I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12178

The Legacy Lead Deposition in Soils and Its Impact on Cognitive Function in Preschool-Aged Children in the United States

Karen Clay
Margarita Portnykh
Edson Severnini

FEBRUARY 2019

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12178

The Legacy Lead Deposition in Soils and Its Impact on Cognitive Function in Preschool-Aged Children in the United States

Karen Clay
Carnegie Mellon University and NBER

Margarita Portnykh
Carnegie Mellon University

Edson Severnini
Carnegie Mellon University and IZA

FEBRUARY 2019

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 Legacy Lead Deposition in Soils and Its Impact on Cognitive Function in Preschool-Aged Children in the United States*

Surface soil contamination has been long recognized as an important pathway of human lead exposure, and is now a worldwide health concern. This study estimates the causal effects of exposure to lead in topsoil on cognitive ability among 5-year-old children. We draw on individual level data from the 2000 U.S. Census, and USGS data on lead in topsoil covering a broad set of counties across the United States. Using an instrumental variable approach relying on the 1944 Interstate Highway System Plan, we find that higher lead in topsoil increases considerably the probability of 5-year-old boys experiencing cognitive difficulties such as learning, remembering, concentrating, or making decisions. Living in counties with topsoil lead concentration above the national median roughly doubles the probability of 5-year-old boys having cognitive difficulties. Nevertheless, it does not seem to affect 5-year-old girls, consistent with previous studies. Importantly, the adverse effects of lead exposure on boys are found even in counties with levels of topsoil lead concentration considered low by the guidelines from the U.S. EPA and state agencies. These findings are concerning because they suggest that legacy lead may continue to impair cognition today, both in the United States and in other countries that have considerable lead deposition in topsoil.

JEL Classification: N52, Q53, Q56, R11, I15, I18, I25, I28

Keywords: legacy lead in soil, cognition, pre-school children

Corresponding author:
Edson Severnini
Carnegie Mellon University
4800 Forbes Avenue
Pittsburgh PA 15213
E-mail: ersevernini@gmail.com

* We thank Guy Michaels for sharing his data on the 1944 Highway Plan, and Howard Mielke for sharing the lead emissions data. Karen Clay, Margarita Portnykh, and Edson Severnini gratefully acknowledge financial support from Heinz College at Carnegie Mellon University, and Margarita Portnykh gratefully acknowledges financial support from the Earhart Foundation, which supported her postdoctoral fellowship at the Bren School at the University of California, Santa Barbara.

1. Introduction

Although lead has been banned from gasoline, paint, and other substances in the United States and many other countries around the world, the legacy of lead use is a critical environmental and public health issue. Surface soil contamination, in particular, has been long recognized as an important pathway of human lead exposure, and is now a worldwide health concern (Mielke et al. 1983, Lanphear et al. 1998, Mielke and Reagan 1998, Mielke 1999, Levin et al. 2008, Jennings 2013, Filippelli et al. 2015, Laidlaw et al. 2016, Mielke 2016). Indeed, elevated lead concentrations in soil are likely to remain high for hundreds of years, and resuspension of lead contaminated soil can increase air lead levels and redistribute lead to nearby areas (Dudka and Adriano 1997, Young et al. 2002, Harris and Davidson 2005, Laidlaw and Filippelli 2008, Laidlaw et al. 2012, Zahran et al. 2013a, Datko-Williams et al. 2014). Emissions by on-road vehicles have played a major role in the accumulation of lead in topsoil (Lagerwerff and Specht 1970, Yassoglou et al. 1987, Mielke 1993, Sutherland 2000, Sheets et al. 2001, Newell and Rogers (2003), Wang and Zhang 2018). Over 10 million metric tons of lead were transferred to the global environment by motor vehicles in the 20th century, and about 60 percent of that total was dispersed in the United States alone (Ethyl Corp. 1984, Mielke 1991, 1993, Mielke and Reagan 1998, Sutherland 2000, Sheets et al. 2001, EPA 2006, Mielke et al. 2010, 2011).[2]

Children face the greatest risk of exposure to lead contaminated soil in the streets and yards around their houses, and in open spaces such as public playgrounds and urban gardens, where they play (Mielke 1999, Hunt et al. 2006, Zahran et al. 2013b, Schwarz et al. 2016). A review of several studies by the U.S. EPA indicated that for every 1000 ppm increase in soil lead concentration, blood lead levels generally increase by approximately 1 to 5 µg/dL in infants and children below 6 years old (EPA 2006).[3] The dose-response relationship, however, seems to be nonlinear: about

[2] Even though both deteriorating exterior lead-based paint and automotive traffic from the days of leaded gasoline are the two major sources of lead in urban soils, extensive studies in a number of American cities found the highest soil-lead concentrations in the central sections of each city, where traffic and population density are greatest (Mielke 1991, 1993, Sutherland 2000, Sheets et al. 2001, EPA 2006). "[T]he age of housing did not seem to be a major factor, which suggests that the impacts of Pb-based paint may be dominated by historic emissions of leaded gasoline additives" (EPA 2006, p. 3-20). In our analysis we control for the age of the housing.

[3] Lead-formulated aviation gasoline, used in a large fraction of piston-engine aircraft, has also been shown to affect child blood lead levels, even adjusting for other known sources of lead exposure. Over half of the current flow of lead into the atmosphere in the United States is attributable to aviation fuel (Zahran et al. 2017).

1.4 μg/dL per 100 ppm for soil lead concentrations below 100 ppm, but only 0.32 μg/dL per 100 ppm for soil lead above 300 ppm (Mielke et al. 2007). After an exhaustive analysis of existing epidemiological studies on the health effects of low level lead, the U.S. National Toxicology Program (NTP 2012) noted that "young children (ages 1-5 years) consistently have higher blood Pb levels than do older children, likely due to hand-to-mouth behavior in this age group" (p.9), and concluded that "[i]n children, there is *sufficient* evidence that blood Pb levels <5 μg/dL are associated with increased diagnosis of attention-related behavioral problems, greater incidence of problem behaviors, and decreased cognitive performance" (NTP 2012, p. xviii).

Environmental and public health agencies worldwide have issued guidance values to monitor exposure to surface soil contamination (Jennings and Petersen 2006, Jennings 2013). For reference, lead is naturally found in soils at very low concentrations – a median of 11 ppm for U.S. agricultural soils (Holmgren et al. 1993). The U.S. EPA has established 400 ppm as a generally protective soil screening level (SSL) for residential soil-lead (EPA 1994), but recommends 100 ppm as a threshold to initiate best management practices to mitigate exposure to lead in soil for gardening-related exposure pathways (EPA 2014). "The basis for the SSL is children playing in lead contaminated soil and other gardening exposures, with the predominant source of exposure from soil ingestion" (EPA 2014, p. 9).

In this study, we focus on topsoil lead concentration considered low by the guidelines from the U.S. EPA and state agencies, and estimate *causal* effects of that exposure on cognitive function of pre-school children. To estimate those effects, we draw on individual level data for 5-year-old children from the U.S. 2000 Census and USGS data on lead in topsoil covering a broad set of counties across the nation, as well as data from a range of other sources. The Census included a question for individuals 5 years and older inquiring whether they had experienced any "cognitive difficulties" that persisted for at least six months such as learning, remembering, concentrating, or making decisions because of a physical, mental, or emotional condition. We focus on children who are age 5 in the Census, because the age range 1-5 years is emphasized by the NTP. Our sample includes children from the 252 largest counties in the United States (100,000 people or more), representing 45 percent of the total population. In the 2000s, the U.S. Geological Survey (USGS) collected data from topsoil (depth of 0 to 5 cm) on soil composition. The soil was analyzed for many components including lead. To control for other factors in the empirical analysis, we use data from other sources that are described in the Materials and Methods section below.

Our analysis builds on previous research linking lead exposure to cognitive function by estimating the causal effect of lead. Since soil contamination is not randomly assigned, studies that attempt to compare mental development outcomes for children exposed to differing levels of topsoil lead may underestimate the effects of lead (Glass et al. 2013, Currie et al. 2014, Dominici et al. 2014, Dominici and Zigler 2017). Estimates of the effects of lead are likely to be biased downward for at least three reasons. First, since environmental amenities are capitalized into housing prices (Chay and Greenstone 2005, Mendelsohn and Olmstead 2009, Currie et al. 2015), families with higher incomes or preferences for cleaner environment might avoid exposure by sorting into locations with better soil quality (Banzhaf and Walsh 2008). Failure to account for this avoidance behavior will lead to underestimation of the effects of soil contamination on cognitive difficulty. Second, topsoil lead levels are higher in urban areas where there are often more educated individuals with better access to health care (Diamond 2016), which can also cause underestimation of the true effects of soil contamination on cognitive difficulty. Third, because soil samples are not necessarily taken from the neighborhoods where children live, there might be a disconnection between average lead in soil measured at the county level and the exposure experienced by those children. Therefore, the relationship between individual cognitive difficulty and average topsoil lead might become weaker, making the estimate of true effect of interest biased towards zero (e.g., Aizer et al. 2018) – this is the so-called attenuation bias.

To estimate the causal effect of exposure to lead through contaminated soil on cognitive difficulty, we use instrumental variable (IV) methods (e.g., Bowden and Turkington 1984, Angrist et al. 1996, Greenland 2000, Angrist and Krueger 2001, Hernan and Robins 2006, Martens et al. 2006). This approach tackles those two issues at the same time, and was used recently to study the effects of lead exposure on fertility (Clay et al. 2018). Besides establishing a clear pathway of exposure, an instrumental variable quasi-experimentally induces a group of children to be exposed to lead in soil and another group to not face that exposure (or to at least experience less exposure). The resulting causal estimate is a local average treatment effect (LATE), the average treatment effect for the children whose "treatment" status was influenced by the instrument (Imbens and Angrist 1994). Because lead emissions by vehicles were an important source of topsoil contamination in the U.S. before the phasedown of lead additives to gasoline in the 1980s, and the eventual ban in 1996, the presence of intense traffic nearby should affect exposure to lead-contaminated soil. We thus exploit the 1944 Interstate Highway System Plan as an instrument, which was designed primarily for military purposes rather than potential economic and environmental outcomes, but has been shown to predict the placement of interstate highways (Baum-Snow 2007, Michaels 2008). Therefore, children growing up in counties recommended to receive a highway by the 1944 plan may experience more exposure to lead in soil in the 2000s than children in counties not supposed to receive any part of the interstate highway system. It is worth explicitly noting why we do not use actual highways as an instrument. Households concerned with environmental amenities might influence the precise location of highways within a county, and might also be more willing to invest in other types of infrastructure such as schools and hospitals, invalidating the quasi-experimental aspect of the potential instrument, and generating the same bias we are attempting to address.

The IV estimates show that increased exposure to lead contaminated soil heightens the probability of boys experiencing cognitive impairment. Those *causal* estimates indicate that living in counties with topsoil lead concentration above the national median roughly doubles the probability of 5-year-old boys having cognitive difficulties. Consistent with previous findings (Jedrychowski et al. 2009, Llop et al. 2013, Khanna 2015) no effects were found for girls. The IV estimate for boys is an order of magnitude larger than the *associational* relationship estimated by standard regression analysis in our sample, highlighting the need to address the biases discussed above.

2. Background

Link Between Lead in the Environment and Lead in the Human Body

Lead primarily enters the body from breathing in dust or chemicals that contain lead or by ingesting food or liquids that contain lead. Once lead reaches the lungs, it goes quickly to other parts of the body via blood stream. Once lead reaches the stomach, some is absorbed into the bloodstream and the remainder is excreted (U.S. Dept. of Health and Human Services 2007). Once in the blood, lead travels to the "soft tissues" and organs such as the liver, kidneys, lungs, brain, spleen, muscles, and heart. After several weeks, most of the lead moves into the bones and teeth. The half-life of lead in blood is approximately 30 days (Griffin et al. 1975, Rabinowitz et al. 1976, Chamberlain et al. 1978). Once it is taken in and distributed to organs, the lead that is not stored in bones leaves the body via urine or feces. Lead, stored in the bones, can be released during pregnancy and transferred to the infant (Gulson et al. 2016). The primary method for determining lead exposure is measurement of blood lead levels. According to the Centers for Disease Control and Prevention (CDC 2017), "The National Toxicology Program [NTP 2012], and the American Academy of Pediatrics [AAP 2016] have concluded that there is *sufficient* evidence for adverse health effects in children and adults at blood lead levels (BLLs) <5 micrograms per deciliter (μg/dL)."

In most settings, however, the data on blood lead levels are not available, and only data on airborne or topsoil lead are available. Because this is the case in this study, we provide evidence on the relationship between air and soil lead levels and blood lead levels. In a review of several studies, EPA concluded that "at relatively low air-Pb concentrations (≤ 2 μg/m3), pediatric blood-Pb levels generally increase by ~2 μg/dL per each 1 μg/m3 increment in air-Pb concentration" (EPA 2006, p. E-7). Regarding soil, besides ingestion, soil lead has been shown to affect young children's blood levels via soil resuspension (Zahran et al. 2013a). Overall, "[i]t has been estimated that for every 1000 ppm increase in soil-Pb concentration, pediatric blood-Pb levels generally increase by ~1 to 5 μg/dL in infants and children <6 years old" (EPA 2006, p. E-8). More recent evidence shows a curvilinear dose-response relationship between lead in soil and in blood: below 100 ppm soil lead children's blood level exposure response is steep (1.4 μg/dL per 100 ppm), while above 300 ppm the blood level exposure response is gradual (0.32 μg/dL per 100 ppm) (Mielke et al. 2007).

Lead Exposure and Cognitive Function and Attention-Related Behavior

Lead is known to be toxic. Since the pioneering work of Byers and Lord (1943), Needleman et al. (1979), Bellinger et al. (1987), McMichael et al. (1988), Needleman et al. (1990), and Baghurst et al. (1992), the relationship between lead exposure and neurobehavioral function has been studied extensively. The National Toxicology Program of the U.S. Department of Health and Human Services published an exhaustive analysis of existing epidemiological studies on the health effects of low level lead, including studies of the effects of lead on cognitive function and attention-related behavior in children (NTP 2012).

Lead has been long shown to affect IQ, both in children and adults (Ferrie et al. 2012). "The conclusion of *sufficient* evidence that decreases in IQ scores in children are associated with blood Pb levels <10 μg/dL measured in early childhood or in concurrent blood Pb samples is based on consistent evidence for decreased IQ across multiple studies and in well-accepted pooled analyses (e.g., Lanphear et al. 2005). Multiple studies (e.g., Baghurst et al. 1992, Bellinger et al. 1992, Min et al. 2009) reported that early-childhood (2-4 years of age) Pb exposure is associated with IQ score in children at later ages. Clear evidence that early- childhood exposure is associated with decreased IQ at later ages is complicated by the high degree of correlation in childhood blood Pb levels over time (e.g., see Dietrich et al. 1993a, Lanphear et al. 2005)" (NTP 2012, p.27).

Exposure to lead also affects other measures of cognitive function. "The conclusion of *sufficient* evidence for decreases in specific measures of cognitive function in children 3 months to 16 years of age with blood Pb levels <5 μg/dL measured in concurrent blood or in early childhood is based on the consistency of effects on multiple measures of cognitive function in multiple studies. Multiple studies (e.g., Min et al. 2009) reported that early-childhood Pb exposure is associated with decreases in cognitive function observed at later ages. However, as discussed for IQ, clear evidence that early-childhood or prenatal Pb exposure is associated with decreases in specific indices of cognitive function at later ages is complicated by the high degree of correlation in childhood blood Pb levels over time (e.g., see Dietrich et al. 1993a, Lanphear et al. 2005)." (NTP 2012, p.31). Lead is associated with brain development. Children who were exposed to lead see a reduction in gray matter in adult (Cecil et al. 2008).

Lead exposure impacts behavior as well (Reyes 2007, 2015)."The conclusion of *sufficient* evidence for a positive association with attention-related behaviors in children at blood Pb levels <5 μg/dL is based on the consistency of effects in these studies and supports effects down to and below 2 μg/dL blood Pb. This conclusion is for an association with attention-related behaviors rather than ADHD alone (…). '[A]ttention-related behaviors' is a more inclusive term that more accurately reflects the support for a range of Pb-associated behavioral changes in the area of attention, of which ADHD is one example on the more severe end of the spectrum of effects. (…) Recent studies found effects on attention-related behaviors after controlling for a large number of confounders, such as socioeconomic variables, sex, race/ethnicity, age of blood Pb measurement, parental education, and tobacco exposure. Several studies reported that blood Pb levels were significantly associated with ADHD even after controlling for potential mediating effects of child IQ (e.g., Nigg et al. 2008, Nigg et al. 2010)." (NTP 2012, p.34)

3. Data

Lead in Topsoil

The data on lead concentration in topsoil are taken from the U.S. Geological Survey (Smith et al. 2013). The USGS survey in the 2000s was designed to study the concentration and spatial distribution of chemical elements and minerals in soils of the conterminous United States. The sampling sites (1 site per 1,600 km^2) were selected based on the generalized random tessellation stratified (GRTS) design, which produces a spatially balanced set of sampling points without adhering to a strict grid-based system. Soils samples were collected from topsoil (depth of 0 to 5 cm). For each sample we know the latitude and longitude where it was taken. SM Figure S1 provides a map of the 4,857 soil sampling sites in the conterminous United States.

To construct the county level data on topsoil lead concentration, we aggregated lead measurements by taking the average of all available lead samples within a county. As a result of this procedure, we may have more than one measurement for a county with a large area, but may not have information for a county with a small area. We end up with 2096 counties with a measurement, as displayed in Figure S2 in SM. To examine the effect of lead on cognitive ability we constructed an indicator variable for whether the lead concentration in a particular county is above or below the national median topsoil lead concentration, calculated using the sample for the whole nation.

As robustness checks, we explore alternative measures of lead exposure. First, we construct topsoil lead concentrations based on the soil sampling sites located in urban areas only. The sample drops considerably, but has enough observations to provide precise estimates of the impacts of interest. Second, we leverage lead emissions from the transportation sector from 1950-1982 as estimated by Mielke et al. 2011. We generate lead emissions per square mile for the 90 urbanized areas available in their dataset, and merge with our county-level dataset.

Cognitive Ability Data

To study the effect of lead on cognitive ability, we use data from the U.S. Census of Population 2000, which inquired whether individuals as young as 5 years old experienced any "cognitive difficulties." This is a binary variable taking the value one whenever the respondent reports cognitive difficulties such as learning, remembering, concentrating, or making decisions because of a physical, mental, or emotional condition. The questionnaire stipulated that the causative health condition must have persisted for at least six months. The sample is restricted to 5-year-old children because we would like to investigate the lead effects before formal education starts, and because it is within the age range described by the NTP (2012) as having the highest rate of soil lead ingestion due to hand-to-mouth behavior (ages 1-5 years). Because the Integrated Public Use Microdata Series (IPUMS) provides county identifiers only for counties with 100,000 population or more due to confidentiality issues, we end up with 257 counties, representing about 45 percent of the U.S. population. These counties are displayed in Figure 2.

Additional Data

In our regression analysis we also deal with observable confounding factors by including state fixed effects and a variety of control variables associated with climatic, economic, demographic, housing, and other county and children's attributes. County Characteristics include average temperature, degree days below 10ºC, degree days above 29ºC, precipitation, latitude, longitude, share of white people, percentage of foreign people, share of people with completed high school, share of people with completed college, share of people below 5 years old, share of people above 65 years old, per capita income and its square, total employment and its square, unemployment rate, percentage of people below the poverty level, number of housing units, median number of rooms per house, share of Democratic votes in the presidential election, nonattainment status for any EPA criteria pollutant. Child and Household Characteristics are child's education status (indicator for attending nursery school to grade 4), race (9 groups), and Hispanic origin (5 categories), and education attainment of the head of the household (12 categories), household poverty status, household income, number of siblings in household, number of rooms in the house, time period when the house was built (9 groups). Climate variables come from the PRISM Climate Data. Economic, demographic, and housing controls come from the U.S. Censuses and the Bureau of Economic Analysis (BEA). Nonattainment status for any criteria pollutant comes from the EPA Green Book (www.epa.gov/green-book).

4. Empirical Strategy

To estimate the *causal* effect of lead in topsoil on cognitive ability, we adopt an instrumental variable approach (e.g., Bowden and Turkington 1984, Angrist et al. 1996, Greenland 2000, Angrist and Krueger 2001, Hernan and Robins 2006, Martens et al. 2006). The equation of interest is

$$Y_{ic} = \alpha + \beta SoilLead_c + X_{ic}\gamma + \eta_s + \varepsilon_{ic}, \quad (1)$$

where Y_{ic} is a binary variable taking the value one if the child *i* living in county *c* is reported to have experienced any cognitive difficulties. $SoilLead_c$ is a binary variable indicating whether soil lead concentration in county *c* is above the national median of topsoil lead levels, X_{ic} represents various county geographical, demographic, economic controls, as well as child and household characteristics such as race, mother's education, household income, house characteristics etc. (the complete list of control variables is described in the table notes), η_s represents state fixed effects, and ε is an error term. We adopt a single pollutant approach to examine the impact of topsoil lead on cognition because, as noted by Dominici et al. (2010), "the results of any regression model become highly unstable when incorporating two or more pollutants that are highly correlated". Nevertheless, the continental-scale USGS soil geochemical survey of the 2000s also collected information on other hazardous chemicals such as cadmium, mercury, and nickel. If a subset of these additional chemicals also affects cognitive ability, such as potentially cadmium (e.g., Ciesielski et al. 2012, Gustin et al. 2018, Lee et al. 2018), *SoilLead* may represent a sufficient statistic of exposure to contaminated soil. To the extent that some of the chemicals, such as cadmium and zinc, may also be added to soils adjacent to roads – the sources being tires and lubricant oils (e.g., Lagerwerff and Specht 1970, Turer et al. 2001, Wuana and Okieimen 2011) – our instrumental variable might capture variation on them as well. Table S1 in SM shows the positive and somewhat large correlation between lead in soil and other chemicals. Tables S2 and S3 in SM show that lead is a more important predictor of cognitive difficulty than any other chemicals.

Our coefficient of interest is β. As explained previously, because there may be important omitted factors affecting the outcome variable that are correlated with *SoilLead*, such as avoidance behavior (Currie et al. 2014), it is likely that $\hat{\beta}_{OLS}$ is biased and inconsistent. In particular, if households avoid exposure more often when lead concentration in soil is higher, and avoidance is negatively related to cognitive difficulty, then the bias should be negative, and $\hat{\beta}_{OLS}$ underestimated. In the extreme, $\hat{\beta}_{OLS}$ could have the "wrong" (negative) sign. In addition, exposure to topsoil lead might be measured with error because of the potential disconnection between where it is measured and where children live, leading to attenuation bias in the OLS estimate.

Instead of directly observing (and controlling for) defensive responses in the estimation of the causal effect of lead on cognition, the strategy pursued in this study is to use instruments that shift lead levels through a well-defined channel but are plausibly unrelated to avoidance behavior and measurement error. We estimate equation (1) using an instrumental variable strategy, using the 1944 Interstate Highway System Plan (see Figure 3) as an instrument for *SoilLead*. By affecting the location of the major highways built with the funds earmarked by the Federal Aid Highway Act of 1956 (Baum-Snow 2007, Michaels 2008), the 1944 plan generates variation in how much lead from gasoline was deposited and historically accumulated in the topsoil. The first stage equation is:

$$SoilLead_c = \theta + \lambda HWPlan1944_c + X_{ic}\delta + \eta_s + v_{ic}, \quad (2)$$

where *HWPlan1944* is an indicator for whether a county was recommended to receive a highway as part of the 1944 Interstate Highway System Plan. The other variables are defined as before, with v as the error term for this equation.

This is an intent-to-treat (ITT) strategy that addresses the unobserved association between lead in soil and avoidance responses. In our ITT approach, we isolate the portion of the cross-sectional variation in lead in topsoil that is related only to the highways that were built following exactly the 1944 plan. This variation should be unrelated to voters' preferences: the design of the 1944 plan was not supposed to reflect local preferences, but rather address primarily national security issues. Therefore, our instrument should satisfy both the relevance condition and the exclusion restriction.[4]

[4] Because our IV strategy relies solely on highway placement to avoid the endogeneity bias from cross-county sorting based on household preferences for environmental quality, but it has been over half a century since the 1944 interstate plan, one could argue that there was ample time for households to re-sort in response to the plan, threatening the plausibility of our exclusion restriction. Three pieces of evidence, however, may support our

As discussed before, measurement error in exposure to topsoil lead concentration might also generate attenuation bias in the OLS estimates of the impact of lead exposure on cognitive ability. To make this point clear, consider the simplified regression model:

$$Y_{ic} = \alpha + \beta SoilLead^{*}_{c} + u_{ic},$$

where *i* and *c* stand for individual and county, respectively. *Y* is an indicator variable for cognitive difficulty as before, $SoilLead^{*}$ is the accurate exposure to lead in topsoil, and *u* is the usual error term. Suppose $SoilLead^{*}$ is not observed, but we do observe *SoilLead*, a mismeasured level of exposure to topsoil lead. Again, the most important mismeasurement does not arise from the soil analysis in the lab, but rather the disconnection between where the soil samples were collected and where the children live.

To fix ideas, let $SoilLead = SoilLead^{*} + e$, where *e* is classical measurement error (i.e., uncorrelated to $SoilLead^{*}$ and *u*). By running a regression of *Y* on the observed *SoilLead* instead of the unobserved $SoilLead^{*}$, the OLS estimator for β will be downward biased. In fact,

$$\hat{\beta} \;=\; \underline{Cov(Y,SoilLead)} \;=\; \underline{Cov(\alpha + \beta SoilLead^{*} + u,\, SoilLead^{*} + e)}$$

research design. First, as pointed out by Michaels (2008), the construction of the Interstate Highway System began immediately after funding was approved in 1956, and by 1975 the system was mostly complete, spanning over 40,000 miles. Consistent with that timing, Baum-Snow (2007) finds that highways led to decreases in central city population and increases in suburban populations between 1950 and 1990, but with most of the suburbanization happening between 1970 and 1980 (see Table 1 in Baum-Snow 2007). Second, although the Clean Air Act Amendments of 1970 may have raised awareness of a number of criteria pollutants in the early years of the 1970s, a NAAQS for lead was only set by EPA in 1978. Moreover, the enforcement of those standards was only effective with the publication of the county nonattainment status for lead in 1991, and only 11 counties were out of attainment in the whole country (EPA 1991). By that time, lead in the air had already been reduced considerably (see Figure 1), and lead in gasoline was close to zero. It is unlikely that households were still relocating based on the highway plan and airborne lead concentration in the early 1990s. Third, early USGS surveys of soil composition were primarily conducted to help the development of agriculture and energy resources (Shacklette and Boerngen 1984). Therefore, although relocation occurred extensively from mid-century to 1990, it is unlikely that lead was among the main reasons. In fact, Boustan (2010) provides causal evidence that most of the post-World War II suburbanization was a "white flight", that is, a response of white households to the influx of southern blacks into northern U.S. cities during the Great Migration. On a different note, Couture and Handbury (2017) document a striking reversal in the fortunes of urban America since 2000, when college-educated population started flocking near city centers in most large U.S. cities. They find that changing preferences of young college graduates for non-tradable service amenities such as restaurants, bars, gyms, and personal services account for more than 50 percent of their growth near city centers. In fact, Edlund et al. (2015) point out that in 1980, housing prices in the main U.S. cities rose with distance to the city center, but by 2010 that relationship had reversed. They argue, however, that it might have been the shrinking leisure of high-income households that propelled centrality to the top of the local amenities list. Greater labor supply of high-income households reduced tolerance for commuting. Thus, it is likely that the unobserved forces underlying sorting patterns in the second half of the 20th century reversed in the 2000s.

$$Var(SoilLead) \qquad\qquad Var(SoilLead^{*} + e)$$

$$= \beta \frac{Var(SoilLead^{*})}{Var(SoilLead^{*}) + Var(e)} \leq \beta.$$

As explained previously, the instrumental variable approach addresses the measurement error issue as well. Indeed, with an instrument *Z* (1944 planned interstate highways) for *SoilLead*, the IV estimator for β will be consistent:

$$\hat{\beta}_{IV} = \frac{Cov(Y,Z)}{Cov(SoilLead,Z)} = \frac{Cov(\alpha + \beta SoilLead^{*} + \varepsilon, Z)}{Cov(SoilLead^{*} + e, Z)}$$

$$= \frac{\beta Cov(SoilLead^{*},Z) + Cov(\varepsilon,Z)}{Cov(SoilLead^{*},Z) + Cov(e,Z)} = \beta \frac{Cov(SoilLead^{*},Z)}{Cov(SoilLead^{*},Z)} = \beta,$$

where $Cov(\varepsilon,Z) = Cov(e,Z) = 0$ provided the instrument is uncorrelated with the regression error term ε and the measurement error *e*.

5. Results

Summary Statistics

Table 1 shows the summary statistics for the main variables used in our analysis. All variables are weighted by the individual weight provided by the U.S. Census 2000. About 3.5 percent of 5-year-old boys were reported to have experienced any cognitive difficulties that persisted for at least six months. For 5-year-old girls, only 1.8 percent were reported to have experienced any cognitive difficulties.

For counties below the national median, the average topsoil lead concentration is approximately 19.47 ppm (median 16.25, standard deviation 11.71). Therefore, going from below to above the national median means doubling the topsoil lead concentration. Notice that these concentrations are considerably below the regulatory guidance values suggested by EPA. “For residential scenarios, OSWER [The Office of Solid Waste and Emergency Response] has established 400 ppm as the screening level for lead in soil” (EPA 2014, p.4).

Topsoil Lead and Cognitive Difficulty

Table 2 shows the effect of lead in topsoil on cognitive difficulties among 5-year-old boys. Cognitive difficulty is defined as any difficulty that persisted for at least six months such as learning, remembering, concentrating, or making decisions, because of physical, mental, or emotional conditions. Column 1 includes state fixed effects and county level characteristics, such as climatic, demographic and economic variables, housing stock information and other county specific variables that may affect the cognitive development of a child. Column 2 also controls for child and household characteristics, such as race, child's parents' educational attainment, household income, etc. Panel A reports the OLS estimates, which are positive, but small and not statistically significant. Panel B presents the first stage relationship between our instrument and topsoil lead. The indicators for Highway Plan 1944 are positive and statistically significant. The coefficient in column 2 is 0.1689. This means that having a highway recommended by the 1944 plan increases the probability of experiencing lead concentrations in topsoil above the national median by 16.89 percent on average. The first stage F-statistics for the excluded instrument are all above the rule-of-thumb 10, suggesting a strong instrumental variable. Panel C reports the results estimated by IV. The IV coefficients on topsoil lead are positive, statistically significant, and stable across specifications. If lead concentration in counties with lead concentration above the median were to decrease to the levels in counties with lead concentration below the median, the probability of experiencing the cognitive difficulties would decrease by approximately 4 percentage points. This is a sizeable effect given that only about 3.5 percent of boys are reported to experience cognitive difficulties in the population represented by our sample.

The IV estimates are much larger than the associational relationship estimated by OLS. This is consistent with the presence of the household avoidance behavior and/or measurement error associated with the potential disconnection between where topsoil lead is measured and where children live, as discussed in the introduction. As a result, the health problems which are caused by topsoil lead are actually more severe than one might think just by looking at the OLS results.

Table 3 presents the results for the subsample of counties in which lead concentration is below EPA guidance values for intervention, providing evidence that the effects we estimate are not driven by a small number of counties with high topsoil lead concentration. Specifically, column 1 repeats the estimates from Table 2 (column 2) for comparison. Column 2 restricts attention to counties with topsoil lead concentration below 100ppm, column 3 reports the results for the subsample of counties with lead concentration below 80ppm, and, finally, column 4 presents the results for the subsample of counties with topsoil lead concentration below 50ppm. As in the previous tables, the OLS estimates are positive, but small and not statistically different from zero. IV estimates, however, all statistically significant, large in economic sense, and similar in magnitudes across different subsamples. Living in counties with topsoil lead concentration above the national median increases the probability of 5-year-old boys having cognitive difficulties by about 3-4 percentage points.

Table 4 explores the alternative measures of lead exposure. For reference, column 1 reports estimates from the main specification from Table 1 (column 2). Column 2 restricts the sample to counties with topsoil measures taken only in residential areas. The estimated effect of lead appears larger in this subsample, but as mentioned before the number of observations is quite small. The last column shows the results if lead emissions from the transportation sector from 1950-1982 per square mile is used to measure lead exposure. The estimated effect is positive and somewhat precise. If lead emission in counties with lead emission above the median were to decrease to the levels in counties with lead emission below the median, the probability of experiencing the cognitive difficulties would decrease by approximately 3 percentage points.

Table 5 reports the effect of lead in topsoil on other difficulties among 5-year-old children. Column 1 repeats the effect for "cognitive difficulty" for comparison. Column 2 presents the effects of topsoil lead on "self-care difficulty", which indicates whether respondents have any *physical* or *mental* health condition that has lasted for at least 6 months and made it difficult for them to take care of their own personal needs, such as bathing, dressing, or getting around inside the home. This does not include temporary health conditions, such as broken bones. Column 3 shows the effect of topsoil lead on "ambulatory difficulty", which indicates whether the respondent has a condition that substantially limits one or more basic *physical* activities, such as walking, climbing stairs, reaching, lifting, or carrying. And, finally, column 4 reports the effect of lead on "vision or hearing difficulty", which indicates whether the respondent has a long-lasting condition of blindness, deafness, or a severe vision or hearing impairment. Panel A reports the OLS estimates and Panel B presents the IV estimates.

Three patterns emerge from this table. First, there is a positive causal relationship between lead concentration in soil and cognitive and self-care difficulties as estimated in the Panel B in specifications 5 and 6 by the instrumental variables approach. Second, the OLS estimates in panel

A are much smaller than the corresponding IV estimates, suggesting the presence of important confounding factors and/or measurement error, which are known to bias the OLS estimates. Third, there is no evidence of causal relationship between lead concentration in soil and ambulatory and vision/hearing difficulties as estimated in Panel B specifications 7 and 8. Broadly, Table 5 shows that there is a causal relationship between lead in topsoil and mental child's development, but not physical development.

According the NTP review, "[t]here is *sufficient* evidence that blood Pb levels <10 μg/dL in children are associated with decreased auditory acuity. Multiple cross-sectional studies reported hearing loss, as indicated by higher hearing thresholds and increased latency of brainstem auditory evoked potentials (BAEPs), in children with blood Pb levels <10 μg/dL" (NTP 2012, p. xxi). Our empirical evidence does not provide support for that finding. Perhaps this arises from the definition of our measure, which combines the hearing issues with the vision problems. This would, indeed, lead to an imprecise estimate.

Table 6 explores the difference in the effect of lead in topsoil on cognitive difficulties between 5-year-old boys and 5-year-old girls. As before, the OLS estimates are presented in Panel A, the IV estimates are reported in Panel B. No effects are found for girls: the IV estimates for girls are much smaller than for boys and not statistically significant. The last column estimates the effect of topsoil lead exposure for both girls and boys altogether. There is a statistically significant difference between the effects on girls versus boys. Even though the F-stat is noisier when including both genders together, the effect for boys are more than four times larger than the effect for girls.

The results are consistent with previous findings from the epidemiological and experimental studies about gender-related differences in susceptibility to chemical exposure, which indicate that exposure to lead affects males more than females. Boys are found to be more susceptible to the negative effects of lead exposure than girls for executive function and reading-based skills, which is consistent with the hypothesis that lead is more deleterious in the frontal areas of the brain as compared to other areas (Khanna 2015). The gender difference in lead susceptibility is also present for low lead levels exposure due to the differences in the various neurotransmitter systems, as suggested by Jedrychowski et al. (2009). By reviewing the literature on the effect of chemicals on health, Llop et al. (2013) concludes that there is a gender-specific health effect of lead, and exposure to this heavy metal seems to affect boys more than girls.

Table 7 explores the role of a potential remediation mechanism: pre-school attendance. For reference, column 1 reports estimates from the main specification from Table 1 (column 2). Column 2 explores the differential effect of lead on children attending pre-school versus children who are not in pre-school. In column 2 the estimated effect is almost three times larger for children not in school. This suggests that pre-school may offset some of the effects of lead exposure on cognitive function, consistent with recent evidence found by Billings and Schnepel (2018). Nevertheless, these estimates should be interpreted with caution because preschool attendance in our setting is a result of parental decisions rather than random assignment.

6. Discussion

The findings from this paper strengthen our understanding of the adverse effects of lead exposure on child's cognitive development. Our estimates suggest that the adverse *causal* effect of lead exposure is large in economic sense. Living in counties with topsoil lead concentration above the national median increases the probability of 5-year-old boys having cognitive difficulties by about 3-4 percentage points.

We also provide new evidence of the damaging effect of lead on cognitive development even in areas with relatively low lead concentration. Previous work by Mielke et al. (1999) and Mielke et al. (2016) suggests that in order to reduce the children's blood level below 10 mg/dL, the soil lead standard should be around 80 ppm. Similarly, to reduce the blood lead level below 5 mg/dL, the lead soil standard should be around 40 ppm. The current guidance from the EPA regarding the management practices for gardening in lead contaminated areas (EPA 2014) considers lead concentration in soil below 100ppm as low risk category with no need for specific action of remediation. Our results indicate the need for further monitoring of urban soils, and potential remediation measures.

A number of regulatory actions were taken to decrease the exposure: use of lead in paint was banned in 1978, lead was banned in plumbing fixtures in 1986, the U.S. manufactures stopped using lead solder in 1991, and, in 1996 the U.S. Clean Air Act banned the sale of leaded fuel for use in on-road vehicles (Kovarik 2005). However, soil is still contaminated, and little has been done to clean it. The estimates of the effect of top soil lead exposure on cognitive development estimated in this paper are important, suggesting that the issue of lead exposure is still present, lead may continue to impair cognition today, both in the United States and in other countries that have significant amounts of lead in topsoil, so further actions may be needed to decrease the topsoil lead concentration.

This study makes important contributions to the literature and policymaking. It adds to the literature on the cognitive development (Sánchez (2017), Chen (2016), Fiorini and Keane (2014), Autor et al. (2014), Spears (2012)) by looking at the effect of an overlooked environmental factor: exposure to lead in topsoil. Our results strengthen the *associational* evidence on the neurodevelopmental effects of lead exposure in children (Lanphear et al. 2005, NTP 2012, AAP 2016, Geier et al. 2017, Delgado et al. 2018, Lee et al. 2018), and provides further evidence that those detrimental effects on cognitive function are *causal* (Rau et al. 2015, Sauve-Syed 2017, Aizer et al. 2018, Billings and Schnepel 2018, Gronqvist et al. 2018). Its main finding is concerning: it indicates that lead may continue to impair cognition today, both in the United States and in other countries that have significant amounts of lead in topsoil (Chen et al. 2015, Tóth et al. 2016). Also, given the low levels of topsoil lead observed in our sample relative to North American regulatory guidance values (Jennings and Petersen 2006, Jennings 2013), and the nonlinearities in the dose-response relationship between soil and blood lead levels (Mielke et al. 2007, Zahran et al. 2011), it corroborates the conclusions of previous studies that there may be room to lower the EPA threshold that calls for further monitoring of urban soils and remediation (Mielke et al. 1999, Zahran et al. 2011, Mielke 2016, Laidlaw et al. 2017). Cost-effective remediation options seem to be available (Martin and Ruby 2004, Wuana and Okieimen 2011, Laidlaw et al. 2017).

As alluded to before, a key advantage of our study is examining the adverse impact of lead exposure in a context broader than previous studies, covering all large metropolitan areas in the United States. Nevertheless, an important limitation of our analysis is that we do not observe children's blood lead levels to estimate the direct relationship between blood lead levels and cognition. Instead, we rely on an indirect measure that has been shown to have a clear relationship with blood lead levels, namely, topsoil lead concentration.

7. Conclusion

This study presents causal evidence on the relationship between topsoil lead exposure and cognitive development among 5-year-old children. Using individual level data on preschool-aged children from the U.S. Census 2000 and an instrumental variable approach leveraging the 1944 Interstate Highway System Plan, we find that higher lead in topsoil increases considerably the probability of 5-year-old boys experiencing cognitive difficulties such as learning, remembering, concentrating, or making decisions. Living in counties with topsoil lead concentration above the national median increases the probability of 5-year-old boys having cognitive difficulties by 4 percentage points. This harmful effect does not seem to extend to 5-year-old girls, potentially due to the natural protection of estrogen. Importantly, the adverse effect of lead exposure on boys are found even in counties with topsoil lead concentration below the EPA guidance values for intervention, highlighting the need for a revision of those guidelines. Overall, the main finding of this study is concerning because it suggests that lead may continue to impair cognition today, both in the United States and in other countries that have lead deposition in topsoil.

References

Aizer, Anna, Janet Currie, Peter Simon, and Patrick Vivier. (2018). "Do Low Levels of Blood Lead Reduce Children's Future Test Scores?" *American Economic Journal: Applied Economics* 10(1): 307-41.

American Academy of Pediatrics (AAP). (2016). "Prevention of Childhood Lead Toxicity," *Pediatrics* 38(1): e20161493 (Policy Statement – Council on Environmental Health).

Angrist, Joshua, D., and Alan B. Krueger. (2001). "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments," *Journal of Economic Perspectives* 15(4): 69-85.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin. (1996). "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association* 91(434): 444-455.

David Autor, David Figlio, Krzysztof Karbownik, Jeffrey Roth, Melanie Wasserman. (2016). School Quality and the Gender Gap in Educational Achievement. American Economic Review 106:5, 289-295

Baghurst, Peter A., Anthony J. McMichael, Neil R. Wigg, Graham V. Vimpani, Evelyn F. Robertson, Russell J. Roberts, and Shi-Lu Tong. (1992). "Environmental Exposure to Lead and Children's Intelligence at the Age of Seven Years – The Port Pirie Cohort Study," *New England Journal of Medicine* 327(18): 1279-1284.

Banzhaf, H. Spencer, and Randall P. Walsh. (2008). "Do People Vote with Their Feet? An Empirical Test of Tiebout," *American Economic Review* 98(3): 843-863.

Baum-Snow, Nathaniel. (2007). "Did Highways Cause Suburbanization?" *Quarterly Journal of Economics* 122(2): 775-805.

Bellinger, David C., Alan Leviton, Christine Waternaux, Herbert Needleman, and Michael Rabinowitz. (1987). "Longitudinal Analyses of Prenatal and Postnatal Lead Exposure and Early Cognitive Development," *New England Journal of Medicine* 316(17): 1037-1043.

Billings, Stephen B., and Kevin T. Schnepel. (2018). "Life after Lead: Effects of Early Interventions for Children Exposed to Lead," *American Economic Journal: Applied Economics* 10(3): 315-44.

Boustan, Leah P. (2010). "Was Postwar Suburbanization "White Flight"? Evidence from the Black Migration," *Quarterly Journal of Economics* 125(1): 417-443.

Bowden, Roger J., and Darrell A. Turkington. (1984). *Instrumental Variables*. Cambridge, U.K.: Cambridge University Press.

Byers, Randolph K., and Elizabeth E. Lord. (1943). "Late Effects of Lead Poisoning on Mental Development," *American Journal of Diseases of Children* 66(5): 471-494.

Cecil KM, Brubaker CJ, Adler CM, Dietrich KN, Altaye M, Egelhoff JC, et al. (2008) Decreased Brain Volume in Adults with Childhood Lead Exposure. PLoS Med 5(5): e112. https://doi.org/10.1371/journal.pmed.0050112

Centers for Disease Control and Prevention (CDC). (2017). "Lead Toxicity: What Are Possible Health Effects from Lead Exposure?" Available at https://www.atsdr.cdc.gov/csem/csem.asp?csem=34&po=10. Accessed on April 29, 2018.

Chay, Kenneth Y., and Michael Greenstone. (2005). "Does Air Quality Matter? Evidence from the Housing Market," *Journal of Political Economy* 113(2): 376-424.

Chamberlain, A.C., M.J. Heard, P. Little, D. Newton, A.C. Wells, and R.D. Wiffen. (1978). "Investigations into Lead from Motor Vehicles". Harwell, United Kingdom: United Kingdom Atomic Energy Authority. Report no. AERE-9198.

Chen, Haiyang, Yanguo Teng, Sijin Lu, Yeyao Wang, and Jinsheng Wang. (2015). "Contamination Features and Health Risk of Soil Heavy Metals in China," *Science of The Total Environment* 512-513: 143-153.

Chen, Wen-Yi. (2016). "On the Relationship Between Economic Conditions Around the Time of Birth and Late Life Cognitive Abilities: Evidence from Taiwan," *Economics and Human Biology* 22: 126-139.

Ciesielski, Timothy, Jennifer Weuve, David C. Bellinger, Joel Schwartz, Bruce P. Lanphear, and Robert O. Wright. (2012). "Cadmium Exposure and Neurodevelopmental Outcomes in U.S. Children," *Environmental Health Perspectives* 120(5): 758-763.

Clay, Karen, Margarita Portnykh, and Edson Severnini. (2018). "Toxic Truth: Lead and Fertility," *NBER Working Paper* # 24607.

Couture, Victor, and Jessie Handbury. (2017). "Urban Revival in America, 2000 to 2010," *NBER Working Paper #24084*.

Currie, Janet, Joshua Graff Zivin, Jamie Mullins, and Matthew Neidell. (2014). "What Do We Know About Short- and Long-Term Effects of Early-Life Exposure to Pollution?" *Annual Review of Resource Economics* 6(1): 217-247.

Currie, Janet, Lucas W. Davis, Michael Greenstone, and W. Reed Walker. (2015). "Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings," *American Economic Review* 105(2): 678-709.

Datko-Williams, Laura, Adrien Wilkie, Jennifer Richmond-Bryant. (2014). "Analysis of U.S. Soil Lead (Pb) Studies from 1970 to 2012," *Science of the Total Environment* 468-469: 854-863.

Delgado, Christine F., Mary Anne Ullery, Melissa Jordan, Chris Duclos, Sudha Rajagopalan, Keith Scott. (2018). "Lead Exposure and Developmental Disabilities in Preschool-Aged Children," *Journal of Public Health Management and Practice* 24(2): e10-e17.

Diamond, Rebecca. (2016). "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000," *American Economic Review* 106(3): 479-524.

Dominici, Francesca, Roger D. Peng, Christopher D. Barr, and Michelle L. Bell. (2010). "Protecting Human Health from Air Pollution: Shifting from a Single-Pollutant to a Multi-Pollutant Approach," *Epidemiology* 21(2): 187-194.

Dominici, Francesca, Michael Greenstone, and Cass R. Sunstein. (2014). "Particulate Matter Matters," *Science* 344(6181): 257-259.

Dominici, Francesca, and Corwin Zigler. (2017). "Best Practices for Gauging Evidence of Causality in Air Pollution Epidemiology," *American Journal of Epidemiology* 186(12): 1303-1309.

Dudka, Stanislaw, and Domy C. Adriano. (1997). "Environmental Impacts of Metal Ore Mining and Processing: A Review," *Journal of Environmental Quality* 26(3): 590-602.

Edlund, Lena, Cecilia Machado, and Maria Micaela Sviatschi. (2015). "Bright Minds, Big Rent: Gentrification and the Rising Returns to Skill," *NBER Working Paper # 21729*.

Ethyl Corp. (Statement of the Ethyl Corporation). (1984). Airborne Lead Reduction Act of 1984 (S. 2609) – A Bill to Amend the Clean Air Act with Regard to Mobile Source Emission Control. *Hearings before the Committee on Environment and Public Works*, U.S. Senate, 98th Congress, 2nd Session, June 22.

Ferrie, Joseph, Karen Rolf, and Werner Troesken. (2012). "Cognitive Disparities, Lead Plumbing, and Water Chemistry: Prior Exposure to Water-borne Lead and Intelligence Test Scores Among World War Two Army Enlistees," *Economics and Human Biology* 10: 98-111.

Filippelli, Gabriel M., Martin Risch, Mark A.S. Laidlaw, Deborah E. Nichols, and Julie Crewe. (2015). "Geochemical Legacies and the Future Health of Cities: A Tale of Two Neurotoxins in Urban Soils," *Elementa: Science of the Anthropocene* 3:59.

Fiorini , Mario, and Michael P. Keane. (2014). "How the Allocation of Children's Time Affects Cognitive and Noncognitive Development," *Journal of Labor Economics* 32(4): 787-836.

Geier, David A., Janet K. Kern, and Mark R. Geier. (2017). "Blood Lead Levels and Learning Disabilities: A Cross-Sectional Study of the 2003–2004 National Health and Nutrition Examination Survey (NHANES)," *International Journal of Environmental Research and Public Health* 14(10): 1202.

Glass, Thomas A., Steven N. Goodman, Miguel A. Hernán, and Jonathan M. Samet. (2013) "Causal Inference in Public Health," *Annual Review of Public Health* 34: 61-75.

Greenland, Sander. (2000). "An Introduction to Instrumental Variables for Epidemiologists," *International Journal of Epidemiology* 29(4): 722-729.

Griffin, T.B.; Coulston, R.; Wills, H.; Russel, J.C.; and Knelson, J.H. (1975) "Clinical Studies on Men Continuously Exposed to Airborne Particulate Lead" In: Griffin T.B., Knelson J.G., eds. *Lead*. Stuttgart, West Germany: Georg Thieme Publisher, 221-240.

Gronqvist, Hans, J. Peter Nilsson, and Per-Olof Robling. (2018). "Early Lead Exposure and Outcomes in Adulthood," *Mimeo*.

Gulson, Brian, Karen Mizon, Michael Korsch, Alan Taylor. (2016). "Revisiting mobilisation of skeletal lead during pregnancy based on monthly sampling and cord/maternal blood lead relationships confirm placental transfer of lead," *Archives of Toxicology* 90: 805–816 DOI 10.1007/s00204-015-1515-8

Gustin, Klara, Fahmida Tofail, Marie Vahter, and Maria Kippler. (2018). "Cadmium Exposure and Cognitive Abilities and Behavior at 10 Years of Age: A Prospective Cohort Study," *Environment International* 113: 259-268.

Harris, Allison R., and Cliff I. Davidson. (2005). "The Role of Resuspended Soil in Lead Flows in the California South Coast Air Basin," *Environmental Science & Technology* 39(19): 7410-7415.

Hernán, Miguel A., and James M. Robins. (2006). "Instruments for Causal Inference: An Epidemiologist's Dream?" *Epidemiology* 17(4): 360-372.

Holmgren, G.G.S., M.W. Meyer, R.L. Chaney, and R.B. Daniels. (1993). "Cadmium, Lead, Zinc, Copper, and Nickel in Agricultural Soils of the United States of America," *Journal of Environmental Quality* 22(2): 335-348.

Hunt, Andrew, David L. Johnson, and Daniel A. Griffith. (2006). "Mass Transfer of Soil Indoors by Track-in on Footwear," *Science of the Total Environment* 370(2-3: 360-371.

Jedrychowski, Wieslaw, Frederica Perera, Jeffery Jankowski, Dorota Mrozek-Budzyn, Elzbieta Mroz, Elzbieta Flak, Susan Edwards, Anita Skarupa, and Ilona Lisowska-Miszczyk. (2009). "Gender Specific Differences in Neurodevelopmental Effects of Prenatal Exposure to Very Low-

Jennings, Aaron A., and Elijah J. Petersen. (2006). "Variability of North American Regulatory Guidance for Heavy Metal Contamination of Residential Soil," *Journal of Environmental Engineering and Science* 5(6): 485-508.

Jennings, Aaron A. (2013). “Analysis of Worldwide Regulatory Guidance Values for the Most Commonly Regulated Elemental Surface Soil Contamination,” *Journal of Environmental Management* 118: 72-95.

Khanna, Maya M. (2015). “Boys, Not Girls, Are Negatively Affected on Cognitive Tasks by Lead Exposure: A Pilot Study,” *Journal of Environmental Health* 77(6): 72-77.

Kovarik, William (2005) “Ethyl-leaded gasoline: how a classic occupational disease became an international public health disaster,” Int. J. Occup. Environ. Health 11: 384–397.

Lagerwerff, John V., and A.W. Specht. (1970). “Contamination of Roadside Soil and Vegetation with Cadmium, Nickel, Lead, and Zinc,” *Environmental Science & Technology* 4(7): 583-586.

Laidlaw, Mark A.S., and Gabriel M. Filippelli. (2008). “Resuspension of Urban Soils as a Persistent Source of Lead Poisoning in Children: A Review and New Directions,” *Applied Geochemistry* 23(8): 2021-2039.

Laidlaw, Mark A.S., Sammy Zahran, Howard W. Mielke, Mark P. Taylor, and Gabriel M. Filippelli. (2012). “Re-suspension of Lead Contaminated Urban Soil as a Dominant Source of Atmospheric Lead in Birmingham, Chicago, Detroit and Pittsburgh, USA,” *Atmospheric Environment* 49: 302-310.

Laidlaw, Mark A.S., Gabriel M. Filippelli, Richard C. Sadler, Christopher R. Gonzales, Andrew S. Ball, and Howard W. Mielke. (2016). “Children’s Blood Lead Seasonality in Flint, Michigan (USA), and Soil-Sourced Lead Hazard Risks,” *International Journal of Environmental Research and Public Health* 13(4): 358.

Laidlaw, Mark A.S., Gabriel M. Filippelli, Sally Brown, Jorge Paz-Ferreiro, Suzie M. Reichman, Pacian Netherway, Adam Truskewycz, Andrew S. Ball, and Howard W. Mielke. (2017). “Case Studies and Evidence-Based Approaches to Addressing Urban Soil Lead Contamination,” *Applied Geochemistry* 83: 14-30.

Lanphear, Bruce P., David A. Burgoon, Steven W. Rust, Shirley Eberly, and Warren Galke. (1998). "Environmental Exposures to Lead and Urban Children's Blood Lead Levels," *Environmental Research* 76(2): 120-130.

Lanphear, Bruce P., Richard Hornung, Jane Khoury, Kimberly Yolton, Peter Baghurst, David C. Bellinger, Richard L. Canfield, Kim N. Dietrich, Robert Bornschein, Tom Greene, Stephen J. Rothenberg, Herbert L. Needleman, Lourdes Schnaas, Gail Wasserman, Joseph Graziano, and Russell Roberts. (2005). "Low-Level Environmental Lead Exposure and Children's Intellectual Function: An International Pooled Analysis," *Environmental Health Perspectives* 113(7): 894-899.

Lee, Min-Jing, Miao-Chun Chou, Wen-Jiun Chou, Chien-Wei Huang, Ho-Chang Kuo, Sheng-Yu Lee, and Liang-Jen Wang. (2018). "Heavy Metals' Effect on Susceptibility to Attention-Deficit/Hyperactivity Disorder: Implication of Lead, Cadmium, and Antimony," *International Journal of Environmental Research and Public Health* 15(6): 1221.

Levin, Ronnie, Mary Jean Brown, Michael E. Kashtock, David E. Jacobs, Elizabeth A. Whelan, Joanne Rodman, Michael R. Schock, Alma Padilla, and Thomas Sinks. (2008). "Lead Exposures in U.S. Children, 2008: Implications for Prevention," *Environmental Health Perspectives* 116(10): 1285-1293.

Llop, Sabrina, Maria-Jose Lopez-Espinosa, Marisa Rebagliato, and Ferran Ballester. (2013). "Gender Differences in the Neurotoxicity of Metals in Children," *Toxicology* 311(1-2): 3-12.

Martens, Edwin P., Wiebe R. Pestman, Anthonius de Boer, Svetlana V. Belitser, and Olaf H. Klungel. (2006). "Instrumental Variables: Application and Limitations," *Epidemiology* 17(3): 260-267.

Martin, Todd A., and Michael V. Ruby. (2004). "Review of In Situ Remediation Technologies for Lead, Zinc, and Cadmium in Soil," *Remediation* 14(3): 35-53.

McMichael, Anthony J., Peter A. Baghurst, Neil R. Wigg, Graham V. Vimpani, Evelyn F. Robertson, and Russell J. Roberts. (1988). "Port Pirie Cohort Study: Environmental Exposure to Lead and Children's Abilities at the Age of Four Years," *New England Journal of Medicine* 319(8): 468-475.

Mendelsohn, Robert, and Sheila Olmstead. (2009). "The Economic Valuation of Environmental Amenities and Disamenities: Methods and Applications," *Annual Review of Environment and Resources* 34: 325-347.

Michaels, Guy. (2008). "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System," *Review of Economics and Statistics* 90(4): 683-701.

Mielke, Howard W. (1991). "Lead in Residential Soils: Background and Preliminary Results of New Orleans," *Water, Air, & Soil Pollution* 57-58(1): 111-119.

Mielke, Howard W. (1993). "Lead Dust Contaminated USA Communities: Comparison of Louisiana and Minnesota," *Applied Geochemistry* 8(Supplement 2): 257-261.

Mielke, Howard W. (1999). "Lead in the Inner Cities: Policies to Reduce Children's Exposure to Lead may be Overlooking a Major Source of Lead in the Environment," *American Scientist* 87(1): 62-73.

Mielke, Howard W. (2016). "Nature and Extent of Metal-Contaminated Soils in Urban Environments (Keynote Talk)," *Environmental Geochemistry and Health* 38(4): 987-999.

Mielke, Howard W., Jana C. Anderson, Kenneth J. Berry, Paul W. Mielke, Rufus L. Chaney, and Meredith Leech. (1983). "Lead Concentrations in Inner-city Soils as a Factor in the Child Lead Problem," *American Journal of Public Health* 73(12): 1366-1369.

Mielke, Howard W., Chris R. Gonzales, M.K. Smith, and Paul W. Mielke Jr. (1999). “The Urban Environment and Children's Health: Soils as an Integrator of Lead, Zinc, and Cadmium in New Orleans, Louisiana, U.S.A.” *Environmental Research* 81(2): 117-129.

Mielke, Howard W., Chris R. Gonzales, Eric Powell, Morten Jartun, and Paul W. Mielke Jr. (2007). “Nonlinear Association Between Soil Lead and Blood Lead of Children in Metropolitan New Orleans, Louisiana: 2000‑2005,” *Science of The Total Environment* 388(1-3): 43-53.

Mielke, Howard W., Mark A.S. Laidlaw, and Chris R. Gonzales. (2011). “Estimation of Leaded (Pb) Gasoline’s Continuing Material and Health Impacts on 90 US Urbanized Areas,” *Environment International* 37(1): 248-257.

Mielke, Howard W., Mark A.S. Laidlaw, and Chris R. Gonzales. (2010). “Lead (Pb) Legacy from Vehicle Traffic in Eight California Urbanized Areas: Continuing Influence of Lead Dust on Children’s Health,” *Science of the Total Environment* 408(19): 3965-3975.

Mielke, Howard W., Mark A.S. Laidlaw, and Chris R. Gonzales. (2011). “Estimation of Leaded (Pb) Gasoline’s Continuing Material and Health Impacts on 90 US Urbanized Areas,” *Environment International* 37(1): 248-257.

National Toxicology Program (NTP). (2012). “NTP Monograph on Health Effects of Low-Level Lead,” U.S. Department of Health and Human Services, NTP Monograph. 1. xiii, xv-148.

Needleman, Herbert L., Charles Gunnoe, Alan Leviton, Robert Reed, Henry Peresie, Cornelius Maher, and Peter Barrett. (1979). “Deficits in Psychologic and Classroom Performance of Children with Elevated Dentine Lead Levels,” *New England Journal of Medicine* 300(13): 689-695.

Needleman, Herbert L., Alan Schell, David Bellinger, Alan Leviton, and Elizabeth N. Allred. (1990). “The Long-Term Effects of Exposure to Low Doses of Lead in Childhood – An 11-Year Follow-up Report,” *New England Journal of Medicine* 322(2): 83-88.

Newell, Richard G., and Kristian Rogers. (2003). “The U.S. Experience with the Phasedown of Lead in Gasoline,” *Resources for the Future Discussion Paper*.

Rabinowitz, Michael, George W. Wetherill, and Joel D. Kopple. (1976). “Kinetic Analysis of Lead Metabolism in Healthy Humans,” *Journal of Clinical Investigation* 58(2): 260-270.

Rau, Tomás, Sergio Urzúa, and Loreto Reyes. (2015). “Early Exposure to Hazardous Waste and Academic Achievement: Evidence from a Case of Environmental Negligence,” *Journal of the Association of Environmental and Resource Economists* 2(4): 527-63.

Reyes, Jessica Wolpaw. (2007). “Environmental Policy as Social Policy? The Impact of Childhood Lead Exposure on Crime,” *B.E. Journal of Economic Analysis and Policy (Contributions)* 7(1): Article 51.

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

Sánchez, Alan. (2017). “The Structural Relationship Between Early Nutrition, Cognitive Skills and Non-cognitive Skills in Four Developing Countries,” *Economics and Human Biology* 27(Part A): 33-54.

Sauve-Syed, Jessica. (2017). “Lead Exposure and Student Performance: A Study of Flint Schools,” *Mimeo*.

Schwarz, Kirsten, Richard V. Pouyat, and Ian Yesilonis. (2016). “Legacies of Lead in Charm City’s Soil: Lessons from the Baltimore Ecosystem Study,” *International Journal of Environmental Research and Public Health* 13(2): 209.

Shacklette, Hansford T., and Josephine G. Boerngen. (1984). “Element Concentrations in Soils and Other Surficial Materials of the Conterminous United States,” *U.S. Geological Survey Professional Paper* 1270.

Sheets, Ralph W., Joseph R. Kyger, Richard N. Biagioni, Shelly Probst, Ron Boyer, Karl Barke. (2001). "Relationship Between Soil Lead and Airborne Lead Concentrations at Springfield, Missouri, USA," *Science of the Total Environment* 271(1-3): 79-85.

Smith, David B., William F. Cannon, Laurel G. Woodruff, Federico Solano, James E. Kilburn, and David L. Fey. (2013). "Geochemical and Mineralogical Data for Soils of the Conterminous United States," *U.S. Geological Survey Data Series 801*, 19 p., http://pubs.usgs.gov/ds/801/.

Spears, Dean. (2012). "Height and Cognitive Achievement Among Indian Children," *Economics and Human Biology* 10: 210-219.

Sutherland, Ross A. (2000). "Depth Variation in Copper, Lead, and Zinc Concentrations and Mass Enrichment Ratios in Soils of an Urban Watershed," *Journal of Environmental Quality* 29(5): 1414-1422.

Tóth, Gergely, Tamás Hermann, Gábor Szatmári, and László Pásztor. (2016). "Maps of Heavy Metals in the Soils of the European Union and Proposed Priority Areas for Detailed Assessment," *Science of The Total Environment* 565: 1054-1062.

Turer, Dilek, J. Barry Maynard, and J. John Sansalone. (2001). "Heavy Metal Contamination in Soils of Urban Highways Comparison Between Runoff and Soil Concentrations at Cincinnati, Ohio," *Water, Air, and Soil Pollution* 132(3-4): 293-314.

U.S. Department of Health and Human Services. (2007). Toxicological Profile for Lead. Available at https://www.atsdr.cdc.gov/toxprofiles/tp13.pdf, accessed on April 20, 2018.

U.S. Environmental Protection Agency (EPA). (1994). "Revised Interim Soil Lead Guidance for CERCLA Sites and RCRA Corrective Action Facilities," Office of Solid Waste and Emergency Response, Washington, D.C. Directive 9355.4-12.

U.S. Environmental Protection Agency (EPA). (2006). "Air Quality Criteria for Lead (Final): Volume I and II," Office of Research and Development, National Center for Environmental Assessment-RTP Division, EPA/600/R-5/144aF.

U.S. Environmental Protection Agency (EPA). (2014). "Technical Review Workgroup Recommendations Regarding Gardening and Reducing Exposure to Lead-Contaminated Soils," Office of Solid Waste and Emergency Response, Washington, D.C. Directive 9200.2-142.

Wang, Meie, and Haizhen Zhang. (2018). "Accumulation of Heavy Metals in Roadside Soil in Urban Area and the Related Impacting Factors," *International Journal of Environmental Research and Public Health* 15(6): 1064.

Wuana, Raymond A., and Felix E. Okieimen. (2011). "Heavy Metals in Contaminated Soils: A Review of Sources, Chemistry, Risks and Best Available Strategies for Remediation," *ISRN Ecology*: 1-20.

Yassoglou, N., C. Kosmas, J. Asimakopoulos, and C. Kallianou. (1987). "Heavy Metal Contamination of Roadside Soils in the Greater Athens Area," *Environmental Pollution* 47(4): 293-304.

Young, Thomas M., Deo A. Heerama, Gorkem Sirin, and Lowell L. Ashbaugh. (2002). "Resuspension of Soil as a Source of Airborne Lead near Industrial Facilities and Highways," *Environmental Science & Technology* 36(11): 2484-2490.

Zahran, Sammy, Howard W. Mielke, Stephan Weiler, and Christopher R. Gonzales. (2011). "Nonlinear Associations Between Blood Lead in Children, Age of Child, and Quantity of Soil Lead in Metropolitan New Orleans," *Science of The Total Environment* 409(7): 1211-1218.

Zahran, Sammy, Mark A. S. Laidlaw, Shawn P. McElmurry, Gabriel M. Filippelli, and Mark Taylor. (2013a). "Linking Source and Effect: Resuspended Soil Lead, Air Lead, and Children's

Blood Lead Levels in Detroit, Michigan," *Environmental Science & Technology* 47(6): 2839-2845.

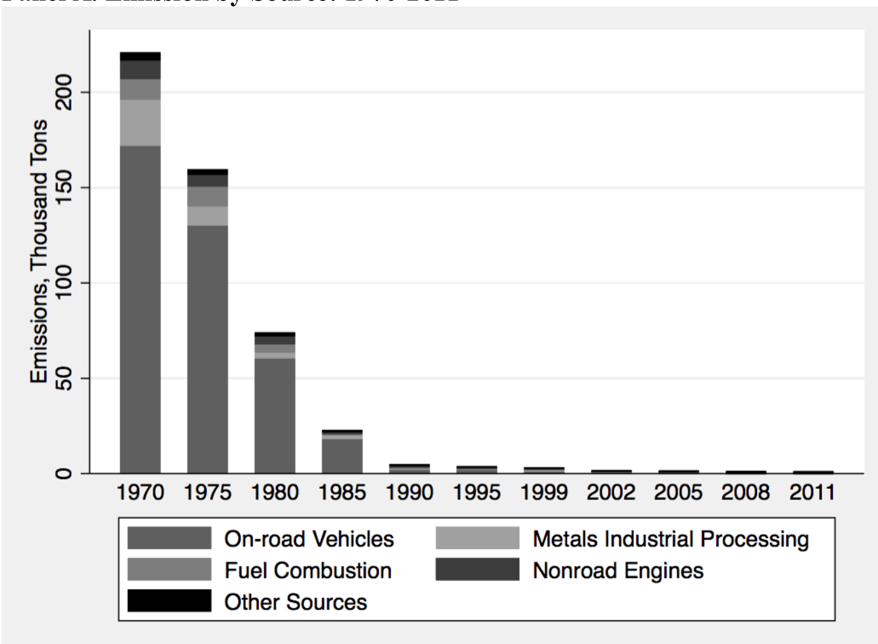
Zahran, Sammy, Howard W. Mielke, Shawn P. McElmurry, Gabriel M. Filippelli, Mark A.S. Laidlaw, and Mark P. Taylor. (2013b). "Determining the Relative Importance of Soil Sample Locations to Predict Risk of Child Lead Exposure," *Environment International* 60: 7-14.

Zahran, Sammy, Terrence Iverson, Shawn P. McElmurry, and Stephan Weiler. (2017). "The Effect of Leaded Aviation Gasoline on Blood Lead in Children," *Journal of the Association of Environmental and Resource Economists* 4(2): 575-610.

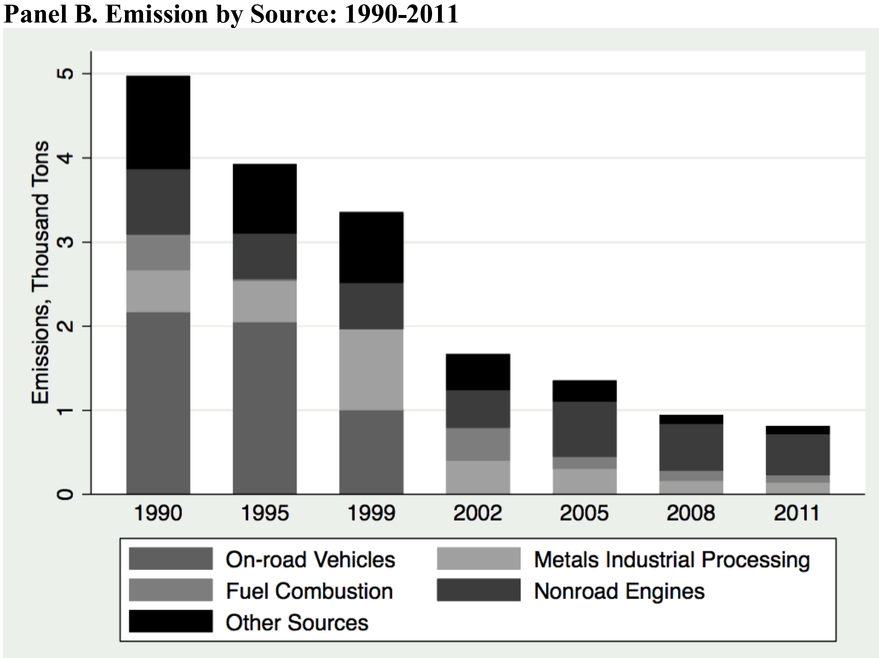
Figures and Tables

Figure 1 – Anthropogenic Lead Emissions in the U.S. by Source Category, 1970-2011

Panel A. Emission by Source: 1970-2011

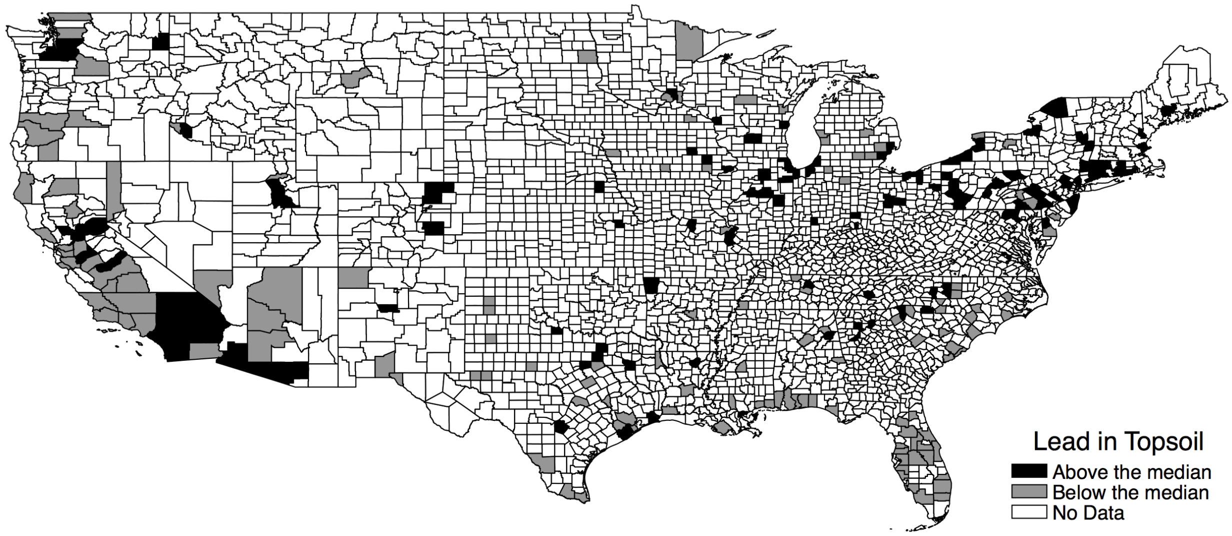


Panel B. Emission by Source: 1990-2011



Notes: Data are taken from the U.S. EPA, 2014. Emissions inventory data presented for years that allow reliable estimation of long-term trends. Changes shown reported for 1970-2011 include both emissions changes and methods changes. While the trends displayed in the figure are generally representative, actual changes from year to year could have been larger or smaller than those reported here.

Figure 2 – Topsoil Lead: Counties in Our Sample



Notes: This figure shows the lead concentration (mg/kg) in topsoil, at a depth of 0-5 cm for the counties in our sample. We have a total of 252 counties. Data are taken from U.S. Geological Survey. Darker color represents the lead concentration above the median, lighter color represents the lead concentration below the median.

Figure 3 – Routes of the Recommended Interregional Highway System: "1944 Plan"

Notes: This figure shows the 1944 Interstate Highway System Plan Map (Michaels 2008). In 1941, President Roosevelt appointed a National Interregional Highway Committee to design a interregional highway system addressing three policy goals (Michaels, 2008): (i) to improve the connection between major metropolitan areas in the U.S., (ii) to serve U.S. national defense, and (iii) to connect with major routes in Canada and Mexico. Congress acted on these recommendations in the Federal-Aid Highway Act of 1944. In our analysis, we refer to the plan recommended by that committee as the "1944 plan".

Table 1 – Summary Statistics

Variables	Obs.	Mean	Std. Dev.	Min.	Max.
Panel A. 5-Year-Old Boys					
Cognitive difficulty	39,463	0.0347	0.1829	0	1
Self-care difficulty	39,463	0.0131	0.1136	0	1
Ambulatory difficulty	39,463	0.0100	0.0994	0	1
Vision or hearing difficulty	39,463	0.0089	0.0939	0	1
Topsoil Lead Concentration (ppm)	39,463	28.855	18.480	1.70	138.05
Panel B. 5-Year-Old Girls					
Cognitive difficulty	37,643	0.0178	0.1322	0	1
Self-care difficulty	37,643	0.0087	0.0926	0	1
Ambulatory difficulty	37,643	0.0077	0.0873	0	1
Vision or hearing difficulty	37,643	0.0077	0.0876	0	1
Topsoil Lead Concentration (ppm)	37,643	29.077	18.783	1.70	138.05

Notes: Table presents the summary statistics for the main variables used in our analysis. All variables are weighted by the individual weight provided by the U.S. Census 2000.

Table 2 – Topsoil Lead and Cognitive Difficulty in 5-year-old Boys

Panel A. Cognitive Difficulty - OLS		
Variables	(1)	(2)
Topsoil Lead	0.0026	0.0029
	(0.004)	(0.004)
R-squared	0.0037	0.0114
Panel B. First Stage - Topsoil Lead		
Variables	(3)	(4)
HWPlan1944	0.1697***	0.1689***
	(0.051)	(0.051)
R-squared	0.6196	0.6202
First Stage F Stat	11.03	10.79
Underidentification Test (P-value)	0.0159	0.0168
Panel C. Cognitive Difficulty - IV		
Variables	(5)	(6)
Topsoil Lead	0.0442**	0.0414**
	(0.019)	(0.019)
State Fixed Effects	x	x
County Characteristics	x	x
Child and Household Characteristics		x
Observations	39,463	39,463

Notes: This table presents the OLS, first stage, and IV estimates of the effect of topsoil lead on cognitive difficulty, defined as any difficulties such as learning, remembering, concentrating, or making decisions, because of physical, mental, or emotional conditions. The sample is restricted to 5-year-old boys from the Census 2000. Topsoil lead is the county-level average of all soil samples collected by the USGS in the 2000s. HWPlan1944 is an indicator for whether a county was recommended to receive a highway as part of the 1944 Interstate Highway System Map. County Characteristics include average temperature, degree days below 10C, degree days above 29C, precipitation, latitude, longitude, share of white people, percentage of foreign people, share of people with completed high school, share of people with completed college, share of people below 5 years old, share of people above 65 years old, per capita income and its square, total employment and its square, unemployment rate, percentage of people below the poverty level, number of housing units, median number of rooms per house, share of Democratic votes in the presidential election, nonattainment status for any EPA criteria pollutant. Child and Household Characteristics are child's education status (indicator for attending nursery school to grade 4), race (9 groups), and Hispanic origin (5 categories), and education attainment of the head of the household (12 categories), household poverty status, household income, number of siblings in household, number of rooms in the house, time period when the house was built (9 groups). Regressions are weighted by person weight provided by the Census Bureau. Standard errors are clustered at the state level and are in parentheses. The null hypothesis of underidentification test is that the excluded instruments are not "relevant", i.e., not correlated with the endogenous regressors. A rejection of the null indicates that the model is identified. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 3 – Topsoil Lead and Cognitive Difficulty: Sample Restrictions

Dep. Var.: Cognitive Difficulty	Main Results: All Counties	Dropping 3 counties with Topsoil Lead >100ppm	Dropping 8 counties with Topsoil Lead >80ppm	Dropping 24 counties with Topsoil Lead >50ppm
Panel A. Other Samples - OLS	(1)	(2)	(3)	(4)
Variables				
Topsoil Lead	0.0029	0.0032	0.0031	0.0030
	(0.004)	(0.004)	(0.004)	(0.004)
R-squared	0.0114	0.0117	0.0117	0.0118
Panel B. Other Samples - IV	(5)	(6)	(7)	(8)
Variables				
Topsoil Lead	0.0414**	0.0390**	0.0440**	0.0302*
	(0.019)	(0.018)	(0.021)	(0.018)
State Fixed Effects	x	x	x	x
County Characteristics	x	x	x	x
Child and Household Characteristics	x	x	x	x
Observations	39,463	39,008	38,684	35,261
First Stage F Stat	10.79	9.718	7.659	6.411
Underidentification Test (P-value)	0.0168	0.0179	0.0218	0.0292

Notes: This table presents the effect of topsoil lead on cognitive difficulty, defined as any difficulties such as learning, remembering, concentrating, or making decisions, because of physical, mental, or emotional conditions, if counties with high lead concentration are dropped from the analysis. Topsoil lead is the county-level average of all soil samples collected by the USGS in the 2000s. HWPlan1944 is an indicator for whether a county was recommended to receive a highway as part of the 1944 Interstate Highway System Map. Column 1 repeats the estimates from the Table 2 column 2 for comparison. Column 2 presents the effects if three counties with Topsoil Lead >100ppm are dropped: Chester, PA, Jefferosn, TX, and St. Louis, MO. Column 3 presents the estimates if the following eight counties with Topsoil Lead >80ppm are dropped: Chester, PA, Jefferson, TX, St. Louis, MO, Rockland, NY, Westmoreland, PA, La Crosse, WI, Schuylkill, PA, and La Porte, IN. Column 4 restricts the sample further by dropping the counties with Topsoil Lead >50ppm. There are 24 such counties. All specifications include controls for County Characteristics: average temperature, degree days below 10C, degree days above 29C, precipitation, latitude, longitude, share of white people, percentage of foreign people, share of people with completed high school, share of people with completed college, share of people below 5 years old, share of people above 65 years old, per capita income and its square, total employment and its square, unemployment rate, percentage of people below the poverty level, number of housing units, median number of rooms per house, share of Democratic votes in the presidential election, nonattainment status for any EPA criteria pollutant and Child and Household Characteristics: child's education status (indicator for attending nursery school to grade 4), race (9 groups), and Hispanic origin (5 categories), and education attainment of the head of the household (12 categories), household poverty status, household income, number of siblings in household, number of rooms in the house, time period when the house was built (9 groups). Regressions are weighted by person weight provided by the Census Bureau. Standard errors are clustered at the state level and are in parentheses. The null hypothesis of underidentification test is that the excluded instruments are not "relevant", i.e., not correlated with the endogenous regressors. A rejection of the null indicates that the model is identified. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 4 – Topsoil Lead and Cognitive Difficulty: Alternative Measures

Panel A. Cognitive Difficulty - OLS			
Variables	(1)	(2)	(3)
Topsoil Lead	0.0029		
	(0.004)		
Urban Topsoil Lead		0.0228	
		(0.050)	
Vehicle Lead Emissions			0.0045
			(0.004)
R-squared	0.0114	0.0193	0.0118
Panel B. Cognitive Difficulty - IV			
Variables	(4)	(5)	(6)
Topsoil Lead	0.0414**		
	(0.019)		
Urban Topsoil Lead		0.0778**	
		(0.038)	
Vehicle Lead Emissions			0.0289*
			(0.015)
State Fixed Effects	x	x	x
County Characteristics	x	x	x
Child and Household Characteristics	x	x	x
First Stage F Stat	10.79	9.117	6.612
Underidentification Test (P-value)	0.0168	0.217	0.159
Observations	39,463	3,262	28,092

Notes: This table presents the OLS, first stage, and IV estimates of the effect of topsoil lead on cognitive difficulty, defined as any difficulties such as learning, remembering, concentrating, or making decisions, because of physical, mental, or emotional conditions. The sample is restricted to 5-year-old boys from the Census 2000. Topsoil lead in columns 1 is the county-level average of all soil samples collected by the USGS in the 2000s. In column 2 restricts attention to more urban counties with topsoil lead measurements taken in the developed areas. Column 3 uses vehicles lead emissions as a measure of lead exposure. In columns 1 and 3 County Characteristics include average temperature, degree days below 10C, degree days above 29C, precipitation, latitude, longitude, share of white people, percentage of foreign people, share of people with completed high school, share of people with completed college, share of people below 5 years old, share of people above 65 years old, per capita income and its square, total employment and its square, unemployment rate, percentage of people below the poverty level, number of housing units, median number of rooms per house, share of Democratic votes in the presidential election, nonattainment status for any EPA criteria pollutant. In columns 2 County Characteristics include degree days below 10C, degree days above 29C, share of white people, share of people with completed high school, per capita income and its square, total employment and its square, nonattainment status for any EPA criteria pollutant. Child and Household Characteristics in all columns are child's education status (indicator for attending nursery school to grade 4), race (9 groups), and Hispanic origin (5 categories), and education attainment of the head of the household (12 categories), household poverty status, household income, number of siblings in household, number of rooms in the house, time period when the house was built (9 groups). Regressions are weighted by person weight provided by the Census Bureau. Standard errors are clustered at the state level and are in parentheses. The null hypothesis of underidentification test is that the excluded instruments are not "relevant", i.e., not correlated with the endogenous regressors. A rejection of the null indicates that the model is identified. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 5 – Topsoil Lead and Other Difficulties in 5-year-old Boys

Panel A. Other Difficulties - OLS	**Cognitive**	**Self-care**	**Ambulatory**	**Vision/Hearing**
Variables	(1)	(2)	(3)	(4)
Topsoil Lead	0.0029	0.0038*	0.0024	0.0000
	(0.004)	(0.002)	(0.002)	(0.001)
R-squared	0.0114	0.0067	0.0060	0.0067
Panel B. Other Difficulties - IV	**Cognitive**	**Self-care**	**Ambulatory**	**Vision/Hearing**
Variables	(5)	(6)	(7)	(8)
Topsoil Lead	0.0414**	0.0177*	0.0069	-0.0078
	(0.019)	(0.011)	(0.009)	(0.012)
State Fixed Effects	x	x	x	x
County Characteristics	x	x	x	x
Child and Household Characteristics	x	x	x	x
Observations	39,463	39,463	39,463	39,463
First Stage F Stat	10.79	10.79	10.79	10.79
Underidentification Test (P-value)	0.0168	0.0168	0.0168	0.0168

Notes: This table presents the OLS and the IV estimates of the effect of topsoil lead on cognitive, self-care, ambulatory and vising/hearing difficulties. The sample is restricted to 5-year-old boys from the Census 2000. Topsoil lead is the county-level average of all soil samples collected by the USGS in the 2000s. HWPlan1944 is an indicator for whether a county was recommended to receive a highway as part of the 1944 Interstate Highway System Map. County Characteristics include average temperature, degree days below 10C, degree days above 29C, precipitation, latitude, longitude, share of white people, percentage of foreign people, share of people with completed high school, share of people with completed college, share of people below 5 years old, share of people above 65 years old, per capita income and its square, total employment and its square, unemployment rate, percentage of people below the poverty level, number of housing units, median number of rooms per house, share of Democratic votes in the presidential election, nonattainment status for any EPA criteria pollutant. Child and Household Characteristics are child's education status (indicator for attending nursery school to grade 4), race (9 groups), and Hispanic origin (5 categories), and education attainment of the head of the household (12 categories), household poverty status, household income, number of siblings in household, number of rooms in the house, time period when the house was built (9 groups). Regressions are weighted by person weight provided by the Census Bureau. Standard errors are clustered at the state level and are in parentheses. The null hypothesis of underidentification test is that the excluded instruments are not "relevant", i.e., not correlated with the endogenous regressors. A rejection of the null indicates that the model is identified. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 6 - Topsoil Lead and Cognitive Difficulty: Boys vs. Girls

Dep. Var.: Cognitive Difficulty	5-Year-Old Boys	5-Year-Old Girls	5-Year-Old Boys & Girls
Panel A. Other Children - OLS	(1)	(2)	(3)
Variables			
Topsoil Lead	0.0029	-0.0014	
	(0.004)	(0.003)	
Topsoil Lead x Boys			0.0092***
			(0.003)
Topsoil Lead x Girls			-0.0077***
			(0.002)
R-squared	0.0114	0.0094	0.0096
P-value (Boys = Girls)			0.0000
Panel B. Other Children - IV	(4)	(5)	(6)
Variables			
Topsoil Lead	0.0414**	0.0054	
	(0.019)	(0.010)	
Topsoil Lead x Boys			0.0403***
			(0.011)
Topsoil Lead x Girls			0.0084
			(0.011)
State Fixed Effects	x	x	x
County Characteristics	x	x	x
Child and Household Characteristics	x	x	x
Observations	39,463	37,643	77,106
First Stage F Stat	10.79	8.875	4.922
Underidentification Test (P-value)	0.0168	0.0258	0.0207
P-value (Boys = Girls)			0.0000

Notes: This table presents the OLS and the IV estimates of the effect of topsoil lead on cognitive difficulty, defined as any difficulties such as learning, remembering, concentrating, or making decisions, because of physical, mental, or emotional conditions. Topsoil lead is the county-level average of all soil samples collected by the USGS in the 2000s. HWPlan1944 is an indicator for whether a county was recommended to receive a highway as part of the 1944 Interstate Highway System Map. County Characteristics include average temperature, degree days below 10C, degree days above 29C, precipitation, latitude, longitude, share of white people, percentage of foreign people, share of people with completed high school, share of people with completed college, share of people below 5 years old, share of people above 65 years old, per capita income and its square, total employment and its square, unemployment rate, percentage of people below the poverty level, number of housing units, median number of rooms per house, share of Democratic votes in the presidential election, nonattainment status for any EPA criteria pollutant. Child and Household Characteristics are child's education status (indicator for attending nursery school to grade 4), race (9 groups), and Hispanic origin (5 categories), and education attainment of the head of the household (12 categories), household poverty status, household income, number of siblings in household, number of rooms in the house, time period when the house was built (9 groups). Regressions are weighted by person weight provided by the Census Bureau. Standard errors are clustered at the state level and are in parentheses. The null hypothesis of underidentification test is that the excluded instruments are not "relevant", i.e., not correlated with the endogenous regressors. A rejection of the null indicates that the model is identified. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Table 7 - Topsoil Lead and Cognitive Difficulty: Potential Remediation

Panel A. Cognitive Difficulty - OLS		
Variables	(1)	(2)
Topsoil Lead	0.0029	
	(0.004)	
Topsoil Lead x Not in School		0.0059
		(0.004)
Topsoil Lead x In School		0.0004
		(0.004)
R-squared	0.0114	0.0115
Panel B. Cognitive Difficulty - IV		
Variables	(3)	(4)
Topsoil Lead	0.0414**	
	(0.019)	
Topsoil Lead x Not in School		0.0664***
		(0.020)
Topsoil Lead x In School		0.0259
		(0.025)
State Fixed Effects	x	x
County Characteristics	x	x
Child and Household Characteristics	x	x
First Stage F Stat	10.79	3.217
Underidentification Test (P-value)	0.0168	0.0655
Observations	39,463	39,463

Notes: This table presents the OLS, first stage, and IV estimates of the effect of topsoil lead on cognitive difficulty, defined as any difficulties such as learning, remembering, concentrating, or making decisions, because of physical, mental, or emotional conditions. The sample is restricted to 5-year-old boys from the Census 2000. Topsoil lead in columns 1 and 2 is the county-level average of all soil samples collected by the USGS in the 2000s. HWPlan1944 is an indicator for whether a county was recommended to receive a highway as part of the 1944 Interstate Highway System Map. In column 3 restricts attention to more urban counties with topsoil lead measurements taken in the developed areas. Column 4 uses lead emissions as a measure of lead exposure. County Characteristics include average temperature, degree days below 10C, degree days above 29C, precipitation, latitude, longitude, share of white people, percentage of foreign people, share of people with completed high school, share of people with completed college, share of people below 5 years old, share of people above 65 years old, per capita income and its square, total employment and its square, unemployment rate, percentage of people below the poverty level, number of housing units, median number of rooms per house, share of Democratic votes in the presidential election, nonattainment status for any EPA criteria pollutant. Child and Household Characteristics are child's education status (indicator for attending nursery school to grade 4), race (9 groups), and Hispanic origin (5 categories), and education attainment of the head of the household (12 categories), household poverty status, household income, number of siblings in household, number of rooms in the house, time period when the house was built (9 groups). Regressions are weighted by person weight provided by the Census Bureau. Standard errors are clustered at the state level and are in parentheses. The null hypothesis of underidentification test is that the excluded instruments are not "relevant", i.e., not correlated with the endogenous regressors. A rejection of the null indicates that the model is identified. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

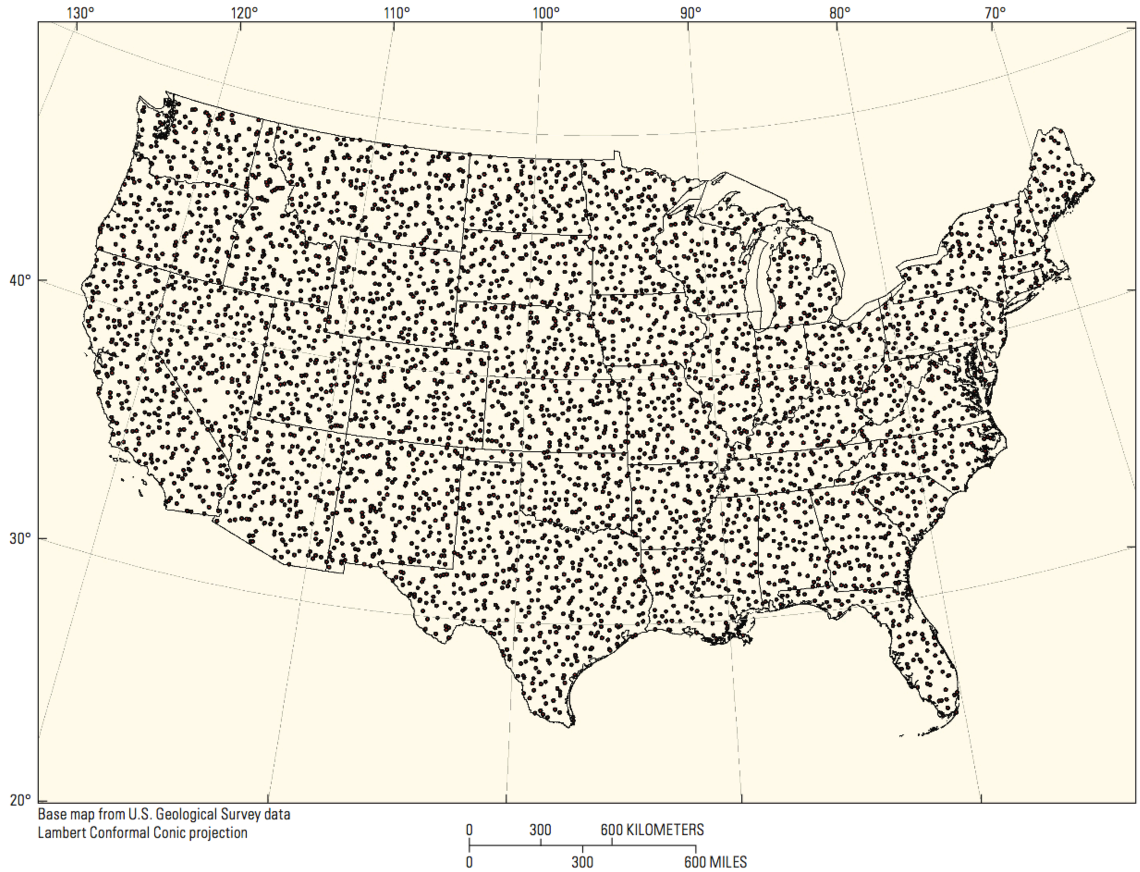
Supplemental Material

The Legacy Lead Deposition in Soils and Its Impact on Cognitive Function in Preschool-Aged Children in the U.S.

Karen Clay, Margarita Portnykh, Edson Severnini

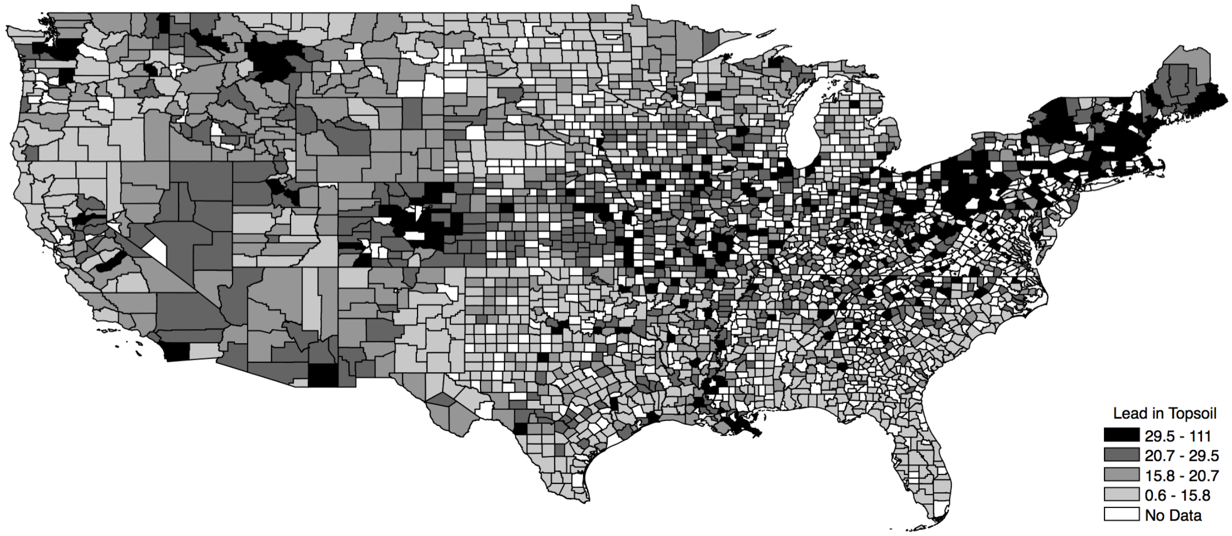
Table of Contents

Figure S1 – Soil Sampling Sites



Notes: This map shows the location of 4,857 soil sampling sites in the conterminous United States. Source: Smith, D.B., Cannon, W.F., Woodruff, L.G., Solano, Federico, Kilburn, J.E., and Fey, D.L., 2013, Geochemical and mineralogical data for soils of the conterminous United States: U.S. Geological Survey Data Series 801, 19 p., http://pubs.usgs.gov/ds/801/.

Figure S2– Lead in Topsoil (mg/kg) in the 2000s



Notes: This figure shows the lead concentration (mg/kg) in topsoil, at a depth of 0-5 cm. Data are taken from U.S. Geological. Lead concentrations in each county are constructing by taking the average of all available lead samples within a county. As a result, there are 2096 counties with a measurement.

Table S1 – Correlation between Hazardous Chemicals in Topsoil

	Lead	Cadmium	Zinc	Mercury
Lead	1			
Cadmium	0.44	1		
Zinc	0.56	0.43	1	
Mercury	0.33	0.08	0.09	1

Notes: This table shows the (weighted) correlation between different hazardous chemicals in topsoil in the 2000s. All variables are indicators for whether a county has the chemical concentration above the national median.

Table S2 – Cognitive Difficulty and Hazardous Chemicals: OLS

Variables	(1)	(2)	(3)	(4)
Lead	0.0042	0.0045	0.0048	0.0062
	(0.004)	(0.004)	(0.004)	(0.004)
Cadmium	-0.0039			-0.0021
	(0.002)			(0.003)
Zinc		-0.0064		-0.0044
		(0.005)		(0.006)
Mercury			-0.0049	-0.0038
			(0.003)	(0.004)
State Fixed Effects	x	x	x	x
County Characteristics	x	x	x	x
Child and Household Characteristics	x	x	x	x
Observations	39,463	39,463	39,463	39,463
R-squared	0.0115	0.0115	0.0115	0.0115

Notes: This table presents the OLS estimates of the topsoil chemicals: lead, cadmium, zinc, and mercury for cognitive difficulty, defined as any difficulties such as learning, remembering, concentrating, or making decisions, because of physical, mental, or emotional conditions. The sample is restricted to 5-year-old boys from the Census 2000. Topsoil lead is the county-level average of all soil samples collected by the USGS in the 2000s. HWPlan1944 is an indicator for whether a county was recommended to receive a highway as part of the 1944 Interstate Highway System Map. County Characteristics include average temperature, degree days below 10C, degree days above 29C, precipitation, latitude, longitude, share of white people, percent of foreign people, share of people with completed high school, share of people with completed college, share of people below 5 years old, share of people above 65 years old, per capita income and its square, total employment and its square, unemployment rate, percent of people below the poverty level, number of housing units, median number of rooms per house, share of Democratic votes in the presidential election, nonattainment status for any EPA criteria pollutant. Child and Household Characteristics are child's education status (indicator for attending nursery school to grade 4), race (9 groups), and hispanic origin (5 categories), and education attainment of the head of the household (12 categories), household poverty status, household income, number of siblings in household, number of rooms in the house, time period when the house was built (9 groups). Regressions are weighted by person weight provided by the Census Bureau. Standard errors are clustered at the state level and are in parentheses. The null hypothesis of underidentification test is that the excluded instruments are not "relevant", i.e., not correlated with the endogenous regressors. A rejection of the null indicates that the model is identified. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.

Table S3 – Cognitive Difficulty and Hazardous Chemicals: OLS and IV

Panel A. Cognitive Difficulty - OLS Variables	(1)	(2)	(3)	(4)
Lead	0.0029			
	(0.004)			
Cadmium		-0.0027		
		(0.003)		
Zinc			-0.0041	
			(0.004)	
Mercury				-0.0031
				(0.003)
R-squared	0.0114	0.0114	0.0114	0.0114
Panel B. Cognitive Difficulty - IV Variables	(5)	(6)	(7)	(8)
Lead	0.0414**			
	(0.019)			
Cadmium		0.0362**		
		(0.017)		
Zinc			0.0646	
			(0.040)	
Mercury				0.1459
				(0.221)
State Fixed Effects	x	x	x	x
County Characteristics	x	x	x	x
Child and Household Characteristics	x	x	x	x
Observations	39,463	39,463	39,463	39,463
First Stage F Stat	10.79	9.132	7.517	0.470
Underidentification Test (P-value)	0.0168	0.0116	0.0542	0.483

Notes: This table presents the OLS, first stage, and IV estimates of the effect of topsoil chemicals: lead, cadmium, zinc, ad mercity on cognitive difficulty, defined as any difficulties such as learning, remembering, concentrating, or making decisions, because of physical, mental, or emotional conditions. The sample is restricted to 5-year-old boys from the Census 2000. Topsoil lead is the county-level average of all soil samples collected by the USGS in the 2000s. HWPlan1944 is an indicator for whether a county was recommended to receive a highway as part of the 1944 Interstate Highway System Map. County Characteristics include average temperature, degree days below 10C, degree days above 29C, precipitation, latitude, longitude, share of white people, percent of foreign people, share of people with completed high school, share of people with completed college, share of people below 5 years old, share of people above 65 years old, per capita income and its square, total employment and its square, unemployment rate, percent of people below the poverty level, number of housing units, median number of rooms per house, share of Democratic votes in the presidential election, nonattainment status for any EPA criteria pollutant. Child and Household Characteristics are child's education status (indicator for attending nursery school to grade 4), race (9 groups), and hispanic origin (5 categories), and education attainment of the head of the household (12 categories), household poverty status, household income, number of siblings in household, number of rooms in the house, time period when the house was built (9 groups). Regressions are weighted by person weight provided by the Census Bureau. Standard errors are clustered at the state level and are in parentheses. The null hypothesis of underidentification test is that the excluded instruments are not "relevant", i.e., not correlated with the endogenous regressors. A rejection of the null indicates that the model is identified. ***, **, and * indicate statistical significance at the 1, 5, and 10 percent levels, respectively.