Airborne Lead Pollution and Infant Mortality*

Karen Clay (Carnegie Mellon University and NBER)
Edson Severnini (Carnegie Mellon University, NBER, and IZA)
Xiao Wang (Carnegie Mellon University)

June 2022

Abstract

This paper uses U.S. Toxic Release Inventory data for 1988-2018 on industrial air lead emissions to provide national IV estimates of the effects of air lead concentration on infant mortality. The causal effect of lead on infant mortality is identified by annual variation in air fugitive lead emissions and wind speed near reporting plants, which together determine local ambient lead concentration. Unlike stack emissions, which occur routinely and may be subject to avoidance behavior, fugitive emissions occur irregularly and unexpectedly. We find a positive and statistically significant relationship between air lead concentration and infant mortality. Estimates by race are imprecise but suggest that lead exposure may be disproportionately affecting nonwhite infants. Cause of death data show that lead increases deaths from low birthweight and sudden unexplained infant death. Back of the envelope estimates indicate that declines in fugitive lead emissions prevented 34-59 infant deaths per year, generating benefits of $313-$533 million annually.

JEL Codes: I120, Q530, Q580
Keywords: Infant Mortality, Airborne Lead Pollution, Toxic Release Inventory (TRI)

*We thank Teevrat Garg, Josh Graff-Zivin, Rhiannon Jerch, Catherine Maclean, and Louis Preonas for insightful suggestions, and seminar participants at Carnegie Mellon University, University of Maryland, Temple University, and University of California–San Diego, and conference participants at the ASSA Meetings, AERE Summer Conference, and Midwest Energy Fest for valuable comments. The authors gratefully acknowledge financial support from the Heinz College at Carnegie Mellon.

1 Introduction

Although air lead emissions and air lead concentration fell with the decline in lead in gasoline, air emissions of lead continue to impact millions in the US and globally. In 1995, global emissions of lead from fuel combustion, metal and cement production, and waste disposal and other sources were estimated to be 30 million tons (Pacyna and Pacyna, 2001). In the US in 1995, emissions from these sources were about 3,400 tons (EPA, 2018). Evidence suggests that emissions remain high globally. In China, lead emissions from these sources are estimated to have tripled from 2001 to 2009 due to increased coal consumption and lead battery production (Li et al., 2012). Further, regulations on lead emissions have caused some plants to relocate to countries with less stringent laws (Tanaka et al., 2021).

While the literature on lead and health is large, we know remarkably little about air lead concentration and infant mortality.[1] Although a small literature provides evidence on water lead and infant mortality, to our knowledge there are no studies that show a causal relationship between air lead and infant mortality.[2] The lack of evidence is surprising, given that the relationship between lead and infant mortality is a key input into regulatory cost-benefit analysis.

This paper provides national IV estimates of the effects of air lead concentration on infant mortality in the United States over the period 1988-2018. The start year is dictated by the creation of the Toxic Release Inventory. It is worth noting that by 1988, nearly all of the decline in lead in gasoline has occurred. To estimate the effects of emissions on infant mortality, we draw on a range of data, including from EPA monitor data on air lead concentration, PM10, and carbon monoxide, and Toxic Release Inventory data on emissions of air lead and other chemicals. Infant mortality, birth outcomes, and maternal characteristics are from the confidential National Vital Statistics data set. Additional data sets include information on wind, weather, and demographic characteristics of the counties.

[1]There is a large epidemiological literature on lead and health, much of which is reviewed in NTP (2012). For the economics literature on the topic, see the extensive discussion in Hollingsworth and Rudik (2021).

[2]On water lead and infant mortality, see Troesken (2008), Clay et al. (2014), and Edwards (2014).

Estimating the causal effect of lead on health is complicated for three possible reasons. First, the impact of airborne lead may be affected by avoidance behavior of individuals. This behavior may include reductions in outdoor activities and investments that improve indoor air quality (Neidell, 2009; Moretti and Neidell, 2011; Ito and Zhang, 2020). Avoidance behavior is likely to be particularly relevant for infant health, because pregnant women have strong incentives to engage in avoidance. Without considering avoidance behavior, the effect of lead on infant mortality may be underestimated. Second, the estimation is confounded by many factors that may also generate omitted variable bias, including emissions of other pollutants, weather conditions, and impacts of industrial activities on local employment and economic conditions (Ruhm, 2000; Currie and Schmieder, 2009; Agarwal et al., 2010; Heutel and Ruhm, 2013; Knittel et al., 2016). Third, measurement error due to the usual disconnection between where air lead is measured and where individuals reside may generate attenuation bias (Moretti and Neidell, 2011; Knittel et al., 2016; Deryugina et al., 2019).

Our identification overcomes these challenges by instrumenting for EPA air lead concentration with TRI fugitive lead emissions interacted with wind speed near the emitting plants and including a rich set of controls. Variation in fugitive lead emissions and wind speed is shown to determine lead exposure. While stack lead emissions occur routinely and may be subject to avoidance behavior, fugitive lead emissions are unintended and intermittent (EPA, 1994). Wind speed near emitting plants provides another layer of plausible exogeneity. The unpredictable annual variations in fugitive lead and wind speed make it difficult to engage in avoidance behavior. We provide evidence that fugitive lead emissions and wind speeds are strongly linked to air lead concentration readings of EPA monitors. Further, we provide evidence that fugitive lead emissions interacted with wind speed has low predictive power for county socioeconomic characteristics and mothers' characteristics.

We find a positive and statistically significant relationship between air lead concentration and infant mortality. Accounting for the timing of lead exposure in utero increases the magnitude of the estimates. Estimates by race are imprecise but suggest that lead exposure may be disproportionately affecting nonwhite infants. Cause of death data show that lead increases deaths from low birthweight and sudden unexplained infant death.[3] Back of the envelope estimates indicate that declines in fugitive lead emissions prevented 34-59 infant deaths per year, generating benefits of $313-$533 million annually. Our estimates imply that the full decline in air lead concentration from all sources prevented 218 infant deaths per year, generating benefits of $2.0 billion per year. This is likely a lower bound of the total health value of emission reduction given evidence that lead exposure is harmful to young children and other vulnerable groups (Klemick et al., 2020; Hollingsworth and Rudik, 2021).

This paper makes two main contributions to the literature on lead and health. First, it provides nationwide causal effects of airborne lead on infant mortality. Because the air lead concentration in our context arises from ongoing industrial sources, our estimates could be used in future regulatory impact analysis by EPA and other government agencies. Previous studies have reported historical effects of lead pipes on infant mortality (Troesken, 2008; Clay et al., 2014; Edwards, 2014), or effects of gasoline-driven airborne lead on elderly mortality across 75 US counties hosting NASCAR races (Hollingsworth and Rudik, 2021). Second, it contrasts the extensive margin (mortality) effects with the intensive margin effects of lead on infant health. Prior literature has shown detrimental effects of lead on the incidence of low birth weight or preterm birth (Dave and Yang, 2020; Bui et al., 2021; Tanaka et al., 2021), but not on infant mortality.

Our findings also add to the economics literature on the impacts of the overall TRI emissions on infant health and education (Currie and Schmieder, 2009; Agarwal et al., 2010; Persico, 2020; Persico and Venator, 2021), the effects of lead on human capital accumulation (Aizer et al., 2018; Billings and Schnepel, 2018; Aizer and Currie, 2019; Clay et al., 2019; Hollingsworth et al., 2020; Gronqvist et al., 2020; Gazze et al., 2021), and the effects of lead on fertility (Grossman and Slusky, 2019; Clay et al., 2021).

The rest of the paper proceeds as follows. Section 2 discusses the background information

[3] Sudden unexplained infant death includes sudden infant death syndrome (SIDS), accidental suffocation and strangulation on bed (ASSB), and other unexplained deaths.

on global use of lead, the Toxic Release Inventory, and the literature on lead and infant health. Section 3 describes our data. Section 4 describes our empirical strategy. Section 5 presents results, and Section 6 concludes.

2 Background

2.1 Toxic Release Inventory

The TRI was created by the Emergency Planning and Community Right-to-know Act (EPCRA) in 1986. The Toxic Release Inventory was a response to chemical releases in Bhopal in 1984 and in West Virginia in 1985. The EPCRA required plants meeting certain criteria to annually report their emissions to the EPA for public disclosure through the TRI beginning in 1987. Lead was included in the original set of chemicals and so plants reported emissions beginning in 1987.[4] The EPA brought enforcement actions for non-reporting (Marchi and Hamilton, 2006). Analyses have found that TRI reporting is generally accurate (Brehm and Hamilton, 1996; Natan and Miller, 1998; Marchi and Hamilton, 2006).

Plants separately report stack and fugitive emissions for each chemical, including lead. Stack emissions are all releases "to the air that occur through confined vents, ducts, pipes, or other confined air stream." Most plants use air pollution control devices to reduce stack emissions. These devices can include cooling towers, scrubbers, and bag houses that separate lead and other heavy metals from the exhaust. Lead collected by those devices may be recycled, transferred to offsite treatments, or emitted via wastewater or landfill, which generates substitution between air, water, and land emissions. The remaining lead in exhaust is emitted via stacks or other confined air streams.[5] Fugitive emissions are "all releases that are not released through confined vents, ducts, pipes, or other confined air stream." Some examples of fugitive air emissions include: i) leaks from operating machinery; ii) emissions

[4]Initially reporting was less accurate, particularly in 1987, which was the first reporting year. Thus we follow the literature and begin our analysis in 1988.

[5]Definitions are from * https://www.epa.gov/trinationalanalysis/air-releases.

from opening doors or panels of machinery; iii) air emissions as the result of spills; and iv) emission from the handling of ash.

Figure 1 shows the downward trend in stack and fugitive air lead emissions, Appendix Figure A.3 compares airborne with water and land emissions and recycled lead, and Appendix Figure A.1 shows emissions by industry.[6] Fugitive emissions are around one-third of air lead emissions, with stack emissions making up two-thirds. The vertical lines indicate changes to the database in 1998 and 2001, when seven industries were added and the reporting threshold was lowered.[7] Four industry groups – lead manufacturing, other metal manufacturing, ceramics manufacturing, and paint and pigment manufacturing – account for more than 90% of the total air lead emissions by all TRI-reporting facilities.[8]

2.2 Vectors of Lead Exposure

Although pregnant and nursing mothers and infants were exposed to lead through a number of vectors between 1988 and 2018, apart from industrial emissions these vectors did not vary or did so very slowly. For example, lead in soil and paint changed slowly or not at all. Lead in soil is a reflection of past deposition and local geology. Lead in paint is a function of the age of the housing stock. The federal government banned lead paint for housing in 1978, but some states had banned it earlier. Thus most housing with lead paint was built before 1960. Maternal and infant exposure tends to be through ingestion of soil, paint chips, or dust that includes lead from these sources or through breathing aerosolized dust.

[6] Pressure from environmental groups may have contributed to declines (Maxwell et al., 2000). Avoidance of nonattainment designation under NAAQS, which occurred in January 1992 for the 1978 lead standards, may also have contributed to declines.

[7] In 1998, metal mining, coal mining, electric utilities, hazardous waste disposal, chemical wholesalers, petroleum terminals, and solvent recovery services were added to the list for reporting. They account for 14% of lead emissions after 1998. In 2001, the reporting threshold for lead was significantly lowered. Figure A.2 shows that the number of reporting firms increased dramatically, but their contribution to the reported TRI lead emissions was small.

[8] The main activities of the lead manufacturing plants include extracting lead from lead ore or lead-bearing scrap materials (e.g., used lead-acid batteries) through high-temperature smelting and refining work. Iron, copper, and other metal manufacturing plants passively process lead contained in raw materials or in the coke and oil for combustion. Ceramics manufacturing uses lead compounds in glazing and paint and pigment foundries use lead as quick driers (EPA, 2020).

Lead in water also changed very slowly. Lead in water is a function of the age of the housing stock and historical factors that drove the use of lead pipes for water in specific locations. The 1986 Safe Drinking Water Amendments required the use of lead free plumbing in public water systems. The 1991 Lead and Copper rule set limits on lead in tap water and required water systems to survey their use of corrosion control. Major changes to water systems such as those that occurred in Washington DC in 2001 and in Flint Michigan in 2014 can affect leaching of lead from lead pipes into water. These events have, however, been rare.[9]

By 1988, emissions due to lead in gasoline had fallen dramatically. This decline had been driven by two factors. The first was the requirement that new cars have catalytic converters beginning in 1974. Cars with these converters required unleaded gasoline. Over time leaded gasoline as a share of all gasoline fell. The second was regulatory decreases in allowable lead levels in leaded gasoline, which began in 1979 and reached it final level of 0.1 g/gallon in 1988.[10] As a result of these changes, air emissions from on-road vehicles fell from 171.96 in 1970 to 0.42 thousand tons in 1990. In 1990, air emissions from metals industrial processing was 2.17 and chemical production and petroleum processing was 1.1 thousand tons.

Although it may be possible to partially or fully avoid some vectors of exposure, avoidance is likely to be particularly difficult for fugitive lead. Lead exposure from water, soil, and paint can be to varying degrees be avoided by not living in older housing stock, through testing of water, soil, and paint, and by remediation if levels are high. Stack lead may be partially avoidable if residents are aware of general smokestack emissions and stay inside during certain times to avoid them. Residents are, however, unlikely to be aware of fugitive emissions, which are intermittent and not stack based, and the role that wind speed plays in the dispersion of lead.

[9]Washington DC and Genesee County Michigan are not in our IV sample.

[10]In 1996, the use of lead in gasoline for on-road vehicles was banned entirely in the United States.

2.3 Lead and Infant Mortality

Lead is known to cause adverse health effects across a range of exposures. At very high levels, exposure to lead can cause lead poisoning. Lead exposure is also harmful at lower levels. The large epidemiological literature on the health effects of low level lead exposure is comprehensively reviewed in NTP (2012). Lead adversely impacts the neurological, immune, cardiovascular, and renal systems.Mason et al. (2014) reviews the neuropsychological effects of lead toxicity. Adverse effects of lead come through at least three channels – morphological, pharmacological, and indirect effects. Morphologically, lead disrupts or alters development of the nervous system both prenatally and after birth. Pharmacologically, lead substitutes for calcium and zinc, disrupting or altering operation of the nervous system. Indirect effects come from lead's effects on other bodily systems.

One causal study and numerous epidemiological studies show relationships between air lead and adult mortality. Hollingsworth and Rudik (2021) use the switch in racing fuel from leaded to unleaded in NASCAR and ARCA races to examine the causal effect of lead on elderly mortality. They find that having a leaded race in a county in a given year increased the elderly all-cause mortality rate, with much of the change coming from cardiovascular mortality and ischemic heart disease. Epidemiological evidence supports a link between lead and cardiovascular mortality that appears to be driven at least in part by lead's impact on blood pressure.[11]

To our knowledge, there are no studies that show a causal relationship between air lead and infant mortality. A small literature provides evidence on water lead and infant mortality. Employing data from Massachusetts towns in 1900, Troesken (2008) compares infant death rates in cities that used lead water pipes to rates in cities that used nonlead pipes. In the average town in 1900, the use of lead pipes increased infant mortality by 25 to 50 percent.

[11]Recent work by Lanphear et al. (2018) using NHANEs data suggests that lead may account for as much as 18% of all cause mortality and larger shares of cardiovascular mortality. While not a causal analysis, the estimates are worth noting because of their large size. Earlier papers including Pirkle et al. (1985), Lustberg and Silbergeld (2002), Menke et al. (2006) also found a relationship between lead and mortality.

Edwards (2014) shows that spikes in water lead in Washington DC due to the switch from chlorine to chloramine are associated with higher fetal death rates. Using data from 1900-1920, Clay et al. (2014) provide causal evidence on water lead and infant mortality, leveraging differences in use of lead pipes for water and differences in the acidity of water sources.[12] Animal studies also support the link between water lead and infant mortality (Aprioku and Siminialayi, 2013).

The mechanism through which lead causes infant mortality is poorly understood. Lead may be adversely affecting the part of the brain related to respiration in infants. Respiratory severity scores have been linked to mortality in low birthweight infants(Shah et al., 2020). There has been speculation in the medical literature about the link between lead and SIDS (Lyngbye et al., 1985; Erickson et al., 1983).

3 Data

Data on industrial fugitive and stack emissions of lead and other chemicals are from the U.S. Toxic Release Inventory (TRI). TRI covers 650 chemicals up to year 2018. The TRI includes 189 chemicals that are on the EPA's hazardous air pollutants (HAP) list. The EPA states that "[h]azardous air pollutants (HAPs) are those pollutants that are known or suspected to cause cancer or other serious health effects, such as reproductive effects or birth defects, or adverse environmental effects." The chemicals include 100 developmental toxicants which are thought to affect reproductive success or to affect fetal, infant, or child development. Chemicals can be in multiple categories; for example, lead is listed as both HAP and a developmental toxin. [13]

[12] Currie and Schmieder (2009) uses TRI data, leveraging the difference between fugitive and stack emissions. As part of a larger analysis of TRI emissions and their impact on birth outcomes, they find wrong signed estimates of lead on infant mortality. A related economics literature uses aggregate TRI data to examine infant mortality. Agarwal et al. (2010) uses national data from 1989-2002 and finds elasticity of infant deaths with respect to TRI air emissions is 0.03.

[13] We use information from the TRI-chemical Hazard Information Profiles to identify developmental toxins. See https://www.epa.gov/toxics-release-inventory-tri-program/tri-supplemental-documentation for more information on developmental toxins. See https://www.epa.gov/haps/what-are-hazardous-air-pollutants

Lead monitor data are from the EPA's Air Quality System (AQS). The AQS provides daily-level monitoring data on lead pollution measured in micrograms per cubic meter of air ($\mu g/m^3$). The number of lead monitors varies over time. AQS monitor data on PM and CO are also used in some specifications.

Wind data are from the National Oceanic and Atmospheric Administration's Global Surface Summary of the Day (GSOD). The GSOD provides daily-level monitoring data on wind speed and direction. Because wind is highly local, we restrict the sample to plants within 10 miles of a wind monitor.[14] County-level climate data are from the National Climatic Data Center.

Data on infant health are from the National Vital Statistics system of the National Center for Health Statistics (NCHS).[15] The NCHS provides information on infant birth and death, birth weight (in grams), and gestation weeks. It also provides data on mothers' characteristics including age, race, ethnicity, and education. County-level characteristics are from the census.

Our sample includes 127 counties that have plants with lead emissions that are within 2 miles of EPA lead monitors and within 10 miles of a wind monitor.[16] In 1990, 26% of the US population lived in these counties. This sample accounts for 21% of lead emissions and 1.6% of TRI total emissions. Appendix Figure A.4 shows the geographic distribution of the counties used for the IV sample. Appendix Figure A.5 shows the distribution of wind speeds for the IV sample. Appendix Table A.1 presents birth-weighted county-level summary statistics for the IV sample.[17] We present reduced form analysis for the IV sample.

for more details on HAP chemicals.

[14]Expanding the radius does not include many more plants in the sample, but makes the estimates noisier.

[15]There is a public sample and a restricted sample. Our analysis uses the restricted sample, which covers reported births and infant deaths in all counties.

[16]Analysis below shows that plant lead emissions are only detectable at EPA lead monitors within 2 miles.

[17]Appendix Table **??** provides additional detail on the construction of the sample.

4 Empirical Strategy

As discussed earlier, there are three challenges in identifying the effect of air lead concentrations on infant mortality. First, downward bias may arise if mothers engage in avoidance behavior based on observed pollution. Second, omitted variable bias may arise if local economic activity affect infant health through economic impacts such as on local employment and family income. This is because the level of pollution emitted is partially determined by the production scale of the plants, which is correlated with the economic impacts. Third, measurement error arising from the usual disconnection between where air lead is measured and where individuals live may generate to attenuation bias.

Our empirical strategy overcomes these challenges by comparing the effect on infant mortality of air lead concentrations caused by fugitive lead emissions in more and less windy years. Stack emissions are predictable, are often emitted from high smokestacks, may be subject to avoidance behavior and are correlated with production scale and economic impact. In contrast, fugitive emissions are intermittent, often occur closer to the ground, may be less observable to people than the smoke from tall stacks and are relatively uncorrelated (0.5) with stack emissions. Wind speed affects how far lead emissions disperse and year to year variations in wind speed are plausibly exogenous.

4.1 Air Lead Emissions and Air Lead Concentrations

To examine the relationship between lead emissions, wind, and distance between emitting plant and EPA monitor, we begin by estimating the following model:

$$\begin{aligned} AirLead_{imt} = {} & (F_{imt} \times Wind_{it} \times Distance_{im})\phi^F \\ & + (S_{imt} \times Wind_{it} \times Distance_{im})\phi^S + Wind_{it}\phi_w + \eta_{im} + \lambda_{rt} + \nu_{imt}. \end{aligned} \tag{1}$$

The dependent variable $AirLead_{imt}$ is the annual mean of ambient lead concentration readings from monitor m linked to plant i in year t. F_{imt} and S_{imt} are the fugitive and stack lead emissions from plant i linked to monitor m in year t, respectively. We interact the emission variables with a fourth order polynomial for wind speed and a vector of indicator variables, $Distance_{imt}$, for the distance from plant i to monitor m. We include the distance dummies for < 0.2 mile, 0.2-0.5 mile, 0.5-1 mile, 1-2 miles, 2-3 miles, 3-5 miles, and 5-10 miles.[18] To capture differences across plants, monitors and years, we include plant-by-monitor fixed effects η_m and region-by-year fixed effects λ_{rt}. Standard errors are clustered at the monitor level.

Having established the appropriate distance that EPA monitors detect air lead emissions – which as we show below is 2 miles – we estimate the following model:

$$\begin{aligned} AirLead_{ct} = {} & \delta^F F_{ct} + Wind_{ct}\delta_w + (Wind_{ct} \times F_{ct})\delta_w^F \\ & + \delta^S S_{ct} + (Wind_{ct} \times S_{ct})\delta_w^S + Chem_{ct}\delta_c + (Wind_{ct} \times Chem_{ct})\delta_{wc} \\ & + \eta_c + \lambda_{rt} + Media_{ct}\psi + Z_{ct}\pi + \omega_{ct} \end{aligned} \tag{2}$$

where $AirLead_{ct}$ is air lead concentration in county c in year t. The key explanatory variables are F_{ct}, denoting the aggregated fugitive lead emissions from plants in county c in year t, and its interaction with $Wind_{ct}$, a fourth order polynomial for wind speed.[19] We control for the stack lead emissions (S_{ct}), fugitive and stack emissions of other TRI-reported chemicals ($Chem_{ct}$), and their interactions with wind. The regression includes county fixed effects (η_c) and region-by-year fixed effects (λ_{rt}) to control for time-invariant determinants of and region-specific trends in infant mortality over time. We also control for waterborne and landborne lead emissions ($Media_{ct}$).[20]

To address concerns on omitted variable bias, we control for a rich set of factors that have been linked to infant health (Z_{ct}): county socioeconomic characteristics (population

[18] We also run a model similar to Currie et al. (2015), which uses continuous variables for the distance. The results are qualitatively similar.

[19] In the analysis below, we compare first stage and IV results with third, fourth, fifth and sixth order polynomials.

[20] We only include onsite emissions. The offsite emissions via water and landfill mainly happened on waste treatment facilities that locate far from the original neighborhood of the manufacturing facilities. But controlling for offsite emissions does not affect our results.

density, percent white, percent age 25 and older with high school degree, median household income, percent manufacturing employment, and employment rate); climate variables (county-average annual precipitation and temperature); mothers' characteristics (percent white, percent Hispanic, percent high school degree, and percent aged over 35); and linear trends of baseline mortality rate from 1980 to 1986. Regressions are weighted by the number of live births. Robust standard errors are clustered at the county-level to adjust for arbitrary heteroskedasticity and within-county serial correlation.[21]

4.2 Air Lead Concentrations and Infant Mortality

To measure the effect of air lead concentrations on infant mortality, we estimate the following equation:

$$\begin{aligned} InfMort_{ct} = {} & \beta^{AL} AirLead_{ct} + \beta^{F} F_{ct} + Wind_{ct}\beta_w \\ & + \beta^{S} S_{ct} + (Wind_{ct} \times S_{ct})\beta^{S}_{w} + Chem_{ct}\beta_c + (Wind_{ct} \times Chem_{ct})\beta_{wc} \\ & + \eta_c + \lambda_{rt} + Media_{ct}\psi + Z_{ct}\pi + \epsilon_{ct} \end{aligned} \tag{3}$$

where $InfMort_{ct}$ is infant mortality in county c in year t.[22] We instrument $AirLead_{ct}$ with the interactions between fugitive lead emissions and a quartic polynomial in wind speed ($Wind_{ct} \times F_{ct}$). Fugitive lead, wind speed, and the other control variables are the same as in the first stage.

The exclusion restriction is that *interactions* between fugitive lead emissions and wind speed will affect infant mortality only through their effect on air lead concentrations. We allow for direct effects of fugitive lead emissions and wind speed on infant mortality, as well as direct effects of stack lead emissions and interactions between them and wind, and the other control variables. Because it is only the combination of intermittent lead fugitive emissions with wind strength that is excluded from the equation of interest, we believe this

[21]In robustness checks, we also provide results using spatial standard errors.

[22]We also examine premature birth (< 36 weeks), birth weight, and low birthweight (< 2500 grams).

is a plausible assumption.

5 Results

5.1 Air Lead Emissions and Concentrations

In this section, we present evidence on a number of points that are relevant to identification of the effects of air lead concentrations on infant mortality. First, we show that wind speed affects the relationship between fugitive emissions and ambient lead concentrations within 2 miles of the plant. In contrast, wind speed has very little effect on the relationship between stack emissions and ambient lead concentrations. Second, we document that a greater share of fugitive emissions in overall lead emissions is associated with a higher daily standard deviation in ambient lead concentration, which supports the intermittency of fugitive emissions. Third, we show that fugitive lead emissions has predictive power for air lead concentration even when we include a range of controls. Fourth, we provide evidence that a fourth order polynomial provides a better fit than higher or lower order polynomials. Fifth, we show that fugitive lead interacted with wind does not predict either county or maternal characteristics.

Figure 2 shows that wind speed has a positive effect on the relationship between fugitive emissions and ambient lead concentrations within 2 miles of the plant, but has very little effect on the relationship between stack emissions and ambient lead concentrations (see Appendix Figure A.6). It plots the marginal effect of fugitive lead emissions ($\hat{\phi}^F$) on air lead concentration as a function of wind speed near the plants for different distance ranges from plant to monitor. Each panel of Figure 2 presents the distance-specific wind gradient, showing how the marginal effect of fugitive lead emissions changes with higher average wind speed. Fugitive lead emissions have a large effect on air lead concentration within 0.2 mile from the plants, and the effect fades when wind gets stronger. From 0.5 to 1 mile, there is weak but significant effect at wind speed between 5.5 and 6.5 knots. From 1 to 2 miles, fugitive lead emissions have a strong effect when wind speed is over about 7.5. In Appendix

Figure A.6, stack lead emissions have a weak effect on air lead concentration within 0.5 miles from plants at high wind and little impact on areas beyond 0.5 miles. This is because stack lead emissions occur more continuously and higher in the air than fugitive emissions and so are more disperse irrespective of wind speed. The remaining analysis focuses on EPA air lead monitors that are within 2 miles of a plant.

Wind speed is important, because it affects the share of the county population that is exposed to fugitive lead emissions. When local wind speed is low, fugitive lead emissions only affect the neighborhoods extremely close to the plants. The 1990 block-group level data show less than 0.9% of county population living within 0.2 miles from the plants. When local wind speed is high, fugitive lead emissions affect many more people. 25.6% of the population in our sample counties lived within 2 miles of a lead emitting plant. Although 25.6% may seem high, these counties have high population density and multiple lead emitting plants. On average in 1990 there were about 7 lead emitting plants per county.[23]

Table 1 documents that a greater share of fugitive emissions in overall lead emissions is positively and statistically significantly associated with the daily standard deviation in ambient lead concentration. This supports the intermittency of fugitive emissions.

Figure 3 highlights that fugitive lead emissions have predictive power for air lead concentration even when we include a rich set of controls. Figure 3 plots the highly nonlinear relationship between fugitive lead emissions ($\hat{\phi}^F$) and air lead concentration as a function of wind speed for counties with that have plants within 2 miles of an EPA air lead monitor. Appendix Figure A.8 shows the graphs as we move from the most parsimonious specification to specifications with richer sets of controls. The F-statistics are all at 38.0 or higher, and the graphs are very similar.

Appendix Figure A.9 demonstrates that a fourth order polynomial is a parsimonious model to capture the nonlinear relationship between fugitive lead emissions ($\hat{\phi}^F$) and air lead concentration. It plots the marginal effects of a specification that includes a full set

[23] A 2 mile circle around a plant covers 12.6 square miles.

of controls as we move from a first order polynomial to a fourth order polynomial.[24] The F-statistic is much higher for the fourth order polynomial (41.2) than for the third order (17.6), which suggests that the addition of a fourth order term provides benefits. Although they are not graphed, it is worth noting that the fifth and sixth order polynomials have lower F-statistics (31.4, 26.3), suggesting that there are not benefits to adding more terms beyond the fourth order term.

Appendix Tables A.2 and A.3 show that fugitive lead interacted with wind does not predict either county or maternal characteristics. The F-statistics are all below 4, with the exception of population density, where the F statistic is below 6. We provide a range of additional robustness checks in section 5.3.

5.2 IV Effects of Air Lead Concentrations on Infant Mortality

Table 2 shows that higher levels of air lead concentrations cause higher infant mortality.[25] The coefficient on air lead concentration in the IV specification is positive and statistically significant across all five specifications. The decline in air lead concentration in the sample is 0.13. This decrease in air lead concentration would decrease infant mortality by 0.218 per thousand live births or about 2.8% of mean infant mortality. We discuss deaths averted due to changes in fugitive lead emissions and all deaths averted due to declines in air lead concentration further in Section 5.4.

Table 3 shows that the estimates in Table 2 are robust to different ages at death and to adjusting for the timing of exposure. Two-thirds of infant deaths occur within the first month. For these infants, a large share of lead exposure will have come in utero. Column 1 present the results for deaths within the first month, and column 2 replicates our preferred results from column 5 of Table 2 for deaths within the first year. The coefficient on air

[24]This is not surprising given the complexity in modelling air pollution due to atmospheric turbulence (e.g., Nieustadt and van Dop, 1982; Raputa and Lezhenin, 2020).

[25]Figure 3 shows the birth weighted first stage. Appendix Figure A.10 graphs the reduced form relationship between fugitive lead emissions and infant mortality as a function of wind speed for the IV sample. The pattern is generally consistent with Figure 3. The reduced form shows a positive effect at higher wind speeds where larger shares of the population are affected.

lead concentration for deaths within the first month in column 1 is positive and statistically significant, and the magnitude of the coefficient is smaller than in column 2.

One concern with the analysis in columns 1 and 2 is that many infants may have been exposed to lead in the previous year. For example, an infant that is born in January and dies in January experienced nearly all of its in utero exposure in the previous year.In columns 3 and 4, which restrict attention to infant deaths from April to December, the coefficient estimates are higher, which suggests that there is some attenuation bias due to mismeasured exposure. For mortality in the first month in column 3, all of the infants spent at least the third trimester in the current year. The median infant in this sample was born in mid-August and so spent part of the first trimester and all of the second and third trimesters in the current year. Compared to the results in columns 1 and 2, the point estimates in columns 3 and 4 are larger and the coefficient for deaths within one month is more statistically significant. This suggests that using current year exposure for infants with significant exposure in the previous year is causing some attenuation bias.

In columns 5 and 6, which restrict attention to infant deaths from July to December, the coefficient estimates are similar or higher than in columns 3 and 4. For mortality in the first month in column 5, all of the infants spent at least the second and third trimester in the current year. The median infant in this sample was born at the very beginning of October and so spent all three trimesters in the current year. The point estimate in column 5 is nearly identical to the point estimate in column 3. The point estimate in column 6 is larger than the point estimate in column 4, although the two are not statistically significantly different. This may reflect the fact that some infants who died in April through June at ages beyond one month had significant exposure in the previous year. Dropping them may be further reducing attenuation bias.

Table 4 suggests that lead exposure may be disproportionately affecting nonwhite infants at the margin. The coefficient on lead in column 1 for nonwhite infants is substantially larger than in column 2 for white infants. The implied effects for a decrease of 0.13 are 0.40 per

1000 live births for nonwhite infants and 0.18 for white infants, which are 3.4% and 3.0% of their means. Columns 3 and 4 suggest that these differences are present in the first month. The reasons for these differences are unclear. Some of the difference appears to reflect a differential likelihood of living within 2 miles of a lead plant. It is worth noting that because of the large standard errors, the two point estimates are not statistically significantly different than one another.

Tables 5 and 6 draw on data on cause of death to better understand the mechanisms through which lead is causing mortality. In both tables, air lead concentration is statistically significantly related to three causes: low birthweight; sudden infant death syndrome (SIDS) and accidental suffocation and strangulation on bed (ASSB); and other causes of death. The coefficients on air lead concentration are small and not statistically significant for the other four causes: congenital anomalies; respiratory; other conditions originating in the perinatal period; and homicide.

Before discussing the low birthweight mortality results, it is useful to examine the average effect of lead exposure on the probability of low birthweight and other birth outcomes. Table 7 shows the IV effect of lead on the likelihood of low birthweight and prematurity are positive but not significant. The implied effects are very small. The coefficients on premature and low birthweight are positive, which suggests that higher air lead may lead to worse birth outcomes. The implied effects are 0.17 and 0.10 per 1000 live births, which are 0.28% and 0.13% of the mean. The coefficient on birthweight is positive, but the implied effect is 0.24 grams.

Taken together, Tables 5-7 suggest that in places with higher air lead concentration, although the incidence of low birthweight is not statistically significantly higher, low birthweight infants are more likely to die. In column 2 of both tables, higher air lead concentration is causally related to death related to low birthweight. The coefficients are nearly identical in magnitude (0.312 at one month and 0.310 at one year) across Tables 5 and 6, which examine infant mortality in the first year and first month. We would expect the magnitudes to be similar, since low birthweight infants generally die in the first month.

Because of inconsistent classification, sudden infant death syndrome and accidental suffocation and strangulation on bed are often combined with other unexplained deaths into a single category called SUID, or sudden unexpected infant death. SUID deaths tend to peak in months 1-4 (Moon et al., 2016) As we noted earlier, there has been speculation in the medical literature about the link between lead and SIDS (Lyngbye et al., 1985; Erickson et al., 1983). In 1994, a number of organizations including the American Academy of Pediatrics launched the Back-to-Sleep Campaign to address SIDS. The inclusion of region x year fixed effects should control for introduction of this campaign.

Tables 5 and 6 provide evidence of a causal link between air lead concentration and incidence of SUID. Consistent with the evidence on the timing of SUID, the coefficient is much smaller for mortality in the first month (0.076) than it is for mortality in the first year (0.414).

Other deaths are also strongly related to air lead concentration at one year and at one month. We are in the process to doing further analysis of causes within this subgroup.

5.3 Further Robustness Checks

Appendix Table A.4 shows the results are robust to adding individual groups of chemicals and particulate matter (PM10) and carbon monoxide (CO) as controls. Recall that our estimates in Table 2 shifted very little from column 4 to column 5 with the addition of three groups of chemicals – developmental chemicals, nondevelopmental chemicals, and HAPs – all interacted with wind. Consistent with this in Appendix Table A.4, the F-stat is quite similar across the first three columns as we include controls for developmental chemicals in column 1, add nondevelopmental chemicals in column 2, and add HAPs in column 3. The third column is our base model from Table 2. Column 4 adds controls for PM10 and column 5 adds controls for CO. Unfortunately, not all monitors have data for PM10 and CO, so the sample sizes are smaller in these columns. Despite the fact that the sample in column 5 is about one-third smaller than in column 3, the coefficient on air lead concentration is nearly identical – 1.725 vs. 1.676.

One possible concern is that lead is co-emitted with other chemicals and so the coefficient on lead captures the effect of lead and other chemicals. Column 6 adds metals emissions (excluding zinc), which partially overlap with HAPs but are particularly likely to be co-emitted with lead given the nature of the industries in our sample. Compared to the previous columns, the F-stat is somewhat lower but still above 10. The coefficient on air lead concentration is also higher, although not statistically significantly different than previous estimates.

In column 7, we see the estimate is still positive but noisier with the inclusion of zinc. We cannot rule out that it is statistically similar to our main estimate either. Because zinc tends to be co-emitted with lead – ores tend to have both – and air fugitive lead and air fugitive zinc are relatively highly correlated (0.58), we may have a multicollinearity issue in the estimation. That might be a reason for the imprecise estimate. Although our estimates for lead may capture the effect of zinc and lead, the adverse effects are very likely driven by lead. Lead is a toxin and has adverse effects on a range of bodily functions. In contrast, zinc is essential for a wide range of enzymatic and structural functions. The EPA's RSEI model assigns a high toxicity score to lead inhalation (23,000) and a low toxicity score to zinc (100). Consistent with this the CDC has extensive guidance on reducing lead exposure, while guidance for zinc primarily involves ensuring that there is adequate nutritional intake.[26]

Appendix Table A.5 presents a series of additional robustness checks. Column 1 replicates our preferred specification from column 5 of Table 2. Column 2 drops counties with plants that always report zero fugitive lead emissions. Column 3 controls for emissions from other non-lead emitting plants in the county. Column 4 requires that all counties have at least 10 monitor-years of data. It is well-known that EPA monitoring data are quite unbalanced due the the entry and exit of pollution monitors over time. In columns 5 and 6, we

[26]The CDC has guidance on occupational exposure to very high levels of zinc fumes.

shorten the sample period by five and ten years. Much of the variation occurs in the early part of the sample period, so the question is whether the coefficients differ with a shorter sample period. The coefficients on air lead concentration in columns 2-6 are similar to the baseline estimate in column 1 in sign, magnitude, and significance.

5.4 Infant Deaths Averted

In Table 8, we use our estimates from Tables 2 and the regression underlying Appendix Figure A.10 to do back of the envelope calculations of the number of infant deaths averted. For the IV and reduced form, we use county-specific realized declines in fugitive lead emissions to estimate the effects on air lead concentration and on infant mortality. To reduce the reliance on any one year, the comparison is between average county-specific fugitive emissions over 1988-1991 and 2015-2018.

In the IV specification, the realized decline in fugitive lead emissions over these two periods implies a fall in air lead of 0.035 $\mu g/m^3$,[27] and a fall in infant mortality of 0.059 per 1,000 live births.[28] In the reduced form specification for the IV sample, the realized decline implies a fall in infant mortality of 0.034 per 1,000 live births. As usual, the IV estimates are larger than the reduced form estimates.

Table 8 summarizes the infant deaths averted and the value of these lives saved. The annual number of births in the sample counties was approximately 1 million in 2015-2018. Depending on the specification, the implied number of deaths averted in these counties per year is 34-59. At the EPA valuation of \$9.2 million per death averted in 2018 USD, the benefits of infant lives saved are \$313-533 million per year.[29] The full decline in air lead concentration from all sources in our sample counties was 0.130 $\mu g/m^3$.[30] This implies a fall in infant mortality of 0.218 per 1,000 live births or 218 deaths averted per year. The benefits

[27] For reference, the average air lead concentration in our sample is $0.08\mu g/m^3$.

[28] 0.059=0.035*1.676, where 1.676 is the coefficient on air lead concentration in column 5 of Table 2

[29] The value of a statistical life that the EPA uses is \$7.4 million (2006 USD). This is \$9.2 million (2018 USD).

[30] 0.130 is the difference between the air lead concentration for 1988-1991 and the air lead concentration for 2015-2018 for the IV sample.

of infant lives saved are $2.0 billion per year.

6 Conclusion

This paper provides national IV estimates of the effects of air lead concentration on infant mortality in the United States over the period 1988-2018. Our identification overcomes the challenges associated with avoidance behavior, omitted variable bias, and measurement error by instrumenting for EPA air lead concentration with TRI fugitive lead emissions interacted with wind speed near the emitting plants and including a rich set of controls. Variation in fugitive lead emissions and wind speed is shown to determine lead exposure. While stack lead emissions occur routinely and may be subject to avoidance behavior, fugitive lead emissions are unintended and intermittent (EPA, 1994). The unpredictable annual variations in fugitive lead and wind speed make it difficult to engage in avoidance behavior. We provide evidence that fugitive lead emissions interacted with wind speed has low predictive power for county socioeconomic characteristics and mothers' characteristics. We provide evidence that fugitive lead emissions and wind speeds are strongly linked to air lead concentration readings of EPA monitors.

We find a positive and statistically significant relationship between air lead concentration and infant mortality. Accounting for the timing of lead exposure in utero increases the magnitude of the estimates. Estimates by race are imprecise but suggest that lead exposure may be disproportionately affecting nonwhite infants. Cause of death data show that lead increases deaths from low birthweight and sudden unexplained infant death. Back of the envelope estimates indicate that declines in fugitive lead emissions generated benefits of $313-$533 million annually and the full decline in air lead concentration from all sources generated benefits of $2.0 billion per year. These are only the benefits from avoided infant deaths. Given that lead exposure causes morbidity and mortality in other populations and has lasting impacts on child development, the benefits are likely substantially larger.

Returning to the broader picture, air emissions of lead from industry and some other sectors such as aviation continue to impact millions in the US and globally. These new estimates can inform investments in reductions in air lead emissions. In the U.S., industrial firms and the aviation industry still emit hundreds of thousands of pounds of lead into the air. A recent Unicef report found that 1 in 3 children worldwide had blood lead levels above 5 g/DL (Burki, 2020; Rees and Fuller, 2020). While the report notes that elevated blood lead levels are due to range of channels of exposure, air lead emissions are an important contributor.

References

Agarwal, Nikhil, Chanont Banternghansa, and Linda T.M. Bui, "Toxic exposure in America: Estimating fetal and infant health outcomes from 14 years of TRI reporting," *Journal of Health Economics*, 2010, *29* (4), 557–574.

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

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

Aprioku, JS and IM Siminialayi, "Maternal lead exposure and pregnancy outcome in Wistar albino rats," *Journal of Toxicology and Environmental Health Sciences*, 2013, *5* (10), 185–193.

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

Brehm, John and James T Hamilton, "Noncompliance in Environmental Reporting : Are Violators Ignorant, or Evasive, of the Law?," *Journal of Political Science*, 1996, *40* (2), 444–477.

Bui, Linda TM, Ron Shadbegian, Alicia Marquez, Heather Klemick, and Dennis Guignet, "The Effects of Short-Term, In Utero Lead Exposure on Birth Outcomes by Trimester: Quasi-Experimental Evidence from NASCAR's Deleading Policy," *U.S. Environmental Protection Agency (EPa) National Center for Environmental Economics (NCEE) Working Paper #21-02*, 2021.

Burki, Talha, "Report says 815 million children have high blood lead levels," *The Lancet*, 2020, *396* (10248), 370.

Clay, Karen, Margarita Portnykh, and Edson Severnini, "The legacy lead deposition in soils and its impact on cognitive function in preschool-aged children in the United States," *Economics & Human Biology*, 2019, *33*, 181–192.

_ , _ , **and** _ , "Toxic Truth: Lead and Fertility," *Journal of the Association of Environmental and Resource Economists*, 2021, *8* (5), 975–1012.

_ , **Werner Troesken, and Michael Haines**, "Lead and Mortality," *Review of Economics and Statistics*, 2014, *96* (3), 458–470.

Currie, Janet and Johannes F. Schmieder, "Fetal exposures to toxic releases and infant health," *American Economic Review: Papers & Proceedings*, 2009, *99* (2), 177–183.

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

Dave, Dhaval M. and Muzhe Yang, "Lead in Drinking Water and Birth Outcomes: A Tale of Two Water Treatment Plants," *NBER Working Paper #27996*, 2020.

Deryugina, Tatyana, Garth Heutel, Nolan H. Miller, David Molitor, and Julian Reif, "The mortality and medical costs of air pollution: Evidence from changes in wind direction," *American Economic Review*, 2019, *109* (12), 4178–4219.

Edwards, Marc, "Fetal death and reduced birth rates associated with exposure to lead-contaminated drinking water," *Environmental science & technology*, 2014, *48* (1), 739–746.

EPA, "Estimating and Controlling Fugitive Lead Emissions from Industrial Sources," Technical Report 1994.

— , "Lead Emissions," Technical Report, US Environmental Protection Agency 2018.

— , "TOXICS RELEASE INVENTORY Guidance for Reporting the Dioxin and Dioxin-like Compounds Category," Technical Report 2020.

Erickson, Marilyn M, Alphonse Poklis, George E Gantner, Allan W Dickinson, and Laura S Hillman, "Tissue mineral levels in victims of sudden infant death syndrome I. Toxic metals—lead and cadmium," *Pediatric research*, 1983, *17* (10), 779–784.

Gazze, Ludovica, Claudia L. Persico, and Sandra Spirovska, "The Long-Run Spillover Effects of Pollution: How Exposure to Lead Affects Everyone in the Classroom," *NBER Working Paper #28782*, 2021.

Gronqvist, Hans, J. Peter Nilsson, and Per-Olof Robling, "Understanding How Low Levels of Early Lead Exposure Affect Children's Life Trajectories," *Journal of Political Economy*, 2020, *128* (9), 3376–3433.

Grossman, Daniel S. and David J.G. Slusky, "The Impact of the Flint Water Crisis on Fertility," *Demography*, 2019, *56* (6), 2005–2031.

Heutel, Carth and Christopher J. Ruhm, "Air Pollution and Procyclical Mortality," 2013.

Hollingsworth, Alex and Ivan Rudik, "The effect of leaded gasoline on elderly mortality: Evidence from regulatory exemptions," *American Economic Journal: Economic Policy*, 2021, *13* (3), 345–73.

— , **Mike Huang, Ivan J. Rudik, and Nicholas J. Sanders**, "Lead Exposure Reduces Academic Performance: Intensity, Duration, and Nutrition Matter," *NBER Working Paper # 28250*, 2020.

Ito, Koichiro and Shuang Zhang, "Willingness to pay for clean air: Evidence from air purifier markets in China," *Journal of Political Economy*, 2020, *128* (5), 1627–1672.

Klemick, Heather, Henry Mason, and Karen Sullivan, "Superfund cleanups and children's lead exposure," *Journal of Environmental Economics and Management*, 2020, *100*, 102289.

Knittel, Christopher R., Douglas L. Miller, and Nicholas J. Sanders, "Caution, drivers! children present: Traffic, pollution, and infant health," *Review of Economics and Statistics*, 2016, *98* (2), 350–366.

Lanphear, Bruce P, Stephen Rauch, Peggy Auinger, Ryan W Allen, and Richard W Hornung, "Low-level lead exposure and mortality in US adults: a population-based cohort study," *The Lancet Public Health*, 2018, *3* (4), e177–e184.

Li, Qian, Hongguang Cheng, Tan Zhou, Chunye Lin, and Shu Guo, "The estimated atmospheric lead emissions in China, 1990–2009," *Atmospheric Environment*, 2012, *60*, 1–8.

Lustberg, Mark and Ellen Silbergeld, "Blood lead levels and mortality," *Archives of internal medicine*, 2002, *162* (21), 2443–2449.

Lyngbye, Troels, Ole Noerby Hansen, Lilian Vangbert, and Philippe Grandjean, "Lead as a cause of SIDS," *New England Journal of Medicine*, 1985.

Marchi, Scott De and James T. Hamilton, "Assessing the accuracy of self-reported data: An evaluation of the toxics release inventory," *Journal of Risk and Uncertainty*, 2006, *32* (1), 57–76.

Mason, Lisa H, Jordan P Harp, and Dong Y Han, "Pb neurotoxicity: neuropsychological effects of lead toxicity," *BioMed research international*, 2014, *2014*.

Maxwell, John W, Thomas P Lyon, and Steven C Hackett, "Self-regulation and social welfare: The political economy of corporate environmentalism," *The Journal of Law and Economics*, 2000, *43* (2), 583–618.

Menke, Andy, Paul Muntner, Vecihi Batuman, Ellen K Silbergeld, and Eliseo Guallar, "Blood lead below 0.48 μmol/L (10 μg/dL) and mortality among US adults," *Circulation*, 2006, *114* (13), 1388–1394.

Moon, Rachel Y, Robert A Darnall, Lori Feldman-Winter, Michael H Goodstein, Fern R Hauck, Task Force on Sudden Infant Death Syndrome et al., "SIDS and other sleep-related infant deaths: evidence base for 2016 updated recommendations for a safe infant sleeping environment," *Pediatrics*, 2016, *138* (5).

Moretti, Enrico and Matthew Neidell, "Pollution, health, and avoidance behavior: Evidence from the ports of Los Angeles," *Journal of Human Resources*, 2011, *46* (1), 154–175.

Natan, Thomas E. and Catherine G. Miller, "Are toxics release inventory reductions real?," *Environmental Science and Technology*, 1998, *32* (15).

Neidell, Matthew, "Information, avoidance behavior, and health: The effect of ozone on asthma hospitalizations," *Journal of Human Resources*, 2009, *44* (2), 450–478.

Nieustadt, Frans T.M. and Han van Dop, *Atmospheric Turbulence and Air Pollution Modelling*, Dordrecht, Holland: D. Reidel Publishing Company, 1982.

NTP, "NTP Monograph on health effects of low-level lead. US Department of Health and Human Services," 2012.

Pacyna, Jozef M and Elisabeth G Pacyna, "An assessment of global and regional emissions of trace metals to the atmosphere from anthropogenic sources worldwide," *Environmental reviews*, 2001, *9* (4), 269–298.

Persico, Claudia L., "Can Pollution Cause Poverty? The Effects of Pollution on Educational, Health and Economic Outcomes," *IZA Discussion Paper #12965*, 2020.

_ **and Joanna Venator**, "The Effects of Local Industrial Pollution on Students and Schools," *Journal of Human Resources*, 2021, *56* (2), 406–445.

Pirkle, James L, Joel Schwartz, J Richard Landis, and William R Harlan, "The relationship between blood lead levels and blood pressure and its cardiovascular risk implications," *American journal of epidemiology*, 1985, *121* (2), 246–258.

Raputa, V.F. and A.A. Lezhenin, "Estimation of the Altitude of Smoke Plumes from Satellite Images," *Atmospheric and Oceanic Optics*, 2020, *33*, 539–544.

Rees, Nicholas and Richard Fuller, *The toxic truth: children's exposure to lead pollution undermines a generation of future potential*, UNICEF, 2020.

Ruhm, Christopher J., "Are recessions good for your health?," *Quarterly Journal of Economics*, 2000, *115* (2), 617–650.

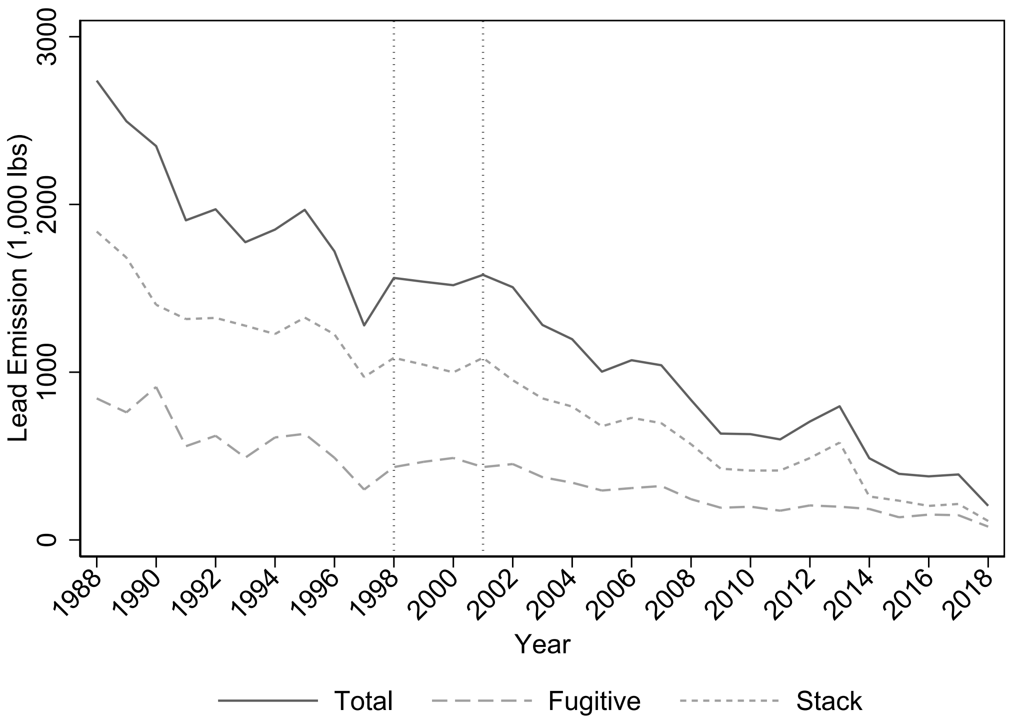
Shah, Shetal I, David Aboudi, Edmund F La Gamma, and Heather L Brumberg, "Respiratory Severity Score greater than or equal to 2 at birth is associated with an increased risk of mortality in infants with birth weights less than or equal to 1250 g," *Pediatric Pulmonology*, 2020, *55* (12), 3304–3311.

Tanaka, Shinsuke, Kensuke Teshima, and Eric Verhoogen, "North-South Displacement Effects of Environmental Regulation: The Case of Battery Recycling," 2021.

Troesken, Werner, "Lead Water Pipes and Infant Mortality at the Turn of the Twentieth Century," *Journal of Human Resources*, 2008, *43* (3), 553–575.

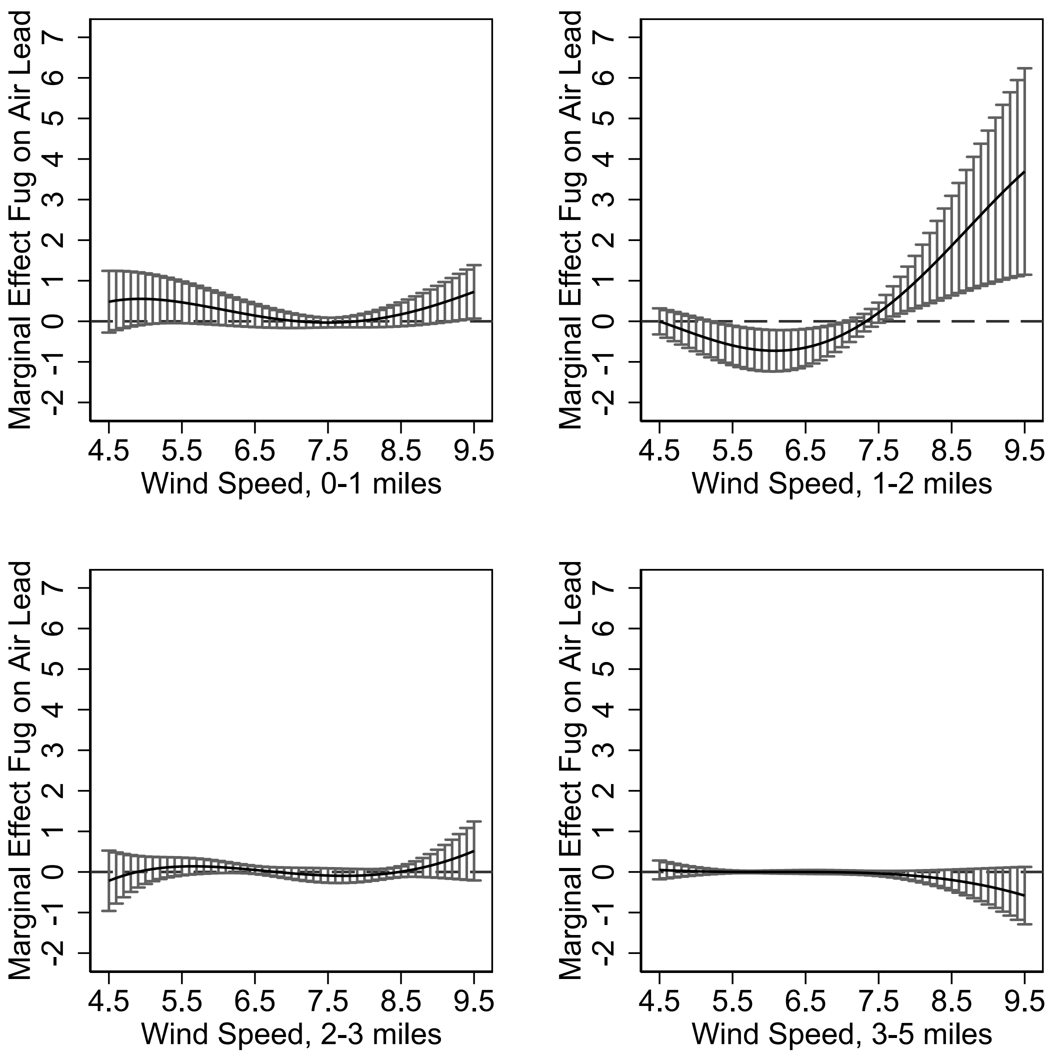
Tables and Figures

Figure 1: Trends in Fugitive and Stack Emissions



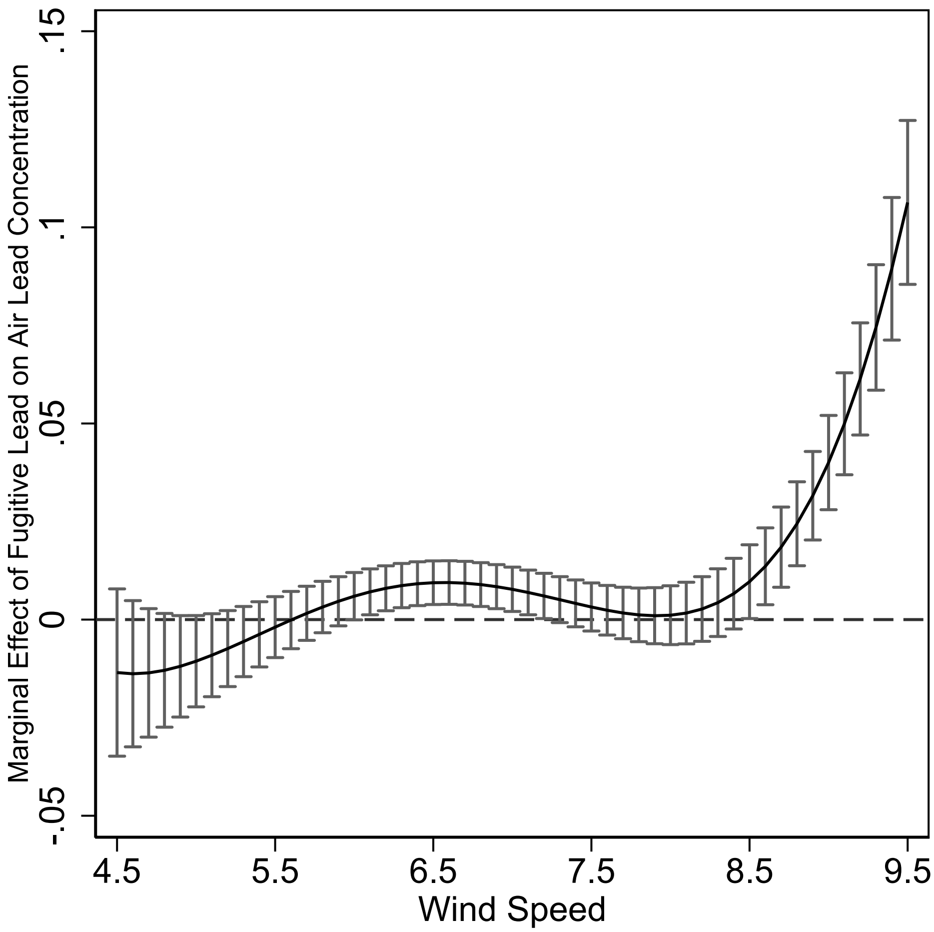
Notes: This figure shows the trend of fugitive, stack, and total air lead emissions reported by TRI plants during 1988 to 2018. The vertical lines mark year 1998 when seven additional industries were added to TRI and year 2001 when the threshold for lead reporting was significantly lowered. Appendix Figure A.2 shows the number of reporting plants and changes that their inclusion have on reported totals. Appendix Figure A.3 shows trends in airborne, waterborne, landborne, and recycled lead.

Figure 2: Effect of Fugitive Lead Emissions on Air Lead Concentration by Wind Speed



Notes: This figure reports the estimated ϕ^F in equation 1 – the marginal effect of fugitive lead emissions (in 10,000 pounds) on ambient lead concentration (in μg/m^3) as a function of wind speed (in knots) for different distance ranges from monitors to plants. Wind speed is captured by weather monitoring stations within 10 miles of each plant. Figure A.7a and A.7b show the distribution of monitors over distance to plants and over wind speed, respectively. There is little impact of fugitive lead on ambient lead concentration shown beyond 2 miles, so we did not show the panel for 5 to 10 miles for simplicity.

Figure 3: Air Fugitive Lead and Air Lead Concentration with Controls



Notes: The figure plots the marginal effect of fugitive lead emissions (F_{ct}) of plants on the air lead concentration readings at lead monitors. The figure is obtained from the first-stage regression of the IV estimation: $AirLead_{ct} = \beta^F F_{ct} + \beta^F_w HighWind_{ct} \times F_{ct} + \delta_w WindSpeed_{ct} + \beta^S S_{ct} + \beta^S_w WindSpeed_{ct} \times S_{ct} + \delta Chem_{ct} + \delta_w WindSpeed_{ct} \times Chem_{ct} + \eta_c + \lambda_{st} + Media_{ct}\psi + Z_{ct}\phi + \epsilon_{ct}$, where $WindSpeed_{ct}$ denotes polynomials (linear, quadratic, cubic, quartic) of average wind speed near the plants.

Table 1: Fugitive Lead Emissions and Daily Variation of Ambient Lead Concentration

	Dep Var: S.D. Daily Ambient Lead Concentration				
	(1)	(2)	(3)	(4)	(5)
High Frac Fugitive Lead	0.144***	0.140***	0.139***	0.138***	0.138***
	(0.046)	(0.042)	(0.043)	(0.042)	(0.042)
Mean Dep Var	0.227	0.227	0.227	0.227	0.227
Adjusted R^2	0.692	0.694	0.696	0.696	0.696
Monitor-Year	3015	3015	3015	3015	3015
Monitors	352	352	352	352	352
Counties	127	127	127	127	127
Monitor, Region-Year FE	Y	Y	Y	Y	Y
Air, Water, Land Lead		Y	Y	Y	Y
Socio-economic			Y	Y	Y
Climate Var				Y	Y
Other Emissions					Y

Notes: This table reports the results for regressing standard deviation of daily ambient lead concentration within 2 miles of lead plants on the fraction of fugitive over total air lead emissions of the plants. Variable *High Frac Fugitive Lead* is an indicator for monitor-years that have above-median fraction of fugitive over total air lead. Mean number of days for calculating the standard deviation of monitoring data in a year is 213. We control for monitor fixed effects, region-by-year fixed effects, and the level of air, water, land-borne lead emissions from plants in all specifications. We test for robustness including county socio-economic characteristics (population density, percent white, percent high school degree, median household income, percent manufacturing employment, and employment rate), county climate (annual total precipitation, annual average temperature, and wind speed), and other toxic emissions from lead and non-lead plants in the county. Monitor-years with positive air lead emissions and ambient lead concentration are included for the regressions. Standard errors are clusters at county level.

Table 2: IV Estimates of Air Lead Concentration and Infant Mortality

	(1) IMR	(2) IMR	(3) IMR	(4) IMR	(5) IMR
Air Lead Concentration	1.998***	1.753***	1.717***	1.685***	1.676***
	(0.429)	(0.453)	(0.447)	(0.395)	(0.435)
KPFstat	41.983	40.252	38.029	45.036	41.232
DepMean	7.718	7.718	7.718	7.718	7.718
County-Year	1553	1553	1553	1553	1553
Counties	127	127	127	127	127
County,Region-by-Year FE	Y	Y	Y	Y	Y
Base IMR		Y	Y	Y	Y
Socioeconomic,Mother		Y	Y	Y	Y
Climate Var			Y	Y	Y
Water,Land Lead				Y	Y
Other Chem					Y

Notes: Baseline IMR is the county infant mortality rate averaged over 1980 to 1986 (the year prior to the start of TRI). Controls on other chemicals include air fugitive and stack emissions of developmental toxins and other TRI reported chemicals. Controls on socio-economic characteristics include population density, percent white, percent high school degree, median household income, percent manufacturing employment, and employment rate at county level. Controls on mothers' characteristics include county-average percent white, percent Hispanic, percent of high school degree, and percent of mothers aged over 35. Controls on climate variables include annual total precipitation and annual average temperature at the county level. Regressions are weighted by the number of births. Standard errors clustered at county level. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 3: IV Estimates By Age at Death and Timing of Birth

	(1) IMR 1m	(2) IMR 1y	(3) IMR AD1m	(4) IMR AD1y	(5) IMR JD1m	(6) IMR JD1y
Air Lead Concentration	0.728**	1.676***	1.155***	2.128***	1.143***	2.660***
	(0.323)	(0.435)	(0.336)	(0.430)	(0.387)	(0.530)
KPFstat	41.232	41.232	40.884	40.884	40.337	40.337
DepMean	5.104	7.718	5.09	7.565	4.988	7.453
County-Year	1553	1553	1553	1553	1553	1553
Counties	127	127	127	127	127	127
All Controls	Y	Y	Y	Y	Y	Y

Notes: The dependent variable is infant mortality in the first month or year for infants in columns 1-2. The dependent variable is infant mortality in the first month or year for infants born April to December in columns 3-4. The dependent variable is infant mortality in the first month or year for infants born July to December in columns 5-6. All controls are the controls from column 5 of Table 2. Regressions are weighted by the number of births. Standard errors clustered at county level. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 4: IV Estimates By Race of Mother

	(1) IMR1yrnwh	(2) IMR1yrwh	(3) IMR1mnwh	(4) IMR1mwh
Air Lead Concentration	3.068**	1.415***	1.977**	0.478*
	(1.212)	(0.361)	(0.998)	(0.256)
KPFstat	39.646	41.711	39.646	41.711
DepMean	11.778	6.067	8.346	4.044
CountyYear	1552	1552	1552	1552
Counties	127	127	127	127
AllControls	Y	Y	Y	Y

Notes: The dependent variable is infant mortality in the first year for infants born to nonwhite (IMR1yrnwh) and white (IMR1yrwh) in columns 1 and 2. The dependent variable is infant mortality in the first month for infants born to nonwhite (IMR1mnwh) and white (IMR1mwh) in columns 3 and 4. The mean of the dependent variables differs from the summary statistics in Appendix Table A.1. In the summary statistics, the observations are weighted by all births, while here they are weighted by race-specific births. All controls are the controls from column 5 of Table 2. Regressions are weighted by the number of births to nonwhite and white mothers. Standard errors clustered at county level. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 5: IV Estimates By Cause, 1 year

	(1) IMR ca	(2) IMR lbw	(3) IMR resp	(4) IMR pp	(5) IMR sidsassb	(6) IMR homi	(7) IMR other
Air Lead Concentration	0.044	0.310**	0.103	-0.027	0.414*	0.010	0.503***
	(0.102)	(0.139)	(0.130)	(0.179)	(0.229)	(0.026)	(0.182)
KPFstat	26.079	26.079	26.079	26.079	26.079	26.079	26.079
DepMean	1.43	1.225	.635	1.824	.825	.075	1.571
CountyYear	1470	1470	1470	1470	1470	1470	1470
Counties	121	121	121	121	121	121	121

Notes:ca = congenital anomalies; lbw = low birthweight; resp = respiratory; pp = other conditions originating in perinatal period; sidsassb = sudden infant death syndrome and accidental suffocation and strangulation on bed; homi = homicide; other = all other causes. All controls are the controls from column 5 of Table 2. Regressions are weighted by the number of births to nonwhite and white mothers. Standard errors clustered at county level. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 6: IV Estimates By Cause, 1 month

	(1) IMR ca	(2) IMR lbw	(3) IMR resp	(4) IMR pp	(5) IMR sidsassb	(6) IMR homi	(7) IMR other
Air Lead Concentration	0.040	0.312**	0.081	-0.010	0.076***	-0.011	0.260**
	(0.087)	(0.142)	(0.110)	(0.172)	(0.025)	(0.008)	(0.128)
KPFstat	26.079	26.079	26.079	26.079	26.079	26.079	26.079
DepMean	1.018	1.203	.556	1.731	.059	.008	.973
CountyYear	1470	1470	1470	1470	1470	1470	1470
Counties	121	121	121	121	121	121	121

Notes: ca = congenital anomalies; lbw = low birthweight; resp = respiratory; pp = other conditions originating in perinatal period; sidsassb = sudden infant death syndrome and accidental suffocation and strangulation on bed; homi = homicide; other = all other causes. All controls are the controls from column 5 of Table 2. Regressions are weighted by the number of births to nonwhite and white mothers. Standard errors clustered at county level. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table 7: IV Estimates of Other Infant Health Outcomes

	(1) Premature	(2) Bthweight	(3) Lowbw
Air Lead Concentration	1.315	1.853	0.779
	(2.583)	(7.104)	(2.281)
KPFstat	41.232	41.232	41.232
DepMean	60.804	3297.398	76.858
CountyYear	1553	1553	1553
Counties	127	127	127
AllControls	Y	Y	Y

Notes: This table reports regressions on premature (gestation weeks < 36) per 1,000 live births, birth weight (in grams), and low birth weight ($<$ 2,500g) per 1,000 live births. Means of the dependent variables are reported under the coefficients. Regressions control for the full set of other controls described in Table 2. Standard errors are clustered at county level.

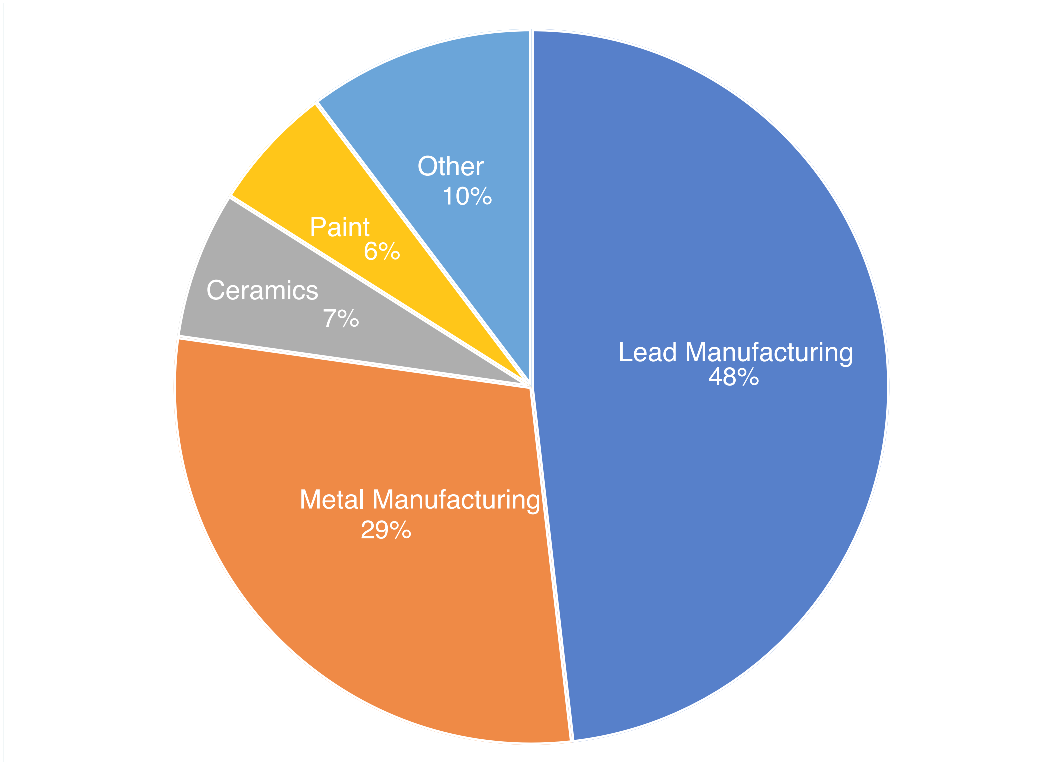
Table 8: Back of the Envelope Calculations

	(1) Annual Infant Deaths Averted	(2) Value in 2018$
IV, Fug Lead	59	$533 million
Red Form, Fug Lead	34	$313 million
IV, All Lead	218	$2 billion

Notes: The IV, Fugitive lead estimates are generated by using the average of fugitive lead from 1988-1991 and 2015-2018 in the first stage to generate the reduction in air lead concentration. This is then multiplied by the coefficient on air lead concentration in column 5 of Table 2 and by 1 million to get the annual infant deaths averted. The reduced form estimates the average of fugitive lead from 1988-1991 and 2015-2018 to generate the reduction in infant deaths based on the regression underlying Appendix Figure A.10. The IV, all lead estimates are generated by using the average decline in air lead concentration from 1988-1991 and 2015-2018. This is multiplied by the coefficient on air lead concentration in column 5 of Table 2 and by 1 million to get the annual infant deaths averted. Deaths are multiplied by the EPA valuation of $9.2 million per death averted (2018 USD) to get values.

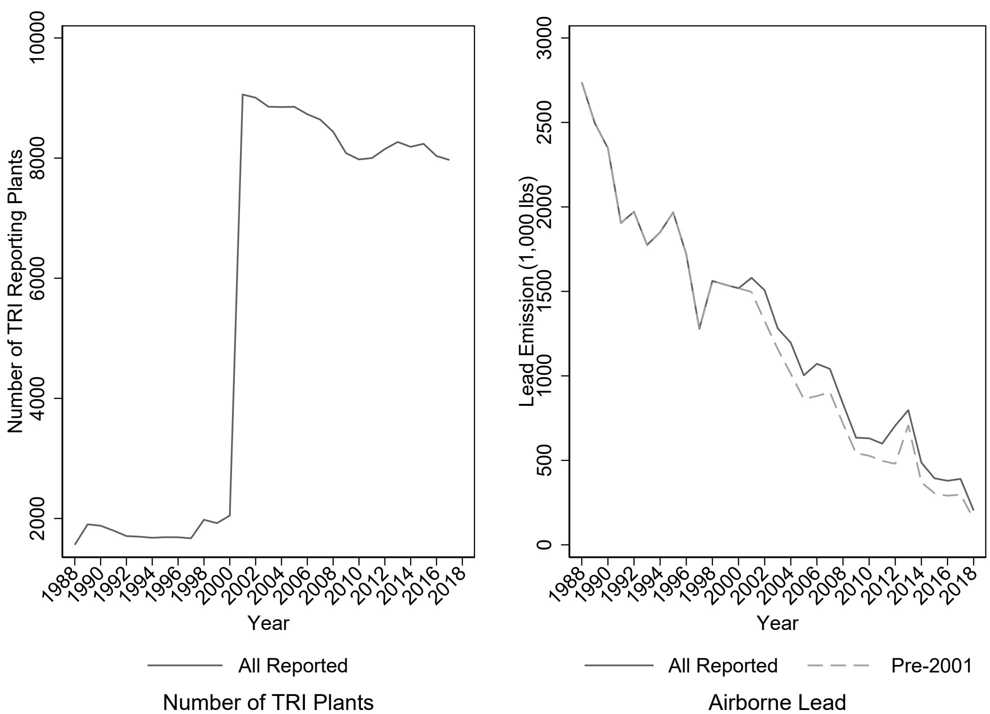
A Appendix Figures and Tables

Figure A.1: Industry Distribution of Air Lead Emissions of Facilities



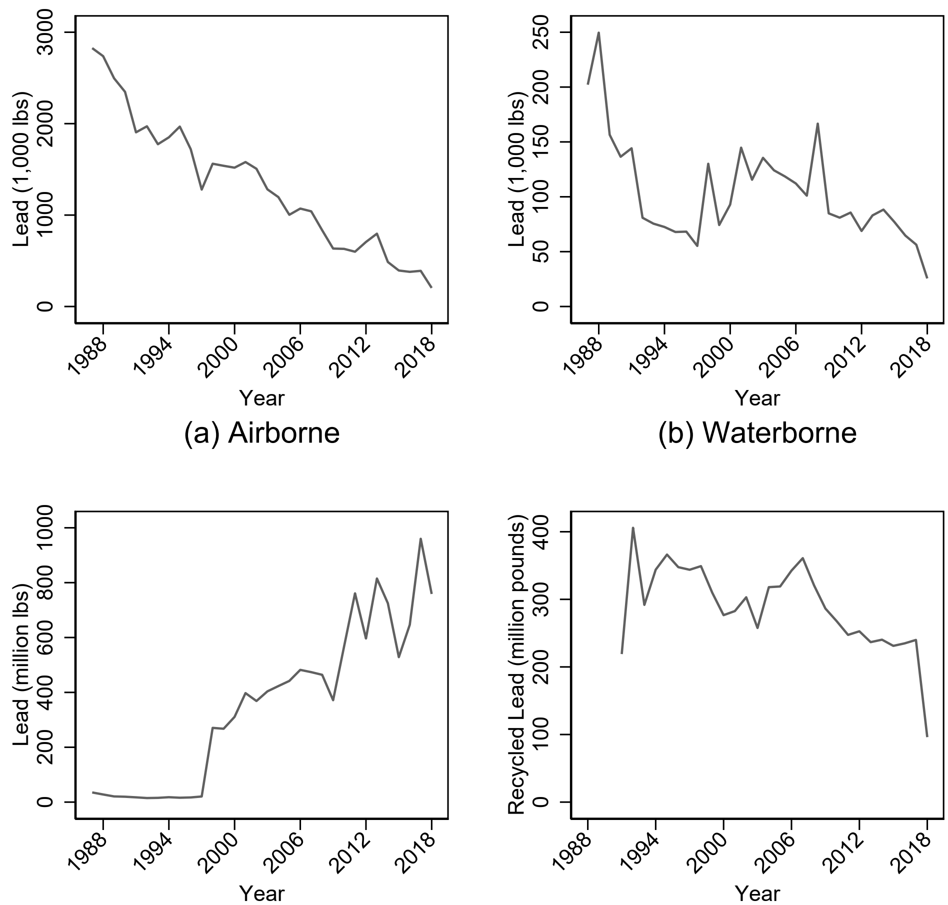
Notes: This pie-chart shows the industry distribution of airborne lead emissions (sum over time) by the sampling industrial facilities. The emissions include both fugitive and stack lead emissions. Calculates are weighted by the number of births in a county.

Figure A.2: Changes in Reporting in 2001



Notes: This figure shows trends in the number of reporting plants and their effect on air lead emissions in our sample during 1988 to 2018.

Figure A.3: Lead Emissions from Air, Water, and Land-borne Sources and Recycled Lead



(a) Airborne

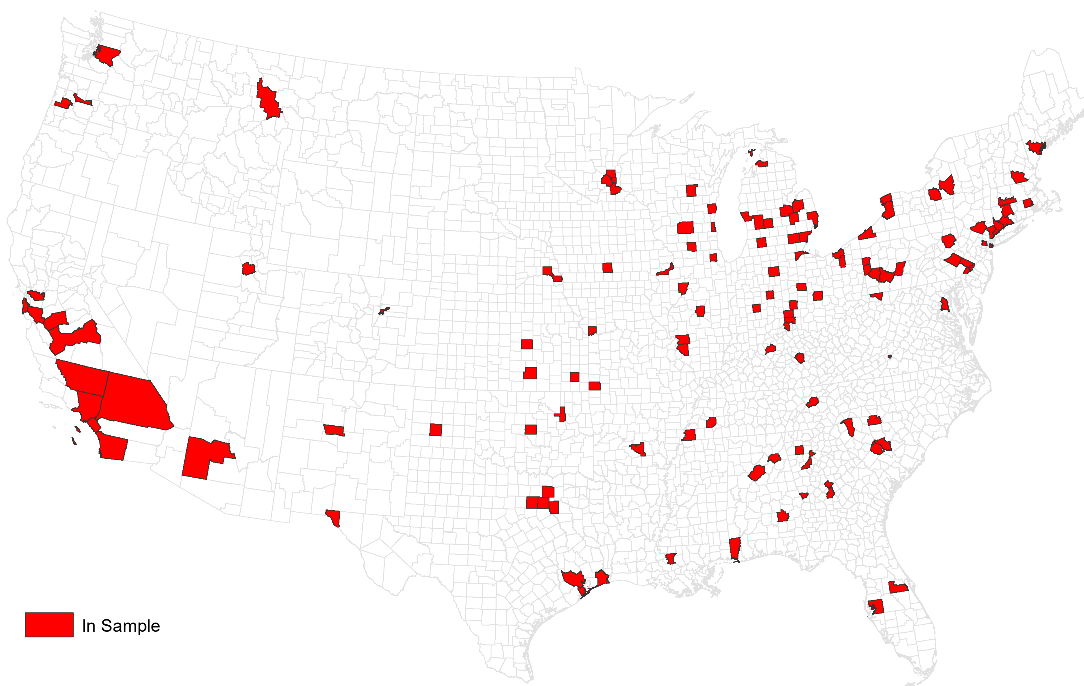
(b) Waterborne

(c) Landborne

(d) Recycled

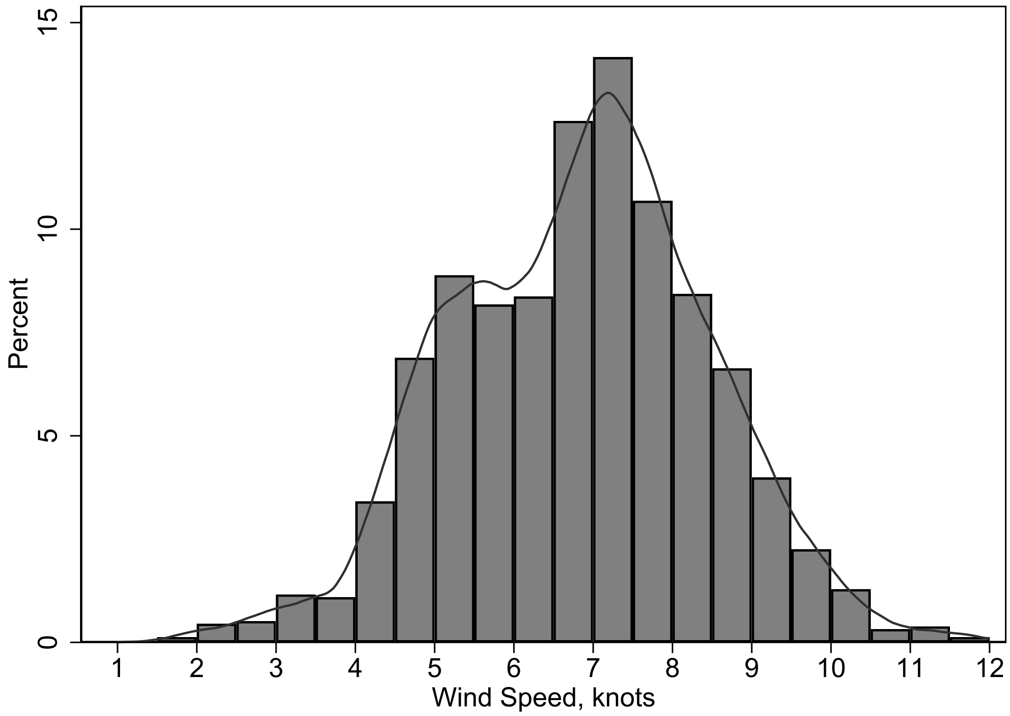
Notes: Figure (a) to (c) plot the trends of air-, water-, and land-borne lead emissions by TRI plants. Figure (d) plots the lead recycled from production waste. Data on recycled lead started from 1991, following the Pollution Prevention Act (1990) that expanded TRI to include additional information on toxic chemicals in waste and on source reduction methods.

Figure A.4: Geographic Distribution of Counties in the IV Sample



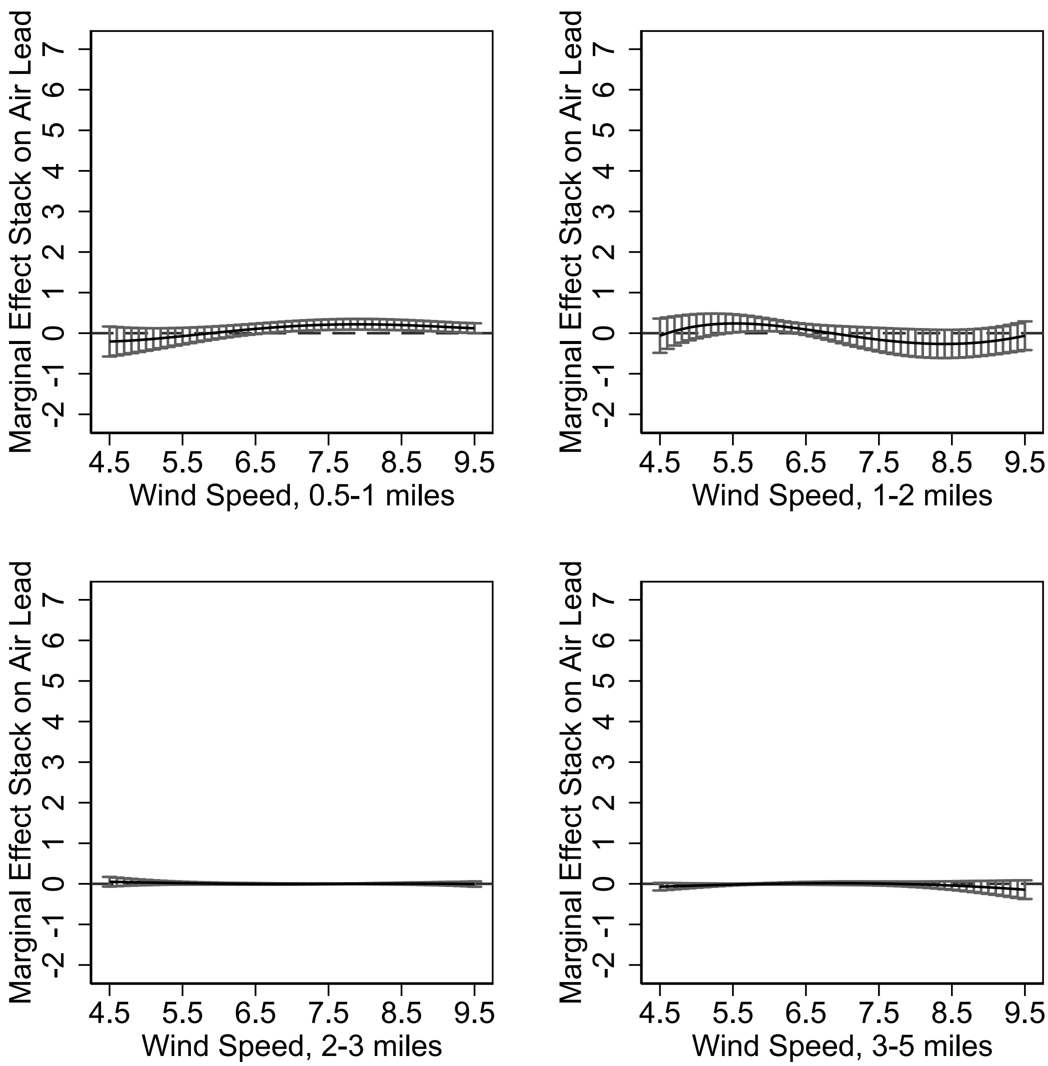
Notes: This map presents locations of the 127 counties in our sample.

Figure A.5: Histogram of Wind Speeds



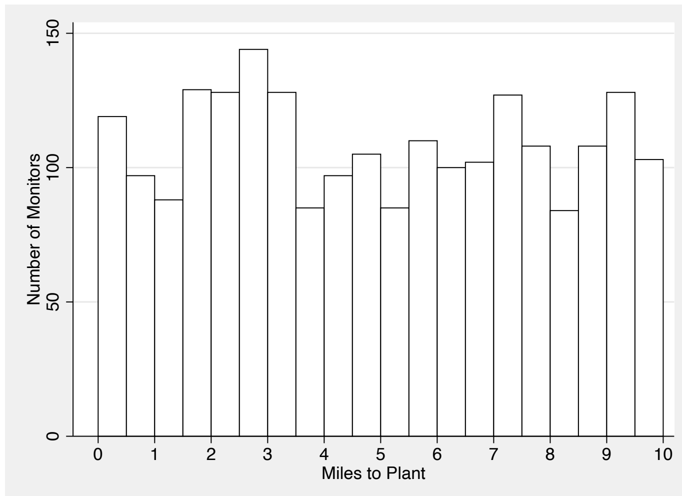
Notes: This figure plots the distribution of wind speeds.

Figure A.6: Effect of Stack Lead Emissions on Air Lead Concentration by Wind Speed

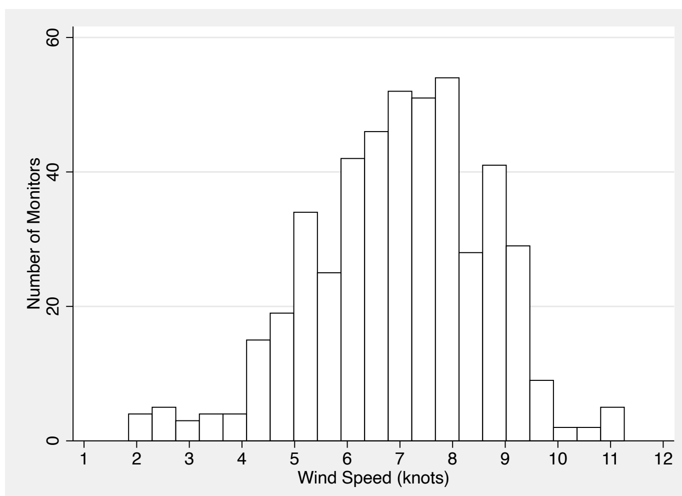


Notes: This figure reports the estimated coefficient ϕ^S in equation 1 - the marginal effect of stack lead emissions on ambient lead concentration - as a function of wind speed (in knots) for different distance ranges from monitors to plants.

Figure A.7: Number of Lead Monitors by Distance to Plants and by Wind Speed



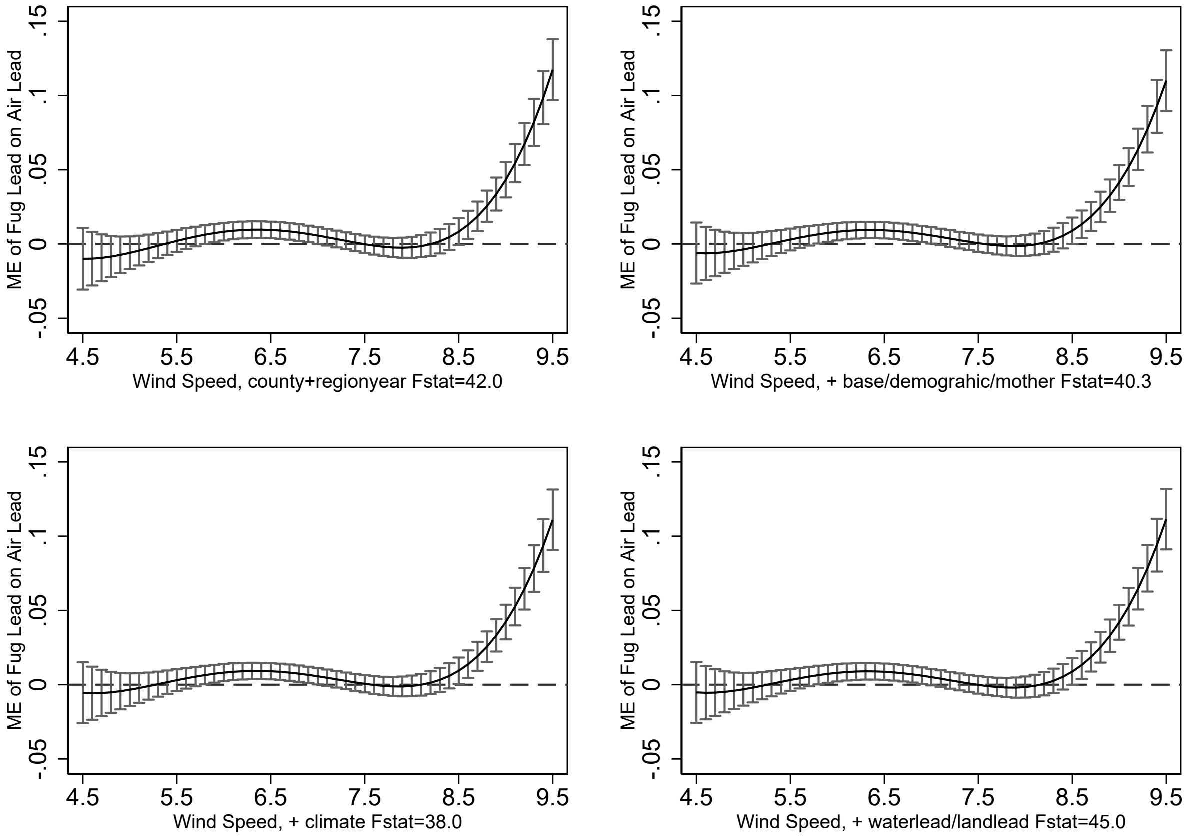
(a) By Distance to Lead Plants



(b) By Wind Speed (within 10 miles)

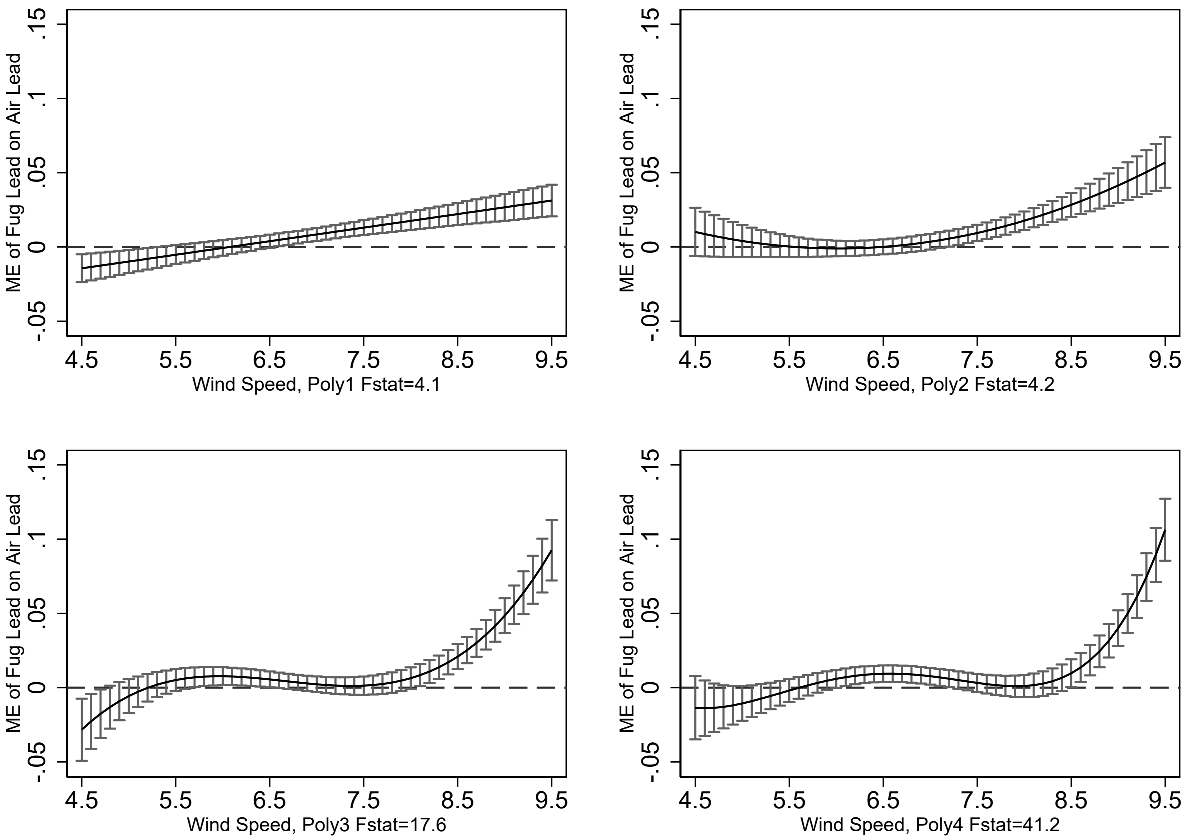
Notes: Figure A.7a plots the number of monitors in 0.5 mile increments over miles from monitor to plant. Figure A.7b plots the number of monitors in 1 knot increments over the average wind speed within 10 miles from the plant.

Figure A.8: First Stage Adding Controls



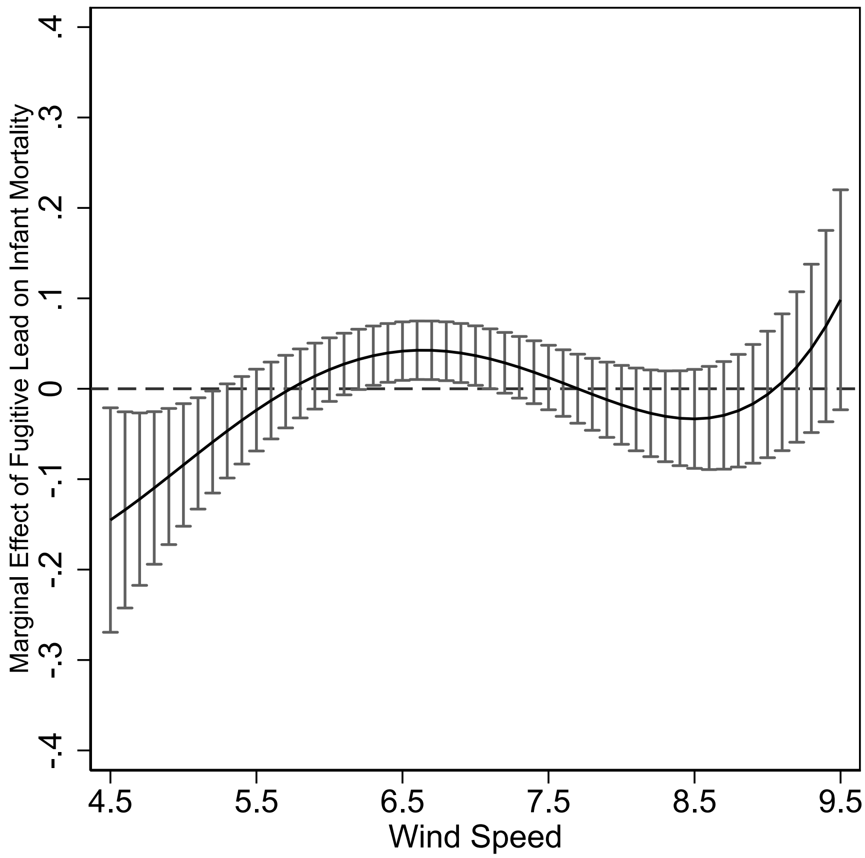
Notes: The figure plots the marginal effect of fugitive lead emissions (F_{ct}) of plants on the air lead concentration readings at lead monitors as different sets of controls are added. These first stages correspond to columns 1-4 in Table 2.

Figure A.9: First Stage Polynomial



Notes: The figure plots the marginal effect of fugitive lead emissions (F_{ct}) of plants on the air lead concentration readings at lead monitors as the wind polynomial is increased from 1 to 4. These first stages include the full set of controls from columns 5 in Table 2.

Figure A.10: Reduced Form Specification



Notes: This figure reports the marginal effect of fugitive lead emissions on infant mortality as a function of wind speed (in knots). The specification is the reduced form version of the specification in column 5 of Table 2.

Table A.1: County Summary Statistics

	mean	sd
IMR 1yr, per 1,000	7.72	2.72
IMR 1mo, per 1,000	5.10	1.86
IMR Nonwh, per 1,000	11.10	4.90
IMR White, per 1,000	6.11	1.68
Premature, per 1,000	60.80	22.65
Birthweight, grams	3297.40	62.06
Low Bthwt, per 1,000	76.86	15.79
Births	51784.52	59008.90
Air Fug Lead, 1,000 lbs	1.71	3.79
Air Stack Lead, 1,000 lbs	4.65	12.48
Air Lead Conc.	0.08	0.22
Windspeed, knots	6.47	1.53
County Pop Density	2335.09	3693.82
County HH Income	48594.42	11813.09
County Pct Mfg Employ	14.05	4.97
County Pct Employ	91.62	2.76
County Pct White	66.07	13.86
County Pct HSchool	89.22	5.63
Mother White	0.74	0.14
Mother Hispanic	0.33	0.26
Mother Age over 35	0.13	0.05
Mother High School	0.71	0.17
Avg Temp, C	15.08	4.08
Avg Precip, Mm	797.92	471.79
Water Lead, 1,000 lbs	0.34	1.20
Land Lead, 1,000 lbs	35.53	268.35
Air Fug Dev, 1,000 lbs	39.69	113.32
Air Stack Dev, 1,000 lbs	96.69	236.71
Air Fug NDev, 1,000 lbs	274.40	751.98
Air Stack NDev, 1,000 lbs	767.85	1326.10
Air Fug HAP, 1,000 lbs	92.01	246.04
Air Stack HAP, 1,000 lbs	237.59	670.35
Observations	1553	

Table A.2: Prediction of County Characteristics

	(1) MFemp	(2) Emp	(3) Popdens	(4) Hinc	(5) Pwhite	(6) Phighsch
KPFstat	3.772	1.235	5.659	3.905	1.139	0.552
CountyYear	1553	1553	1553	1553	1553	1553
Counties	127	127	127	127	127	127
Allcontrols	Y	Y	Y	Y	Y	Y

Notes: The dependent variable is listed in the column header: (MFemp) percent in manufacturing employment; (Emp) percent employed; (Popdens) population density per square mile; (Hinc) median household income; (Pwhite) percent white; (Phighsch) percent with high school education over people above age 25. All controls are the controls from column 5 of Table 2, but excludes the dependent variable. F-statistics are the joint significance of fugitive lead interacted with the wind variables. Regressions are weighted by the number of births. Standard errors clustered at county level. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.3: Prediction of Mothers' Characteristics

	(1) Pmwhite	(2) Pmhisp	(3) Pmolder35	(4) Pmhighsch
KPFstat	1.568	0.630	0.783	0.686
CountyYear	1553	1553	1553	1553
Counties	127	127	127	127
Allcontrols	Y	Y	Y	Y

Notes: The dependent variable is listed in the column header: (Pmwhite) percent mothers white; (Pmhisp) percent mothers hispanic; (Pmolder35) percent mothers older than 35; (Pmhighsch) percent mothers with high school education. All controls are the controls from column 5 of Table 2, but excludes the dependent variable. F-statistics are the joint significance of fugitive lead interacted with the wind variables. Regressions are weighted by the number of births. Standard errors clustered at county level. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.4: Other Fugitive Chemicals, PM10, and CO

	(1) IMR	(2) IMR	(3) IMR	(4) IMR	(5) IMR	(6) IMR	(7) IMR
Air Lead Conc.	1.529***	1.546***	1.676***	1.901**	1.725**	2.343**	0.940
	(0.458)	(0.411)	(0.435)	(0.804)	(0.842)	(0.915)	(1.640)
KPFstat	35.778	34.202	41.232	34.107	19.051	11.120	1.497
DepMean	7.718	7.718	7.718	7.782	7.753	7.718	7.718
CountyYear	1553	1553	1553	1282	1046	1553	1553
Counties	127	127	127	105	81	127	127
Allothercontrols	Y	Y	Y	Y	Y	Y	Y
DevChem	Y	Y	Y	Y	Y	Y	Y
NonDevChem		Y	Y	Y	Y	Y	Y
HAP			Y	Y	Y	Y	Y
PM10				Y	Y		
CO					Y		
Metal2						Y	Y
Zinc							Y

Notes: This table reports the results of adding controls for other chemicals and pollutants. The dependent variable of all regressions are infant mortality rate within the first year of births. All other controls are the controls from column 5 of Table 2, but excludes other chemicals. F-statistics are the joint significance of fugitive lead interacted with the wind variables. Regressions are weighted by the number of births. Standard errors clustered at county level. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.

Table A.5: Robustness Checks

	(1) IMR Base	(2) IMR Nozerofug	(3) IMR OtherPlantE	(4) IMR 10years	(5) IMR To2013	(6) IMR To2008
Air Lead Concentration	1.676***	1.695***	1.537***	1.603***	1.798***	1.728***
	(0.435)	(0.441)	(0.455)	(0.527)	(0.471)	(0.493)
KPFstat	41.231	41.278	32.545	28.347	41.259	38.382
CountyYear	1553	1534	1550	1153	1344	1112
Counties	127	122	127	57	126	106
Allcontrols	Y	Y	Y	Y	Y	Y

Notes: This table reports the results of several robustness exercises. The dependent variable of all regressions are infant mortality rate within the first year of births. Column 1 is the baseline (col. 5, Table 2). All controls are the controls from column 5 of Table 2. Column 2 drops counties that report stack emissions but always report zero fugitive emissions. Column 3 controls for chemical emissions of other non-lead emitting plants in the county. Column 4 only includes counties with at least 10 years of lead monitor data in the sample. Columns 5 and 6 shorten the sample from 1988-2018 to 1988-2013 and to 1988-2008. F-statistics are the joint significance of fugitive lead interacted with the wind variables. Regressions are weighted by the number of births. Standard errors clustered at county level. *** denotes statistical significance at the 1 percent level, ** at the 5 percent level, and * at the 10 percent level.