I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13133

The Spillover Effects of Pollution: How Exposure to Lead Affects Everyone in the Classroom

Ludovica Gazze
Claudia Persico
Sandra Spirovska

APRIL 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13133

The Spillover Effects of Pollution: How Exposure to Lead Affects Everyone in the Classroom

Ludovica Gazze
University of Chicago

Claudia Persico
American University and IZA

Sandra Spirovska
University of Wisconsin-Madison

APRIL 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

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

Evidence shows that lead-exposed children are more disruptive and have lower achievement. However, we know less about how lead-exposed children affect the learning environment of their classroom peers. We estimate these spillover effects using new data on children's blood lead levels (BLLs) matched to all education data in North Carolina. We compare siblings who attend the same school, but whose school-grade cohorts differ in the proportion of children with elevated BLLs. We find that having more lead-exposed peers is associated with lower test scores and graduation rates, increased suspensions and dropping out of school, and a decrease in college intentions.

JEL Classification: Q53, I24, I14

Keywords: pollution, education, health, peers

Corresponding author:
Claudia Persico
Department of Public Administration and Policy
School of Public Affairs
American University
4400 Massachusetts Ave, NW
Washington, DC 20016-8070
USA
E-mail: cpersico@american.edu

* 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 also grateful to the Institute for Research on Poverty at the University of Wisconsin-Madison and American University's School of Public Affairs and Provost for their generous financial support of this research. All errors are our own. The North Carolina school and blood lead levels data is restricted under contract and is not available for sharing with the public. We will share our .do files and a readme file for replication purposes.

I. Introduction

Lead poisoning is extremely costly to children and society, as it can cause children to engage in risky behavior (Aizer and Currie 2018; Reyes 2015) and have poor academic achievement (Aizer et al. 2015; Ferrie, Rolf, and Troesken 2012; Grönqvist, Nilsson, and Robling forthcoming). Large effects on student achievement, school suspensions, and crime can occur at blood lead levels (BLLs) as low as 1–2 micrograms per deciliter of blood (Aizer et al. 2018; Feigenbaum and Mueller 2016; National Toxicology Program 2012). Moreover, lead may affect children's health and disability status (Gazze 2016). These negative effects of lead exposure are costly to children, families and society in terms of health, reduced tax revenues and increased expenditure on special education and crime (Korfmacher 2003; Reyes 2014). Although children's blood lead levels have declined over time (Aizer et al. 2018), recent estimates suggest that half a million young children are still poisoned by lead each year. Low-income children are up to 12 times more likely to have elevated blood-lead-levels (EBLLs) (CDC 2005; 2014), and African American children aged 1 to 5 are more than twice as likely to be lead-poisoned than their white peers (CDC 2005).

So far, the economic and public health literature has largely focused on estimating the effects of lead poisoning on directly exposed children. However, these children have daily interactions with peers that may also be affected by lead poisoning. In this paper, we ask whether there are spillover effects of lead poisoning on children who are exposed to school peers with high blood lead levels, even if the children themselves might not be directly exposed to lead. Because children exposed to lead are more disruptive, have lower achievement and engage in risky behavior, lead exposure and its effects on child health may spill over to affect everyone in the classroom. These spillover effects of lead exposure are an unexplored mechanism through which social context, pollution and built environment affect schools and children's outcomes. While hundreds of billions of dollars are spent annually on public education in the United States, many children might not benefit at the fullest from this education because of where they grow up. By finding that there are externalities of lead exposure, this paper shows that environmental hazards are a key factor contributing to human capital accumulation even for children who are not themselves exposed to these hazards.

Rigorously estimating peer effects is challenging because peers influence each other simultaneously, so it is unclear whether a disruptive child is causing his fellow classmates to misbehave, or whether the classmates cause him 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., lead 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 who attended the same school a few years apart and whose cohorts happen to randomly differ in the proportion of children with elevated blood lead levels in their grade-cohort. Our preferred specification includes sibling, school, grade, birth month, and year fixed effects, and controls for a broad set of child and cohort demographic characteristics and health status. Because elevated BLLs have been linked to behavioral incidents in school, criminality, and lower test scores, we use them as a proxy for peers with potential behavioral, health, and academic problems. This methodology avoids the reflection problem because a child cannot affect the BLLs of her peers, but elevated BLLs could affect children negatively, which in turn may affect peers. Furthermore, including siblings 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 classrooms with fewer lead-poisoned children.

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 0.54 percent of a standard deviation decrease in test scores and a 1.3 percentage point increase in the likelihood of suspension from school, which represents a 17 percent increase in suspensions. We also find a 0.06 percentage point increase in the likelihood that a child drops out of school, which represents a 30 percent increase in the dropout rate. In addition, having one additional lead-exposed peer in a child's elementary school classroom of 20 is associated with 1 percent of a standard deviation lower average test scores in eighth grade and a 0.12 percentage point increase in the likelihood that the unexposed child will ever drop out of school. We also find that having more lead-exposed children in a child's elementary school classroom is associated with a lower likelihood of ever graduating high school, while having more middle school peers with elevated BLLs is associated with a reduction in plans to attend a four-year college and a decrease in the likelihood of taking the SAT.

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 are able to assess the relative costs of direct effects and peer effects of lead exposure. 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.[1] We show that exposure to lead-poisoned peers in elementary school can have long term consequences, including dropping out of high school, even if those children who were not

[1] Carrell, Hoekstra, and Kuka (2018) also show that peers exposed to domestic violence impacts wages and educational attainment.

themselves exposed to lead. Estimating these broad negative effects, as well as understanding how the organization of schools can mitigate or exacerbate them, is crucial to design policies that effectively 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.[2] Our findings of long-term effects of exposure to lead-poisoned children in elementary school may explain why longer stints in high-poverty and high-pollution neighborhoods have persistent effects (Chetty, Hendren, and Katz 2016; Chetty and Hendren 2018). Low-income children are more likely to live near sources of toxic waste (Currie, Greenstone, and Moretti 2011; Persico, Figlio, and Roth 2020), and neighborhood characteristics have been shown to contribute significantly to health disparities, such as disparities in asthma rates (Alexander and Currie 2017). Our paper presents another channel through which inequalities in prevalence of toxins and pollutants at the neighborhood levels contribute to the persistence of inequality in the US.

II. Background

Lead Exposure

Ingestion or inhalation of lead dust or lead-contaminated water causes lead poisoning, which, if severe, may induce widespread brain damage that causes encephalopathy, coma, and sometimes death (Meyer, McGeehin, and Falk 2003). Small children are especially exposed to lead-contaminated dust from paint, and lead-contaminated soil due to normal hand-to-mouth activity (Fee 1990). Moreover, lead is most damaging to small children: they absorb and retain more lead than adults and their neurological development is particularly susceptible to

[2] See Graham (2018) for a review.

neurotoxins (see, e.g., McCabe 1979; Meyer, McGeehin, and Falk, 2003). Even low levels of lead exposure have been linked to negative educational and behavioral outcomes (Aizer et al. 2018; Blackowicz et al. 2016; Feigenbaum and Muller 2016; Ferrie, Rolf, and Troesken 2012; Grönqvist, Nilsson, and Robling forthcoming), with early life exposure to pollution having worse effects, including cognitive disabilities (Persico, Figlio, and Roth 2020; Gazze 2016). Reyes (2014) estimates that lead costs $200 billion for a single birth-year cohort.

Peer Effects in the Classroom

Peer effects may work through different channels, both positively and negatively. Children teaching each other is an example of a positive peer effect, while factors such as disruptive classroom behavior can negatively affect the learning of all children in a classroom (Carrell and Hoekstra 2010; Figlio 2007; Hoxby 2000a). For example, using the random assignment of roommates in college Sacerdote (2001) finds that college roommates can influence college grade point averages positively or negatively. Studies suggest a variety of mechanisms linking peer composition and academic outcomes, including differential curricular offerings and instructional practices in classrooms with higher average ability (Lee, Bryk, and Smith 1990; Jackson 2013); social dynamics operating in a student's reference group (Coleman 1961; Hoxby 2000a); and low performing students being unable to keep up with their higher-achieving peers (Imberman, Kugler and Sacerdote 2009; Owens 2010). [3] Peers might also draw disproportionately on a teacher's time and influence the classroom culture and standards.

One strand of the literature examines how low-performing and disruptive children affect peers. Carrell and Hoekstra (2010) find that having an additional peer who was exposed to domestic violence (and is more likely to be disruptive) leads to a decrease in peer test scores and

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

an increase in suspensions from school. Hoxby (2000a) finds that having more low achieving peers leads to lower achievement. Fletcher (2010) finds that the addition of children with learning disabilities in a classroom affects peer test scores negatively. Figlio (2007) also finds that boys with more feminine sounding names are more likely to be disruptive and negatively affect peers' achievement and behavior.

However, much less is known about the long-term consequences of having disruptive peers, and 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) also find that a higher percentage of high school classmates with college-educated mothers decreases the likelihood of dropping out of school and increases college attendance. Bobonis and Finan (2009) find that the PROGRESA program in Mexico had positive spillover effects on non-eligible peers to attend college. Black, Devereux, and Salvanes (2013) find that a higher proportion of girls in ninth grade reduces mean educational attainment, but lowers teen birth rates. It also reduces the likelihood of selecting an academic track for college.

In this paper, we show that disruptive children who were exposed to lead are associated with long-run negative outcomes for their peers, and that the types of outcomes observed vary with the timing of exposure to lead-exposed peers. Thus, 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

We use rich, population-level data on every child attending public school in North Carolina from 1997 to 2017 that link individual-level education outcomes to blood lead test records. To our knowledge, this is the first state-level data set linking individual BLLs to detailed schooling records that allow the matching of siblings and students to classrooms. These unique data include home address identifiers that allow us to match siblings and to follow them longitudinally as they attend school. The data also include detailed information on student demographic characteristics, such as race and socioeconomic status, the schools and classrooms children attended, test scores, suspensions, high school dropout and completion information, college intentions, as well as teacher characteristics. We link siblings based on home address location. 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 also test for whether error in sibling matches affect the results in Section VD. We provide more information on the linkage performed by the North Carolina Education Research Data Center (NCERDC), our sibling identification algorithm, and other details of our variable construction and estimation in the Data Appendix.

For our outcomes, we use the average of standardized mathematics and reading test scores and construct an indicator for receiving an in-school or out-of-school suspension, as well the number of days the child was suspended during the year. We also construct an indicator for whether the child drops out that year based on codes in the North Carolina education data. For our long-term outcomes, we use indicators for graduation and ever dropping out, as well as college intentions and whether the student took the SAT.

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, type of blood draw (capillary or venous), test result in µg/dL, and the child's identifier and address. In our main analysis, 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 take-up to be disproportionally 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 sample. We then compute the share of child's peers with EBLLs using all children in the cohort or classroom at 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 blood lead tests' coverage increases over time, the incidence of lead poisoning decreases.

Sample Description

Table 1 presents summary statistics for our original dataset (3.2 million children, Column 1) and our analysis sample (1.6 million children, Column 2). The Data Appendix details our sample selection criteria. 36.4 percent of children in our analysis sample have a blood lead test,

[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.

and 9.73 percent have at least one test greater or equal than 5 µg/dL (Column 3). Children with EBLLs are more likely to be black, be economically disadvantaged (ED), as measured by an indicator for receiving free or reduced-price lunch, and have teachers without Master's degrees. The average cohort in our sample includes 195 children. Children who spend at least one elementary school year in a cohort with above median share of lead-exposed children (or >10.9 percent of cohort peers) 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 sibling fixed effects, thus assuaging concerns of omitted variable bias due to these group differences.

IV. Identification Strategy

Rigorously estimating peer effects is methodologically difficult and has proven challenging due to limitations of existing data sets. First, peers influence each other simultaneously, so it is unclear whether a disruptive child is causing his fellow classmates to misbehave, or whether the classmates cause him 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 come from similar neighborhoods and 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 achievement is a causal effect of the peers or the reason the individuals joined the peer group/school (Carrell and Hoekstra 2010; Hoxby 2000a). Third, if children are performing poorly, this could be caused by unobserved factors that affect them and their peers simultaneously.

We solve the reflection problem by finding a suitable proxy for peer ability: lead exposure. Consistent with the literature on lead exposure and academic outcomes, Figure 2 shows that being exposed to lead is strongly associated with worse academic achievement and more suspensions, and an increase in the likelihood of dropping out of high school in our sample.[5] Previous research has proxied for peer ability and behavior using preexisting measures such as peers' race and gender (Hoxby and Weingarth 2006; Hoxby 2000b; Lavy and Schlosser 2007), boys with feminine-sounding names (Figlio 2007), peers' retention status (Lavy, Paserman, and Schlosser 2012), peers' disability (Fletcher 2010), or peers who have experienced 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 so long as a student's peers do not cause their own elevated blood lead levels.

We compare siblings whose grade cohorts randomly happen to have different proportions of children with EBLLs. This methodology avoids the reflection problem because a child cannot affect the BLLs of her peers, but EBLLs affect children negatively. Furthermore, including siblings 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-exposed children and socioeconomic status. Thus, idiosyncratic variation in the BLLs of siblings' cohorts offers a plausibly exogenous form of variation to examine the spillover effects of lead and effects of peer quality more broadly.

[5] Our companion paper (Gazze, Persico, and Spirovska 2020) replicates and expands those findings by estimating the effects of early childhood lead exposure on short- and long-run outcomes in our sample using a sibling fixed effects model. That paper estimates that having a BLL greater or equal than 5ug/dL is associated with lower average test scores, and a greater likelihood of suspension, repeating a grade or dropping out of school entirely.

We examine whether lead exposure affects peers' test scores and likelihood of suspensions from school, dropping out of school, graduating from high school and college intentions. Our main family and school fixed effects estimation is given by:

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

where Y_{ijsgt} is some outcome for child *i*, 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 EBLLs (greater or equal than 5 ug/dL) 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 have 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, an indicator for whether a child was tested for lead, and an indicator for whether their BLLs were ≥5 ug/dL. The vector S_{sgt} controls for time-varying school-grade characteristics: the percent Black and Hispanic by school-grade-year and the percent of the cohort who are economically disadvantaged by school-grade-year. θ_j is a family fixed effect. τ_g is a grade fixed effect to account for any grade-specific effects. 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.[6] Throughout our analysis we cluster standard errors at the school level to account for arbitrary correlation in the error terms.

[6] 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 the North Carolina Board of Education approved a new 5-level scale that replaced the long-standing 4-level scale. See the data appendix for more details.

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. As lead exposure is most damaging at ages one through six, a child's lead exposure status pre-dates school entrance. Moreover, controlling for sibling fixed effects accounts for omitted variables such as neighborhood characteristics that could confound the estimated effects of peer quality. Controlling for school fixed effects also accounts 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. Introducing more time-varying school and teacher controls allows us to test whether these channels can explain our results. Third, bias could arise if high-quality students systematically select out of schools in years when there are more students with EBLLs. We test for and discuss these mechanisms in Section VD.

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 3 shows the distribution of the within-school-grade-year interquartile range of this variable. Panel A plots the distribution of this variable constructed as the share over the number of children in the cohort, while Panel B plots the share over the number of children with lead tests in the cohort. Despite the distribution appearing fairly 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. Importantly, a comparison of Panels A and B shows that part of the skewness of the distribution is due to some children not having lead tests being counted as not having EBLLs.

V. Results

A. The Short-Term 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 dropping out of school. To fix ideas, Figure 4 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 and dropping out of school. Table 2 presents our main results for the effects of peers with elevated BLLs, which mirror the patterns in Figure 4. Panel A presents the results for the effect of additional cohort peers who are lead-poisoned on a child's outcomes using only sibling, year, and grade fixed effects, as well as all controls described in our primary specification for student, peer and school characteristics. We find that a ten percent increase in the proportion of cohort-level peers with elevated BLLs leads to a 0.2 percent of a standard deviation lower test scores and a one percentage point increase in the likelihood of suspension, compared with siblings. We also find a 0.06 percentage point increase in the likelihood of dropping out of school in that year.

Since some siblings might not attend the same school and there could be selection into schools, Panel B presents the results of our preferred specification in which we also include school fixed effects. 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 0.535 percent of a standard deviation decrease in test scores and a 1.29 percentage point increase in the likelihood of suspension from school, compared with a sibling who attended the same school. This represents a 17 percent increase in suspensions based on a mean suspension rate of 7.5 percent and amounts to about one tenth of an additional school day suspension. We also find that a 10 percent increase in peers with EBLLs is associated with a 0.063 percentage point increase in the likelihood that a child drops out of school in a given year, which represents a 30 percent increase above the mean. Overall the results that control for both sibling and school fixed effects are very similar and, if anything, slightly larger than the results that only control for sibling fixed effects. This suggests that while omitted school quality characteristics might be correlated with the proportion of students with EBLLs, these unobserved characteristics do not drive our estimated peer effects. Appendix Table A1 shows that our results are also robust to using different measures of lead exposure to define disruptive peers.

Panel C of Table 2 presents the estimates of the effect of having more lead-exposed peers in the same *classroom* using sibling, school, grade and year fixed effects and all controls. 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. For example, if students are switching classrooms, they will have more peers overall.[7] We find that peers in the same classroom have an impact on a child's test score that is nearly four times as large as peers in the same cohort but potentially in different classrooms. By contrast, classroom peers have a smaller effect than cohort peers on a child's likelihood of suspension or dropout. One potential explanation is that while peers outside the classroom but in the same cohort do not directly affect a child's ability to learn, children in the same age group might still spend time together at lunch or recess, thus influencing each other's behavior.

In Panel D, 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 8^{th}-grade test scores: an additional student in a class of 20 decreases average eighth grade test scores by one percent of a standard deviation and increases the likelihood of dropping out in eighth grade by 0.1 percentage point, which represents a 50

[7] Children in grades 3-5 do not typically switch classrooms for different subjects, so they are counted as in the same classroom 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.

percent increase. However, elementary school peers do not have a longer-run impact on behavior, suggesting that it is contemporaneous exposure to misbehaving peers that increases suspensions.

Because peer effects might vary based on the gender of lead-poisoned peers or the age at which children are exposed to disruptive peers, we also examine whether the effects of lead-exposed peers differ in elementary school and middle school and by peer gender. Table 3 presents the effects of contemporaneous exposure to lead-poisoned boys versus girls (Panel A), children in elementary school (Panel B) and middle school (Panel C). We find that exposure to male peers with elevated BLLs has the biggest impact on test scores and suspensions from school, whereas female peers have larger effects on dropping out of school. We also find that exposure to peers with elevated BLLs affects test scores in elementary and middle school similarly. However, the likelihood of being suspended from school or dropping out of school if one has a higher proportion of disruptive peers is much higher in middle school. This makes sense since middle school children are much more likely to drop out or be suspended.

B. Long Term Effects of Peers Exposed to Lead

We next examine whether a child's lead-exposed peers affect that child's long-run outcomes. Table 4 presents estimates of the long-run effects of being exposed to more lead-poisoned peers in elementary school and middle school on average, including all of the individual-level and cohort controls and fixed effects in our primary specification, as well as the average share of non-White peers in elementary school and middle school, and the average share of economically disadvantaged peers in elementary and middle school.

Having one additional lead-exposed peer in a child's elementary school classroom of 20 is associated with a 0.12 percentage point increase in the likelihood that the student will ever drop out of school, which represents a 9 percent increase on the mean dropout rate. We also find that having one additional lead-exposed child in a child's elementary school classroom is associated with a 0.4 percentage point decrease in the likelihood of ever graduating high school. A ten percent increase in middle school classroom peers with elevated BLLs (i.e., about 2 more lead-poisoned peers per classroom) is associated with a 4 percentage point reduction in plans to attend a four-year college and a 4.93 percentage point decrease in the likelihood of taking the SAT. However, we also observe that a ten percent increase in middle school peers with elevated BLLs is associated with a 5.7 percentage point increase in the likelihood of plans to attend a community college, implying that there may be a substitution effect wherein students swap four-year college plans for community college plans.

While the effects on test scores we find are smaller than those obtained by Carrell, Hoekstra, and Kuka (2018), the effects we find on college going are similar in magnitude. Carrell and colleagues find that adding one male peer with domestic violence decreases four-year college going by 1.4 percentage points and decreases test scores by 3 percent of a standard deviation, whereas we find that adding an additional lead-poisoned peer would lead to a 2 percentage point reduction in attending a four year school and a 1 percent of a standard deviation lower test scores. Nevertheless, because we lack precise college enrollment data, we are unable to determine whether the students followed through on their college plans.

C. Heterogeneity of Estimated Effects

Because exposure to lead-poisoned peers could interact with a child's background to shape their outcomes, we next examine the results by several demographic subgroups of interest. For example, different socioeconomic subgroups might have different access to resources, such as academic help outside of school, that might ameliorate the effects of peers with EBLLs. Table 5 presents our preferred estimation for several subgroups: 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 (Panels G and H).

We find evidence of heterogeneous effects of lead-exposed peers on test scores, suspension and drop out for students of different races. White students show the greatest test score losses when exposed to peers with elevated BLLs, while Black students show the largest results for suspensions and dropping out of school. The larger disciplinary impact for Black students may be reflect the fact that Black students are more likely to be suspended in general, so they might be more likely to get in trouble when there are disruptive peers.

We find that students across the socioeconomic spectrum have similar test score losses associated with disruptive peers. However, students who are sometimes listed as economically disadvantaged have larger impacts on the likelihood of being suspended from school and dropping out of school than students who are never or always economically disadvantaged. Appendix Table A2 presents estimates of the effects of lead-poisoned peers for children in schools with different levels of poverty. Specifically, we find stronger peer effects on test scores and suspensions from school in schools with the highest proportion of students who are economically disadvantaged (i.e., the poorest schools). However, the effects on dropping out of school are strongest for children in schools with the lowest proportion of students who are economically disadvantaged (i.e., the wealthiest schools).

In addition, we find that boys and girls are similarly affected by peers with elevated BLLs in terms of test scores. However, boys who have more cohort peers with elevated BLLs have higher rates of suspensions from school and dropping out than their siblings who attended the same school. One potential explanation for these heterogeneous effects could be that lead-poisoned peers of different genders affect their peers differently, as children may spend more time with peers of their same gender. Indeed, Table 3 Panel A shows that male lead-poisoned peers have larger effects in terms both of reducing test scores and increasing suspensions.

D. Additional Threats to Internal Validity

This section discusses potential threats to internal validity and tests that assuage these concerns. If our results are driven by increases in peers' blood lead levels, we would expect students to do worse if they are exposed to a higher percentage of cohort peers with elevated BLLs. We operationalize this by estimating our main model using bins for different percentages of cohort peers with elevated BLLs in bins (0-5% of classroom peers have elevated BLLs, 5%-10% of classroom peers have elevated BLLs, etc.).[8] Figure 5 plots the effect of different percentages of classroom peers on outcomes. As the percentage of peers with elevated BLLs increases, test scores decrease and suspensions from school and dropout rates increase.

However, one might worry that there may be selection in who is tested for lead. Indeed, we do not observe lead exposure for all children in our sample, implying that we measure the share of lead-poisoned children in each cohort with error. If other children who have elevated BLLs are not tested for lead, this would bias our results towards zero since we compute this share over all children in the cohort, irrespectively of whether they have a blood lead test or not. Panel A of Table 6 shows the effects of the share of lead-poisoned children for cohorts in which more than two-thirds of children have lead tests. While the sample size drops dramatically to slightly over 220,000 observations, we obtain a larger effect on test score than in the full sample suggesting attenuation bias due to measurement error could be a concern. However, the

[8] Having 0-5% of classroom peers that have elevated BLLs is the omitted category in the regression.

coefficients on suspensions and dropouts are smaller in this smaller sample, largely because this sample includes more students in early grades.

In addition, one's own lead exposure might be correlated with peers' exposure, and children who are themselves exposed to lead might be driving the results. While we control for a child's own exposure in our main specification, we also examine whether the results differ for children who were not themselves exposed to lead. The results, presented in Panel B of Table 6, suggest that a child's own exposure to lead does not explain our main findings, since the estimates for children who are not exposed to lead are nearly identical to our main estimates in Panel B of Table 2.

Another concern is that our results might be confounded by school characteristics varying over time that are conflated with the percentage of children with EBLLs in each cohort. This could happen if principals shifted resources across cohorts depending on the students' quality. Thus in Panel C of Table 6, we additionally control for school stability, size, and share of teachers with a Master's degree, as well as our main controls (i.e., gender, race, birth month, birth order, economically disadvantaged status, an indicator for whether a child was tested for lead, and an indicator for whether their BLLs were ≥5 ug/dL, the percent Black and Hispanic by school-grade-year, and the percent of the cohort who are economically disadvantaged by school-grade-year). The results are unchanged when controlling for these additional school quality measures.

One might also worry that there could be time trends and selection bias in terms of who is tested for lead over time that could bias our results. In Panel D of Table 6, we control for the share of children who are tested for lead in each school-grade-year cohort and find similar results. Unfortunately, a child's share of peers who are tested for lead exposure is highly correlated with the share of a child's peers who have elevated BLLs (correlation of 0.63) so controlling for the share of a child's peers who are tested for lead exposure slightly shrinks the effects on test scores and behavior.

Our estimates could also be biased if the share of peers with EBLLs in a school-grade-year is systematically correlated with time varying students' or classrooms' characteristics other than those included in Panel C of Table 6. While most common shocks are likely to bias our results towards zero, in order to test for this, we estimate a variety of models that include sibling fixed effects, school fixed effects, and two-way fixed effects. Panel A of Table 7 shows that estimates using school-grade fixed effects are virtually indistinguishable from those in Table 2 that only include sibling, school, grade and year fixed effects. We might also worry that secular trends might bias our results. For example, certain neighborhoods might gentrify, or pollution might worsen, in ways that are both correlated with lead exposure and school outcomes. In Panel B of Table 7, we additionally control for school-grade specific linear time trends to account for these trends, and estimate similar peer effects. Moreover, Figure 1 shows that the incidence of lead poisoning has decreased dramatically since the 1990s. While we control for year fixed effects, we worry that our estimates in Table 2 capture similarly occurring secular trends in test scores, suspensions, and dropout rates. To assuage this concern, in Panel C of Table 6, we additionally control for grade-year fixed effects and find similar results.

One might also be concerned that peer characteristics might be correlated with a child's own characteristics. We investigate this possibility in Table 8 by regressing various child characteristics on the proportion of peers with EBLLs to see if peer EBLLs predict a child's own characteristics. Generally, peer characteristics for those peers with EBLLs have only a very small correlation with a student's own characteristics, and we control for most of the characteristics in

Table 8 in our primary specification. For example, a 10 percent increase in peers with EBLLs is associated with a 0.29 percentage point increase in the likelihood that a child is female, as shown in Column 1 of Table 8. There is no statistically significant association between peer EBLLs and cohort size.

If parents of high-achieving students were to pull their children out of a cohort with particularly disadvantaged students and elevated BLLs, such nonrandom selection could cause poor performance to be misattributed to the presence of more students with elevated BLLs. Column 3 of Table 8 shows that having a 10 percent increase in peers with EBLLs is associated with a 0.1 percentage point decrease in the likelihood of a child being economically disadvantaged. While there is some correlation between a child's characteristics and that of their peers, if children with more lead-exposed peers are less likely to be economically disadvantaged themselves, this implies that the results are not driven by children differentially being in poverty. In addition, Columns 4 and 5 of Table 8 show a 10 percent increase in peers with EBLLs is associated with a 0.6 percentage point decrease in missing both test scores and a slightly higher stability rate, [9] suggesting that children are not differentially switching schools or missing tests if they have more disruptive peers.

Finally, one might worry that blood lead levels are measured with some error. In Appendix Table A1 we show that our results are robust to using different measures of lead exposure to define disruptive peers, although the estimates are noisier when we define EBLLs as BLLs$\geq$10 ug/dL due to the smaller residual variation in the variable after controlling for school, year, and grade fixed effects. Furthermore, we match siblings based on home addresses, which may contain some error, particularly if the building is large and contains multiple housing units.

[9] The stability rate is defined as the total membership in the fall minus the number of children who were in membership more than 90 days later divided by the total number of children.

Thus, Panel A of Appendix Table A3 depicts results in which we limit the sample to census blocks in which more than half of the homes are single family homes. We find that our main results are nearly identical using this sample, suggesting that error in matching siblings does not drive the findings. Because having low income peers is to some extent conflated with having more lead-poisoned peers, we also show that the results are robust to different types of controls for peer poverty status. Panel B of Table A3 also shows the results when controlling for the percentage of peers who are always economically disadvantaged and the percentage of peers who are sometimes economically disadvantaged instead of controlling for the percentage of peers who are economically disadvantaged in that year. These results are quite similar to our main results in Table 2.

VI. Conclusion

This is the first study to date documenting the spillover effects of lead onto peers in schools and classrooms. In addition, we show that peer effects can have long term consequences on human capital formation and reveal some mechanisms through which peer effects may affect long-run human capital outcomes. By comparing siblings attending the same school, we find that a child's own lead exposure spills over to affect other children's test scores, suspensions from school, the decision to drop out of school, and even whether to graduate, take the SAT and attend a four-year college. Each additional child with elevated BLLs in an elementary school classroom of 20 reduces test scores by 1 percent of a standard deviation. We also find that a ten percent increase in peers with elevated BLLs leads to a 9 to 50 percent increase above the mean in dropping out of school. The magnitude of these effects is substantively important, suggesting that the social cost of lead exposure has been drastically underestimated.

We also strong evidence of worse outcomes for children exposed to more lead-poisoned peers despite the fact that their siblings are likely also exposed to other disruptive peers. In addition, if there are additional children with EBLLs who were not tested for lead, this would bias our findings towards zero. Thus, our findings can be interpreted as a lower bound of the effect of having lead-poisoned peers.

While it is difficult to estimate how lead exposure might affect wages based on test scores because of issues with external validity and differences across samples, we attempt a rough back of the envelope calculation for the effect of having one additional lead-exposed peer in a class of 20. We find that being exposed to one additional lead-poisoned peer is associated with $1,878 in lost earnings associated with test scores alone.[10] This estimate does not take into account the additional costs of dropping out of school, not taking the SATs and not getting a four-year college degree. This also does not account for the additional loss of income for the children who were themselves exposed to lead, which has been estimated at $200 billion for a single birth-year cohort (Reyes 2014). Nevertheless, using these figures, we get a cost estimate that is nearly half of the size of the estimate of pollution's effect on long term wages given in Isen et al. (2017), who find that a 10 percent reduction in ambient TSP levels from the Clean Air Act led to a 1 percent increase in mean annual earnings per person (about $4,300 in present value

[10] We take the estimate from Chetty, Friedman, and Rockoff (2014) that a one standard deviation increase in test scores is associated with a 0.12 increase in wages at age 28, conditional on ethnicity, gender, age, lagged suspensions and absences, and indicators for grade repetition, economically disadvantaged, special education, and limited English, cubic polynomials in prior-year math and English scores, interacted with the student's grade level, and classroom level controls, grade and year. If 1 in 20 students is a 5% increase, we multiply that by our estimate on eight grade test scores (-0.2065) = -0.0103. This gives us the implied test score loss of one more child with EBLLs added to a classroom of 20 children over elementary school. If we have a -0.0103 standard deviation decrease in test scores, one more kid with EBLLs would decrease annual earnings at 28 by - 0. 0103 *0.12=0.00124, or 0.124 percent. We then calculate expected earnings at 28 by using the March Current Population Survey data to estimate an age-earnings profile using a non-linear function of age to predict earnings 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). Using these figures, we get a cost estimate very close to the Isen et al. (2017) estimate: we find a $1,878 decrease in lifetime income per person just from peers through test score losses.

terms). This suggests that much of the cost of exposure to pollution might operate through peer effects.

Our results imply several important lessons for policy. First, 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-exposed children improves their outcomes. Thus, offering early interventions for lead might help both lead-exposed children and their peers.

Finally, school segregation by race and socioeconomic status likely exacerbates these peer effects, suggesting that efforts to desegregate students might be beneficial. This work connects with a literature showing that low income schools have some of the largest achievement gaps (e.g., see Reardon 2015). Our results suggest that peer effects and lead exposure are drivers of low performance in high-poverty schools, as well as some of the negative long-run outcomes associated with poverty. Lead exposure and exposure to lead-exposed 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. 2015. "Inequality in Lead Exposure and the Black-White Test Score Gap." Unpublished manuscript.

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.

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.

Blackowicz, M. J., Hryhorczuk, D. O., Rankin, K. M., Lewis, D. A., Haider, D., Lanphear, B. P., & Evens, A. 2016. "The Impact of Low-Level Lead Toxicity on School Performance among Hispanic Subgroups in the Chicago Public Schools." *International journal of environmental research and public health*, 13(8), 774.

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

Graham, B. S. 2018. "Identifying and Estimating Neighborhood Effects." *Journal of Economic Literature*, 56(2): 450–500.

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*.

CDC (Centers for Disease Control and Prevention). 2014. "Lead Screening and Prevalence of Blood Lead Levels in Children Aged 1–2 Years." Child Blood Lead Surveillance System, United States, 2002–2010 and National Health and Nutrition Examination Survey, United States, 1999–2010.

Chetty, R., & Hendren, N. 2018. "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects." *Quarterly Journal of Economics*, 133(3): 1107–1162.

Chetty, R., Friedman, J. N., & Rockoff, J. E. 2014. "Measuring the Impacts of Teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review*. 104(9): 2633–2679.

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.

Coleman, J. S. 1961. "The Adolescent Society; The Social Life of the Teenager and its Impact on Education." New York: Free Press of Glencoe.

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.

Fee, E. 1990. “Public Health in Practice: An Early Confrontation With the ’Silent Epidemic’ of Childhood Lead Paint Poisoning.” *Journal of the History of Medicine and Allied Sciences,* 45: 570–606.

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. Forthcoming. “Understanding How Early Lead Exposure Affect Children’s Life Trajectories.” *Journal of Political Economy*.

Hoxby, C. 2000a. “The Effects of Class Size on Student Achievement: New Evidence from Population Variation.” *Quarterly Journal of Economics*, 115(4): 1239–1285.

Hoxby, C. 2000b. “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.

Korfmacher, K. S. 2003. “Long-Term Costs of Lead Poisoning: How Much Can New York Save by Stopping Lead?” Working Paper: Environmental Health Sciences Center, University of Rochester.

Lavy, V., & Schlosser, A. 2011. "Mechanisms and Impacts of Gender Peer Effects at School." *American Economic Journal: Applied Economics*, 3(2): 1–33.

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.

Lee, V. E., Bryk, A. S., & Smith, J. B. 1990. "High School Organization and its Effects on Teachers and Students: An Interpretive Summary of the Research" in *Choice and Control in American Education*, edited by W. Clune and J. Whitte. New York: Falmer Press.

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

McCabe, E. 1979. "Age and Sensitivity to Lead Toxicity: A Review." *Environmental Health Perspectives*, 29: 29–33.

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.

National Toxicology Program. 2012. "Health Effects of Low-Level Lead." NTP Monograph, U.S. Department of Health and Human Services, 1, xiii, xv–148.

Owens, A. 2010. "Neighborhoods and Schools as Competing and Reinforcing Contexts for Educational Attainment." *Sociology of Education*, 83: 287–311.

Persico, C., Figlio, D., and Roth, J. Forthcoming. Forthcoming. "The Developmental Consequences of Superfund Sites." *Journal of Labor Economics.*

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 EBLLs	(4) Children without EBLLs	(5) Children in sample with more peers with EBLLs in elementary school at least one year (above median share)	(6) Children in sample with fewer peers with EBLLs in elementary at least one year (below median share)
Cohort size	196	195	174	198	171	226
Percent of teachers with an MA degree	0.355	0.354	0.333	0.357	0.338	0.375
Percent economically disadvantaged	0.402	0.424	0.536	0.409	0.506	0.315
Stability rate	0.960	0.961	0.957	0.961	0.958	0.964
Percent Black	0.274	0.266	0.334	0.256	0.306	0.213
Percent Hispanic	0.095	0.099	0.101	0.099	0.107	0.089
Percent with a BLL test	0.325	0.364	1	0.278	0.464	0.231
N Students	3,193,781	1,591,339	154,901	1,436,438	822,160	769,179

Notes: The table present summary statistics for selected variables in our sample. Observations are at the student-year level. 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 with elevated blood lead levels (EBLLs), and Column 4 shows means for children without elevated blood lead levels. Column 5 shows means for children whose share of elementary school peers with elevated BLLs was above the median share at the grade-year level, while Column 6 shows means for children whose shares was below the median.

Table 2: 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) Dropped out that year
	Panel A: Cohort Peers with Sibling Fixed Effects			
Share of peers with BLLs over 5 ug/dL	-0.0232^{+} (0.0138)	0.1150*** (0.0059)	0.6315*** (0.0765)	0.0059*** (0.0008)
	Panel B: Cohort Peers with Sibling and School Fixed Effects			
Share of peers with BLLs over 5 ug/dL	-0.0535*** (0.0136)	0.1285*** (0.0059)	0.7683*** (0.0693)	0.0063*** (0.0009)
	Panel C: Peers in the Same Classroom with Sibling and School Fixed Effects			
Share of peers with BLLs over 5 ug/dL	-0.1748*** (0.0172)	0.0256*** (0.0068)	0.1142^{+} (0.0648)	0.0043*** (0.0008)
Observations	7,444,395	8,058,285	8,058,285	8,058,285
N Students	1,555,711	1,587,165	1,587,165	1,587,165
	Panel D: Average Exposure in Classrooms Grades 3-5 on 8th Grade Outcomes			
Share of peers with BLLs over 5 ug/dL	-0.2011** (0.0624)	0.0301 (0.0287)	0.2465 (0.3882)	0.0209^{+} (0.0108)
Observations	160,750	162,073	162,073	162,073
N Students	159,666	160,877	160,877	160,877
Mean of outcome	0.0639	0.0754	0.3386	0.0020

Notes: The table reports the effect of a child's share of peers with EBLLs on the child's school outcomes. Each cell reports results from a separate regression. All regressions include cohort and individual controls, as well as family, birth month, birth order, grade, 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. Panels B and C add school fixed effects. Individual controls include indicators for gender, race, economically disadvantaged status, whether the student has a blood lead level test, and whether the average of the child's blood lead level tests is above 5 ug/dL. Cohort controls include the share of nonwhite peers and the share of peers who are economically disadvantaged at the school-grade-year level. Panels A and B use the share of peers with maximum BLLs over 5 ug/dL *at the school-grade-year level* as the main explanatory variable, while Panel C uses the share of peers with maximum BLLs over 5 ug/dL *at the classroom level.* The outcome mean, number of observations, and number of students are from the regressions in Panel B. Standard errors are clustered at the school level. Panel D restricts the sample to students in 8th grade and uses the average share of peers with BLLs over 5 ug/dL in elementary school as the main explanatory variable. Cohort controls include the share of 8th grade and the average share of elementary school peers who are non-white or economically disadvantaged. Standard errors are in parentheses and clustered at the school level.

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 3: Heterogeneity by Peer Gender and Age of Exposure

Dependent Variable:	(1) Average Test score	(2) Any Suspensions	(3) Days suspended	(4) Dropped Out
	Panel A: Heterogeneity by Gender of Lead-Poisoned Peers			
Share of male peers with BLLs over 5 ug/dL	-0.0816***	0.1553***	1.0859***	0.0048***
	(0.0192)	(0.0088)	(0.1449)	(0.0013)
Share of female peers with BLLs over 5 ug/dL	-0.0221	0.0985***	0.4133***	0.0080***
	(0.0202)	(0.0094)	(0.1195)	(0.0014)
Observations	7,444,395	8,058,285	8,058,285	8,058,285
N Students	1,555,711	1,587,165	1,587,165	1,587,165
Mean of outcome	0.0639	0.0754	0.3386	0.0020
	Panel B: Elementary School Cohort Grades 3-5			
Share of peers with BLLs over 5 ug/dL	-0.0523**	0.0298***	0.1120***	-0.0000
	(0.0177)	(0.0051)	(0.0261)	(0.0001)
Observations	3,489,872	3,554,389	3,554,389	3,554,389
N Students	1,304,129	1,321,392	1,321,392	1,321,392
Mean of outcome	0.0586	0.0348	0.0941	0.0000
	Panel C: Middle School Cohort Grades 6-8			
Share of peers with BLLs over 5 ug/dL	-0.0600***	0.1318***	0.9559***	0.0008*
	(0.0178)	(0.0092)	(0.1089)	(0.0003)
Observations	4,596,470	4,682,144	4,682,144	4,682,144
N Students	1,364,514	1,378,731	1,378,731	1,378,731
Mean of outcome	0.0716	0.0913	0.4165	0.0002

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 reports the effect of a child's share of male and female peers with elevated blood lead levels on the child's school outcomes. Panels B and C restrict the sample to students in elementary or middle school grades. 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, whether the student has a blood lead level test, and whether the average of all blood lead level tests is above 5 ug/dL. Cohort controls include the average share of elementary school peers who are non-white or economically disadvantaged. Standard errors are in parentheses and clustered at the school level.
$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 4: Long-Run Outcomes of Exposure to Peers with Elevated BLLs by Timing of Exposure

	(1)	(2)	(3)	(4)	(5)
Dependent Variable:	Ever graduated	Ever dropped out	Intention to Attend a 4-Year College	Intention to Attend a Community College	Took SAT
Share of peers with BLLs over 5 ug/dL in Elementary School	-0.0944^{+}	0.0237*	0.0368	-0.0363	0.0479
	(0.0495)	(0.0110)	(0.0661)	(0.0767)	(0.0715)
Share of peers with BLLs over 5 ug/dL in Middle School	0.1742	-0.0439	-0.4029^{+}	0.5721*	-0.4929^{+}
	(0.1741)	(0.0267)	(0.2439)	(0.2549)	(0.2694)
Mean of outcome	0.9514	0.0135	0.3294	0.2701	0.3571
N Students	20,198	159,594	38,787	38,787	38,638

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. Each column reports results from a separate regression. We restrict the sample to students observed in 8th grade. 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 cohort and individual controls, as well as family, birth month, birth order, and year fixed effects, and fixed effects for the elementary school the student attended. Individual controls include indicators for gender, race, economically disadvantaged status, whether the student has a blood lead level test, and whether the average of the child's blood lead level tests is above 5 ug/dL. Contemporaneous cohort controls include the share of peers with BLLs over 5 ug/dL, the share of nonwhite peers, and the share of peers who are economically disadvantaged at the school-grade-year level. We also control for the average share of elementary and middle school peers that are nonwhite or on subsidized lunch. Standard errors are in parentheses and clustered at the school level.

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 5: Heterogeneity by Demographic Subgroups

Dependent Variable:	(1) Average Test score	(2) Any Suspensions	(3) Days Suspended	(4) Dropped out
	Panel A: White, non-Hispanic students			
Share of peers with	-0.0534**	0.0894***	0.4636***	0.0053***
BLLs over 5 ug/dL	(0.0175)	(0.0050)	(0.0440)	(0.0009)
	Panel B: Black non-Hispanic students			
Share of peers with	-0.0282	0.1622***	1.2027***	0.0081***
BLLs over 5 ug/dL	(0.0182)	(0.0108)	(0.1351)	(0.0015)
p-val. black=white	0.32	0.00	0.00	0.10
	Panel C: Hispanic students			
Share of peers with	0.0283	0.0819***	0.1554	-0.0017
BLLs over 5 ug/dL	(0.0360)	(0.0112)	(0.1111)	(0.0018)
p-val. hisp=white	0.04	0.54	0.01	0.00
	Panel D: Never Economically Disadvantaged students			
Share of peers with	-0.0240	0.0538***	0.2110***	0.0046***
BLLs over 5 ug/dL	(0.0203)	(0.0042)	(0.0392)	(0.0007)
	Panel E: Sometimes Economically Disadvantaged students			
Share of peers with	-0.0277	0.0993***	0.7934***	0.0163***
BLLs over 5 ug/dL	(0.0185)	(0.0074)	(0.1429)	(0.0020)
p-val. some=never	0.89	0.00	0.00	0.00
	Panel F: Always Economically Disadvantaged students			
Share of peers with	-0.0403+	0.0721***	0.3211*	-0.0012
BLLs over 5 ug/dL	(0.0219)	(0.0110)	(0.1282)	(0.0011)
p-val. always=never	0.59	0.12	0.41	0.00
	Panel G: Girls			
Share of peers with	-0.0551***	0.0942***	0.4603***	0.0059***
BLLs over 5 ug/dL	(0.0167)	(0.0054)	(0.0538)	(0.0011)
	Panel H: Boys			
Share of peers with	-0.0574***	0.1427***	1.0068***	0.0101***
BLLs over 5 ug/dL	(0.0164)	(0.0078)	(0.1188)	(0.0014)
p-val. boys=girls	0.92	0.00	0.00	0.02

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. Each cell reports results from a separate 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 whether the student has a blood lead level test, and whether the average of the child's blood lead level tests is above 5 ug/dL. Depending on the specification, additional individual controls include gender, race, and economically disadvantaged status, while cohort controls include the share of nonwhite peers and the share of peers who are economically disadvantaged at the school-grade-year level. Standard errors are clustered at the school level. $^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 6: Results for cohorts with high shares of children with BLLs and controlling for additional school quality variables

Dependent Variable:	(1) Average Test scores	(2) Any Suspensions	(3) Days Suspended	(4) Dropped out
Panel A: >66% Tested				
Share of peers with	-0.1170***	0.0331+	0.2232	-0.0028*
BLLs over 5 ug/dL	(0.0343)	(0.0172)	(0.1723)	(0.0013)
Observations	737,395	750,540	750,540	750,540
N Students	226,678	228,834	228,834	228,834
Mean of outcome	-0.2015	0.1141	0.4497	0.0007
Panel B: Restricted to Students who Do Not Have Elevated Blood Lead Levels				
Share of peers with	-0.0523***	0.1302***	0.7726***	0.0086***
BLLs over 5 ug/dL	(0.0143)	(0.0057)	(0.0721)	(0.0009)
Observations	6,569,962	7,102,657	7,102,657	7,102,657
N Students	1,396,468	1,427,378	1,427,378	1,427,378
Mean of outcome	0.1081	0.0688	0.2962	0.0018
Panel C: Controlling for additional School-level Variables (school stability, size, and share of teachers with an MA)				
Share of peers with	-0.0545***	0.1278***	0.7506***	0.0067***
BLLs over 5 ug/dL	(0.0137)	(0.0059)	(0.0677)	(0.0008)
Observations	7,298,127	7,905,565	7,905,565	7,905,565
N Students	1,546,545	1,579,581	1,579,581	1,579,581
Mean of outcome	0.0647	0.0758	0.3408	0.0020
Panel D: Controlling for the Share of Students Tested for Lead				
Share of peers with	-0.0230	0.0878***	0.5688***	0.0077***
BLLs over 5 ug/dL	(0.0167)	(0.0070)	(0.0807)	(0.0010)
Observations	7444395	8058285	8058285	8058285
N Students	1555711	1587165	1587165	1587165
Mean of outcome	0.0639	0.0754	0.3386	0.0020

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 cohorts with at least 66% of the students tested (at the school-grade-year level). Panel B restricts the sample to students whose maximum blood lead level reading is below 5 ug/dL. Panel C replicates Panel B of Table 2 using additional controls for school size, stability rate, and the share of teachers with a master's degree. 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, whether the student has a blood lead level test, and whether the average of the child's blood lead level tests is above 5 ug/dL. Cohort controls include the share of nonwhite peers and the share of peers who are economically disadvantaged at the school-grade-year level. Standard errors are in parentheses and clustered at the school level. The results in Panel D additionally control for the share of students by school cohort who are tested for lead.

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table 7: Results with Different Fixed effects and Time Trends

	(1)	(2)	(3)	(4)
Dependent Variable:	Average Test scores	Any suspension	Days Suspended	Drop out
Panel A: School-Grade Fixed Effects				
Share of peers with BLLs over 5 ug/dL	-0.0555***	0.1293***	0.7463***	0.0063***
	(0.0138)	(0.0060)	(0.0668)	(0.0008)
Observations	7,444,300	8,058,207	8,058,207	8,058,207
N Students	1,555,704	1,587,163	1,587,163	1,587,163
Panel B: School-Grade-Specific Linear Time Trends				
Share of peers with BLLs over 5 ug/dL	-0.0728***	0.1096***	0.7447***	0.0086***
	(0.0131)	(0.0069)	(0.0760)	(0.0009)
Observations	7,444,300	8,058,207	8,058,207	8,058,207
N Students	1,555,704	1,587,163	1,587,163	1,587,163
Panel C: Grade-Year Fixed Effects				
Share of peers with BLLs over 5 ug/dL	-0.0306*	0.0402***	0.4516***	0.0033***
	(0.0149)	(0.0066)	(0.0783)	(0.0008)
Observations	7,444,380	8,058,285	8,058,285	8,058,285
N Students	1,555,711	1,587,165	1,587,165	1,587,165
Mean of outcome	0.0639	0.0754	0.3386	0.0020

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, whether the average of the child's blood lead level tests is above 5 ug/dL, share of nonwhite peers, and the share of peers who are economically disadvantaged at the school-grade-year level. All regressions include family, birth month, and birth order fixed effects. Panel A includes school-grade fixed effects. Panel B includes school-grade-specific linear time trends. Panel C includes school and grade-year fixed effects. Standard errors are in parentheses and clustered at the school level.
$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

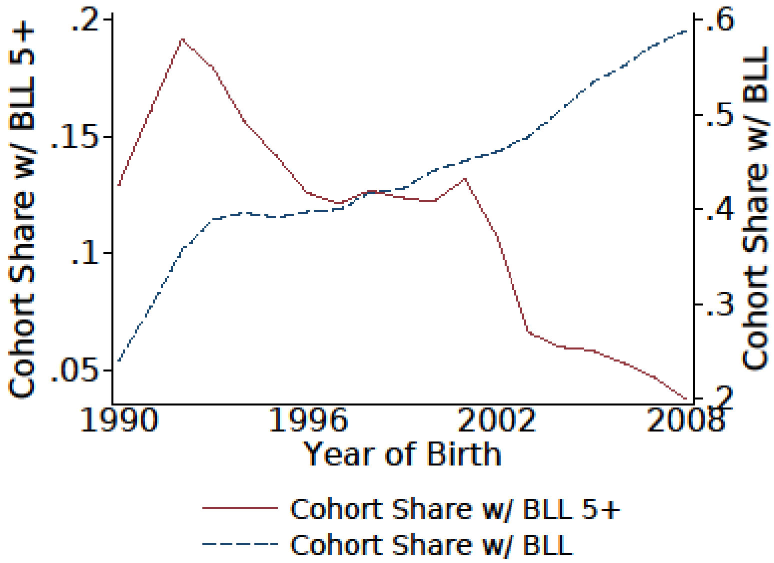
Table 8: Correlation Between Share of Peers with Blood Lead Levels at or above 5ug/dL and a Child's Own Characteristics

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
Share of peers with BLLs over 5 ug/dL	0.0292***	-3.4983	-0.0133**	-0.0641***	0.0044**
	(0.0063)	(4.9362)	(0.0046)	(0.0125)	(0.0016)
Observations	1,587,165	1,587,165	1,191,117	1,587,165	1,587,165
Mean of Outcome	0.4911	194.9144	0.4734	0.0733	0.9606

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, whether the average of the child's blood lead level tests is above 5 ug/dL, share of nonwhite peers, 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.

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

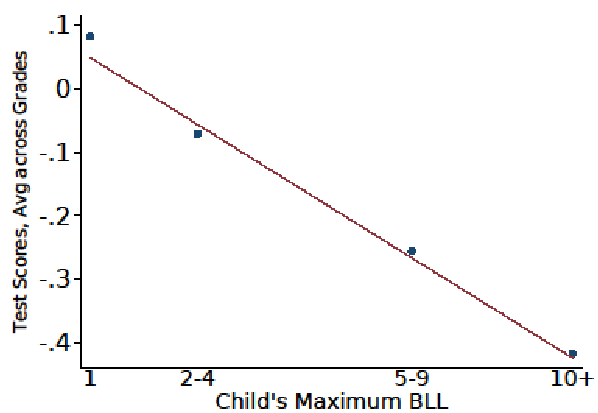
Figure 1: Share of Children with Blood Lead Levels at or above 5ug/dL by Birth Cohort and Share of Children with Blood Lead Tests by Cohort



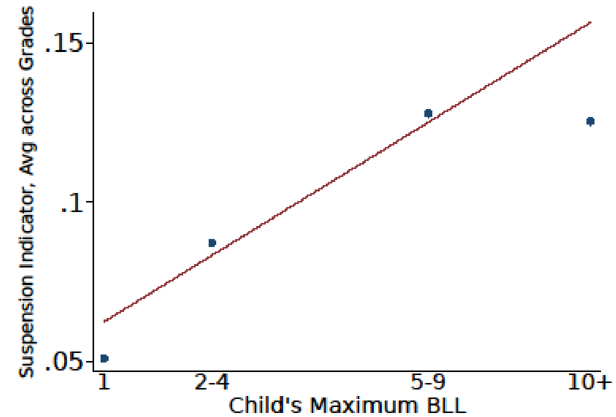
Notes: The figure plots the share of children in a school-grade-year cohort with at least one blood lead test (blue dashed line) and with a blood lead level of at least 5ug/dL (red solid line)

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

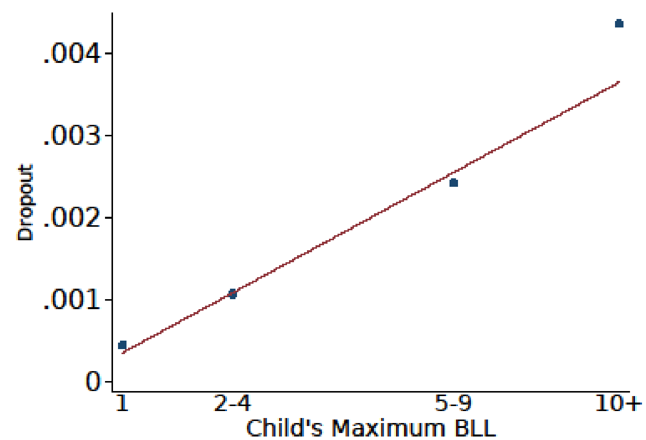
Panel A: Test scores



Panel B: Suspension Likelihood



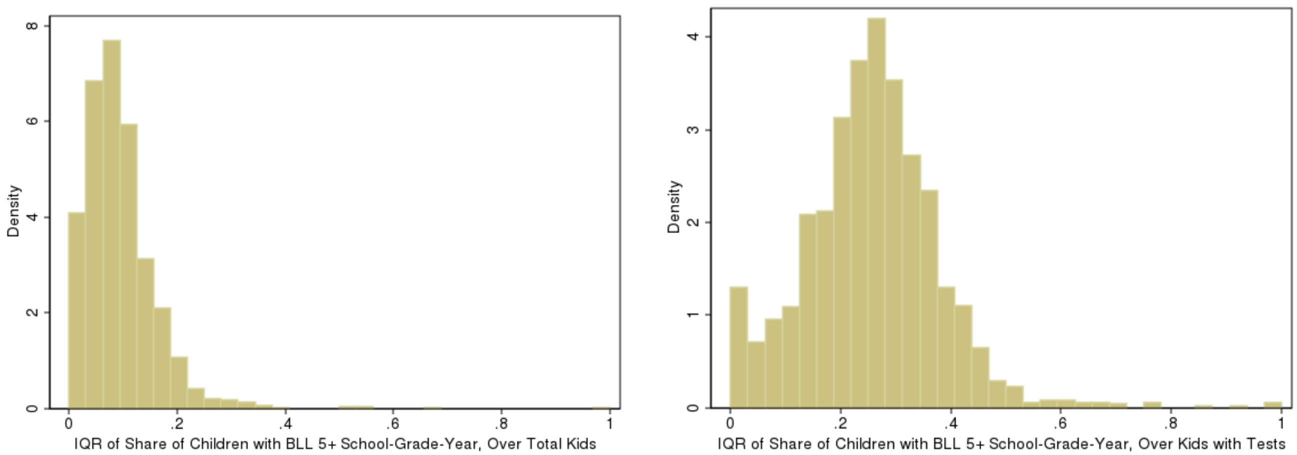
Panel C: Dropout



Notes: The figure plots average test scores (Panel A), suspension rates (Panel B), and dropout rates (Panel C) by students' blood lead levels and adds the line of best fit.

Figure 3: Identifying Variation: Interquartile Range of Share of Children with Blood Lead Levels at or above 5ug/dL within School-Year

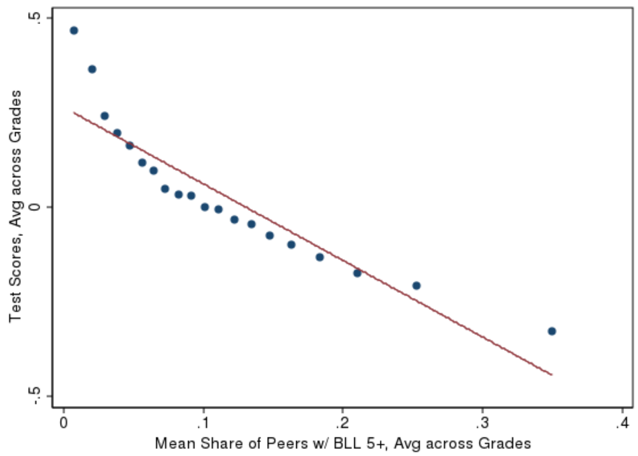
Panel A: All Children at Denominator   Panel B: Only Tested Children at Denominator



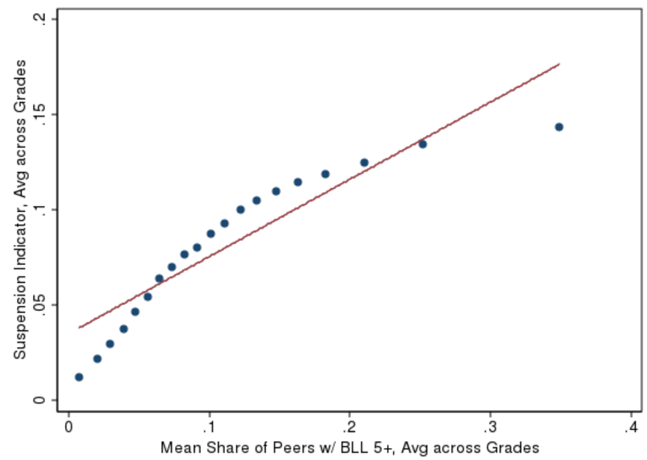
Notes: We compute the share of children in a school-grade-year with blood lead levels at or above 5ug/dL using all children (Panel A) and children with blood lead tests (Panel B). We then compute the interquartile range of this variable at the school-year level. The figure plots the distribution of the school-year level interquartile range of the share of children in a school-grade-year with blood lead levels at or above 5ug/dL.

Figure 4: The Relationship Between Peers' Blood Lead Levels and Test Scores, Suspensions, and Dropping out of School

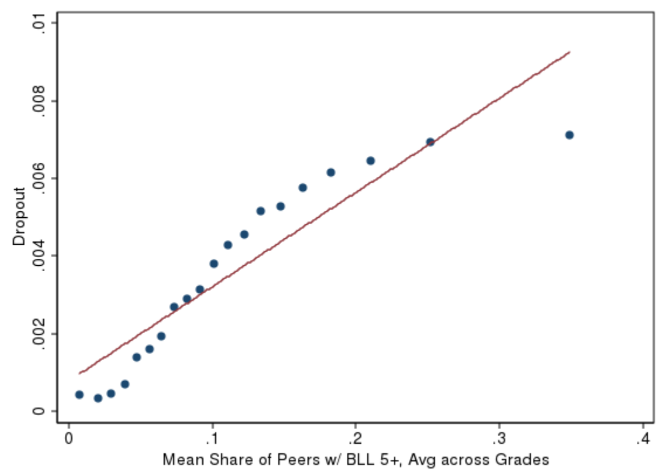
Panel A: Test scores



Panel B: Suspension Likelihood

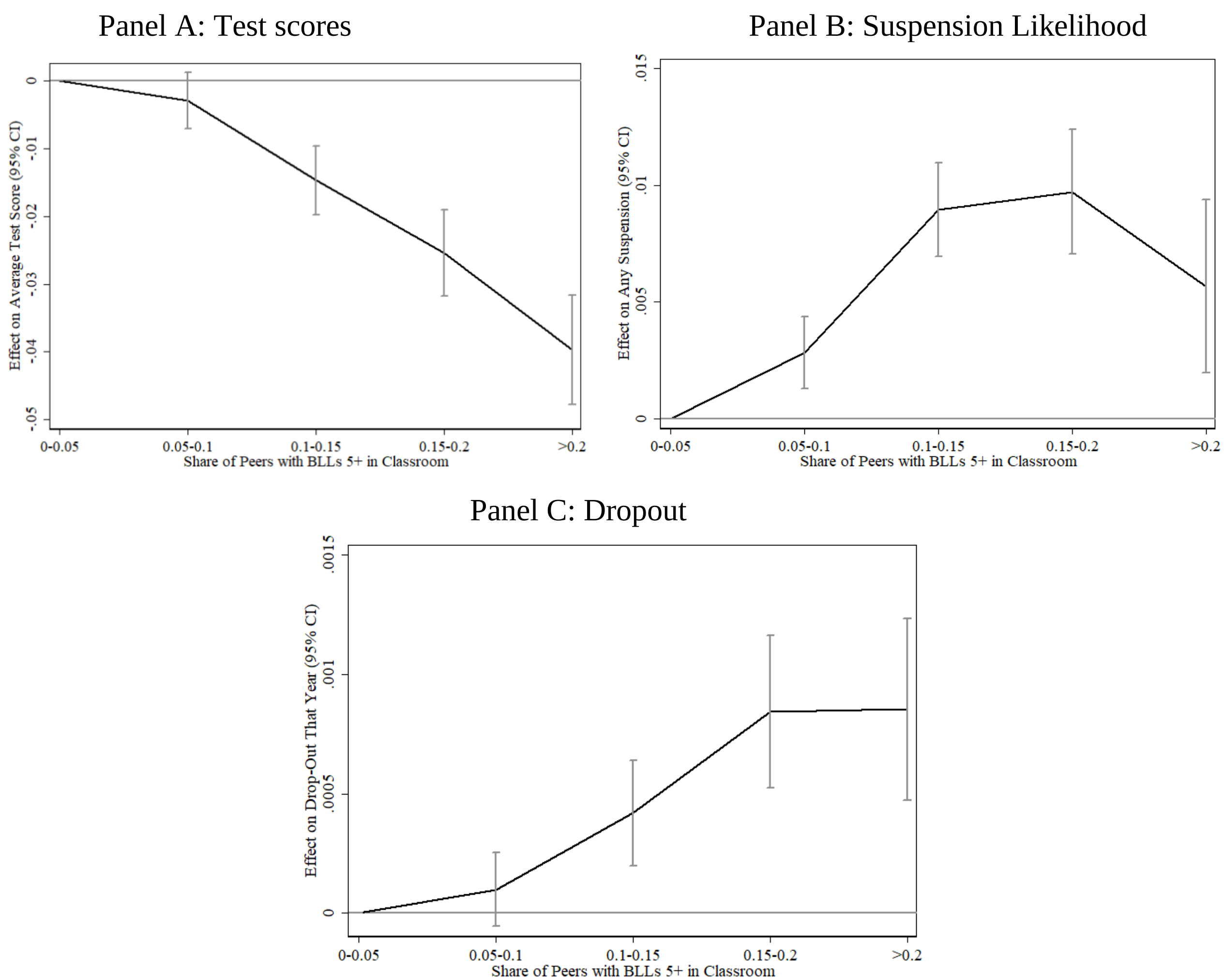


Panel C: Dropout



Notes: The figure plots average test scores (Panel A), suspension rates (Panel B), and dropout rates (Panel C) by quintiles of students' share of peers with blood lead levels at or above 5ug/dL and adds the line of best fit.

Figure 5: Dosage for Peers' Blood Lead Levels and Test Scores, Suspensions, and Dropping out of School



Notes: Each figure plots non-parametric estimates of the effect of having different proportions (binned) of peers with BLLs 5+ in a child's classroom on average test scores (Panel A), suspension rates (Panel B), and dropout rates (Panel C). The omitted category is an indicator for share of peers with BLLs 5+ that is lower than 0.5. We control for all fixed effects and controls in our primary specification (which includes sibling, school, year and grade fixed effects and a variety of individual and demographic controls by cohort). 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: Alternative measures of BLL

	(1) Average Test score	(2) Any Suspensions	(3) Days suspended	(4) Dropped Out
	Panel A: Mean BLL			
Share of peers with BLLs over 5 ug/dL	-0.0569***	0.1261***	0.8906***	0.0119***
	(0.0151)	(0.0069)	(0.0876)	(0.0011)
	Panel B: Max BLL is over 10 ug/dL			
Share of peers with BLLs over 5 ug/dL	-0.0707+	0.0564**	1.1817***	0.0304***
	(0.0372)	(0.0201)	(0.2488)	(0.0028)
Observations	7,444,395	8,058,285	8,058,285	8,058,285
N Students	1,555,711	1,587,165	1,587,165	1,587,165
Mean of outcome	0.0639	0.0754	0.3386	0.0020

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 peers with average BLL above 5 ug/dL. Panel B uses the share of peers with average BLL over 10 ug/dL. Each cell reports results from a separate regression. All regressions include cohort and individual controls, as well as family, birth month, birth order, 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, whether the student has a blood lead level test, and whether the average of the child's blood lead level tests is above 5 ug/dL. Cohort controls include the share of nonwhite peers and the share of peers who are economically disadvantaged at the school-grade-year level. Standard errors are clustered at the school level.

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A2: Heterogeneity by School-Level Demographics

Dependent Variable:	(1) Average Test score	(2) Any Suspensions	(3) Days Suspended	(4) Dropped out
Panel A: Schools in Lowest Tercile of Share Students who are Economically Disadvantaged				
Share of peers with BLLs over 5 ug/dL	-0.0172	0.0378***	0.3587***	0.0341***
	(0.0357)	(0.0095)	(0.0990)	(0.0030)
Panel B: Schools in Middle Tercile of Share Students who are Economically Disadvantaged				
Share of peers with BLLs over 5 ug/dL	-0.0241	0.0929***	0.6056***	0.0005
	(0.0261)	(0.0093)	(0.0996)	(0.0013)
Panel C: Schools in Highest Tercile of Share Students who are Economically Disadvantaged				
Share of peers with BLLs over 5 ug/dL	-0.0471*	0.1085***	0.5169***	-0.0037***
	(0.0191)	(0.0092)	(0.0897)	(0.0009)

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. Each cell reports results from a separate 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, whether the student has a blood lead level test, and whether the average of the child's blood lead level tests is above 5 ug/dL. Cohort controls include the share of nonwhite peers and the share of peers who are economically disadvantaged at the school-grade-year level. Standard errors are clustered at the school level.

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

Table A3: Alternative Specifications and Controls

	(1) Average Test score	(2) Any Suspensions	(3) Days suspended	(4) Dropped Out
Panel A: >50% of Homes in Census Block are Single Family				
Share of peers with BLLs over 5 ug/dL	-0.0582**	0.0969***	0.9882***	0.0080***
	(0.0223)	(0.0096)	(0.1290)	(0.0014)
Panel B: Controlling for Percent of Children Always Economically Disadvantaged and Sometimes Economically Disadvantaged				
Share of peers with BLLs over 10 ug/dL	-0.0292*	0.1014***	0.6440***	0.0061***
	(0.0137)	(0.0059)	(0.0760)	(0.0009)
Observations	7,444,395	8,058,285	8,058,285	8,058,285
N Students	1,555,711	1,587,165	1,587,165	1,587,165
Mean of outcome	0.0639	0.0754	0.3386	0.0020

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 two different specifications. Panel A limits the sample to census blocks in which more than 50 percent of the homes are single family homes. Panel B controls for percent of children who are always listed as ED and the percent of children who are sometimes listed as ED at the school-grade-year level. Each cell reports results from a separate regression. All regressions include cohort and individual controls, as well as family, birth month, birth order, 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, whether the student has a blood lead level test, and whether the average of the child's blood lead level tests is above 5 ug/dL. Cohort controls include the share of nonwhite peers and the share of peers who are economically disadvantaged at the school-grade-year level. In Panel B, we replace the share of peers who are economically disadvantaged with the share of peers who are always listed as ED and the percent of children who are sometimes listed as ED. Standard errors are clustered at the school level.

$^{+}$ $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$

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.[11] 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

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

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 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.