DOI: 10.1111/cdev.13949

EMPIRICAL ARTICLE

CHILD DEVELOPMENT

Environmental inequality and disparities in school readiness: The role of neurotoxic lead

Jared N. Schachner[1] | Geoffrey T. Wodtke[2]

[1]Price School of Public Policy, University of Southern California, Los Angeles, California, USA

[2]University of Chicago, Chicago, Illinois, USA

Correspondence
Jared N. Schachner, University of Southern California—Price School of Public Policy, 635 Downey Way, Los Angeles, CA 90089-3331, USA.
Email: jschachn@usc.edu

Funding information
Division of Social and Economic Sciences, Grant/Award Number: 2015613; Social Sciences and Humanities Research Council of Canada, Grant/Award Number: 435-2018-0736; Mansueto Institute for Urban Innovation at the University of Chicago; Price Center for Social Innovation at USC

Abstract

Developmental science has increasingly scrutinized how environmental hazards influence child outcomes, but few studies examine how contaminants affect disparities in early skill formation. Linking research on environmental inequality and early childhood development, this study assessed whether differences in exposure to neurotoxic lead explain sociodemographic gaps in school readiness. Using panel data tracking a representative sample of 1266 Chicago children (50% female, 16% White, 30% Black, 49% Hispanic, $\mu_{age}=5.2$ months at baseline, collected 1994–2002), analyses quantified the contribution of lead contamination to class and racial disparities in vocabulary skills and attention problems at ages 4 and 5. Results suggested that lead contamination explains 15%–25% and 33%–66% of the disparities in each outcome, respectively, although imprecise estimates preclude drawing firm inferences about attention problems.

Contaminants in the physical environment, such as heavy metals, synthetic chemicals, and other forms of pollution, are widely held to exert a harmful influence on child development (Trentacosta et al., 2016; Trentacosta & Mulligan, 2020). Ecological systems theory and prior research in environmental epidemiology both indicate that exposure to such contaminants within the home, neighborhood, or broader community poses major risks to the health and development of young children (Aizer et al., 2018; Bronfenbrenner & Morris, 2006; Evens et al., 2015; Wodtke et al., 2022). It is also widely known that exposure to many of these hazards is sharply stratified by race and class, with minority groups and children of lower socioeconomic status disproportionately living in contaminated environments (Mohai et al., 2009; Muller et al., 2018). Because health hazards in the physical environment harm child development and are unequally distributed across sociodemographic groups, environmental inequalities likely engender developmental disparities, including class and racial gaps in academic skills, which are among the most debated and decried features of American society.

Despite growing recognition that environmental contaminants may contribute to the etiology of disparities in child development, empirical research "integrating developmental science and the fields of environmental health and toxicology is rare" (Trentacosta et al., 2016: 229). Instead, a focus on the family, early childhood education, and primary schools still predominates in the large, interdisciplinary literature examining class and racial gaps in key developmental outcomes (Duncan & Magnuson, 2005, 2011; Jencks & Phillips, 1998; Reardon, 2016). Although differences in children's families, schools, and early educational environments are undoubtedly important, they do not appear to fully explain disparities

Abbreviations: BLL, blood lead level; CBCL-AP, Child Behavior Checklist attention problems subscale; CDPH, Chicago Department of Public Health; PHDCN, Project on Human Development in Chicago Neighborhoods; PPVT, Peabody Picture Vocabulary Test.

in cognitive and socioemotional skills (Duncan & Magnuson, 2005; Sastry & Pebley, 2010).

If environmental hazards account for these residual disparities, exposure during early childhood is likely pivotal. Skill formation occurs in a cumulative and highly nonlinear manner, with the steepest learning rates and greatest sensitivities to environmental inputs observed during children's earliest years of life (Heckman & Mosso, 2014; Shonkoff & Phillips, 2000). Relatedly, class and racial gaps in cognitive skills emerge between ages 2 and 5, and once formed, they change relatively little thereafter (Duncan & Magnuson, 2011; Heckman, 2006; von Hippel et al., 2018). Thus, the most important environmental determinants of sociodemographic gaps in development are likely to be found early during the course of childhood.

In this study, we asked whether early life exposure to environmental contaminants contributes to class and racial differences in school readiness measured at ages 4 and 5. Among the many different hazards that disproportionately afflict poor and minority children, lead poisoning, in particular, may play an important explanatory role for several reasons. First, exposure to lead, even at very low levels, significantly impairs children's cognitive development, especially when it occurs during early childhood (Aizer et al., 2018; Lanphear et al., 2005). Second, owing to the strong and persistent link between neighborhood segregation and lead contamination, Black, Hispanic, and low-income children are at substantially greater risk of subclinical lead poisoning than are White and more affluent children (Egan et al., 2021; Teye et al., 2021; Wodtke et al., 2022). We therefore hypothesized that differences in exposure to lead-contaminated environments during early childhood can explain, at least in part, class and racial disparities in school readiness.

The developmental effects of subclinical lead toxicity

Lead is a neurotoxic heavy metal that, when ingested, disrupts many different biological processes within the central nervous system, including energy metabolism, neurotransmitter release, protein synthesis, and genetic regulation (Lidsky & Schneider, 2003). Findings from a great variety of research designs show that lead exposure, even at very low levels, is linked with cognitive impairments among children (Aizer et al., 2018; Canfield et al., 2003; Evens et al., 2015; Sorensen et al., 2019). For example, Canfield et al. (2003) found that an increase in blood lead level (BLL) from 1 to 10 micrograms per deciliter (μg/dL) predicts a decline in performance on the Stanford-Binet Intelligence Test of nearly one-half of a standard deviation among children under 5 years old. Lead exposure also harms many different socioemotional and behavioral outcomes among children: a similar change in BLL (1 to 10 μg/dL) is linked with increases in impulsivity, anxiety and depression, antisocial behavior, and the risk of developing an attention deficit disorder (Goodlad et al., 2013; Sampson & Winter, 2016, 2018).

Although lead poisoning can be dangerous at any age, the consequences of subclinical toxicity are greatest among those exposed when they are very young. Infants and toddlers ingest more water and breathe more air per unit of body weight, and they also more frequently engage in hand-to-mouth behavior, which facilitates the consumption of lead in water, soil, and household dust. Once ingested, lead is absorbed more efficiently by the bodies of infants and toddlers, and because their central nervous systems are still developing, exposure to lead can impose more severe and more lasting damage (Lidsky & Schneider, 2003).

The unequal distribution of lead hazards

The risk of lead exposure is much greater for non-White and socioeconomically disadvantaged children. Nationally, BLLs are between 20% and 50% higher for Black compared to White children, and they are independently associated with family income (Teye et al., 2021). In Chicago, the mean BLL for Black children born between 1994 and 1998 was 5.25 μg/dL, compared to 4.16 and 3.72 μg/dL for Hispanic and White children, respectively. There is also a steep gradient in child BLLs across levels of parental education (Evens et al., 2015). Sampson and Winter (2016) found that over half of Chicago children in predominantly Hispanic neighborhoods and two-thirds of children in predominantly Black neighborhoods had elevated BLLs (≥6 μg/dL) during the 1990s, compared to only one-fifth of children in predominantly White neighborhoods. Owing in part to Chicago's lead remediation efforts, BLLs have declined over time. Yet racial differences in exposure remain pronounced: as of the 2010s, roughly three-quarters of the neighborhoods with the worst lead hazards in Chicago remained predominantly Black. Similar patterns have been documented in New York (Lanphear et al., 1998), the Carolinas (Aelion et al., 2013; Miranda et al., 2009), Louisiana (Campanella & Mielke, 2008), and Rhode Island (Aizer et al., 2018).

Racial and class disparities in exposure to neurotoxic lead arise from several different processes. Government and corporate powerbrokers have long sought to exploit cheaper land costs and less onerous community opposition within low-income, racially segregated neighborhoods when selecting sites to construct polluting infrastructure (Bullard, 2000; Elliott & Frickel, 2013). Even though their use was banned decades ago, leaded paint and plumbing are also concentrated in neighborhoods with a greater number of racial minorities and more low-income families because these areas tend to contain a housing stock that is older and suffers from greater disrepair (Cox et al., 2011; Jacobs et al., 2002). Residential selection processes reproduce disparities in lead exposure over time, as White, affluent households are better equipped to avoid areas with known health hazards (Crowder & Downey, 2010). These disparities reflect not only contemporary resource inequality and housing market discrimination but also the historical practices of redlining, neighborhood steering, and de jure segregation, which have stymied wealth accumulation and residential mobility among Blacks and other racial minorities for generations (Pais et al., 2012; Rothstein, 2018).

Sociodemographic gaps in child skills

Because early life exposure to lead impairs cognitive development and differs sharply by race and class, it may partly explain disparities in school readiness. However, the large body of research examining these disparities typically elides this possibility. Instead, most prior studies implicate class and racial differences in school, child care, and family environments.

An emerging consensus suggests that the gap in academic skills between Black versus White students is currently 0.6–0.8 SDs (Hanushek et al., 2020; Reardon, 2011). The gap between Hispanic versus White students is similar (Duncan & Magnuson, 2005). A consensus on the magnitude of class-based gaps in academic skills has proven more elusive, but estimates generally indicate that children in the bottom quintile of socioeconomic indices, which rank families based on parental education, occupation, or income, score between 0.7 and 1.0 SDs lower on reading and math tests than children in the top quintile (Hanushek et al., 2020; Reardon, 2011).

To explain these gaps, research highlighting disparities in child care and education contends that residential segregation consigns a large number of low-income, minority children to inferior daycare centers, preschools, and elementary classrooms. Scholars have proposed several specific mechanisms by which these educational contexts generate disparities in children's skills, including differences in the quality of teachers, instruction, and facilities (Hanushek & Rivkin, 2009; Reardon, 2016). The degree of disciplinary punitiveness and the level of teachers' expectations may also contribute to these disparities, given evidence that the concentration of Black and Hispanic children is positively associated with the former (Owens & McLanahan, 2020) and negatively associated with the latter (McKown & Weinstein, 2008), among a variety of other school inequalities. Moreover, differences in school contexts are intertwined with the effects of environmental hazards. For example, a recent study showed that higher concentrations of lead-poisoned children in a school are linked with a greater number of class disruptions, generating negative spillovers for other students (Gazze et al., 2021).

Another line of research highlighting differences in family environments argues that affluent, well-educated parents' disproportionate investments of time and resources into their children's early cognitive development account for the class-based achievement gap (Guryan et al., 2008; Lareau, 2011; Senechal & LeFevre, 2002). These differences may also account for racial achievement gaps due to the strong associations between income, educational attainment, and race in the United States (Duncan & Magnuson, 2005, 2011; Sastry & Pebley, 2010). Prior empirical studies indicate that unequal family environments are an important driver of gaps in academic skills, with differences in a wide range of household characteristics, including family structure, wealth, and parenting practices, explaining about half of these disparities.

Although differences in schools, early childhood education, and family environments are undoubtedly important contributors to class and racial gaps in academic skills, they do not account for these gaps in their entirety, leaving open the possibility that lead contamination may also play an explanatory role. Yet, this hypothesis has rarely been tested. While many studies have separately examined disparities in lead exposure, disparities in academic skills, or effects of lead exposure on academic skills, as outlined previously, few have analyzed the link between them. For example, Miranda et al. (2009) documented racial differences in BLLs and showed that lead exposure predicts lower fourth-grade reading scores in administrative data from North Carolina, but they stopped short of quantifying the degree to which achievement disparities are explained by these patterns.

There is one noteworthy exception. Sorensen et al. (2019) used county-level data to estimate the association between (i) the total number of children in a county with BLLs ≥10 μg/dL and (ii) the size of racial achievement gaps in the same county. They found that elevated blood lead incidence rates have a modest link with the Hispanic-White achievement gap, and no link with the Black-White gap. However, this study's exclusive reliance on aggregate, county-level data not only invites biases that afflict ecological associations but also obscures the most pronounced disparities in lead exposure, which lie mainly within rather than between counties (Manduca & Sampson, 2021).

The present study

In the present study, we examined whether differences in exposure to lead-contaminated environments during early childhood can explain, at least in part, class and racial disparities in school readiness. To this end, we used longitudinal survey data from a diverse sample of 1266 children in Chicago, who were followed from birth through school entry and assessed for several indicators of school readiness, including their vocabulary skills and difficulties with paying attention. We focused on these indicators because language is the symbolic system through which children learn about the world around them and acquire most other abilities and because the early emergence of attention problems can interfere with future learning in classroom settings. We linked our survey data from children and their families to ecological measures of lead contamination from the Chicago Department of Public Health (CDPH), and we then used machine learning methods to estimate how sociodemographic gaps in school readiness would change if exposures to lead were equalized at low levels for all groups. Because our analysis evaluated hypotheses deduced from prior research linking exposure to neurotoxic lead, residential segregation, and cognitive development, it is more confirmatory than exploratory. Nevertheless, our reliance on observational data and untestable identification assumptions precludes a purely confirmatory interpretation, and the study may be better characterized as laying the groundwork for future confirmatory analyses adopting experimental or quasi-experimental designs to evaluate whether lead abatement reduces skill disparities.

METHOD

Participants

We used data from the Project on Human Development in Chicago Neighborhoods (PHDCN; Earls et al., 2007), a longitudinal study of children living in Chicago. The PHDCN sampled children using a multistage area probability design. In the first stage, neighborhood clusters—defined as groups of 1–3 contiguous census tracts—were selected from within a set of predefined sampling strata designed to ensure representative sociodemographic coverage. In the second stage, block groups were randomly selected from within neighborhood clusters, and then for each block group, all dwellings within it were canvased to assess whether a child with a qualifying birthdate was present. In total, 1666 children in a birth cohort under age 1 were identified as eligible, and their families were recruited for the study with monetary and other in-kind incentives (e.g., $5–20 per interview, free passes to local attractions, and a monthly prize lottery). Ultimately, $n=1266$ of these children participated, yielding a global response rate of about 76%. They composed our analytic sample, which reflects the sociodemographic diversity of children born in Chicago at the turn of the 21st century and included roughly equal proportions of male and female children; over 80% were non-White, with about half identified as Hispanic and 30% identified as Black. Nearly 90% were raised by a caregiver without a bachelor's degree.

We focused on data from Chicago because the city is highly stratified by race and class, and it suffers from extensive lead hazards. The PHDCN's sample design ensures that our findings are generalizable to the population of children born in the city during the mid-1990s. To the extent that Chicago resembles other urban centers in the United States, our findings may be more broadly relevant. We revisit the limits of generalizability in the discussion section.

Procedure

Data on sampled children and their families were collected through three in-person surveys fielded in 1994–1997 (baseline), 1997–1999 (wave 2), and 1999–2002 (wave 3). When these surveys were fielded, members of our analytic sample were under age 1 at baseline, between ages 2 and 3 at wave 2, and between ages 4 and 5 at wave 3, with a few exceptions. We match these survey data with information on areal lead contamination drawn from the CDPH blood lead surveillance system. Since 1993, the Illinois Lead Program has mandated blood-lead screening for all young children in Chicago. Healthcare providers in the city must order screening for those aged 3 or under, and parents are required to provide proof of testing for their child upon enrollment in daycare or kindergarten. The CDPH compiles these test results into a surveillance database containing a tested child's name, date of birth, and home address, as well as information about each test, including the laboratory where it was conducted, the collection date, and the result. Over two million tests have been compiled by the CDPH, with extensive and representative coverage for the population of children born in Chicago (Evens et al., 2015).

Measures

We assessed class and racial differences in two indicators of school readiness—vocabulary skills and attention problems—and in a measure of areal lead contamination.

Vocabulary skills

Vocabulary skills were measured using the Peabody Picture Vocabulary Test, 3rd Edition (PPVT; Dunn, 1997), which was administered to sample members in English at wave 3 of the PHDCN. Scores on the PPVT possess desirable psychometric properties, including high internal consistency (0.92–0.98), split-half reliability (0.86–0.97), test–retest reliability (0.91–0.94), and criterion validity, as they strongly correlate (0.62–0.91) with other tests of verbal abilities (Williams & Wang, 1997). In some studies, these properties appear to hold even among racially, socioeconomically, and linguistically diverse subpopulations, providing little evidence of bias in the PPVT across class or racial lines (Rock & Stenner, 2005; Washington & Craig, 1999; Williams & Wang, 1997). However, several other studies indicate that the possibility of class and racial bias in the PPVT cannot be ruled out (Finneran et al., 2020; Gopaul-McNicol et al., 1998; Helms, 2006)—a point to which we return in the discussion. We standardized PPVT scores to have zero mean and unit variance, where higher values denote a more advanced receptive vocabulary in Standard American English.

Attention problems

Attention problems were also measured at wave 3 of the PHDCN, using the Child Behavior Checklist attention problems subscale (CBCL-AP; Achenbach & Rescorla, 2000). The CBCL-AP is based on parental responses to questions designed to screen for behaviors linked to attention deficit disorders (e.g., the extent to which a child "cannot concentrate," "can't pay attention for long," or "can't sit still"). Caregivers indicate whether statements have been "not true," "somewhat or sometimes true," or "very true or often true" of their child over the past 2 months. These responses were coded zero, one, and two, respectively, and then summed across items to yield a total score. We standardized this score to have zero mean and unit variance, with higher values denoting more severe attention problems. The CBCL-AP also possesses high test–retest reliability (0.78) and criterion validity, as it predicts clinician assessments of psychopathology, other measures of behavior problems, and academic achievement (Achenbach & Rescorla, 2000). The instrument, however, was originally normed on a sample predominated by White, middle-class children, and psychometric evidence of its validity and reliability across diverse subpopulations in the United States is more limited (Konold et al., 2003). Nevertheless, in a study of measurement equivalence by race, family income, and language among a sample of preschool-age children in Chicago, Gross et al. (2006) found that all of the CBCL subscales were "largely equivalent" across groups. In the PHDCN specifically, Cronbach's alpha for the CBCL-AP subtest was .68.

Lead contamination

Our exposure of interest was environmental lead contamination. We measured the degree to which a child's residential context is contaminated by lead based on the CDPH blood-lead surveillance data. These data were used to estimate the local average BLL among children under age 6, separately for each year and neighborhood cluster, as defined by the PHDCN. To improve precision, we smoothed these estimates over time within neighborhoods using kernel regression, which involved computing a weighted "moving average" of BLLs within each neighborhood cluster (Simonoff, 1996). We then matched these smoothed estimates to sample members in the PHDCN at each year in which they were interviewed. The local average BLL provides an ecological proxy for the degree to which a neighborhood was contaminated by lead from all sources. Although not completely free of error, this proxy measure is likely quite accurate because it is typically based on hundreds of individual BLLs per year within each neighborhood cluster, sometimes including multiple tests per child. Averaging results over many children and multiple testing occasions should reduce the influence of measurement errors in any single test.

Covariates

We also incorporated a large number of covariates both to estimate the disparities of interest and to control for confounding of the link between lead contamination and school readiness. These covariates encompass a combination of characteristics measured at baseline and time-varying factors measured at each survey wave. The baseline characteristics included a child's gender and race, family size, the age and education level of a child's primary caregiver, and homeownership status. Gender was coded as a binary variable, 1 for female and 0 otherwise, based on parents' responses to a survey question written in the 1990s that did not distinguish between a child's gender identity and biological sex assigned at birth. Race was expressed as a series of binary variables, also drawn from parent survey responses, that capture whether a child is White, Black, Hispanic, or another race. Family size denoted the number of individuals living in a child's household. The age of a child's primary caregiver was measured in years, and their education level was expressed as a series of binary variables encoding whether they did not complete high school, earned a high school diploma, attended some college, or earned a college degree. Homeownership was captured using an indicator for whether a child's family owned their residence. Although some of these covariates changed over time (e.g., homeownership), we treated them as time-invariant because the PHDCN only measured them at baseline.

The time-varying covariates measured at each wave included household income, the marital and employment status of a child's primary caregiver, family receipt of public assistance, the home language environment, and the socioeconomic composition of a child's neighborhood. Household income is measured in nominal dollars, which we adjusted for inflation using the Consumer Price Index–Research Series and then normalized using the natural log transformation. Marital and employment status were captured with binary variables denoting whether a child's primary caregiver was married or employed at least part time. Public assistance receipt was coded as a binary variable, where 1 denoted the family received cash transfers and 0 otherwise, as was our measure of the home language environment, which denoted whether a child's primary caregiver spoke mainly English at home. The socioeconomic composition of a child's neighborhood was a composite index based on a principal components analysis of the poverty rate, the proportion of adult residents with less than a high school education, the proportion of female-headed households, and the proportion of residents who identify as non-White. Higher values on this index reflect neighborhoods with more disadvantaged populations.

Data analytic plan

First, we performed descriptive analyses, testing the null hypothesis that the covariates, exposure, and outcomes are equal across class and racial subgroups, using robust F and Wald tests. Next, to evaluate whether differences in lead contamination explain disparities in school readiness, we employed a counterfactual approach, which entails estimating how gaps in an outcome of interest between groups would change in a hypothetical scenario where the distribution of another variable is equalized among them (Lundberg, 2022). More specifically, we compared *observed gaps* between sociodemographic groups in vocabulary skills and attention problems to a set of *counterfactual gaps* under a hypothetical intervention to equalize lead contamination at a low level for all children. Differences between the observed versus counterfactual gaps signal that lead contamination is a driver of disparities in school readiness.

Estimands

The observed gap between groups in an outcome can be defined as follows:

$$\mu(x, x') = E(Y_i | X_i = x) - E(Y_i | X_i = x'),$$

where Y_i denotes the observed outcome of child i, X_i denotes membership in a particular group, and $E(\cdot)$ is the expectation operator. With this notation, $\mu(x, x')$ represents the population average difference in an outcome observed between two different social groups, defined by $X_i = x$ versus $X_i = x'$, where x and x' just represent two different values, or categories, of X_i. In this study, it captured, for example, the mean difference in vocabulary skills between Black and White children in Chicago.

We compared the observed gap to a set of counterfactual gaps that involve hypothetical interventions to equalize areal lead exposures at a specific stage of early childhood. These can be expressed as follows:

$$\eta_t(x, x', a_t) = E(Y_i(a_t) | X_i = x) - E(Y_i(a_t) | X_i = x'),$$

where $Y_i(a_t)$ denotes a child's potential outcome had they been exposed at time t to the level of lead contamination given by a_t, possibly contrary to fact. The *observed outcome*, Y_i, is assumed to equal the *potential outcome*, $Y_i(a_t)$, for the single exposure that a child does in fact experience; all other potential outcomes are counterfactuals and thus are not observed. Substantively, $\eta_t(x, x', a_t)$ represents the average difference in an outcome that would persist between two different social categories, defined by $X_i = x$ versus $X_i = x'$, if the exposure of interest at time t, A_{it}, were equalized at the level a_t. In the present study, it captures, for example, the mean difference in vocabulary skills between Black and White children under a hypothetical intervention that equalizes the degree of lead contamination in their residential environments at a particular stage of early childhood—either infancy ($t = 1$), the toddler years ($t = 2$), or the preschool years ($t = 3$).

We also considered counterfactual gaps based on hypothetical interventions cumulatively throughout early childhood rather than at a single point in time. The cumulative gaps can be expressed as:

$$\theta(x, x', \underline{a}) = E(Y_i(\underline{a}) | X_i = x) - E(Y_i(\underline{a}) | X_i = x'),$$

where $\underline{a} = (a_1 a_2 a_3)$ denotes a sequence of exposures to different levels of lead contamination over time, and where $Y_i(\underline{a})$, by extension, denotes a child's potential outcome had they been exposed to a sequence of residential contexts with levels of lead contamination given by $\underline{a}$. Substantively, $\theta(x, x', \underline{a})$ represents the average difference in an outcome that would persist between two social categories if the exposure of interest were equalized at all time points. It captures, for example, the mean difference in vocabulary skills between Black compared to White children under a hypothetical intervention to equalize exposures to lead-contaminated environments from birth through school entry.

The point-in-time estimand, $\eta_t(x, x', a_t)$, and the cumulative estimand, $\theta(x, x', \underline{a})$, provide subtly different but complementary information about the effects of lead contamination on gaps in school readiness. The point-in-time estimand is more immediately relevant to policy, as it corresponds with a class of interventions that would be easier to practically implement. Because families with children are residentially mobile, it is more difficult to design an intervention that would lower lead exposures throughout early childhood rather than lower them at a single point in time. The cumulative estimand, by contrast, is more immediately relevant for evaluating the explanatory role of lead contamination in the etiology of academic gaps, insofar as exposure to lead-contaminated environments is in fact quite stable throughout early childhood in Chicago. Including point-in-time and cumulative estimands together enables our analysis to address a wider range of theoretical and policy considerations.

Both estimands have other important advantages. In particular, they involve manipulations to the exposure of interest, A_{it}, but not to group membership, X_i. They therefore obviate both the conceptual and methodological complications that arise in analyses of race and class disparities, whenever demographic categories are themselves equated with "causal exposures" (Gangl, 2010). They also differ from other quantities that researchers often report in analyses of sociodemographic gaps. A common approach to analyzing these disparities involves estimating the observed gap across groups first unconditionally and then second conditionally on a set of other variables—for example, by adding them to a regression model or, relatedly, by using regression-based decompositions, such as the Kitigawa–Oaxaca–Blinder method (Kitagawa, 1955). Although widely used, this approach does not, in general, properly isolate the effects of manipulating any one explanatory variable on the sociodemographic gaps of interest. By contrast, the counterfactual gaps outlined previously capture exactly how intervening on a single intermediate variable—lead contamination—may impact racial and class disparities in school readiness.

Identification

Unlike observed gaps, counterfactual gaps involve unobserved quantities that can only be identified under conditional independence restrictions, known as "ignorability" assumptions. Specifically, $\eta_t(x, x', a_t)$ can be identified under the assumption that the potential outcomes of exposure at time t, $Y_i(a_t)$, are conditionally independent of a child's observed exposure at time t, A_{it}, given their history of exposures through time $t-1$, covariate history through time t, and group membership, X_i (Lundberg, 2022). The assumption can be written as:

$$Y_i(a_t) \perp A_{it} \mid X_i, \underline{A}_{it-1}, \underline{C}_{it},$$

where $\perp$ denotes statistical independence, $\underline{A}_{it-1}$ denotes a child's exposure history through time $t-1$, and $\underline{C}_{it}$ denotes a child's covariate history through time t. Substantively, it requires that the effect of exposure to a lead-contaminated environment at time t on the outcomes of interest must not be confounded by any unobserved factors. If this assumption holds at every time point, such that $Y_i(\underline{a}) \perp A_{it} \mid X_i, \underline{A}_{it-1}, \underline{C}_{it}\ \forall t$, then the cumulative gap, $\theta(x, x^*, \underline{a})$, can also be identified from the observed data.

These are strong assumptions that the PHDCN is unlikely to fully satisfy. They would be violated, for example, if unobserved factors like parental drug abuse or access to health services affect both lead exposure and child development, net of our measured controls. We attempted to satisfy this assumption approximately by adjusting for the most powerful joint predictors of residential selection and child outcomes. In addition, we performed a sensitivity analysis to assess whether our key inferences were robust to possible biases from unobserved confounding.

Estimation

We estimated counterfactual gaps by combining g-computation—a generic algorithm for imputing potential outcomes—with supervised machine learning (Lundberg, 2022). In our application, the g-computation algorithm proceeded as follows:

1. we fitted a model for the conditional mean of the observed outcome, which can be expressed as $E(Y_i \mid X_i, \underline{C}_{it}, A_{it}) = h(X_i, \underline{C}_{it}, A_{it})$, where $h(\cdot)$ is some function of the predictors and $\underline{A}_{it-1}$ has been subsumed into $\underline{C}_{it}$, here and henceforth, for notational simplicity;
2. we used this model to impute potential outcomes under exposure a_t by setting $A_{it} = a_t$ for all sample members, while leaving the other predictors at their observed values, and then computing $\widehat{h}(X_i, \underline{C}_{it}, a_t)$;
3. we estimated the counterfactual gap by averaging the imputed potential outcomes, separately among sample members in the subgroups for which $X_i = x$ versus $X_i = x'$, and then taking the difference between them—that is, we computed $\widehat{\eta}_t(x, x', a_t) = \frac{1}{n_x} \sum_{i:X_i=x} \widehat{h}(X_i, \underline{C}_{it}, a_t) - \frac{1}{n_{x'}} \sum_{i:X_i=x'} \widehat{h}(X_i, \underline{C}_{it}, a_t)$.

Under the ignorability assumption outlined previously and provided that $h(X_i, \underline{C}_{it}, A_{it})$ is correctly specified, g-computation is consistent for the counterfactual gaps of interest.

Because g-computation hinges on identifying the correct functional form of $h(X_i, \underline{C}_{it}, A_{it})$, attention to model specification is important. Typically, $h(X_i, \underline{C}_{it}, A_{it})$ is assumed to possess a particular parametric form, such as a conventional linear regression, but the true form of this model remains unknown and may be highly complex, especially when the set of predictors, $\{X_i, \underline{C}_{it}, A_{it}\}$, is high-dimensional. For example, the relation between lead exposure and school readiness may be nonlinear or it may differ across levels of the covariates (Muller et al., 2018). Thus, conventional parametric models may yield a poor approximation to the true conditional expectation function, leading to bias. We mitigated this problem by employing methods of supervised machine learning, which construct models for $h(X_i, \underline{C}_{it}, A_{it})$ via data-adaptive algorithms.

We began by adopting a conventional approach to estimation and first fitted a linear model with the following form:

$$h^{\mathrm{lm}}\left(X_i, \underline{C}_{it}, A_{it}\right) = \sum_x \beta_{0x} \mathbf{1}\left(X_i = x\right) + \beta_1^T \underline{C}_{it}, + \beta_2 \ln\left(A_{it}\right),$$

where $\mathbf{1}\left(X_i = x\right)$ is an indicator function equal to 1 when its argument is true and 0 otherwise. This equation models the outcome as a linear function of the covariates through time t, $\underline{C}_{it}$, which includes lagged exposures, and then the natural log of the exposure at time t, $\ln\left(A_{it}\right)$, with separate intercept terms for the subgroups of interest, denoted by X_i. We used the natural log of A_{it} to accommodate a potentially nonlinear relation between lead contamination and cognitive outcomes (Lanphear et al., 2005).

Next, we implemented a set of tree-based machine learning algorithms, including recursive partitioning (Breiman et al., 1984) and random forests (Breiman, 2001). Recursive partitioning creates a so-called "regression tree" by repeatedly dividing the sample into subgroups, or nodes, based on binary partitions of the predictors. The algorithm initiates by considering all possible binary partitions and then selects the single partition that minimizes the sum of the squared deviations around the mean of the outcome in the two resulting nodes. This procedure is then repeated, where the nodes created by each new partition are subsequently split themselves until the algorithm reaches a stopping criterion. The final regression tree gives a set of predicted values equal to the average outcome within each of the terminal nodes. Recursive partitioning is capable of approximating complex forms of nonlinearity and interaction in conditional mean functions, but the method also tends to overfit the data and, relatedly, to produce estimates that are imprecise owing to their high variance.

Random forests are an extension of recursive partitioning that use random subspace selection and bootstrap aggregation to overcome these limitations. The method involves creating an ensemble of many regression trees and combining them to yield improved predictions. Each tree in the ensemble is constructed using (i) a random sample of observations selected with replacement from the observed data and (ii) a random subset of predictors selected as candidates for partitioning at each step of the algorithm. The random forest yields a regression surface equal to the average of the predictions from each of its constituent trees. By averaging over many complex trees, random forests can yield accurate predictions while minimizing the risk of overfitting, thus reducing variance.

Both of these methods require a set of hyperparameters that control the fitting process and determine the stopping criterion for the algorithm. For example, random forests require selecting the minimum number of sample members allowed in a terminal node, the number of predictors to sample for partitioning at each step, and the total number of trees to construct for the ensemble. We constructed an ensemble of 200 trees for the random forests and then tuned the other hyperparameters for both methods using a grid search and $k = 10$-fold cross-validation.

Finally, we combined the results from all of the previous estimation methods using a stacking algorithm known as a "super learner" (van der Laan et al., 2007). This approach generates a weighted average of the results given by our linear model and each of the different tree-based methods such that the combined predictions may provide greater accuracy than those from any one method taken by itself while being guaranteed, at minimum, to perform no worse than the most accurate single method considered by the algorithm.

To estimate counterfactual gaps under a manipulation to cumulative rather than point-in-time lead exposures, it is important to properly account for the dynamic process by which families select into and out of different residential contexts over time. In this setting, time-varying confounders of future residential choices may themselves be outcomes of exposures in the past, and conventional methods of covariate adjustment, such as naively conditioning on these variables in a model for the conditional mean of the outcome, may be biased, even when the ignorability assumptions are all satisfied (Elwert & Winship, 2014).

We therefore used the method of residual balancing to adjust for confounding in our analysis of cumulative gaps (Zhou & Wodtke, 2020). Residual balancing proceeds in two steps. First, the time-varying confounders at each time point are residualized with respect to the observed past, which means that the confounders are set equal to the difference between their observed values and their fitted values from regression models in which they are predicted as a function of all variables measured at an earlier point in time. Second, outcome models similar to those outlined previously are fit using a set of weights constructed so that (i) the residualized confounders are not only orthogonal to the observed past but also to all future exposures in the weighted data and (ii) the variability of the weights, as measured by their relative entropy, is minimized. By fitting models with weights that balance the time-varying confounders across future exposures, conditional on the observed past, this approach appropriately adjusts for dynamic selection.

We estimated class and racial gaps in school readiness under a hypothetical intervention to equalize the local average BLL at 4 μg/dL. We focus on equalizing BLLs at 4 μg/dL because this is a level below recent monitoring guidelines set by the Centers for Disease Control and Prevention where there is common support among all subgroups of interest in the PHDCN. Our focal racial gaps compared White, Black, and Hispanic children only, as we lacked sufficiently large samples of other groups. Our focal class gaps compared children with and without a caregiver who completed college, as educational attainment is widely viewed as the most stable and reliably measured proxy for "social class." Moreover, there is a strong link between post-secondary education and the many other dimensions of "social class," including income and occupational status (Hout, 2012).

Missing values for all variables were simulated using multiple imputation with 50 replications, and estimates were combined across complete datasets following Rubin (1987). Within each dataset, we computed standard errors with the nonparametric bootstrap for complex survey data (Rao & Wu, 1988), which accounted for stratification and clustering in the PHDCN's sample design.

RESULTS

Table 1 presents sample characteristics from the PHDCN, separately by race, together with results of robust F-tests assessing the null hypothesis that covariate means are equal across racial groups. Comparing White,

TABLE 1 Project on Human Development in Chicago Neighborhoods sample characteristics by race.

	White (n=206)		Black (n=383)		Hispanic (n=615)		
Covariates	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	p
Primary caregiver education							
Less than high school	0.17	—	0.37	—	0.59	—	<.001
High school diploma	0.09	—	0.15	—	0.14	—	.121
Some college	0.30	—	0.41	—	0.23	—	<.001
College graduate	0.44	—	0.07	—	0.04	—	<.001
Female	0.45	—	0.48	—	0.50	—	.379
Homeowner	0.54	—	0.12	—	0.13	—	<.001
Family size	4.21	1.45	5.72	2.28	5.37	2.02	<.001
PCG age	31.19	5.48	26.49	8.12	26.50	6.17	<.001
Household income (log)							
Wave 1	10.58	0.72	9.45	1.13	9.47	0.97	<.001
Wave 2	10.71	0.85	9.61	1.09	9.69	0.91	<.001
Wave 3	10.85	0.74	9.87	1.00	9.94	0.83	<.001
PCG employed							
Wave 1	0.60	—	0.42	—	0.37	—	<.001
Wave 2	0.61	—	0.52	—	0.48	—	.040
Wave 3	0.61	—	0.62	—	0.55	—	.123
PCG married							
Wave 1	0.79	—	0.22	—	0.58	—	<.001
Wave 2	0.77	—	0.27	—	0.63	—	<.001
Wave 3	0.78	—	0.28	—	0.65	—	<.001
PCG on public assistance							
Wave 1	0.13	—	0.58	—	0.39	—	<.001
Wave 2	0.13	—	0.55	—	0.18	—	<.001
Wave 3	0.06	—	0.37	—	0.11	—	<.001
English spoken at home							
Wave 1	0.87	—	0.99	—	0.42	—	<.001
Wave 2	0.92	—	1.00	—	0.46	—	<.001
Wave 3	0.99	—	0.99	—	0.45	—	<.001
Neighborhood disadvantage							
Wave 1	−1.00	0.50	0.45	0.60	−0.03	0.54	<.001
Wave 2	−1.05	0.48	0.49	0.62	−0.08	0.59	<.001
Wave 3	−1.06	0.49	0.44	0.61	−0.14	0.59	<.001

Note: Results are combined across 50 multiply imputed datasets. p-Values come from cluster- and heteroscedasticity-robust F-tests of the null hypothesis that the covariate means are equal across racial groups.

Black, and Hispanic children reveals large inequalities, with Black children generally more disadvantaged than Hispanic children, who were in turn more disadvantaged than White children on most measures. For example, Black children in Chicago were more likely than both White and Hispanic children to be part of a family with an unmarried primary caregiver, who received public assistance and lived in a highly disadvantaged neighborhood.

Table 2 presents similar comparisons across levels of parental education, revealing stark class disparities as well. Compared with children whose primary caregiver completed a college degree, children whose primary caregiver did not were more likely to live in lower-income families that did not own a home, had unmarried or unemployed parents, and resided in disadvantaged neighborhoods. Overall, race and class inequalities are chasmic in Chicago.

Observed gaps in school readiness

Table 3 presents estimates of observed class and racial gaps in vocabulary skills and attention problems. Recall that higher scores on the PPVT represent more advanced vocabulary skills, whereas higher scores on the CBCL-AP represent more severe attention problems.

Large gaps in vocabulary skills and attention problems are apparent. Black and Hispanic children scored about 0.8 and 1.4 SDs below White children on our measure of vocabulary skills, respectively. Black and Hispanic children also exhibited more attention problems, scoring more than 0.2 SDs higher than White children. Observed gaps in children's vocabulary skills and attention problems by parental education are similarly pronounced. Children whose primary caregiver does not have a college degree scored about 1.1 SDs below children with a college-educated caregiver on our measure of vocabulary skills and about 0.3 SDs higher on our measure of attention problems.

The explanatory role of lead contamination

Figure 1 contains maps of Chicago depicting the spatial distribution of lead contamination and our focal sociodemographic groups in 1997 when the PHDCN's first wave was concluding. It shows that lead hazards were concentrated on Chicago's South and West sides, which were also predominantly Black and highly segregated. The city's Hispanic population was concentrated in neighborhoods that abut these areas and also suffered from high levels of lead contamination, while the college-educated population was concentrated mainly on the North side of the city, where lead hazards were less severe.

TABLE 2 Project on Human Development in Chicago Neighborhoods sample characteristics by parental education.

	PCG without college degree (n=1110)		PCG with college degree (n=156)		
Covariates	**Mean**	**SD**	**Mean**	**SD**	p
Race					
White	0.10	—	0.58	—	<.001
Black	0.32	—	0.17	—	.007
Hispanic	0.53	—	0.16	—	<.001
Other	0.04	—	0.09	—	.096
Female	0.50	—	0.46	—	.377
Homeowner	0.15	—	0.58	—	<.001
Family size	5.43	2.12	4.05	1.30	<.001
PCG age	26.72	6.95	31.93	4.70	<.001
Household income (log)					
Wave 1	9.50	1.05	10.75	0.54	<.001
Wave 2	9.70	0.99	10.95	0.65	<.001
Wave 3	9.95	0.90	11.11	0.52	<.001
PCG employed					
Wave 1	0.38	—	0.75	—	<.001
Wave 2	0.49	—	0.72	—	<.001
Wave 3	0.57	—	0.65	—	.068
PCG married					
Wave 1	0.46	—	0.84	—	<.001
Wave 2	0.50	—	0.87	—	<.001
Wave 3	0.53	—	0.87	—	<.001
PCG on public assistance					
Wave 1	0.45	—	0.04	—	<.001
Wave 2	0.32	—	0.02	—	<.001
Wave 3	0.21	—	0.01	—	<.001
English spoken at home					
Wave 1	0.65	—	0.84	—	<.001
Wave 2	0.69	—	0.93	—	<.001
Wave 3	0.70	—	0.96	—	<.001
Neighborhood disadvantage					
Wave 1	0.04	0.70	−0.74	0.65	<.001
Wave 2	0.01	0.74	−0.77	0.66	<.001
Wave 3	−0.04	0.73	−0.82	0.68	<.001

Note: Results are combined across 50 multiply imputed datasets. p-Values come from cluster- and heteroscedasticity-robust F-tests of the null hypothesis that the covariate means are equal across education levels.

The strong link between race and the risk of living in a lead-contaminated neighborhood is confirmed by Figure 2, which shows kernel density plots describing racial differences in lead exposure and Wald tests of the null hypothesis that mean exposures are equal across groups. Black children exhibited the highest risk of lead exposure, followed by Hispanic and then White children, who had the lowest risk. This pattern held across every PHDCN wave, although overall levels of, and racial

TABLE 3 Observed gaps in vocabulary skills and attention problems.

	PPVT scores		CBCL-AP scores	
Contrast	**Gap (SE)**	***p***	**Gap (SE)**	***p***
Black versus White	−.806 (.097)	<.001	.279 (.107)	.012
Hispanic versus White	−1.391 (.099)	<.001	.232 (.095)	.017
No college versus college	−1.068 (.102)	<.001	.340 (.087)	<.001

Note: Gaps are reported in SD units. Results are combined across 50 multiply imputed datasets. p-Values come from cluster- and heteroscedasticity-robust F-tests of the null hypothesis that the gap is equal to zero.

Abbreviations: CBCL-AP, Child Behavior Checklist attention problems subscale; PPVT, Peabody Picture Vocabulary Test.

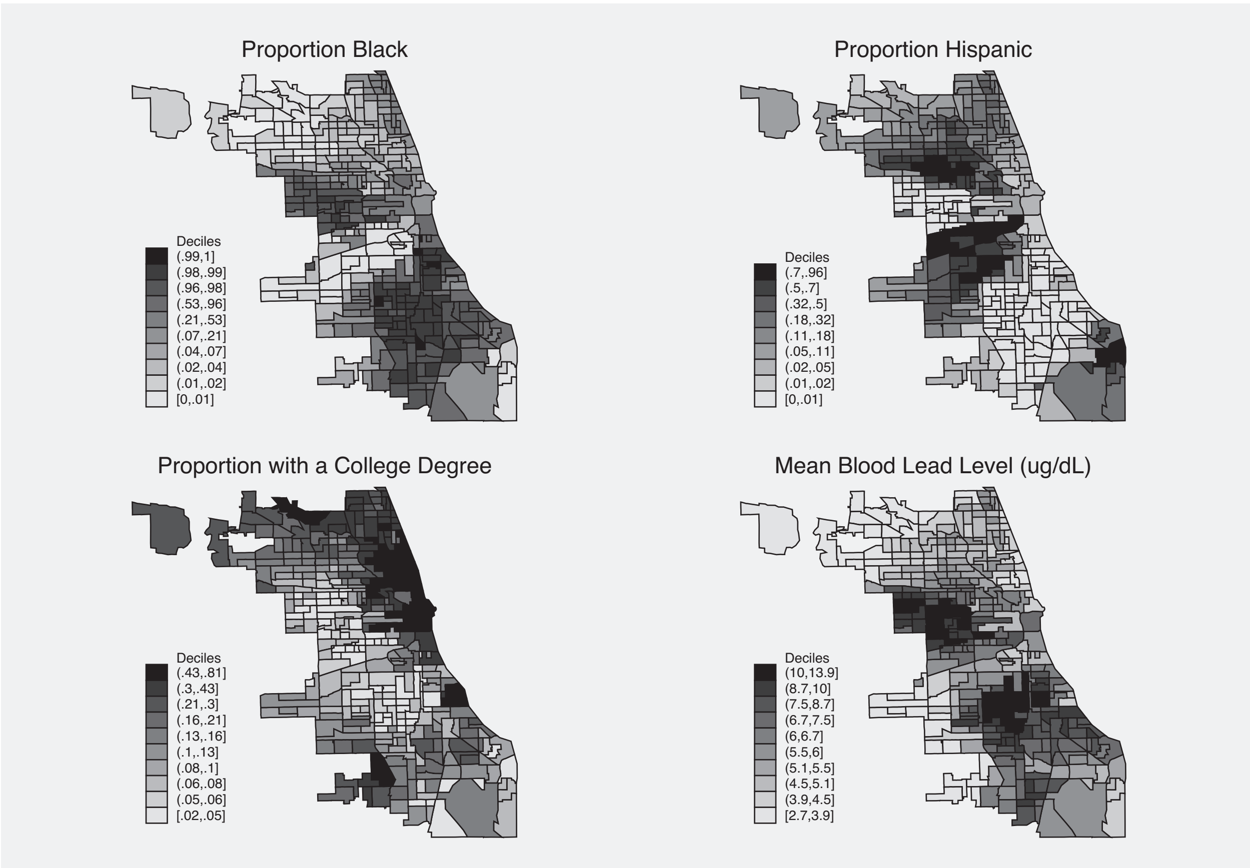

FIGURE 1 Spatial distribution of demographic groups and lead contamination.

differences in, lead contamination declined throughout the study.

Figure 3 summarizes differences in lead exposure by parental education and also shows Wald tests of equal means between these groups. At each stage of early childhood and also cumulatively over time, children without a college-educated caregiver were more likely to live in neighborhoods with high levels of lead contamination than were children whose parents had a college degree. Despite these differences, however, a nontrivial number of children with college-educated parents lived in areas with high levels of lead contamination, as indicated by the long right tail of these distributions and their sizable overlap.

Table 4 presents estimates of differences between counterfactual and observed gaps in vocabulary skills. Results from different estimation procedures are presented across the table's columns, and p-values are from Wald tests of the null hypothesis gauging whether the difference between a counterfactual gap and the corresponding observed gap is equal to zero. We interpret these p-values as a single number summary of evidence against the null hypothesis (Wasserstein & Lazar, 2016).

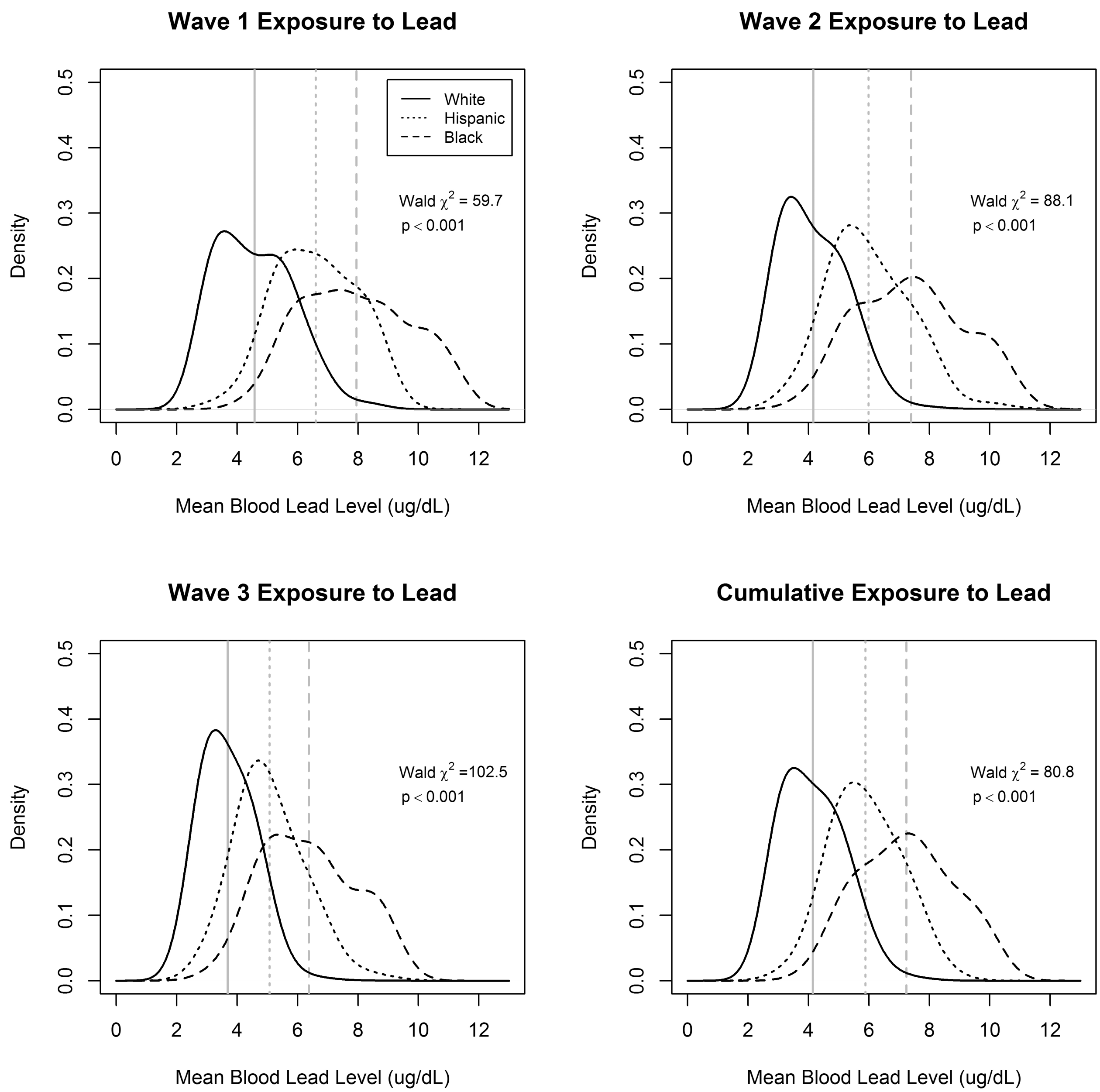

FIGURE 2 Racial differences in exposure to lead-contaminated environments.

In general, these results indicate that lead contamination plays at least some role in explaining both class and racial gaps in vocabulary skills. Specifically, they suggest that equalizing lead exposures at a relatively low level would reduce gaps in vocabulary skills, especially when this type of intervention occurs earlier during the course of development or cumulatively over time. For example, estimates from the super learner indicate that equalizing exposure to lead hazards during infancy would shrink the Black-White gap by about 0.20 SDs, from approximately −0.80 SDs, the observed level of the gap, to −0.60 SDs, the estimated level after equalizing lead exposure. Similarly, the Hispanic-White gap would shrink by about 0.23 SDs, and the gap between children with and without college-educated parents would shrink by 0.15 SDs. Small *p*-values provide considerable evidence against the possibility that these differences are actually equal to zero in the target population. Recall also that large *p*-values do not signal that a null hypothesis is true but rather that we merely lack the evidence to rule out this possibility, which may stem from the high variance of certain estimates, like those from the regression trees.

Figure 4 summarizes the observed and counterfactual gaps separately, as estimated by the super learner. These results also illustrate that the largest reductions to gaps in vocabulary skills occur under hypothetical interventions on lead contamination during infancy or cumulatively throughout early childhood. When exposure to lead-contaminated environments is equalized during infancy,

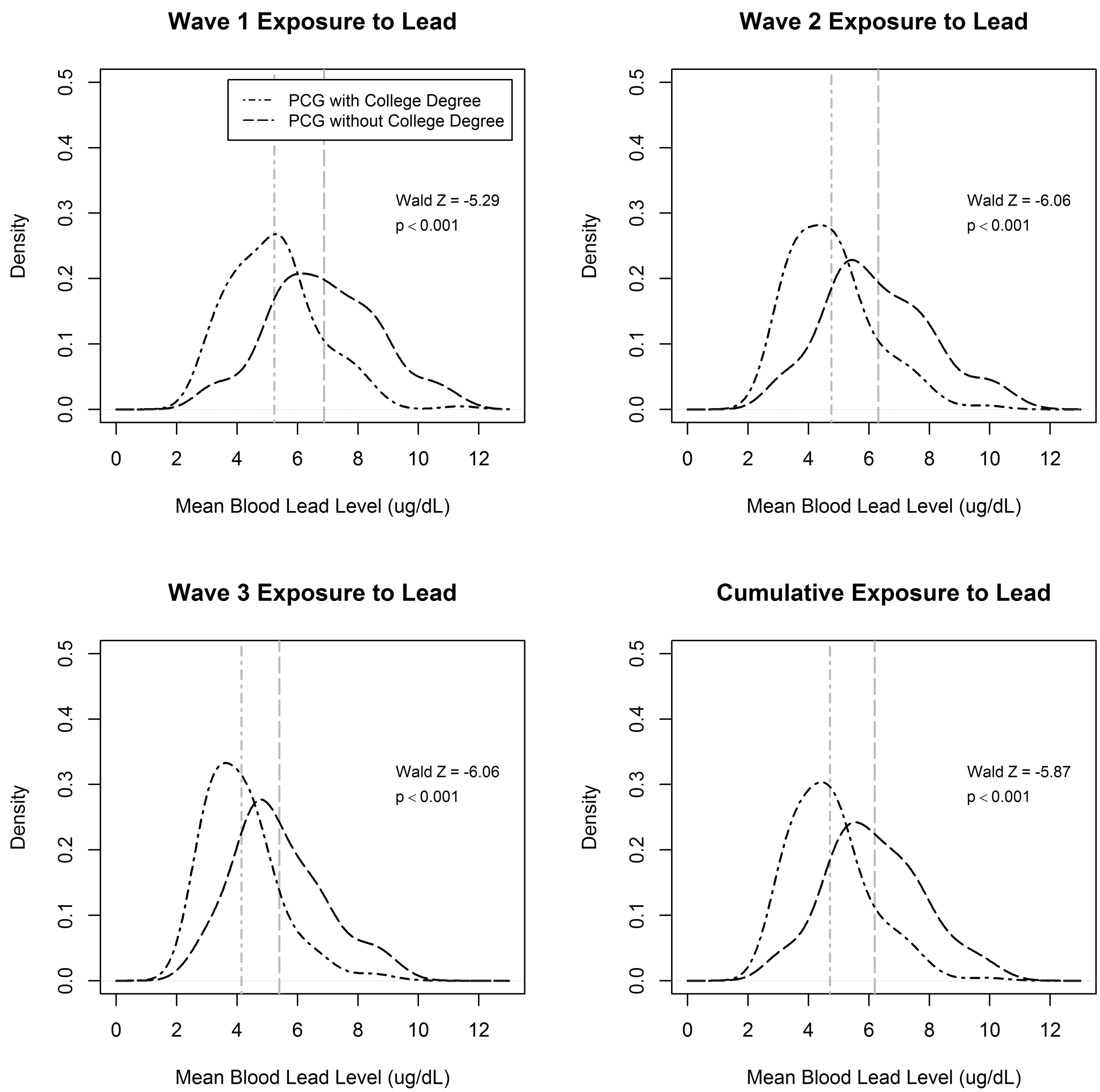

FIGURE 3 Class differences in exposure to lead-contaminated environments.

for example, the Black-White gap in vocabulary skills is estimated to shrink by about 25%, the Hispanic-White gap by about 15%, and the education gap by about 15%.

Table 5 presents estimates of differences between counterfactual and observed gaps in attention problems. Results from the different estimation procedures are somewhat irregular and, in many cases, highly imprecise. Point estimates from the super learner suggest that equalizing lead exposures at a relatively low level during infancy would cause the Black-White gap to decline by 0.18 SDs, from an observed level of 0.28 SDs to an estimated level of 0.10 SDs after equalizing lead hazards. Similarly, equalizing lead hazards during infancy is estimated to reduce the Hispanic-White gap by about 0.12 SDs and the parental education gap by about 0.09 SDs. However, it is important to note that because of their imprecision and instability, these estimates provide relatively little evidence against the null hypothesis that the observed and counterfactual gaps are equal. Note that this imprecision is exacerbated for estimates at waves 2 and 3 by collinearity with the past, as identifying these effects requires controlling for exposure histories that tend to be highly correlated with contemporaneous exposures to lead. For example, the correlations between measures of lead contamination taken at wave 1 and the same measures taken later at waves 2 and 3 are 0.87 and 0.81, respectively, in the PHDCN.

Figure 5 summarizes the observed and counterfactual gaps in attention problems separately, as estimated by the super learner. Equalizing exposures to lead hazards

TABLE 4 Estimated differences between counterfactual and observed gaps in vocabulary skills.

Contrast	Linear model Est. (SE)	*p*	Regression tree Est. (SE)	*p*	Random forest Est. (SE)	*p*	Super learner Est. (SE)	*p*
Eq. wave 1 exposure								
Black versus White	.179 (.085)	.035	.159 (.170)	.350	.235 (.076)	.002	.204 (.070)	.004
Hispanic versus White	.122 (.058)	.035	.275 (.396)	.488	.337 (.108)	.002	.225 (.095)	.018
No college versus college	.086 (.043)	.046	.246 (.228)	.281	.218 (.067)	.001	.150 (.060)	.012
Eq. wave 2 exposure								
Black versus White	.117 (.130)	.368	.242 (.202)	.231	.182 (.062)	.003	.142 (.074)	.055
Hispanic versus White	.075 (.082)	.360	.210 (.130)	.106	.176 (.052)	.001	.113 (.056)	.044
No college versus college	.056 (.063)	.374	.245 (.142)	.085	.158 (.045)	<.001	.096 (.046)	.037
Eq. wave 3 exposure								
Black versus White	−.104 (.143)	.467	.141 (.141)	.317	.085 (.058)	.143	−.025 (.080)	.755
Hispanic versus White	−.061 (.084)	.468	.144 (.072)	.046	.081 (.033)	.014	−.001 (.047)	.983
No college versus college	−.050 (.069)	.468	.174 (.110)	.114	.095 (.036)	.008	.011 (.044)	.803
Eq. cumulative exposure								
Black versus White	.166 (.149)	.265	.426 (.349)	.222	.377 (.147)	.010	.243 (.135)	.072
Hispanic versus White	.110 (.109)	.313	.251 (.394)	.524	.399 (.164)	.015	.209 (.138)	.130
No college versus college	.106 (.096)	.270	.362 (.276)	.190	.326 (.117)	.005	.187 (.102)	.067

Note: Estimates are reported in SD units. Standard errors are estimated from 200 bootstrap samples. Results are combined across 50 multiply imputed datasets. *p*-Values come from two-sided Wald tests of the null hypothesis that the difference between the counterfactual and observed gap is equal to zero.

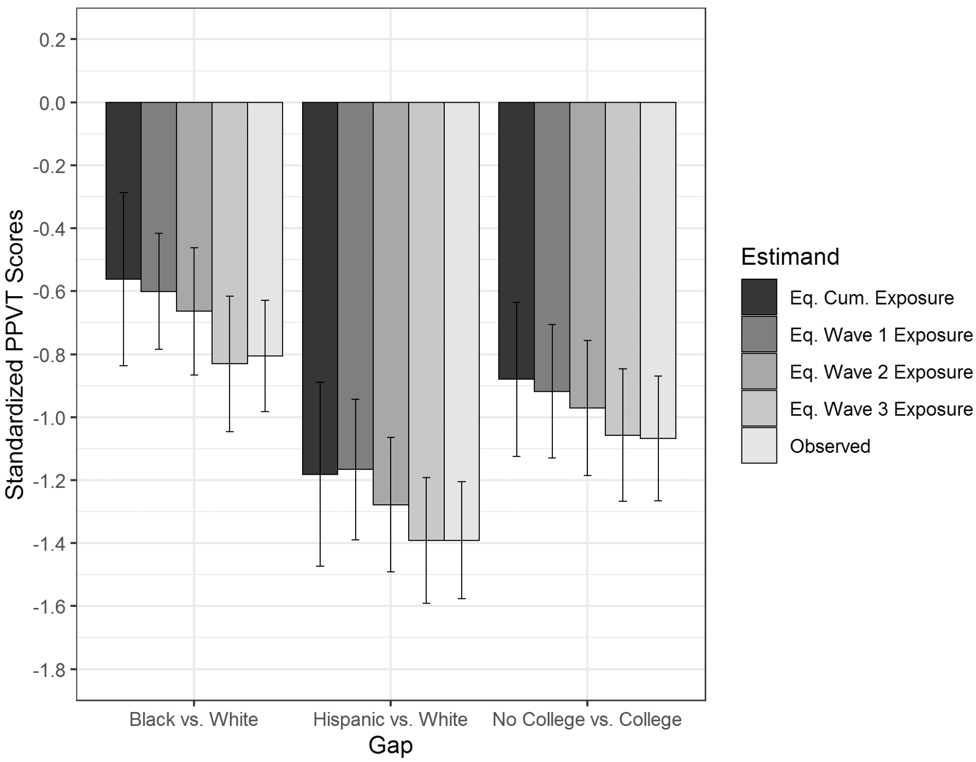

FIGURE 4 Observed and counterfactual gaps in vocabulary skills. PPVT, Peabody Picture Vocabulary Test.

during infancy or cumulatively over time is estimated to reduce the Black-White gap in attention problems by roughly two-thirds, the Hispanic-White gap by roughly half, and the class gap by about one-third. However, the figure also illustrates the imprecision of these estimates, as indicated by the wide confidence intervals

TABLE 5 Estimated differences between counterfactual and observed gaps in attention problems.

Contrast	Linear model Est. (SE)	*p*	Regression tree Est. (SE)	*p*	Random forest Est. (SE)	*p*	Super learner Est. (SE)	*p*
Eq. wave 1 exposure								
Black versus White	−.208 (.139)	.135	−.099 (.186)	.595	−.030 (.086)	.727	−.181 (.101)	.073
Hispanic versus White	−.141 (.096)	.142	−.077 (.136)	.571	.020 (.064)	.755	−.118 (.074)	.111
No college versus college	−.100 (.072)	.165	−.154 (.149)	.301	−.019 (.061)	.756	−.091 (.063)	.149
Eq. wave 2 exposure								
Black versus White	.175 (.216)	.418	−.058 (.269)	.829	.085 (.081)	.294	.147 (.127)	.247
Hispanic versus White	.112 (.138)	.417	−.071 (.174)	.683	.060 (.067)	.370	.094 (.086)	.274
No college versus college	.082 (.103)	.426	−.139 (.179)	.437	.023 (.060)	.702	.061 (.073)	.403
Eq. wave 3 exposure								
Black versus White	−.170 (.187)	.363	−.017 (.223)	.939	.017 (.085)	.841	−.108 (.114)	.344
Hispanic versus White	−.100 (.110)	.363	−.090 (.151)	.551	−.019 (.069)	.783	−.076 (.074)	.304
No college versus college	−.081 (.091)	.373	−.139 (.148)	.348	−.036 (.064)	.574	−.070 (.066)	.289
Eq. cumulative exposure								
Black versus White	−.294 (.210)	.162	−.100 (.320)	.755	.020 (.181)	.912	−.187 (.181)	.302
Hispanic versus White	−.136 (.153)	.374	−.068 (.226)	.763	.028 (.131)	.831	−.082 (.128)	.522
No college versus college	−.177 (.124)	.154	−.164 (.234)	.483	−.037 (.139)	.790	−.132 (.129)	.306

Note: Estimates are reported in SD units. Standard errors are estimated from 200 bootstrap samples. Results are combined across 50 multiply imputed datasets. *p*-values come from two-sided Wald tests of the null hypothesis that the difference between the counterfactual and observed gap is equal to zero.

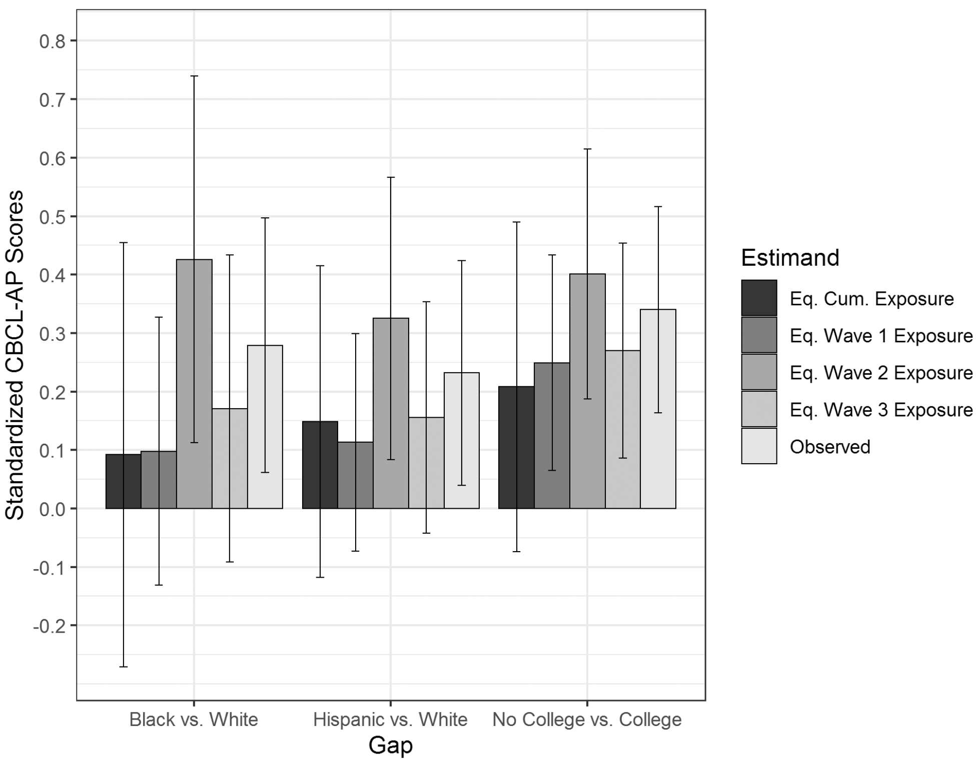

FIGURE 5 Observed and counterfactual gaps in attention problems. CBCL-AP, Child Behavior Checklist attention problems subscale.

and instability over different points in time, which precludes us from drawing any firm inferences about the link between lead contamination and gaps in attention problems.

Ancillary analyses

We performed a number of ancillary analyses to assess the robustness of our results. Part A of the Supporting

Information presents gaps comparing (i) children with and without a caregiver who completed high school and (ii) children in families with above versus below median incomes at baseline. For (i), results suggest that equalizing lead exposures at a relatively low level during infancy would significantly reduce gaps in vocabulary skills but not in attention problems. For (ii), equalizing lead exposures would not appreciably diminish gaps in vocabulary skills or attention problems. Part B of the Supporting Information presents counterfactual gaps under interventions equalizing average BLLs at levels lower than 4 μg/dL, and the estimates are nearly identical to those from the main text based on interventions equalizing average BLLs at 4 μg/dL. Part C of the Supporting Information reports estimates based on an alternative measure of areal lead contamination—specifically, an elevated blood lead incidence rate. These results also closely align with those from the main text, which are based on the average BLL in a neighborhood cluster.

Our estimates of counterfactual gaps only have a causal interpretation under strong assumptions that unobserved factors do not confound the effects of lead contamination on vocabulary skills and attention problems. In Part D of the Supporting Information, we present a sensitivity analysis gauging whether our key inferences are robust to bias arising from hypothetical patterns of unobserved confounding. Our conclusions about the impact of equalizing lead exposures on gaps in vocabulary skills would remain intact under moderate, but not extreme, levels of unobserved confounding. Because estimates for the impact of equalizing lead exposures on gaps in attention problems are so imprecise, our inferences about them, or rather the lack thereof, are insensitive to a wide range of confounding patterns.

DISCUSSION

In this study, we examined whether class and racial gaps in school readiness—measured at ages 4 and 5—are explained by differences in environmental lead contamination. Data from a cohort of Chicago children born in the mid-1990s suggested that these disparities would decline if exposures to lead hazards were equalized at a low level during infancy or cumulatively throughout early childhood. Specifically, we estimated that class and racial gaps in vocabulary skills would shrink by roughly 0.15–0.25 SDs, or about 15%–25%, if lead contamination were equalized, and we find that these results were fairly robust to different modeling choices and to potential violations of key identification assumptions. We also estimated that gaps in attention problems might shrink by roughly 0.1–0.2 SDs, or about 33%–66%, after equalizing lead contamination at a low level, but these results were too imprecise and unstable to draw any firm inferences from them.

To put the size of these estimates in perspective, consider, for example, that participation in the Infant Health and Development Program shrank income-based gaps in cognitive ability by an estimated 75%–100% (Duncan & Sojourner, 2013). However, the confidence intervals around these estimates are congruent with reductions as small as 0%–50%, and the program's calculated cost per household in current dollars ($40,000–$65,000; Karoly et al., 2005; WSIPP, 2019) is large compared to estimates of lead-paint remediation programs' cost per housing unit ($10,000–$12,000; HUD, 2018; Korfmacher, 2003). Overall, then, our results suggest that equalizing lead contamination may yield reductions in skill disparities that are one-seventh to one-third as large as those arising from an intensive early childhood intervention but at only a quarter of the price.

These findings have implications for public policies aimed at attenuating class and racial divergence in early skill formation. Many interventions intended as antidotes for academic disparities have mirrored the scholarly literature's focus on schools, the family, and early childhood education (Berends et al., 2019; Heckman et al., 2017; Reid et al., 2019). Our study suggests that these interventions, which are already making headway in closing academic gaps, could be even more effective were they paired with services that account for environmental inequalities during early childhood. Policymakers should therefore consider interventions aimed at remediating environmental hazards as another possible means for diminishing academic disparities.

To this end, the case of Chicago is informative. In the 1990s, roughly half of young children in the city had elevated BLLs, but less than two decades later, lead exposures declined owing partly to a remediation program that entailed intensive testing, targeted abatement of lead-based paint, and nutritional improvements among children (Sampson & Winter, 2016). Although we lack the data needed to precisely quantify the effects of this program on academic disparities in the city, our results suggest that it likely worked to suppress them, even if differences in lead exposure persist. Promising cases like Chicago, and discouraging ones like the Flint Water Crisis (Campbell et al., 2016), have spurred the federal government to scale up lead remediation efforts nationally. The funds allocated to this end in the Infrastructure Investment and Jobs Act of 2021 likely fall short of the resources needed to fully address the problem (McCormick & Lutz, 2021), but these investments are encouraging nonetheless, and future research should carefully assess their effects on developmental disparities.

Our findings also have implications for ecological systems theory. Consistent with Trentacosta et al. (2016) and prior work highlighting early sensitivities to environmental inputs (e.g., Heckman & Mosso, 2014; Shonkoff & Phillips, 2000), the present study underscores the need to further explore links between toxic contaminants, child development, and the emergence of disparities in cognitive and socioemotional skills. Future research in this area would ideally incorporate data on families, preschools, and environmental hazards to examine how each set of factors jointly contributes to class and racial gaps in development during early childhood.

Finally, our study also introduces new methods for analyzing developmental disparities. Many studies of early childhood interventions analyze the effects of a manipulation to some "input" (e.g., preschool or parenting practices) on individual outcomes, group disparities in the input, or group disparities in outcomes, without directly tying these phenomena together. The "gap-closing" analytic framework (Lundberg, 2022) will allow future research in developmental science to directly quantify how class, racial, or other group disparities in outcomes would change under different types of childhood interventions to equalize an input.

Despite these contributions, our study has nontrivial limitations, including its narrow temporal and geographic scope. Data from children born in the mid-1990s, when BLLs were higher and more unequal than they are today, can only reveal so much about the potential effects of current remediation plans, like those funded by the federal infrastructure bill. Moreover, our study's setting—Chicago—may represent an extreme case, as the city has suffered from particularly stark patterns of class and racial inequality and from particularly severe lead hazards. Estimates from this analysis, then, may provide an upper bound on the degree to which lead contamination drives gaps in school readiness relative to other places and more recent cohorts. We speculate that similar results may hold in other highly stratified U.S. cities, especially those within the Rust Belt that have not yet remediated their lead hazards. Conversely, in rural areas, in Sunbelt cities with newer housing and infrastructure, or in communities that have already implemented ambitious remediation programs, the link between lead contamination and disparities in school readiness may be weaker. Future research assessing more recent data and a broader geographic scope is needed to firmly establish limits on the generalizability of our theoretical model and empirical findings.

Our study also suffers from measurement limitations. We rely on an imperfect proxy for the degree of lead contamination within a child's residential environment, and measurement error in this variable may lead to attenuation bias in our gap-closing estimates (VanderWeele, 2015). Future research might attempt to replicate our findings using alternative measures of lead contamination based on environmental samples of the air, water, soil, or dust in and around households, although these may suffer from errors of their own (Phillis, 2021; Sarnat et al., 2007).

Our measures of school readiness are imperfect as well. Although the PPVT and CBCL-AP generally possess desirable psychometric properties, these measures may suffer from racial, ethnic, or class biases that are difficult to detect, and neither instrument was designed alongside a "justice-oriented, anti-racist validity framework" (Randall et al., 2022). Such biases could inflate our estimates of gaps in school readiness; however, they would only distort our estimates of the difference between counterfactual and observed gaps if the magnitude of any bias were correlated with lead contamination. Beyond differences in validity across diverse subpopulations, the CBCL-AP may also suffer from errors due to inaccurate parental reporting. Future studies should therefore extend our work by examining a broader and more defensible set of outcomes.

Finally, our inferences about the link between lead contamination and disparities in school readiness hinge on strong assumptions about the absence of unobserved confounding. Despite our best efforts to control for the most powerful joint predictors of lead exposure and school readiness, the threat of bias due to unmeasured confounders remains ubiquitous in observational studies. As we show in Supporting Information, the magnitude of confounding bias would have to be fairly large to invalidate our key inferences about the explanatory role of lead contamination, but we cannot rule out this possibility empirically. Thus, it will be important for future studies to buttress our findings with data from designs that can consistently estimate the impact of lead exposure on skill disparities under weaker assumptions.

As public concerns about racial inequality, poverty, and environmental injustice amplify, rigorous scholarship on the connections between these entangled social maladies and on means for redressing them will be essential. In this context, mitigating the methodological limitations outlined above and clarifying the specific environmental mechanisms driving durable academic disparities is a crucial imperative for the next generation of research in developmental science.

ACKNOWLEDGMENTS

The research was supported by the Mansueto Institute for Urban Innovation at the University of Chicago and the Price Center for Social Innovation at USC, as well as grants from the Social Sciences and Humanities Research Council of Canada (Num. 435-2018-0736) and the U.S. National Science Foundation (Num. 2015613). We thank Robert J. Sampson, Alix Winter, Stephen Raudenbush, and participants in the Inequality Workshop at Saint Anselm College for helpful feedback.

DATA AVAILABILITY STATEMENT

Some of the data on which this study is based are restricted access and were obtained from the Chicago Department of Public Health (CDPH) under special contractual arrangements designed to protect sensitive health information. The CDPH disclaims responsibility for any analysis, interpretations, or conclusions drawn from them. Although the data necessary to reproduce the analyses presented here are not publicly accessible, the analytic code necessary to reproduce the analyses is publicly accessible and provided at https://github.com/gtwodtke/skills_gaps_lead, as are the materials (e.g., links to codebooks and user manuals for key variables and data sources) necessary to attempt to replicate the findings and instructions for securely accessing the data used in this analysis via their custodial institutions. The analyses presented here were not preregistered.

ORCID

Jared N. Schachner https://orcid.org/0000-0002-4177-6986
Geoffrey T. Wodtke https://orcid.org/0000-0001-6424-6040

REFERENCES

Achenbach, T. M., & Rescorla, L. (2000). *Manual for the ASEBA preschool-forms & profiles*. University of Vermont.

Aelion, C. M., Davis, H. T., Lawson, A. B., Cai, B., & McDermott, S. (2013). Associations between soil lead concentrations and populations by race/ethnicity and income-to-poverty ratio. *Environmental Geochemistry and Health*, *35*, 1–12. https://doi.org/10.1007/s10653-012-9472-0

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*, 307–341. https://doi.org/10.1257/app.20160404

Berends, M., Waddington, R. J., & Schoenig, J. (2019). *School choice at the crossroads: Research perspectives*. Routledge.

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees*. Champan and Hall/CRC.

Bronfenbrenner, U., & Morris, P. A. (2006). The bioecological model of human development. In R. M. Lerner & W. Damon (Eds.), *Handbook of child psychology* (pp. 793–828). John Wiley & Sons.

Bullard, R. D. (2000). *Dumping in Dixie: Race, class, and environmental quality* (3rd ed.). Routledge.

Campanella, R., & Mielke, H. W. (2008). Human geography of New Orleans' high-lead geochemical setting. *Environmental Geochemistry and Health*, *30*, 531–540. https://doi.org/10.1007/s10653-008-9190-9

Campbell, C., Greenberg, R., Mankikar, D., & Ross, R. (2016). A case study of environmental injustice: The failure in Flint. *International Journal of Environmental Research and Public Health*, *13*, 951. https://doi.org/10.3390/ijerph13100951

Canfield, R. L., Henderson, C. R., Cory-Slechta, D. A., Cox, C., Jusko, T. A., & Lanphear, B. P. (2003). Intellectual impairment in children with blood lead concentrations below 10 mg per deciliter. *The New England Journal of Medicine*, *348*, 1517–1526. https://doi.org/10.1056/NEJMoa022848

Cox, D. C., Dewalt, G., O'Haver, R., & Salatino, B. (2011). *American Healthy Homes Survey: Lead and arsenic findings (HERO technical report no. 1812739)*. U.S. Department of Housing and Urban Development.

Crowder, K., & Downey, L. (2010). Interneighborhood migration, race, and environmental hazards. *American Journal of Sociology*, *115*, 1110–1149. https://doi.org/10.1086/649576

Duncan, G. J., & Magnuson, K. (2005). Can family socioeconomic resources account for racial and ethnic test score gaps? *The Future of Children*, *15*, 35–54. https://doi.org/10.1353/foc.2005.0004

Duncan, G. J., & Magnuson, K. (2011). The nature and impact of early achievement skills, attention skills, and behavior problems. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity?* (pp. 47–69). Russell Sage Foundation.

Duncan, G. J., & Sojourner, A. J. (2013). Can intensive early childhood intervention programs eliminate income-based cognitive and achievement gaps? *Journal of Human Resources*, *48*, 945–968. https://doi.org/10.3368/jhr.48.4.945

Dunn, L. M. (1997). *PPVT-III: Peabody Picture Vocabulary Test*. American Guidance Service.

Earls, F. J., Brooks-Gunn, J., Raudenbush, S. W., & Sampson, R. J. (2007). *Project on Human Development in Chicago Neighborhoods*. Inter-University Consortium for Political and Social Research. https://doi.org/10.3886/ICPSR13717.v1

Egan, K. B., Cornwell, C. R., Courtney, J. G., & Ettinger, A. S. (2021). Blood lead levels in U.S. children ages 1–11 years, 1976–2016. *Environmental Health Perspectives*, *129*, 1–11. https://doi.org/10.1289/EHP7932

Elliott, J. R., & Frickel, S. (2013). The historical nature of cities: A study of urbanization and hazardous waste accumulation. *American Sociological Review*, *78*, 521–543. https://doi.org/10.1177/0003122413493285

Elwert, F., & Winship, C. (2014). Endogenous selection bias: The problem of conditioning on a collider variable. *Annual Review of Sociology*, *40*, 31–53. https://doi.org/10.1146/annurev-soc-071913-043455

Evens, A., Hryhorczuk, D., Lanphear, B. P., Rankin, K. M., Lewis, D. A., Forst, L., & Rosenberg, D. (2015). The impact of low-level lead toxicity on school performance among children in the Chicago Public Schools. *Environmental Health*, *14*, 14–21. https://doi.org/10.1186/s12940-015-0008-9

Finneran, D. A., Heilmann, J. J., Moyle, M. J., & Chen, S. (2020). An examination of cultural linguistic influences on PPVT 4 performance in African American and Hispanic preschoolers from low income communities. *Clinical Linguistics & Phonetics*, *34*, 242–255. https://doi.org/10.1080/02699206.2019.1628811

Gangl, M. (2010). Causal inference in sociological research. *Annual Review of Sociology*, *36*, 21–47. https://doi.org/10.1146/annurev.soc.012809.102702

Gazze, L., Persico, C., & Spirovska, S. (2021). *The spillover effects of pollution: How exposure to Lead affects everyone in the classroom*. National Bureau of Economic Research. https://doi.org/10.3386/w28782

Goodlad, J. K., Marcus, D. K., & Fulton, J. J. (2013). Lead and attention-deficit/hyperactivity disorder (ADHD) symptoms: A meta-analysis. *Clinical Psychology Review*, *33*, 417–425. https://doi.org/10.1016/j.cpr.2013.01.009

Gopaul-McNicol, S.-a., Reid, G., & Wisdom, C. (1998). The psychoeducational assessment of Ebonics speakers: Issues and challenges. *Journal of Negro Education*, *67*, 16–24. https://doi.org/10.2307/2668236

Gross, D., Fogg, L., Young, M. A., & Ridge, A. (2006). The equivalence of the Child Behavior Checklist across parent race/ethnicity, income level, and language. *Psychological Assessment*, *18*, 313–323. https://doi.org/10.1037/1040-3590.18.3.313

Guryan, J., Hurst, E., & Kearney, M. (2008). Parental education and parental time with children. *Journal of Economic Perspectives*, *22*, 23–46. https://doi.org/10.1257/jep.22.3.23

Hanushek, E., Peterson, P., Talpey, L., & Woessmann, L. (2020). *Long-run trends in the U.S. SES-achievement gap*. National Bureau of Economic Research. https://doi.org/10.3386/w26764

Hanushek, E., & Rivkin, S. G. (2009). Harming the best: How schools affect the black-white achievement gap. *Journal of Policy Analysis and Management*, *28*, 366–393. https://doi.org/10.1002/pam.20437

Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science*, *312*, 1900–1902. https://doi.org/10.1126/science.1128898

Heckman, J. J., Holland, M., Makino, K., Pinto, R., & Rosales-Rueda, M. (2017). *An analysis of the Memphis Nurse-Family Partnership* *program*. National Bureau of Economic Research. https://doi.org/10.3386/w23610

Heckman, J. J., & Mosso, S. (2014). The economics of human development and social mobility. *Annual Review of Economics*, *6*, 689–733. https://doi.org/10.1146/annurev-economics-080213-040753

Helms, J. E. (2006). Fairness is not validity or cultural bias in racial-group assessment: A quantitative perspective. *American Psychologist*, *61*, 845–859. https://doi.org/10.1037/0003-066X.61.8.845

Hout, M. (2012). Social and economic returns to college education in the United States. *Annual Review of Sociology*, *38*, 379–400. https://doi.org/10.1146/annurev.soc.012809.102503

Jacobs, D. E., Clickner, R. P., Zhou, J. Y., Viet, S. M., Marker, D. A., Rogers, J. W., Zeldin, D. C., Broene, P., & Friedman, W. (2002). The prevalence of lead-based paint hazards in U.S. housing. *Environmental Health Perspectives*, *110*(10), A599–A606. https://doi.org/10.1289/ehp.021100599

Jencks, C., & Phillips, M. (1998). *The Black-White test score gap*. Brookings Institution Press.

Karoly, L. A., Kilburn, M. R., & Cannon, J. S. (2005). *Early childhood interventions: Proven results, future promise*. RAND Corporation. https://doi.org/10.7249/MG341

Kitagawa, E. M. (1955). Components of a difference between two rates. *Journal of the American Statistical Association*, *50*, 1168–1194. https://doi.org/10.2307/2281213

Konold, T. R., Hamre, B. L., & Pianta, R. C. (2003). Measuring problem behaviors in young children. *Behavioral Disorders*, *28*, 111–123. https://doi.org/10.1177/019874290302800202

Korfmacher, K. S. (2003). *Long-term costs of lead poisoning: How much can New York save by stopping lead?* Environmental Health Sciences Center, University of Rochester.

Lanphear, B. P., Byrd, R. S., Auinger, P., & Schaffer, S. J. (1998). Community characteristics associated with elevated blood lead levels in children. *Pediatrics*, *101*, 264–271. https://doi.org/10.1542/peds.101.2.264

Lanphear, B. P., Hornung, R., Khoury, J., Yolton, K., Baghurst, P., Bellinger, D. C., Canfield, R. L., Dietrich, K. N., Bornschein, R., Greene, T., Rothenberg, S., Needleman, H., Schnaas, L., Wasserman, G., Graziano, J., & Roberts, R. (2005). Low-level environmental lead exposure and children's intellectual function: An international pooled analysis. *Environmental Health Perspectives*, *113*, 894–899. https://doi.org/10.1289/ehp.7688

Lareau, A. (2011). *Unequal childhoods: Class, race, and family life*. UC Press.

Lidsky, T. I., & Schneider, J. S. (2003). Lead neurotoxicity in children: Basic mechanisms and clinical correlates. *Brain*, *126*, 5–19. https://doi.org/10.1093/brain/awg014

Lundberg, I. (2022). The gap-closing estimand: A causal approach to study interventions that close disparities across social categories. *Sociological Methods & Research*. OnlineFirst. https://doi.org/10.1177/00491241211055769

Manduca, R., & Sampson, R. J. (2021). Childhood exposure to polluted neighborhood environments and intergenerational income mobility, teenage birth, and incarceration in the USA. *Population and Environment*, *42*, 501–523. https://doi.org/10.1007/s11111-020-00371-5

McCormick, E., & Lutz, E. (2021, December 8). 'We're losing IQ points': The lead poisoning crisis unfolding among U.S. children. *The Guardian*.

McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology*, *46*, 235–261. https://doi.org/10.1016/j.jsp.2007.05.001

Miranda, M. L., Kim, D., Reiter, J., Galeano, M. A., & Maxson, P. (2009). Environmental contributors to the achievement gap. *Neurotoxicology*, *30*, 1019–1024. https://doi.org/10.1016/j.neuro.2009.07.012

Mohai, P., Pellow, D., & Roberts, J. T. (2009). Environmental justice. *Annual Review of Environment and Resources*, *34*, 405–430. https://doi.org/10.1146/annurev-environ-082508-094348

Muller, C., Sampson, R. J., & Winter, A. S. (2018). Environmental inequality: The social causes and consequences of lead exposure. *Annual Review of Sociology*, *44*, 263–282. https://doi.org/10.1146/annurev-soc-073117-041222

Owens, J., & McLanahan, S. S. (2020). Unpacking the drivers of racial disparities in school suspension and expulsion. *Social Forces*, *98*, 1548–1577. https://doi.org/10.1093/sf/soz095

Pais, J., South, S. J., & Crowder, K. (2012). Metropolitan heterogeneity and minority neighborhood attainment: Spatial assimilation or place stratification? *Social Problems*, *59*, 258–281. https://doi.org/10.1525/sp.2012.59.2.258

Phillis, M. (2021). New lead testing approach could reveal higher levels in water. *The Associated Press News*. https://apnews.com/article/science-business-health-environment-and-nature-michigan-e8cb1eb34727436a855818cf54e6ee91

Randall, J., Slomp, D., Poe, M., & Oliveri, M. E. (2022). Disrupting White supremacy in assessment: Toward a justice-oriented, antiracist validity framework. *Educational Assessment*, *27*, 170–178. https://doi.org/10.1080/10627197.2022.2042682

Rao, J. N. K., & Wu, C. F. J. (1988). Resampling inference with complex survey data. *Journal of the American Statistical Association*, *83*, 231–241. https://doi.org/10.2307/2288945

Reardon, S. F. (2011). The widening academic achievement gap between the rich and the poor. In G. J. Duncan & R. J. Murnane (Eds.), *Whither opportunity?* (pp. 91–116). Russell Sage Foundation.

Reardon, S. F. (2016). School segregation and racial academic achievement gaps. *Russell Sage Foundation Journal of the Social Sciences*, *2*, 34–57. https://doi.org/10.7758/rsf.2016.2.5.03

Reid, J. L., Melvin, S. A., Kagan, S. L., & Brooks-Gunn, J. (2019). Building a unified system for universal pre-K: The case of New York City. *Children and Youth Services Review*, *100*, 191–205. https://doi.org/10.1016/j.childyouth.2019.02.030

Rock, D. A., & Stenner, A. J. (2005). Assessment issues in the testing of children at school entry. *The Future of Children*, *15*, 15–34. https://doi.org/10.1353/foc.2005.0009

Rothstein, R. (2018). *The color of law*. Liveright Publishing Corporation.

Rubin, D. B. (1987). *Multiple imputation for nonresponse surveys*. Wiley.

Sampson, R. J., & Winter, A. S. (2016). The racial ecology of lead poisoning: Toxic inequality in Chicago neighborhoods, 1995–2013. *Du Bois Review*, *13*, 261–283. https://doi.org/10.1017/S1742058X16000151

Sampson, R. J., & Winter, A. S. (2018). Poisoned development: Assessing childhood lead exposure as a cause of crime in a birth cohort followed through adolescence. *Criminology*, *56*, 269–301. https://doi.org/10.1111/1745-9125.12171

Sarnat, J. A., Wilson, W. E., Strand, M., Brook, J., Wyzga, R., & Lumley, T. (2007). Panel discussion review: Session 1—Exposure assessment and related errors in air pollution epidemiologic studies. *Journal of Exposure Science & Environmental Epidemiology*, *17*, S75–S82. https://doi.org/10.1038/sj.jes.7500621

Sastry, N., & Pebley, A. R. (2010). Family and neighborhood sources of socioeconomic inequality in children's achievement. *Demography*, *47*, 777–800. https://doi.org/10.1353/dem.0.0114

Senechal, M., & LeFevre, J. (2002). Parental involvement in the development of children's reading skill. *Child Development*, *73*, 445–460. https://doi.org/10.1111/1467-8624.00417

Shonkoff, J. P., & Phillips, D. A. (2000). *From neurons to neighborhoods: The science of early childhood development*. National Academies Press.

Simonoff, J. S. (1996). *Smoothing methods in statistics*. Springer.

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*, 179–218. https://doi.org/10.1007/s00148-018-0707-y
Teye, S. O., Yanosky, J. D., Cuffee, Y., Weng, X., Luquis, R., Farace, E., & Wang, L. (2021). Exploring persistent racial/ethnic disparities in lead exposure among American children aged 1–5 years: Results from NHANES 1999–2016. *International Archives of Occupational and Environmental Health*, *94*, 723–730. https://doi.org/10.1007/s00420-020-01616-4
Trentacosta, C. J., Davis-Kean, P., Mitchell, C., Hyde, L., & Dolinoy, D. (2016). Environmental contaminants and child development. *Child Development Perspectives*, *10*, 228–233. https://doi.org/10.1111/cdep.12191
Trentacosta, C. J., & Mulligan, D. J. (2020). New directions in understanding the role of environmental contaminants in child development: Four Themes. *New Directions for Child and Adolescent Development*, *2020*, 39–51. https://doi.org/10.1002/cad.20363
U.S. Department of Housing and Urban Development (HUD). (2018). *Lead hazard control and healthy homes: 2018 summary statement and initiatives*.
van der Laan, M. J., Polley, E. C., & Hubbard, A. E. (2007). Super learner. *Statistical Applications in Genetics and Molecular Biology*, *6*, 1–25. https://doi.org/10.2202/1544-6115.1309
VanderWeele, T. J. (2015). *Explanation in causal inference: Methods for mediation and interaction*. Oxford University Press.
von Hippel, P. T., Workman, J., & Downey, D. B. (2018). Inequality in reading and math skills forms mainly before kindergarten. *Sociology of Education*, *91*, 323–357. https://doi.org/10.1177/0038040718801760
Washington, J. A., & Craig, H. K. (1999). Performances of at-risk, African American preschoolers on the Peabody Picture Vocabulary Test-III. *Language, Speech, and Hearing Services in School*, *30*, 75–82. https://doi.org/10.1044/0161-1461.3001.75
Washington State Institute for Public Policy (WSIPP). (2019). *Infant Health and Development Program (IHDP) public health & prevention: Home- or family-based*.
Wasserstein, R. L., & Lazar, N. A. (2016). The ASA statement on p-values: Context, process and purpose. *The American Statistician*, *70*, 129–133. https://doi.org/10.1080/00031305.2016.1154108
Williams, K. T., & Wang, J. (1997). *Technical references to the Peabody picture vocabulary test* (3rd ed.). American Guidance Service.
Wodtke, G. T., Ard, K., Bullock, C., White, K., & Priem, B. (2022). Concentrated poverty, ambient air pollution, and child cognitive development. *Science Advances*, *8*, 1–19. https://doi.org/10.1126/sciadv.add0285
Wodtke, G. T., Ramaj, S., & Schachner, J. (2022). Toxic neighborhoods: The effects of concentrated poverty and environmental lead contamination on early childhood. *Demography*, *59*, 1275–1298. https://doi.org/10.1215/00703370-10047481
Zhou, X., & Wodtke, G. T. (2020). Residual balancing: A method of constructing weights for marginal structural models. *Political Analysis*, *28*, 487–506. https://doi.org/10.1017/pan.2020

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Schachner, J. N., & Wodtke, G. T. (2023). Environmental inequality and disparities in school readiness: The role of neurotoxic lead. *Child Development*, *94*, e308–e327. https://doi.org/10.1111/cdev.13949