixed effects. Panel D includes family, school and grade-year fixed effects. Standard errors are in parentheses and clustered at the school level. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

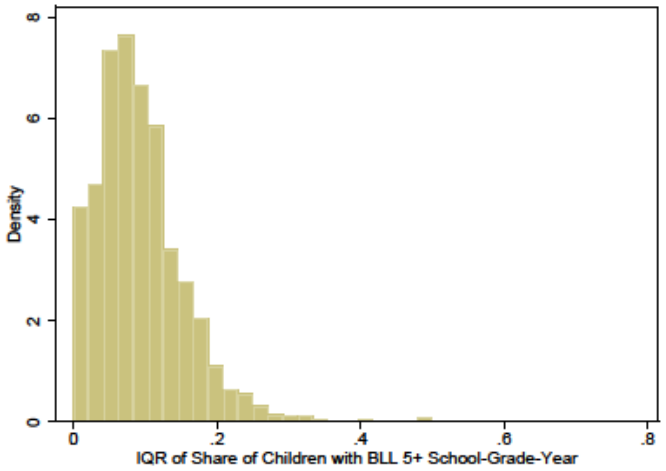
Table A4: Results with Fewer Controls

	Short Run Outcomes		*Long-run Outcomes*			
Dependent Variable:	(1) Average Test Score	(2) Any Suspension	(3) Ever Graduated	(4) Intention to Attend a 4-Year College	(5) Intention to Attend a Community College	(6) Took the SAT
			Panel A: No controls			
Share of peers with BLLs over 5 µg/dL	-0.0578***	0.0870***	-0.0177	-0.0569	0.0811	-0.0773
	(0.0154)	(0.0073)	(0.0241)	(0.0553)	(0.0593)	(0.0560)
Observations	5,572,319	7,083,319	283,032	205,832	205,760	201,783
N Students	1,144,411	1,168,436	283,032	205,832	205,760	201,783
			Panel B: All Controls Except for Share Non-White and Share Educationally Disadvantaged			
Share of peers with BLLs over 5 µg/dL	-0.0247	0.0334***	-0.1689***	-0.1106	0.0466	-0.2164**
	(0.0192)	(0.0082)	(0.0345)	(0.0733)	(0.0790)	(0.0725)
Observations	5,452,009	6,940,254	281,098	204,213	204,141	200,186
N Students	1,135,915	1,161,968	281,098	204,213	204,141	200,186
			Panel C: Family Fixed Effects			
Share of peers with BLLs over 5 µg/dL	-0.0363+	0.0643***	-0.1630***	-0.1078	0.0575	-0.2328**
	(0.0199)	(0.0089)	(0.0377)	(0.0726)	(0.0797)	(0.0745)
Observations	5,452,019	6,940,258	281,302	204,265	204,186	200,216
N Students	1,135,917	1,161,968	281,302	204,265	204,186	200,216

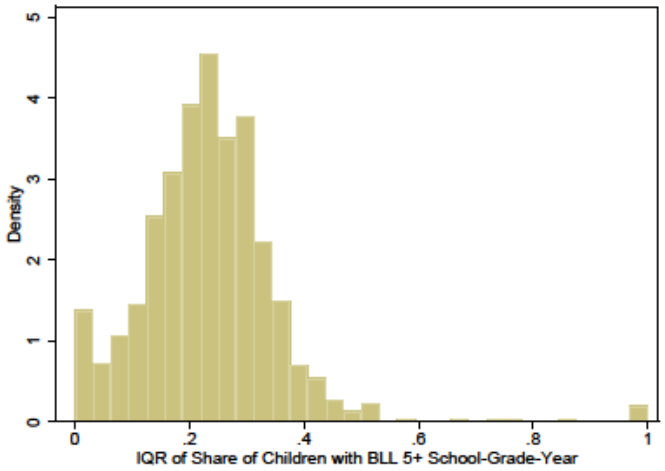
Notes: The table reports the effect of a child's share of peers with elevated blood lead levels on the child's school outcomes. Each cell reports results from a separate regression. All regressions include sibling, birth month, grade, year and birth order fixed effects. Panel A shows our results with no control variables except for our fixed effects and school fixed effects. Panel B includes school fixed effects and controls for gender, race, economically disadvantaged status, whether the student has a blood lead level test, the share of children with a lead test at the school-grade-year level, as well as school size, the stability rate, and the percent of teachers with an MA degree. We omit cohort-level controls for share of non-White peers and share of peers who are economically disadvantaged. Panel C includes our fixed effects together with all controls in our main specification but omits school fixed effects. Cohort and school controls are averaged over elementary and middle school in Columns 3-6. Standard errors are in parentheses and clustered at the school level. + p<0.10, * p<0.05, ** p<0.01, *** p<0.001.

Figure A1: Identifying Variation: Within-School Interquartile Range of Share of Children in 3rd Grade with Elevated Blood Lead Levels

Panel A: All Children as Denominator

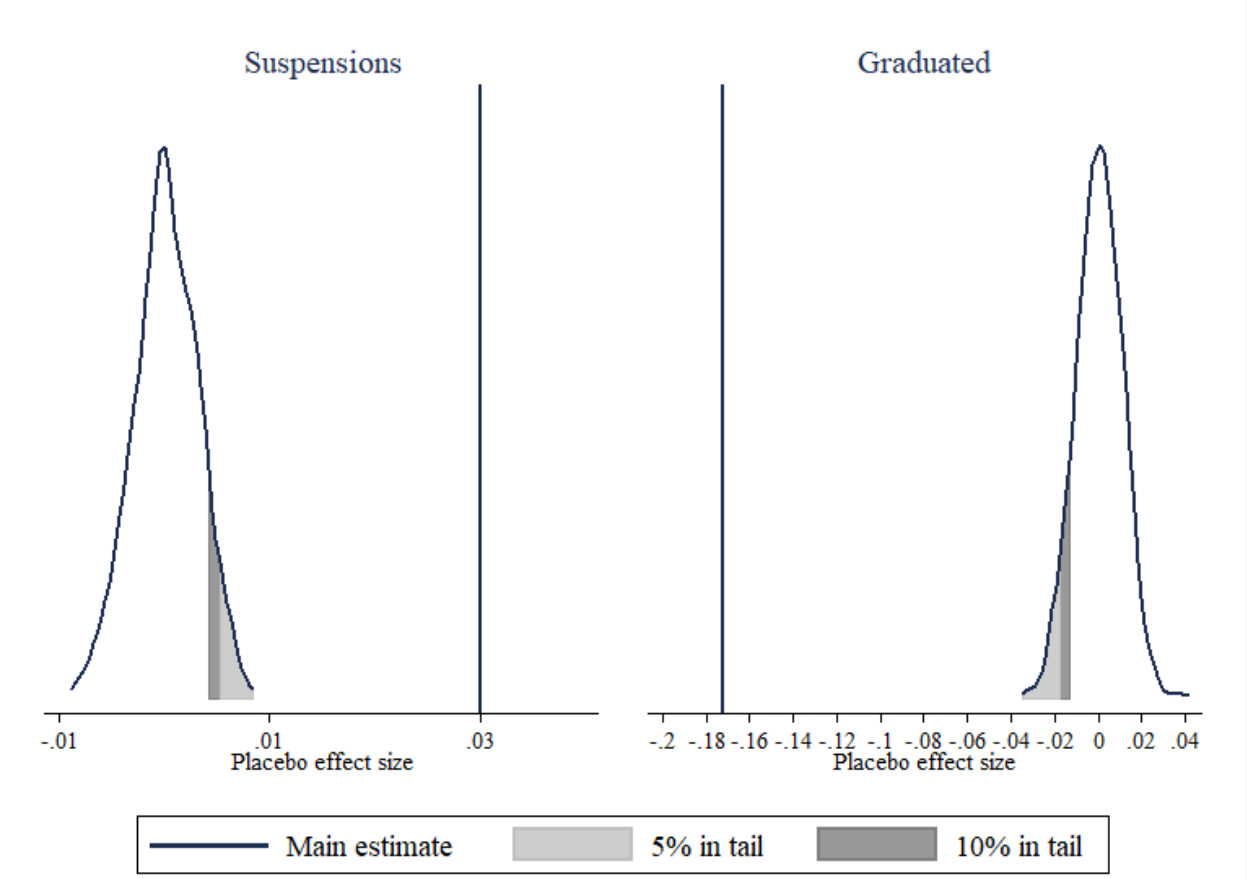


Panel B: Only Tested Children as Denominator



Notes: This figure plots the distribution of the school level interquartile range of the share of children in a school-grade-year with blood lead levels at or above 5µg/dL. We limit the sample to children in 3rd grade and compute the share of children in a school-year with blood lead levels at or above 5µg/dL. We then compute the interquartile range of this variable at the school level. Panel A uses all children and treats missing blood lead levels as zeros, while Panel B uses only children with blood lead tests.

Figure A2: Placebo Estimates



Notes: Distribution of results from 500 placebo tests per outcome. Our main estimates for our preferred specification are represented with a vertical line on the placebo effect size distribution. The lightly shaded gray region is the region of the graph where there is 5% in the tail of the distribution. The darker shaded gray region represents 10% in the tail of the distribution. For each placebo, school-grade cohorts were randomly assigned a percent of peers with EBLLs from the empirically observed distribution and we estimated our main specification.

B: Data Appendix

B1. Data linkage

NCERDC performed the linkage between the education and BLL data according to the following algorithm and anonymized the dataset for us. Appendix Table B1 reports the number of tests matched at each step.

1. Exact match on school district, that is local education agency (LEA), or county and first and last name, date of birth.
2. Exact match on first and last name, date of birth
3. Exact match on LEA or county and first and last name, but allow for mistakes in one of day, month, or year of birth
4. Exact match on LEA or county, last name, and date of birth, allow for close first name or nickname
5. Exact match on LEA or county, first name, and date of birth, allow for close last name
6. Exact match on last name, date of birth, allow for close first name or nickname
7. Exact match on first name, date of birth, allow for close last name
8. Exact match on first and last name, but allow for mistakes in one of day, month, or year of birth
9. Exact match on first and last name

Table B1: Match Results

(1) Match Step	(2) Number of Tests	(3) Share
1	1352623	0.606457
2	431987	0.193684
3	24098	0.010804
4	104751	0.046966
5	190154	0.085257
6	32860	0.014733
7	44963	0.020159
8	5168	0.002317
9	43765	0.019622

Notes: This table reports the additional number of tests matched at each step. Column 1 reports the match step, Column 2 reports the number of standardized tests, and Column 3 reports the share of children with each of these.

B2. Sibling Identification Algorithm

In this data appendix, we describe the algorithm used to identify siblings using students' geocoded home addresses.

There are 4.38 million unique students in the NCERDC data. Of these, about 740,000 do not have a home address and another 640,000 do not have birthday information. Since both home addresses and birthdays are crucial for identifying siblings, we drop these observations when running the linkage algorithm. We also ignore about 660,000 students who never share a home address with another student and therefore do not have siblings in our data.

We further restrict our sample to include home addresses with at most four students in any given year. We do this for several reasons. First, the geocoded address variable provided by NCERDC is based on street address and does not distinguish between different units that share the same street address. This means that students living in different apartment units within the same building appear to be living at the same home address. Because of this, we observe addresses with hundreds of students in a given year, and it is implausible that these students are siblings. Second, we observe that students who share a geocoded address with many other students often move across addresses. We suspect some of these students are in the foster care system and therefore it is difficult to identify their siblings with certainty. Three, according to the 2000 Census, the average number of children per family in North Carolina is 1.75, and thus we are conservative in limiting the number of children living together in any given year to at most four. Four, the algorithm speed is decreasing in the number of students living together in any given year. Thus, we apply our algorithm to addresses with no more than four students in a given year. This selection eliminates about 211,000 students, 80,000 of which always share an address with at least four other children.

We are left with about 2.12 million students on which we run the sibling identifying algorithm. The following steps summarize the process:

1. Identify all students who live together at any point or could be living together by transitivity and assign a tentative family identifier to these students. For example, Ana and Bob are observed living together in some years, Bob and Claire are observed living together in other years, but Ana and Claire are never observed living together. We temporarily assign Ana, Bob, and Claire to the same family.

2. For each potential sibling pair within the temporary families, check if the students are ever observed living at different addresses in the same year and if they are born between 2 and 240 days of each other.[14] If at least one of these holds, the students cannot be siblings. This step produces a dummy variable for each student within the temporary family that equals 1 whenever another student within the temporary family is a potential sibling, and zero otherwise. Table B2 shows a simple scenario for a tentative family with three students where all three can be siblings to one another. In such cases, we assign the temporary family a permanent household identifier.

Table B2

Student	Student1	Student2	Student3
1	1	1	1
2	1	1	1
3	1	1	1

Table B3

Student	Student1	Student2	Student3
1	1	1	0
2	1	1	1
3	0	1	1

3. Table B3 shows a tentative family where not all students can be siblings to one another: student 1 could be a sibling to student 2 but not to student 3, while student 2 could be a sibling to both students 1 and 3. Based on the indices, we conclude there are two potential true sibling groups: either students 1 and 2 are siblings, or students 2 and 3 are siblings. For each potential sibling group, we calculate a score based on the number of years students live together, the number of students in the subgroup, and the span of years for which the students are observed. Specifically:

$$score_g = \frac{\left(\sum_y \sum_{i,j \in g} \mathbb{I}_{j \neq i,y}\right)^2}{N_g} + \frac{\sum_y \sum_{i,j \in g} \mathbb{I}_{j \neq i,y}}{N_y}$$

where *i* and *j* denote students in subgroup *g*, and *y* denotes year. $\mathbb{I}_{j \neq i,y}$ equals 1 if student *i* and student *j* are observed living together in year *y*. $\sum_y \sum_{i,j \in g} \mathbb{I}_{j \neq i,y}$ equals the number of times students in the subgroup live with each other, allowing for double counting. N_g denotes the number of students in subgroup *g*. N_y the is the difference between the first and last year subgroup *g* is observed. For example, if a subgroup is first observed living together in 2000 and last observed in 2005, N_y equals 5. The first term of the index gives more weight to subgroups where students are observed living together more often

[14] We allow students to be born on the same or consecutive days to account for twins.

per student. The second term gives more weight to subgroups observed living together in consecutive years as opposed to many years apart. The subgroup with the highest score is assigned a permanent family identifier, and the step is repeated until all students in the temporary family are assigned a family identifier.

Table B4 shows the distribution of children across family size produced by our algorithm. Almost half of the children have only one sibling (columns 2 and 3), and about 84 percent of families have at most two children (column 5). Dividing the total number of children by the total number of families gives an average number of children per family of 1.80, which is similar to the figure provided by the Census.

Table B4: Distribution of children across family size

(1) Family size	(2) # of children	(3) % of children	(4) # of families	(5) % of families
1	457,796	21.56%	457,796	38.97%
2	1,054,842	49.68%	527,421	44.90%
3	458,760	21.61%	152,920	13.02%
4	127,036	5.98%	31,759	2.70%
5	19,960	0.94%	3,992	0.34%
6	3,798	0.18%	633	0.05%
7	791	0.04%	113	0.01%
8	144	0.01%	18	0.00%
9	45	0.00%	5	0.00%
Total	2,123,172	100%	1,174,657	100%

B3. Sample Selection and Variable Construction

Sample selection criteria: We drop children who are singletons or who live in very large buildings such that we are unable to determine who their siblings are for our main analysis sample. However, all children are used to determine cohort and class size, as well as the percentages of EBLLs, ED students, and Black and Hispanic students by cohort.

Test scores: We standardize mathematics and reading test scores at the grade-year level, and we average the two. When one is missing, we retain the non-missing test score.

Lead exposure definition: Capillary tests are more prone to false positives than venous tests. Thus, to identify lead-poisoned children we used the highest venous test result if available and the highest capillary test result if no venous test was performed.

We construct two measures of peer exposure, one at the cohort, that is school-grade, level and one at the classroom level. To measure class membership we compute the average mathematics classroom size over third through fifth grade.