DISSERTATION

HEALTH AND HUMAN CAPITAL EFFECTS OF LEAD EXPOSURE

Submitted by

Christopher James Keyes

Department of Economics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Fall 2020

Doctoral Committee:

Advisor: Sammy Zahran

David Mushinski
Harvey Cutler
Dale Manning

ABSTRACT

HEALTH AND HUMAN CAPITAL EFFECTS OF LEAD EXPOSURE

The legacy of lead in the United States is complex and intertwined with public health. As concerns over the toxicity of lead increased with time, policy makers responded with a series of national policies aimed at minimizing the risk of lead exposure across society. One such policy, the Clean Air Act (CAA), set a timeline for the removal of lead from gasoline beginning in 1975. This policy would target the anti-knock lead additive tetraethyl-lead (TEL), which was used to boost gasoline octane and improve engine performance (Needleman, 2000; Reyes, 2007). Over the following two decades, the flow of lead entering the environment from automobile emissions decreased precipitously.

This dissertation exploits a natural experiment in lead exposure arising from the differential phase-out of leaded gasoline across states under the CAA. Though the policy was implemented at the national level, enforcement took place at the producer level, creating exogenous variation in lead emissions from automobile exhaust across states and over time. Since lead dust from automobile emissions was a significant source of lead exposure over the period, we leverage this spatial and temporal variation as a quasi-random vector of lead exposure.

Chapter one summarizes the CAA, and the historical significance of the policy as it relates to public health. Using blood lead levels (BLLs) from The Second National Health Nutrition and Exercise Survey (NHANESII) as a bio-marker for lead exposure, this paper models the lead exposure effect of the policy. Combining annual gasoline sales and gasoline lead concentrations at the state level, the steps taken to construct the variables proxying for lead exposure following the CAA are detailed at length. The empirical strategy applied in this chapter is used to identify the causal effect of the phase out on lead exposure, and is carried over in the following two chapters.

Much of the research focusing on the effects of lead exposure emphasize the risk faced by children, who are particularly susceptible to even minute quantities in the first five years of life.

Chapter two tests the hypothesis that lead exposure in childhood impacts cognitive ability and the presence of abnormal latent preferences toward risk and uncertainty in adulthood. Applying the identification strategy detailed in Chapter one, to a nationally representative sample of individuals born during a period of significant reductions in leaded gasoline emissions, we find considerable evidence supporting the causal effect of childhood lead exposure and later in life outcomes. Across a series of tests, we find that BLLs in childhood are a significant predictor of: 1) IQ loss, measured with standardized test scores; 2) increased likelihood of low-IQ outcomes in exposure levels; and 3) increased abnormal risk response across a series of situations involving uncertain outcomes. The results presented in this paper illustrate the significance and persistent affect of early in life lead exposure.

An underappreciated medium of child exposure to flow and legacy sources of lead is in-utero transmission of lead from mother to infant. Transmission of lead to the fetus occurs via diffusion across the placental barrier over the course of a pregnancy. Chapter three estimates the causal effect of maternal lead exposure on birth outcomes during the initial period of the phase out. Results show consistent evidence that fetal exposure to lead through the maternal blood lead pathway significantly depresses infant health. Our findings suggest that an increase in maternal blood lead: 1) decreases infant birthweight; 2) increases the risk of low and very low birthweight; 3) shortens gestation length; 4) increases the risk of prematurity; and 5) increases the risk of a low APGAR score. A back of the envelope calculation of the economic benefits of the phase-out of leaded gasoline through the reduction of healthcare-related costs involved in treating low birthweight infants, are in the tens of billions annually.

It might be tempting to assume that lead exposure is a rear-view problem, at least in the United States, as BLLs in children have fallen since the 1990s, coincident with a series of actions that banned lead from paint, plumbing, food cans and automotive gasoline. However, the flow of lead into the environment continues from various point source polluters as well emissions from aviation gasoline used by an estimated 160,000 piston-engine aircraft (Kessler, 2013). Though the benefits to public health attributable to national policies are immense, the stock of legacy lead and present day flow sources of environmental lead remain a persistent threat to public health.

ACKNOWLEDGEMENTS

I would first like to express my deepest gratitude to my advisor and mentor, Dr. Sammy Zahran, for his guidance and support. If it were not for his encouragement, I may never have set out on this path, and it is with great pleasure I am able to acknowledge in this dissertation the significant impact he has had on my life.

Many thanks to Dr. Dave Mushinski for helping me work through the econometric questions I could not resolve alone, and always leaving me with more questions to explore than when I arrived.

I gratefully acknowledge the funding received to undertake this research from the Economics Graduate Committee research fellowship. Thank you to the Economics Department and Fan family, I am deeply grateful for the funding received through the L.S.and Chuen-mei Fan Scholarship.

Thanks to Dr. Stephan Weiler for his enthusiasm and invaluable feedback during the early stages of my research. I would also like to acknowledge Harvey Cutler and Dale Manning for their input and feedback throughout the writing of this dissertation.

And finally to Olive, for faithfully remaining by my side throughout this endeavor. Without your unwavering support, I may never have made it through this challenging period.

DEDICATION

This dissertation is dedicated to my loving parents, James and Karen Keyes.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Wisdom is not a product of schooling but a lifelong attempt to acquire it.

Albert Einstein

Chapter 1

The Clean Air Act and the phase out of leaded gasoline: The effects of national policy on lead exposure and automobile emissions

1.1 Introduction

As modern science has deepened our understanding of the toxicity of lead, it has motivated national policies aimed at minimizing the public health hazard this element poses to society. This chapter will retrace the path over which the Clean Air Act (CAA) of 1970 would phase out the use of lead additives in automobile gasoline sold to consumers. The direct result of the policy was significant reductions in the volume of atmospheric lead and immense benefits to public health as the risk of lead exposure would decline for all socioeconomic and demographic groups. Though national policies such as the CAA have resulted in significant reductions in the quantities of lead entering the environment from the consumption of products containing lead, the legacy of past use continues to impact people to this day, and will likely extend well into the future.

To identify the causal channel linking gasoline lead to lead exposure, we exploit the phase out of lead from gasoline as an exogenous source of variation in lead emissions at a national scale. Under the CAA, a time line for the removal of lead from gasoline was implemented beginning in 1975. Over the following two decades, lead entering the environment from automobile emissions would fall precipitously. Though the policy was enforced at the national level, the incentive structure for compliance, as well as the characteristics of the petroleum and automobile industries, would result in significant variation in lead emissions across states between 1975 and 1990. Since lead dust resulting from the consumption of leaded gasoline was among the primary vectors of lead expo-

sure in children over this period,[1] we use state and year variation in gasoline lead concentrations to identify our empirical model, and proxy for this exposure channel.

This paper contributes to the body of literature investigating the link between lead exposure and environmental externalities. Modeling the causal relationship between automobile emissions and lead exposure at a national scale, our findings are a novel contribution to prior research in this domain. Results from this analysis show that gasoline lead levels are a statistically significant predictor of blood lead levels during the period following the CAA. At a national level, the estimated marginal effect of reductions in gasoline lead concentrations is a decline in blood lead levels. We show these results persist across a series of sensitivity and robustness tests, as well as for both children and adults. Furthermore, to demonstrate the robustness of this identification strategy, we include the results of a divergent validity test of the identifying assumptions underlying the empirical analysis in this chapter and those that follow. Across this ensemble of evidence, our findings illustrate the considerable improvements in public health following the CAA.

The remainder of the chapter is laid out as follows: Section 1.2 details the widespread use of lead in the United States as it relates to public health and national policies. A summary of the data used and empirical methods are found in Section 1.3. Results are presented in Section 1.4 and the paper closes with a discussion and conclusion as they relate to the following chapters are found in Section 1.5.

1.2 Lead use and national policy

Historically, lead has been used in a variety of applications due to its desirable chemical properties, resulting in widespread use in the manufacturing of goods and productive services (Nevin, 2000; Needleman, 2004; Papanikolaou et al., 2005). Two use cases which accounted for nearly all lead used in the United States over the last century were as an additive to paint and gasoline. Figure 1.1 illustrates the magnitude of demand for these products as total tonnages during the 20^{th} century. Together, the consumption of these products resulted in the distribution of lead across the United States.

[1] See Bellinger and Bellinger (2006), Abadin et al. (2007), Mielke and Reagan (1998) among others.

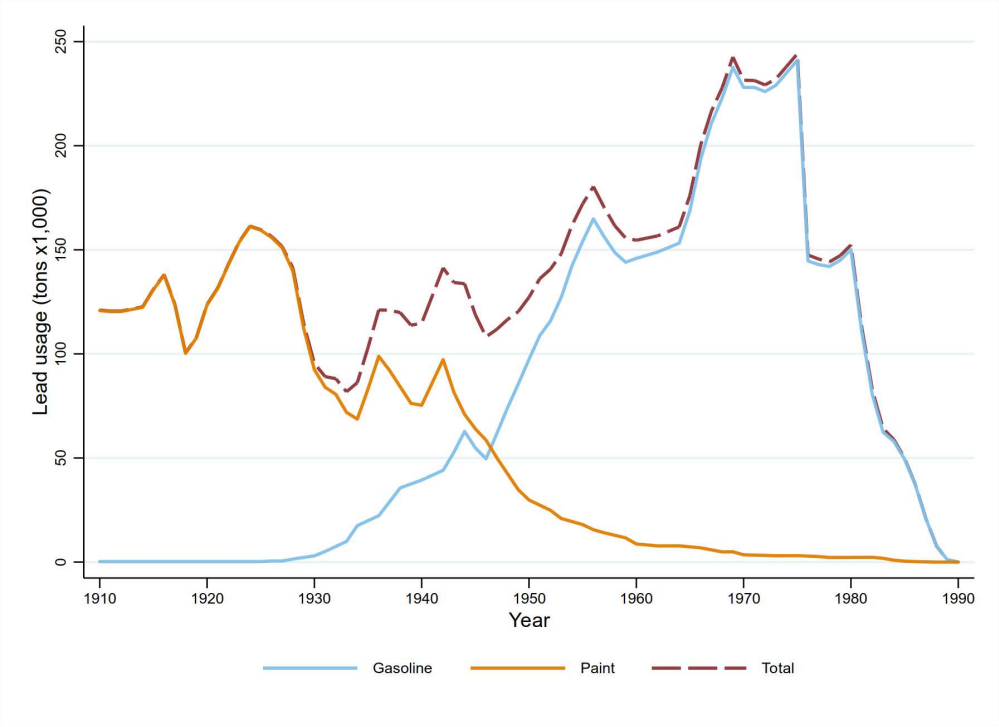


Figure 1.1: US lead use, 20^{th} century (source: Laidlaw and Filippelli (2008))

The period of U.S. history in Figure 1.1 is historically significant for two reasons. First, the emergence of the automobile industry, along with the completion of an interstate highway system, created significant increases in demand for petroleum products, specifically automobile gasoline. At the time, gasoline sold to consumers was formulated with tetraethyl-lead (TEL), an anti-knock additive used to boost octane and improve engine performance (Needleman, 2000; Reyes, 2007). The steady increase in lead production shown in Figure 1.1 between 1930 to 1970 was driven by demand for leaded gasoline and accounts for the majority of the flow of lead entering the environment. Mielke and Reagan (1998) have estimated that the amount of lead used in gasoline between 1929 to 1990 is about as much as the amount used in paint between 1884 to 1989.

Second, this period marked a shift in public opinion with respect to the toxicity of lead and the hazard this element posed to public health. A growing body of evidence linking lead to adverse health outcomes, particularly in children, would ultimately result in policies aimed at minimizing the health risks created by leaded products (Needleman, 2000; Papanikolaou et al., 2005; Bellinger and Bellinger, 2006). The first set of policies targeted leaded paint, and resulted in the removal

of lead from interior paints in 1971, and all paints in 1978 (Mielke and Reagan, 1998). Shortly thereafter, the CAA would target air pollution and regulate lead emissions from the use of leaded gasoline. As evident in Figure 1.1, the dramatic decline in paint beginning around 1925 is largely attributable to the public's awareness of the hazards lead paint created for children, and the decline in gasoline beginning around 1975 attributable to national policy.

1.2.1 Lead exposure and health outcomes

Perhaps the most underappreciated source of lead exposure is contaminated soil or dust. In the United States, the most common channel of non-occupational lead exposure is through the ingestion or inhalation of lead dust (Abadin et al., 2007), and the risk of exposure through this vector of exposure is widespread. Urban soils integrate all dust sources of Pb including lead-based paint (either deteriorated, haphazardly removed by power sanding, sand blasted, scraped without capture, or released by building demolition), lead additives in vehicle fuel emissions, and industrial Pb emissions (Mielke et al., 1999; Mielke and Reagan, 1998; Mielke and Zahran, 2012; Farfel et al., 2005a; Rabito et al., 2007). The combustion of leaded gasoline in automobiles emits Pb particulates, with roughly 75% of the lead in gasoline released in the exhaust (Mielke and Reagan, 1998). In all cases, the Pb dust coats surfaces as a residue and mixes with topsoil. Soil Pb at or near the surface is an exposure risk to humans through direct contact or re-suspension of Pb in contaminated soils during summer periods (Reagan and Silbergeld, 1989; Filippelli et al., 2005; Laidlaw et al., 2005, 2012; Zahran et al., 2013).

The magnitude of automobile emissions as a flow source of atmospheric Pb is illustrated in Figure 1.2, which graphs national automobile lead emissions between 1950 to 1990. Prior to the phase out, the flow of atmospheric Pb from TEL emissions would exceed 200 thousand metric tons annually. In the four year period from 1976 to 1980, identified by the verticle bars in Figure 1.2, all states experienced reductions in gasoline-related Pb emissions of at least 40%, largely driven by reductions in the TEL concentration of gasoline sold. Given the tendency for lead dust to re-mobilize, lead exposure attributable to automobile emissions is not limited to those living in the vicinity of roadways or the source of emissions itself (Zahran et al., 2013; Curci and Masera,

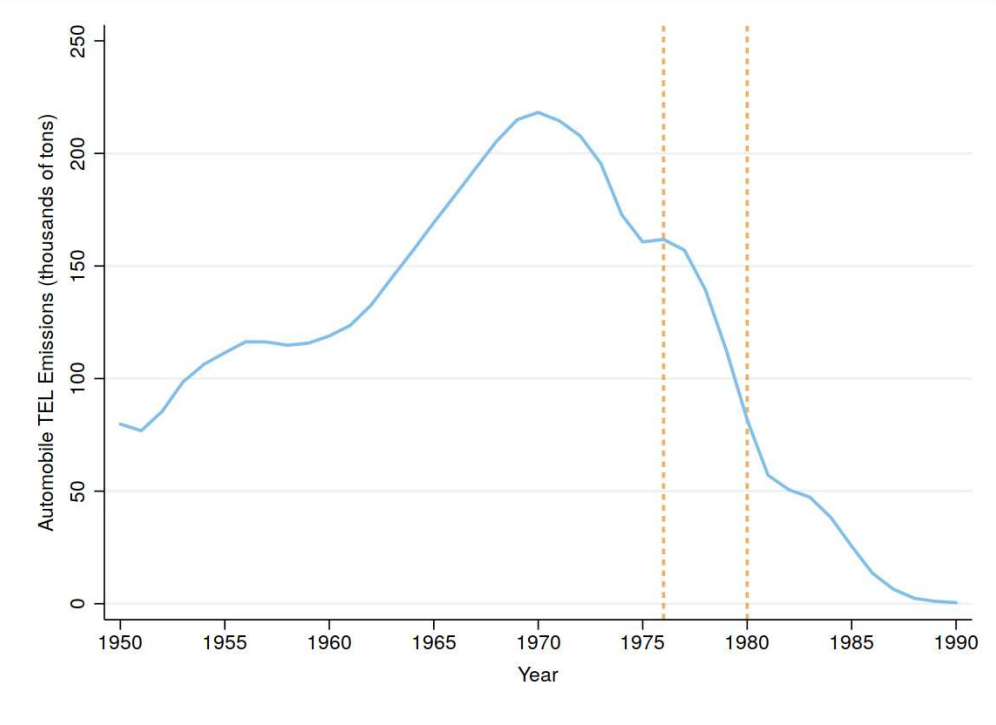


Figure 1.2: US total gasoline lead (TEL) emissions, metric tons

2017). Compared to lead paint, which to date is largely confined to old homes in low-income areas, legacy lead from automobile emissions is not limited to geographic or socioeconomic groups, nor is it simply a rear-view problem.

While both paint and gasoline have left legacy lead, which remains a persistent threat to public health, the risk posed by each source can vary significantly. When comparing the amount of lead absorbed by the body after being ingested, the amount entering the bloodstream is non-linear and inversely related to the size of the particulate consumed (Mielke and Reagan, 1998; Papanikolaou et al., 2005; Bellinger and Bellinger, 2006). In contrast to deteriorating paint, the health hazard posed by the lead dust generated by automobile emissions is of much finer granularity, and readily mobilized once in the body (Mielke and Reagan, 1998; Papanikolaou et al., 2005). So, while paint chips are a high dose source with respect to lead concentrations, there is a growing consensus that lead dust created by automobile emissions is a greater hazard to public health (Needleman, 2004; Toscano and Guilarte, 2005; Nilsson, 2009; Miranda et al., 2011).

The impact of lead entering the body through ingestion or inhalation varies depending on the physical characteristics of the individual, and importantly, one's age. For a given quantity of lead ingested, the percentage which is absorbed into the bloodstream for adults is roughly 6%, and upwards of 50% for children (Clay et al., 2018). Lead which does not make its way into the blood stream is passed in solid waste or urine. While adults pass nearly all the lead entering their bodies within several weeks, less than 40% leaves the body of a child during that time (Abadin et al., 2007). Lead which has entered the bloodstream is then transported throughout the body and deposited in soft tissue, bones, or blood plasma. The half life of lead in the bloodstream is estimated to be 35 days, and is expected to be longer in children (Needleman, 2004; Papanikolaou et al., 2005; Toscano and Guilarte, 2005). Lead in organs and soft tissues has a half-life of approximately 40 days (Papanikolaou et al., 2005), though deposits in the brain have been found to have a half-life of roughly two years (Lidsky and Schneider, 2003). Lead which is deposited in bones and teeth, can remain for up to 30 years, and has been shown to re-mobilize during periods of reduction in bone density and re-enter the bloodstream (Needleman, 2004; Papanikolaou et al., 2005). For example, studies have found elevated blood lead levels (BLLs) in women during pregnancy and while nursing (Rothenberg et al., 1994; Papanikolaou et al., 2005), as well as postmenopausal osteoporosis (Bellinger et al., 1987; Needleman, 2004; Bellinger, 2011).

Though the medical profession has had a basic understanding of the toxicology of lead dating back to antiquity (Needleman, 2004; Toscano and Guilarte, 2005; Bellinger and Bellinger, 2006), recent studies continue to present new evidence on the extent to which even low-level lead exposure impacts health. While lead poisoning occurs at high levels of exposure, typically observed at BLLs in excess of 50 micro-grams per deciliter of (μg/dL) (Abadin et al., 2007), evidence shows that moderate and low-level exposure still poses an immediate hazard, with lasting adverse health outcomes. Researchers have linked conditions such as hypertension, renal impairment, as well as a number of cardiovascular complications in adults to moderate levels of lead exposure, ranging from 20 to 40 μg/dL of blood (Abadin et al., 2007; Bellinger, 2011). In children, the dose-responsiveness of lead and adverse health outcomes is more severe. In 2012, the Centers for Disease Control and Prevention (CDC) updated their BLL reference value for at-risk children. In

response to recommendations made by The Advisory Committee on Childhood Lead Poisoning Prevention, the CDC set what is now the accepted BLL threshold value of 5 μg/dL, half if the previous value, acknowledging that to date "no safe blood lead level in children has been identified" (CDC, 2019). This perspective is echoed in a number of studies which emphasize the marginal impacts of increases in blood lead levels on IQ loss, school performance, and neuropsychological impairment (Bellinger et al., 1987; Needleman and Bellinger, 1991; Toscano and Guilarte, 2005; Bellinger, 2008). Bellinger (2008) finds children with prior acute lead exposure exhibit abnormal behaviors such as irritability, impatience, and aggression which persist after BLLs return to pre-exposure levels. Moreover, preliminary evidence suggests that the intellectual impairments caused by early childhood lead exposure persist into adulthood. Reuben et al. (2017) report that adults raised in the era of widespread leaded gasoline use experienced significant reductions in IQ and deficits in perceptual reasoning and working memory, with a strong positive relationship between childhood blood lead levels and cognitive loss in adulthood.

1.2.2 Variations in lead emissions and the Clean Air Act

Though the Clean Air Act set a national standard, compliance was determined at the producer level. As a result, variations in state level gasoline lead concentrations were influenced by a number of factors relating to the petroleum and automobile industries rather than state specific environmental policy (Reyes, 2015). Such factors include refinery locations, transportation and distribution networks, as well as the age of the automobile stock. These characteristics were used by the Petroleum Administration for Defense (PAD) in 1950 to define 17 distinct districts, across five regions within the United States (Shelton, 1979). The boundaries of each PAD district (PADD) are mapped in Figure 1.3, and reflect the geographic definitions recognized during the phase out. Originally defined during World War II for the purposes of distributing oil throughout the county, the PAD districts shown in Figure 1.3 reflect the connection between petroleum industry resources and proximity to population centers.

Evidence of the geographic and temporal variation in gasoline lead attributable to the enforcement of the CAA is shown in Figure 1.4 which follows average gasoline TEL concentrations by

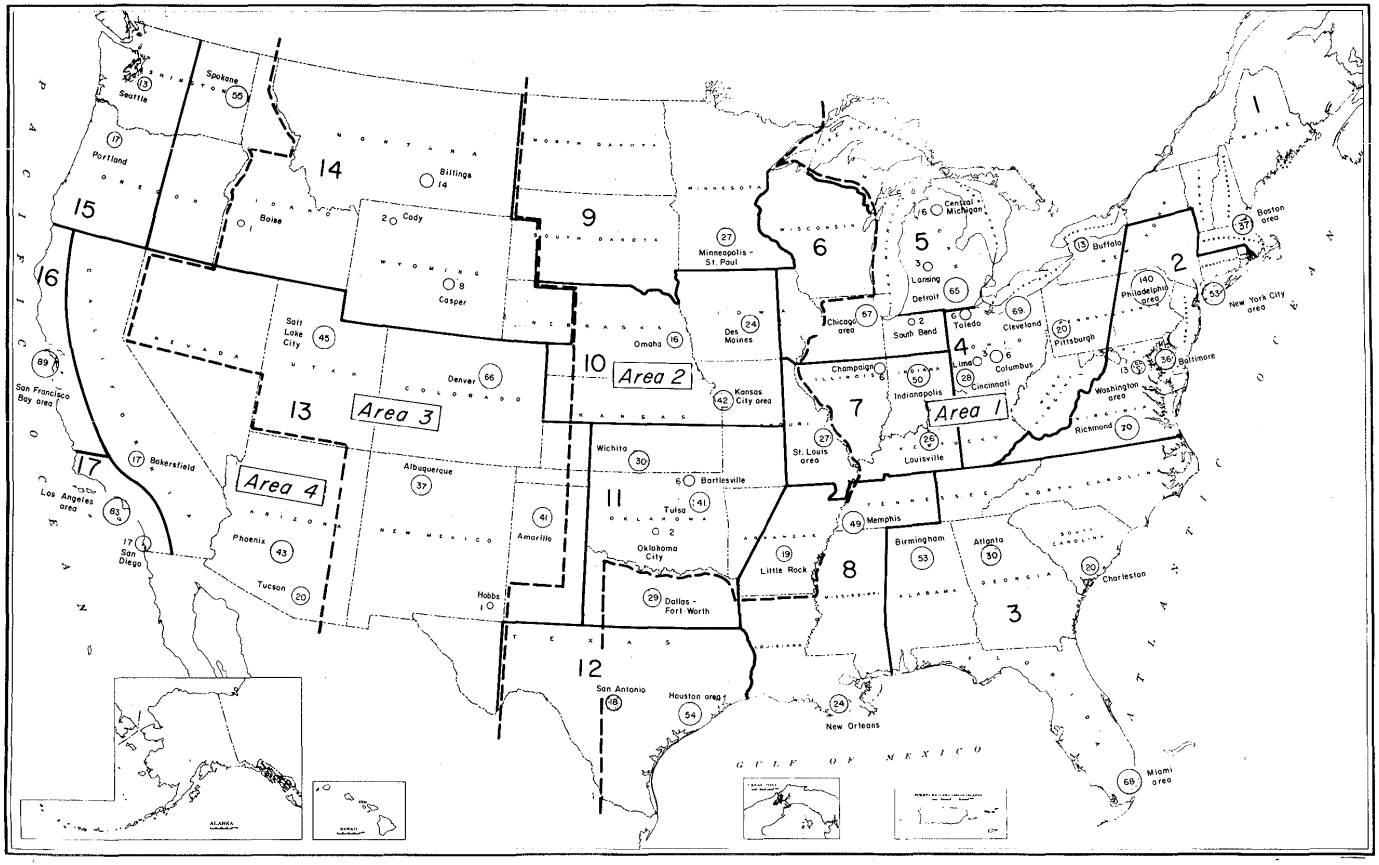


Figure 1.3: Petroleum Administration for Defense (PAD) district boundaries

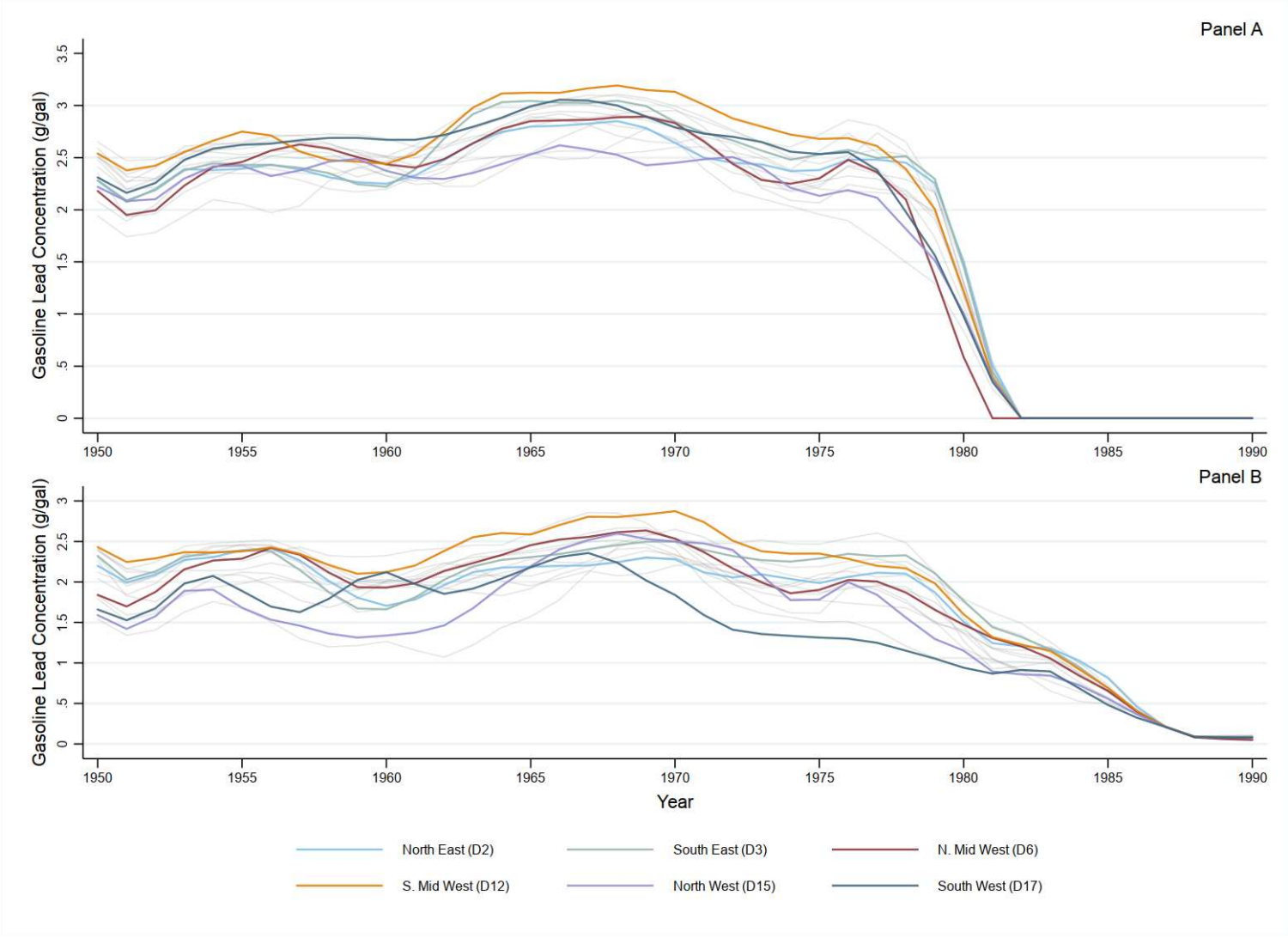


Figure 1.4: Gasoline lead concentrations by PAD district and gasoline grade, 1950-1990

PAD district and leaded grade sold. One PAD district from each of the common geographic regions across the United States has been emphasized for clarity, with remaining PAD districts show in gray. Panel A in Figure 1.4 is restricted to leaded-premium grade gasoline while Panel B is leaded-non-premium gasoline. The onset of the phase out of lead from gasoline in 1975 resulted in the general downward trend in TEL concentrations across all districts and both leaded gasoline grades. Over the four-year period of interest (1976 to 1979), much of the historic phase out of leaded gasoline occurred. Nationally, average gasoline lead concentrations would decline approximately 56% over this period, and 85% between 1976 to 1985. However, the interconnectedness of the petroleum industry to the administrative boundaries shown in Figure 1.3, coupled with the incentive structures of the policy itself, would have a direct affect on gasoline lead concentrations at the local level during the phase out. Authors have argued that together, these forces resulted in random variations in rate of decline in TEL averages across PAD districts during each year of the phase out, evident in Figure 1.4 (Reyes, 2007, 2015). We leverage this quasi-random variation in gasoline lead concentrations, occurring during the years 1976 to 1980, and varying exogenously nationwide, as the identification strategy for estimating the causal relationship between environmental lead from automobile emissions and lead exposure.

1.3 Methods

Based on the findings of Reyes (2007, 2015), petroleum industry experts and historical accounts detailing the characteristics of the petroleum industry, the standards imposed following the CAA induced significant and quasi-random variations in gasoline lead (TEL) concentrations at the state level. Furthermore, given the goal of the CAA and the timeline proposed by the EPA to phase out TEL from gasoline, state level TEL concentrations would vary significantly over time. As gasoline TEL concentrations declined, lead entering the environment from automobile emissions fell. Since automobile emissions were the primary source of lead pollution, the risk of lead exposure would decline in both automobile emissions and gasoline TEL concentrations during the phase out of leaded gasoline. This spatial and temporal variation in gasoline lead concentrations sold to consumers, and ultimately the primary source of lead emissions entering the environment

over this period would have a direct impact on the risk of lead exposure across the U.S.. To estimate the causal effect of the policy shock, we use blood lead data and a time series of annual, state averaged TEL concentrations in the years following the CAA.

1.3.1 Data and measurement

Blood lead

Blood lead measurements are drawn from the Second National Health and Exercise Survey (NHANESII), provided by the CDC. This nationally representative survey was intended to provide a detailed analysis of overall health across the nation (NCHS, 1982). Approximately 20,000 individuals, aged 6 months to 74 years old, were surveyed between February, 1976 and February, 1980. Demographic, socioeconomic, employment and household characteristics were collected alongside a full biochemistry and physiological analysis to deliver a complete inventory of American health at the time.

A subset of respondents were selected for additional testing of hematology and biochemistry, including blood and urine assessments. The selection criteria for laboratory analysis included all children between six months and six years old, along with half of the remaining sample aged 7 to 74 years old (NCHS, 1982). In all, blood was collected and analyzed from 10,049 individuals of the full sample. Descriptive statistics for the sample of NHANESII respondents which were selected and participated in the blood analysis are found in Table 1.1. The outcome of interest is respondent's blood lead levels (BLLs) measured as micro-grams-per-deciliter (μg/dL). Control variables from the data include the respondent's age, sex, racial status, and household income to poverty ratio.[2]

[2]The respondent's age (Age) is measured in years. Indicators for sex and racial status are defined as $Female = 1$ and $Black = 1$ if the respondent identifies as African American. Indicators for income group are constructed from the household income to Federal Poverty Level (FPL) ratio data in the NHANESII data set. We apply the same definitions of Reyes (2015), with the indicator variable Low=1 for household income $< 2 \times$ FPL and Mid=1 for values $\in$ [$2 \times$ FPL, $3 \times$ FPL.

Table 1.1: National Health and Exercise Survey II sample descriptive statistics

	All Ages		Ages 0-6	
Panel A	Mean	Standard Deviation	Mean	Standard Deviation
Demographics				
Age (years)	38.6	20.89	3.4	1.62
Female	51.00%	0.50	0.48%	0.50
Black	12.00%	0.33	18.00%	0.38
Low Income (2x FPL)	34.00%	0.47	50.00%	0.50
Middle Income (2x to 3x FPL)	26.00%	0.44	29.00%	0.45
Lead Variables				
Blood Lead (μg/dL)	13.78	5.88	16.57	7.67
Gasoline Lead (g/gal)	1.46	0.35	1.49	0.35
Sample Size	6,718		2,430	

	Blood Lead (μg/dL)			
Panel B	Mean	Standard Deviation	Mean	Standard Deviation
Gasoline Lead (g/gal)				
Gasoline Lead Q1	12.15	5.02	14.61	6.49
Gasoline Lead Q2	13.5	5.49	14.78	5.73
Gasoline Lead Q3	14.34	6.33	17.33	7.45
Gasoline Lead Q4	15.29	6.18	19.14	9.38

Notes: Gasoline Lead is annual, state, average tetraethyl-lead ($\overline{TEL}$) concentration (g/gal), corresponding to year of laboratory examination and state of residence;

Gasoline lead

To reconstruct a time series of TEL concentrations at the state level we follow the methodology of Reyes (2007, 2015). A collection of petroleum industry records and government publications are used to construct a time series of gasoline lead concentrations along with consumer gasoline sales. To capture the year-over-year effect of the CAA on the amount of lead entering the environment from gasoline consumption, we calculate the variable of interest $\overline{TEL}$ as a weighted average of gasoline sales and TEL concentrations during the phase out. With these data, we are able to quantify reductions in lead emissions through the composition and consumption of gasoline in the years following the CAA. Furthermore, the unique characteristics of this data set captures both the temporal and spatial variations in this hazardous metal at a national level.

The following sections provide a detailed description of each data source. The historical context in which the data was collected and published motivates the methodological decisions made in constructing the variables used in the analysis that follows.

TEL Concentrations

TEL concentrations for both regular and premium gasoline are sourced from the "Petroleum Products Survey: Motor Gasolines," published bi-annually by *The National Institute for Petroleum Research* (NIPER) under the Department of Energy. Gasoline was sampled each year from 17 different PAD districts across the United States. The boundaries of these districts were defined based on refinery locations, distribution networks and infrastructure as well as population centers (Shelton, 1979).

From each PAD district, a dozen samples were collected bi-annually during the summer and winter months. The chemical composition of each sample was analyzed and reported in corresponding "Summer" and "Winter" reports and included PAD district averages.[3] TEL concentrations were reported in grams per gallon (g/gal) for both "leaded non-premium" (regular) and "leaded premium" (premium). Figure 1.5 is an example of the published data, drawn from the Summer 1979 report, showing the PAD district averages for premium gasoline. The published

[3]At this time, we have been unable to attain copies of the "Winter" edition for several years within the period of interest. As such, we only use the TEL concentrations drawn from the "Summer" reports for consistency.

TABLE 4. - MOTOR GASOLINE SURVEY, SUMMER 1979
AVERAGE DATA FOR BRANDS IN EACH DISTRICT--CONTINUED

PREMIUM GASOLINE

DISTRICT NO. AND NAME	NO. OF BRANDS	SAMPLES	GR., ASTM D287 API	SULF., ASTM D1266 WT %	GUM ASTM D381 MG	BENZENE ASTM D3606 VOL. %	LEAD ASTM D526 G/GAL	OCTANE NUMBER RES., ASTM D2699	OCTANE NUMBER MOT., ASTM D2700	OCTANE NUMBER R+M --- 2	RVP, ASTM D323 LB	20V/L ASTM D439 F	DISTILLATION, ASTM D86 TEMPERATURE, F (CORRECTED TO 760 MM HG) IBP	PERCENT EVAPORATED 5	10	20	30	50	70	90	95	EP	RES %	LOSS %
1 NORTHEAST	5	5	61.8	-	-	-	2.81	98.0	90.0	94.0	11.1	125	83	95	106	126	149	212	270	334	366	399	0.9	2.3
2 MID-ATLANTIC COAST	17	68	60.6	0.021	1	0.69	2.59	97.9	89.7	93.8	10.4	131	87	101	113	136	160	215	269	340	371	411	.9	1.6
3 SOUTHEAST	14	45	60.2	-	0	-	2.76	98.0	90.0	94.0	10.1	132	88	102	115	138	162	216	266	340	374	412	.9	1.7
4 APPALACHIAN	9	27	60.3	.022	2	.65	2.52	97.8	90.3	94.0	11.0	127	83	96	110	133	159	213	268	342	372	421	.9	2.5
5 MICHIGAN	5	7	61.9	.038	1	.42	2.36	98.0	89.8	93.9	11.3	127	80	94	109	133	159	216	271	355	395	433	1.4	2.0
6 NORTH ILLINOIS	3	4	61.7	-	2	-	2.05	96.9	89.7	93.4	11.2	126	79	94	110	137	161	211	262	348	390	429	1.2	1.8
7 CENTRAL MISSISSIPPI	10	17	59.9	.031	3	1.09	1.94	97.8	89.3	93.6	10.3	133	87	104	118	142	168	216	261	331	365	407	.8	2.1
8 LOWER MISSISSIPPI	9	15	60.3	-	1	-	2.68	97.7	90.0	93.9	9.9	133	88	104	117	138	162	214	262	333	364	405	.8	1.3
9 NORTH PLAINS	1	1	61.1	.013	1	.67	3.80	98.4	91.0	94.7	9.7	139	88	-	129	158	185	224	262	352	-	428	.5	2.0
10 CENTRAL PLAINS	5	12	64.3	.035	1	.40	2.22	97.8	90.6	94.2	9.4	135	90	106	119	141	165	210	252	329	374	417	.8	1.3
11 SOUTH PLAINS	14	25	61.4	.023	1	.91	2.54	98.2	90.8	94.6	9.9	135	89	105	120	146	173	218	258	328	361	408	.8	1.6
12 SOUTH TEXAS	6	10	60.4	-	1	-	2.20	98.2	89.8	94.1	10.1	132	86	103	114	137	163	215	263	336	368	405	1.0	1.1
13 SOUTH MOUNTAIN STATES	18	81	60.9	.036	1	1.13	2.12	96.8	89.2	93.0	8.7	142	93	109	126	149	172	216	259	332	368	412	.9	1.7
14 NORTH MOUNTAIN STATES	12	39	64.5	.066	1	.33	1.45	96.9	89.3	93.1	9.9	135	93	105	120	145	171	215	254	333	374	413	1.0	3.0
15 PACIFIC NORTHWEST	8	26	59.9	.017	2	1.40	1.58	97.1	89.2	93.2	10.4	129	86	100	113	136	160	208	256	327	361	403	1.0	2.4
16 NORTH CALIFORNIA	10	30	56.3	.011	2	1.69	1.36	97.4	88.8	93.1	8.4	144	96	116	130	153	175	219	268	334	364	417	1.0	1.2
17 SOUTH CALIFORNIA	9	33	58.3	.028	1	1.10	1.59	97.0	88.7	92.9	8.5	142	92	113	127	149	171	215	264	332	361	414	1.0	1.1
AVERAGE			60.8	.028	1	.87	2.27	97.6	89.8	93.7	10.0	133	88	103	117	141	166	215	263	337	371	414	.9	1.8
MINIMUM			56.3	.011	0	.33	1.36	96.8	88.7	92.9	8.4	125												
MAXIMUM			64.5	.066	3	1.69	3.80	98.4	91.0	94.7	11.3	144												
SAMPLES	445																							

Figure 1.5: Motor gasoline survey, summer 1979: Premium grade

data shown in Figure 1.5 are the mean values from the set of samples collected and analyzed for each district. Additional examples from the 1979 report, including both regular and unleaded gasoline grades, are found in Appendix Figure C.1 and Figure C.2 respectively. The example data from the 1979 report has been included as it is representative of the statistics published in the "Petroleum Products Survey: Motor Gasolines" during the period of interest.

Annual Sales of Leaded Gasoline

Prior to 1972, gasoline was sold to consumers in two leaded grades, regular and premium. Formulated to enhance engine performance, premium contained higher TEL concentrations per gallon relative to regular. Unleaded gasoline was introduced in 1972, and was made available to consumers as "unleaded non-premium" (regular-unleaded), though "premium-unleaded" gasoline was not made available until 1981.

Annual gasoline sales for the years 1975-1984 are drawn from the "Yearly Report of Gasoline Sales, by States" published by *Ethyl Corporation*. Ethyl was the primary producer and distributor of the TEL additive used in gasoline from its introduction in the 1920's, until its removal in 1990. In cooperation with the petroleum industry, Ethyl published annual marketing reports detailing the sales volume of various petroleum products available to end users. These reports include monthly, quarterly, and annual sales of automobile gasoline by grade (millions of gallons) for both leaded and unleaded grades. The "Yearly Report of Gasoline Sales, by States" serves as the primary data source for total sales of each grade across all grades between 1975 to 1984. Figure 1.6 graphs national sales volumes of regular and premium gasoline grades between 1950 to 1985, and illustrates the magnitude of sales during this period. At peak demand, Ethyl reported national sales volumes of approximately 60 trillion gallons of regular (non-premium) and 35 trillion gallons of premium annually. If not for the availability of unleaded alternatives and growing sentiment over the health hazards posed by leaded gasoline, these trends may have continued.

Average gasoline lead calculation

Using the gasoline sales data published by Ethyl Corp, and the gasoline lead concentration data published by NIPER, we follow the methodology of Reyes (2015) in constructing annual,

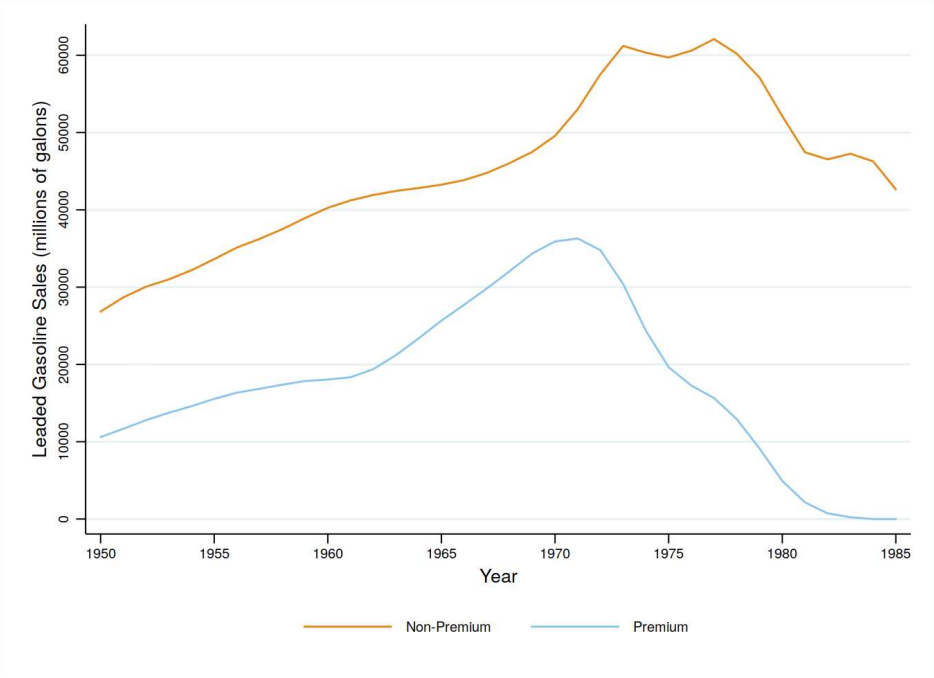


Figure 1.6: US leaded gasoline sales by grade, 1950-1985

state $\overline{TEL}$ concentrations. This metric is used to estimate the underlying lead exposure channel, as it connects the marginal effect of reductions in lead emissions from gasoline consumption to a bio-marker for lead exposure (BLLs).

To construct this metric, we first calculate total gasoline sales (Q_{jt}) for state j in year t, by summing the total sales of each grade sold. Define the set of years under observation as $T = \{1976, ..., 1980\}$ and let J be the set of states.[4] During the phase out, gasoline was sold to consumers in four grades: premium (p), regular (r), premium-unleaded (pu), and regular-unleaded (ru).[5] Let the set of gasoline grades be $G = \{p, r, pu, ru\}$ and define total sales for grade g as q^g_{jt}. Total gasoline sales are then:

$$Q_{jt} = \sum_{g \in G} q^g_{jt} \, , \, \forall \, t \in T, j \in J \tag{1.1}$$

Taking Q_{jt} and q^g_{jt}, the share (s) of g in total sales is:

[4] We restrict analysis to the lower 48 states, excluding Washington D.C.

[5] Noting that both premium (p) and regular (r) leaded gasoline grades along with regular-unleaded (ru) are available over the entire series, premium unleaded (pu) was not introduced until 1982.

$$s_{jt}^{g} = \frac{q_{jt}^{g}}{Q_{jt}} \tag{1.2}$$

By construction, $\sum_{g \in G} s_{jt}^{g} = 1 \; \forall \; t \in T, j \in J$. Using s_{jt}^{g} from Equation 1.2 as weights, $\overline{TEL}$ is calculated as the sum of concentration times the share of total sales across grades. Let tel_{jt}^{g} be the TEL concentration of gasoline grade g in year t and state j, then $\overline{TEL}_{jt}$ is:

$$\overline{TEL}_{jt} = \sum_{g \in G} s_{jt}^{g} \times tel_{jt}^{g}$$

However, since TEL concentrations in both unleaded grades (tel^{pu}, tel^{ru}) are effectively zero, these two terms drop out of the calculation and the final equation simplifies to:

$$\overline{TEL}_{jt} = \; s_{jt}^{p} \times tel_{jt}^{p} + s_{jt}^{r} \times tel_{jt}^{r} \tag{1.3}$$

This calculation is performed for all $j \in J$ and years $t \in T$.

The motivation to use this measure over alternatives is to isolate the effects of the policy shock while maintaining variations in the lead exposure channel at the state level. Since the CAA directly targeted gasoline lead concentrations, impacts of the policy are immediate and distinct in the rate of reduction in $\overline{TEL}$ concentrations over time and across states. Evidence of this can be seen in Figure 1.7 which illustrates $\overline{TEL}$ calculated using Equation 1.3 for all states between the years 1976 to 1985. Interpreted at the state level, Figure 1.7 illustrates the year-over-year decline in TEL as a result of the policy shock. Intra-state variation over time was driven by the specific factors, including age of the automobile stock and consumer preferences, though the policy had a consistent downward trend in $\overline{TEL}$ year-over-year. The inter-state variation in $\overline{TEL}$ is evident when comparing state values by year. Taken together, the affect of the CAA is evident in the quasi-random variations along these two dimensions.

1.3.2 Empirical model

Using BLLs from NHANESII as a bio-marker for lead exposure, the analysis that follows aims to estimate the causal affect of the CAA on reducing lead exposure on a national scale. Again,

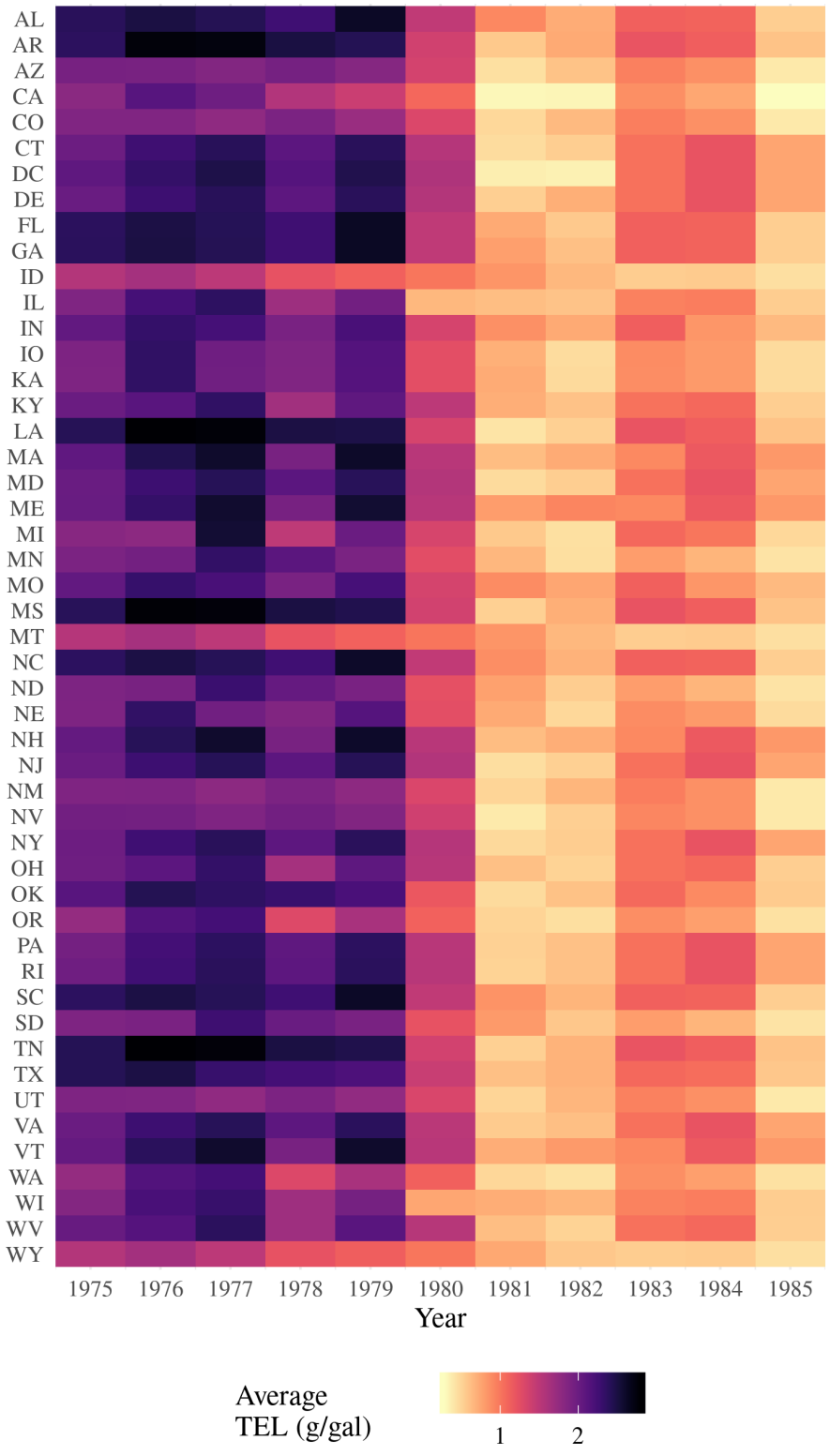


Figure 1.7: Average gasoline lead concentrations by state, 1975-1985

the CAA aimed to improve national air quality by setting standards for emissions of hazardous particulates, including lead. To reduce lead emissions from automobile exhaust, the policy targeted petroleum refiners by regulating gasoline lead concentrations in all grades sold to consumers. Since nearly all of the lead in automobile gasoline is emitted in the exhaust, the policy would reduce and eventually stem the flow of lead into the environment from this source. To estimate the affect of this policy on lead exposure at a national scale, we follow the methodology outlined in Section 1.3.1 and use gasoline lead concentrations as a proxy for lead exposure risk.

To map the affect of the CAA on lead exposure, we restrict our analysis to the years 1976 to 1980 during which the NHANESII survey was conducted and the most dramatic reductions in TEL would occur. The outcome of interest for individual i in state j and year t is BLL_{ijt}, and captures the observed lead exposure. Then, using Equation 1.3 to calculate annual state average gasoline lead concentrations ($\overline{TEL_{jt}}$), the relationship of interest is represented as:

$$BLL_{ijt} = \alpha + \beta_1 \overline{TEL}_{jt} + \boldsymbol{X}'_{ijt}\boldsymbol{\beta} + \boldsymbol{\Gamma}_2 \boldsymbol{M}_t + \boldsymbol{\Gamma}_1 \boldsymbol{Z}_t + \boldsymbol{\Gamma}_3 \boldsymbol{S}_j + \epsilon_{ijt} \tag{1.4}$$

where, $\boldsymbol{X}$ is vector of control variables measuring individual characteristics, including respondent's age, sex, minority status and household income that influence lead exposure outcomes specific to individual i. A series of indicator variables are included and are relative to the NHANESII date of laboratory analysis. $\boldsymbol{M_t}$ and $\boldsymbol{Z_t}$ are fixed effects corresponding to the month and year of the blood draw, with January, 1976 constituting the reference month and year respectively. $\boldsymbol{S_j}$ is the state of residence for i when the blood sample was collected.[6]

The coefficient of analytic interest is β_1, reflecting the marginal change in BLL explained by a change in $\overline{TEL}$. Since BLL is a bio marker for lead exposure, the expectation is that β_1 is negative as lead emissions declined during the period of interest. Evidence of this relationship is shown in Table 1.2, which summarizes the distribution of mean BLL and $\overline{TEL}$ at the national level. The direct effect of the CAA was dramatic year-over-year reductions in lead emissions beginning in

[6]The reference group for state fixed effects is respondents with Alabama as place of residence.

Table 1.2: US lead variables, descriptive statistics

	Mean	Standard Deviation	25^{th} Percentile	Median	75^{th} Percentile	N
Gasoline Lead (g/gal)						
1976-1980	1.35	0.40	1.01	1.4	1.66	245
1976	1.76	0.20	1.61	1.75	1.87	49
1977	1.64	0.20	1.55	1.66	1.79	49
1978	1.30	0.23	1.18	1.33	1.44	49
1979	1.32	0.23	1.17	1.35	1.49	49
1980	0.75	0.01	0.69	0.75	0.82	49
Blood Lead (μg/dL)						
1976-1980	13.99	6.12	10	13	17	10,049
1976	16.11	6.82	12	15	19	2,614
1977	14.2	6.18	10	13	17	2,287
1978	14.12	5.37	10	13	16	2,412
1979	11.87	5.03	8	11	14	2,412
1980	9.64	5.01	7	9	11	315

Notes: Gasoline Lead is national tetraethyl-lead ($\overline{TEL}$) concentrations (g/gal), calculated as weighted average across all gasoline grades; Blood lead is national average BLL (μg/dL) from NHANESII sample;

1976. As $\overline{TEL}$ concentrations declined over the phase out period, similar reductions in BLLs are observed among the NHANESII respondents.

1.4 Results

1.4.1 Evidence of lead effects

The results of fitting Equation 1.4 using linear estimator and the NHANESII data set are presented in Table 1.3. The covariates used follow the same definitions of Reyes (2007, 2015), and are the common set of predictors used in the empirical strategy of Chapters 2 and motivate the analysis of Chapter 3.[7] The dependent variable in Columns 1, 3, 5, and 7 is observed BLL (μg/dL), with

[7]As an additional validity test of the final state gasoline sales and TEL concentration time series data collected, we reproduce the results of Reyes (2015). For reference, Appendix Table C.1 is the replication of Reyes (2015), Table 3. One caveat to note regarding the reconstruction of the data used in our analysis versus that of Reyes (2015) is the time step of gasoline sales. The author was able to collect the complete set of gasoline sales records at the monthly time step. However, available data sources limited our data to annual gasoline sales by grade. As such, we substitute "Year" fixed effects for "Month" fixed effects to control for the seasonality of gasoline sales.

Table 1.3: Average gasoline lead concentration (TEL) effect on blood lead, linear regressions

	(1) Blood Lead	(2) $ln($ Blood Lead $)$	(3) Blood Lead	(4) $ln($ Blood Lead $)$	(5) Blood Lead	(6) $ln($ Blood Lead $)$	(7) Blood Lead	(8) $ln($ Blood Lead $)$
Gasoline Lead (g/gal)	6.763***	0.431***			4.920***	0.361***		
	(0.910)	(0.051)			(0.517)	(0.041)		
Reference (Gasoline Lead Q1)								
Gasoline Lead Q2			2.009**	0.136***			1.928***	0.157***
			(0.798)	(0.045)			(0.475)	(0.040)
Gasoline Lead Q3			3.393***	0.227***			3.226***	0.235***
			(0.685)	(0.036)			(0.571)	(0.046)
Gasoline Lead Q4			6.029***	0.388***			4.294***	0.322***
			(1.090)	(0.062)			(0.488)	(0.039)
Age	-0.313**	-0.015*	-0.318**	-0.015*	0.041***	0.003***	0.041***	0.003***
	(0.139)	(0.008)	(0.140)	(0.008)	(0.010)	(0.001)	(0.010)	(0.001)
Female	-0.485**	-0.031**	-0.446**	-0.029**	-4.203***	-0.307***	-4.211***	-0.308***
	(0.220)	(0.014)	(0.218)	(0.014)	(0.151)	(0.011)	(0.152)	(0.011)
Black	5.326***	0.301***	5.344***	0.302***	2.566***	0.188***	2.636***	0.193***
	(0.487)	(0.020)	(0.486)	(0.020)	(0.305)	(0.019)	(0.294)	(0.018)
Constant	1.364	1.703***	8.789***	2.168***	4.861***	1.841***	8.947***	2.135***
	(1.706)	(0.105)	(1.367)	(0.094)	(0.835)	(0.069)	(0.531)	(0.043)
Observations	2,430	2,430	2,430	2,430	9,148	9,148	9,148	9,148
R^2	0.296	0.359	0.293	0.357	0.240	0.285	0.240	0.283
Ages	0-6	0-6	0-6	0-6	all	all	all	all

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Blood lead is μg/dL; Gasoline Lead ($\overline{TEL}$) is annual, state, average tetraethyl-lead concentrations (g/gal), corresponding to year of laboratory examination and state of residence; Age is respondent's age (years) at the time of examination; Female=1 if respondent is female; Black=1 if respondent is African American;

the natural log of BLL used in the remaining columns. The predictor of interest is gasoline lead concentrations. These values are calculated using Equation 1.4 and are annual, state average $\overline{TEL}$ concentrations capturing the marginal effect of the CAA on the underlying lead exposure channel.

Regressions 1 to 4 of Table 1.3 are restricted to the sample of children under the age of seven years old at the date of examination.[8] Across each of the four specifications, gasoline lead is a significant predictor of blood lead in children. The baseline regression in column one suggests that a 1 g/gal reduction in $\overline{TEL}$ decreases BLL an estimated 6.8 μg/dL (95% CI: 4.9, 8.6). Recall, the current BLL reference value for at risk children, defined by the CDC, is 5 μg/dL (CDC, 2019). When the dependent variable is the natural log of blood lead as shown in Column 2, the results remain highly statistically significant.

To further investigate the lead exposure channel, we convert the continuous $\overline{TEL}$ variable into quartiles and include indicators for the quartile corresponding to the respondent's state of residence at the date of the blood draw. Column 3 shows the results of a linear-linear model when gasoline lead quartiles are introduced. Relative to respondents in states in the first quartile, mean BLLs are an estimated 2 μg/dL higher among those in the second quartile, and an estimated 3.4 and 6 μg/dL higher for those residing in the third and fourth quartiles respectively. Again, these results are mirrored in the log-linear estimates of Column 4 and remain significant at the 99% confidence level.

The results shown in Columns 5 to 8 fit the models used in Columns 1 to 4 on the full NHANESII sample. When the inclusion criteria is relaxed, the estimated effect of changes to gasoline lead on BLL is largely unchanged and remains highly statistically significant. The linear-linear estimates in Column 5 suggest that a one g/gal reduction in $\overline{TEL}$ reduces mean BLL by an estimated 4.9 μg/dL (95% CI: 3.9, 6.0). Comparing the coefficient estimates in Column 1 and Column 5, the underlying physiological factors which place children at greater health risks of lead exposure relative to adults become clear. For example, age is a significant predictor of BLL in both models, with the coefficient estimate for children in Column 1 being non-positive and comparatively large

[8]This inclusion criteria uses the NHANESII sub-sample of children, in which all children aged 6 months through 6 years old were selected for laboratory analysis.

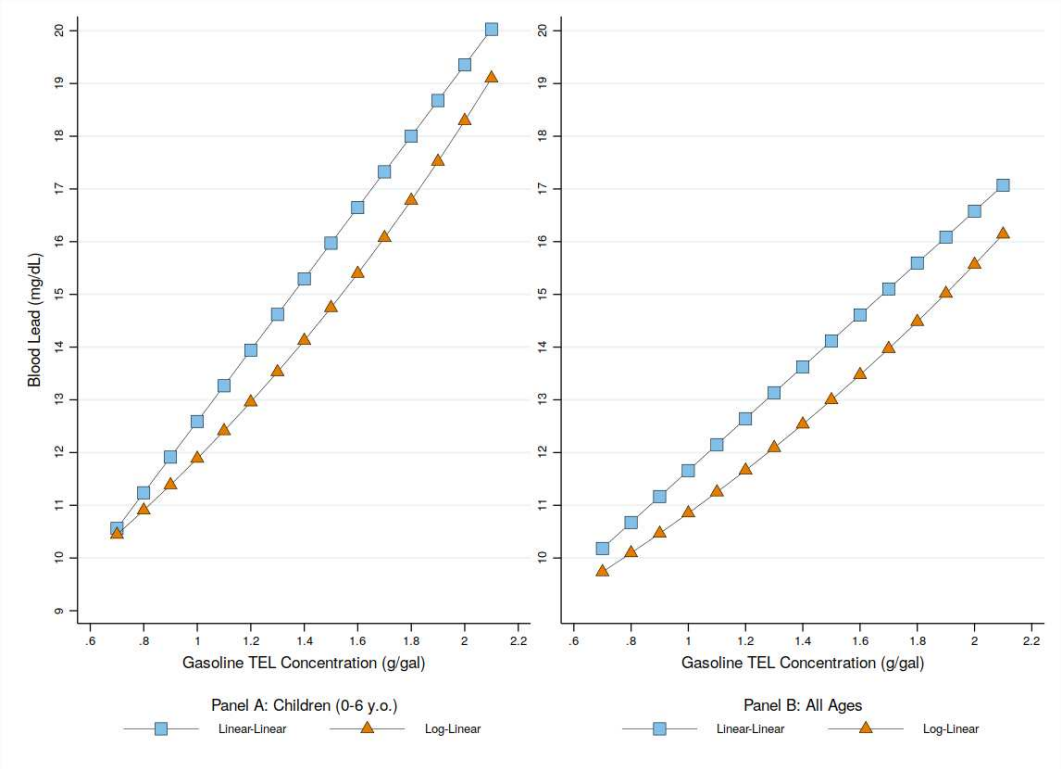


Figure 1.8: Average gasoline lead concentration effect on blood lead

in magnitude relative to the estimate in Column 5. When indicators for gasoline lead quartiles are included in Column 7, the coefficient estimates again reflect the positive, incremental increase in estimated BLLs relative to the first quartile. Here, the estimated level effect on BLLs is an increase of 1.9 and 3.2 μg/dL for those in the second and third quartiles respectively, and is relatively similar to estimated impact on children shown in Column 3. For respondents living in states with gasoline lead concentrations above the 75^{th} percentile, the estimated increase in mean BLL relative to the first quartile is 4.3 μg/dL.

The graphical representation of the dose responsiveness across both sample groups is illustrated in Figure 1.8. Panel A shows the marginal effects using regressions shown in Columns 1 and 2 and are restricted to the children subgroup. Panel B represents the full sample using Columns 5 and 6. Across the two sample groups, the marginal effect estimated across all ages is approximately two thirds of the marginal effect when estimated on the subset of children. This is in line with expectations as the dose responsiveness of BLL and age in children is non-linear with the largest marginal effects occurring between one and three years of age, attributable to repeated hand-to-

mouth behaviors and pica Mielke and Reagan (1998), as well as their physiology at this stage in life (Needleman and Bellinger, 1991).

1.4.2 Test of identifying assumptions

To test the robustness of our identifying assumptions, we apply an indirect test of the gasoline lead – blood lead exposure channel. Recall from Section 1.3.2, the lead exposure channel is proxied for by $\overline{TEL}$ and corresponds to the state and year in which each BLL in the NHANESII data was collected. Since the accurate matching of individual's BLLs to state-year TEL values are essential to the analysis above, we test our identification strategy by randomly assigning the gasoline lead exposure channel.

For each individual i in our NHANESII sample, we construct the following data point for this robustness test. Using the same notation as Equation 1.4, where j is i's state of residence when his blood is measured in year t. Let the set of states in the gasoline lead data set be J and the set of years under observation be T. Then for i, a randomly drawn state $j' \in J$ and year $t' \in T$, where $j' \neq j$ and $t' \neq t$, is used to assign a $\overline{TEL}$ value. Our robustness test estimates:

$$BLL_{ijt} = \alpha + \beta_1 \overline{TEL}_{j't'} + \boldsymbol{X}'_{ijt}\boldsymbol{\beta} + \boldsymbol{\Gamma}_2 \boldsymbol{M}_t + \boldsymbol{\Gamma}_1 \boldsymbol{Z}_t + \boldsymbol{\Gamma}_3 \boldsymbol{S}_j + \epsilon_{ijt} \tag{1.5}$$

where all terms carry from Equation 1.4, and $\overline{TEL}_{j't'}$ is the randomly assigned state, year, TEL concentration capturing the lead exposure channel. While it is plausible that the the findings presented above are simply coincidental, this strategy enforces randomness of the exposure mechanism and indirectly tests the robustness of the gasoline lead effect on BLL. As such, the blood lead – gasoline lead relationship should dissipate for a randomly assigned $\overline{TEL}$, and $\hat{\beta}_1$ should be indistinguishable from zero.

Using the same models of Table 1.3, the results of applying this strategy are shown in Table 1.4. As expected, the estimated direct effect of gasoline lead concentrations on BLLs in both sample groups is insignificant for both the linear-linear and log-linear models. Results are similarly insignificant when gasoline lead is converted to quartiles and fit to the set of children (Columns 3, 4) and the full sample (Columns 7, 8).

Table 1.4: Falsification test of randomized average gasoline lead concentrations on lead exposure

	(1) Blood Lead	(2) $ln($ Blood Lead $)$	(3) Blood Lead	(4) $ln($ Blood Lead $)$	(5) Blood Lead	(6) $ln($ Blood Lead $)$	(7) Blood Lead	(8) $ln($ Blood Lead $)$
Gasoline Lead (g/gal)	0.427	0.018			-0.452	-0.032		
	(1.232)	(0.067)			(0.460)	(0.034)		
Reference (Gasoline Lead Q1)								
Gasoline Lead Q2			-3.024*	-0.163			0.176	0.030
			(1.653)	(0.098)			(0.809)	(0.059)
Gasoline Lead Q3			-0.188	0.004			-0.058	0.011
			(1.778)	(0.114)			(0.924)	(0.071)
Gasoline Lead Q4			-1.115	-0.071			-0.367	-0.018
			(1.622)	(0.091)			(0.640)	(0.047)
Age	-0.293**	-0.014*	-0.294**	-0.014*	0.042***	0.003***	0.042***	0.003***
	(0.143)	(0.008)	(0.140)	(0.008)	(0.010)	(0.001)	(0.010)	(0.001)
Female	-0.444*	-0.028*	-0.440*	-0.029**	-4.243***	-0.310***	-4.240***	-0.310***
	(0.226)	(0.014)	(0.227)	(0.014)	(0.150)	(0.011)	(0.152)	(0.011)
Black	5.421***	0.307***	5.461***	0.310***	2.725***	0.200***	2.718***	0.199***
	(0.498)	(0.019)	(0.451)	(0.017)	(0.261)	(0.016)	(0.252)	(0.016)
Constant	10.673***	2.304***	11.833***	2.354***	12.289***	2.385***	11.970***	2.349***
	(1.371)	(0.084)	(1.715)	(0.113)	(0.521)	(0.040)	(0.885)	(0.067)
Observations	2,430	2,430	2,430	2,430	9,148	9,148	9,148	9,148
R^2	0.267	0.322	0.274	0.328	0.211	0.250	0.210	0.250
Ages	0-6	0-6	0-6	0-6	all	all	all	all

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Blood lead is μg/dL; Gasoline Lead ($\overline{TEL}$) is annual, state, average tetraethyl-lead concentrations (g/gal), corresponding to year of laboratory examination and state of residence; Age is respondent's age (years) at the time of examination; Female=1 if respondent is female; Black=1 if respondent is African American;

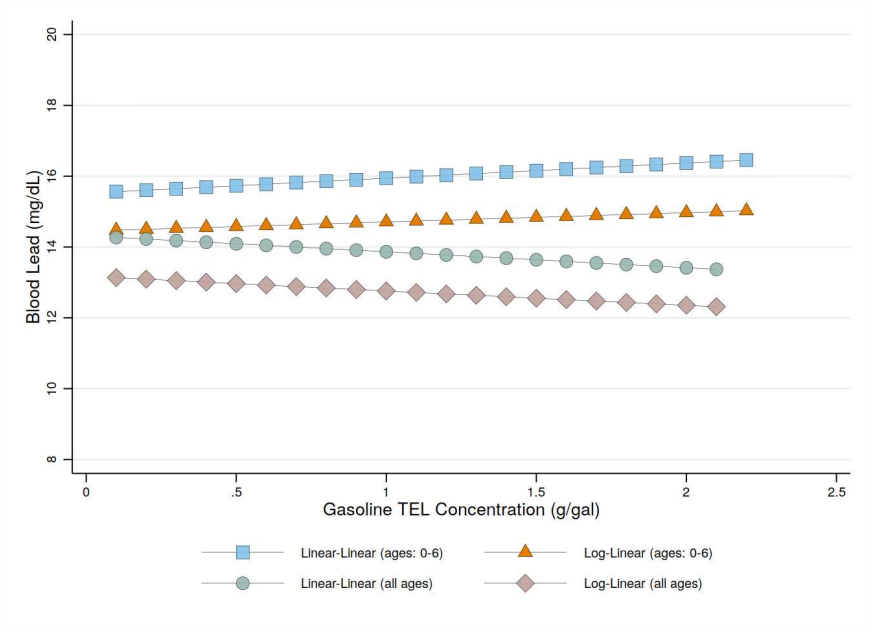


Figure 1.9: Randomized gasoline lead concentration exposure risk on blood lead

The dose response curves estimated using this strategy are shown in Figure 1.9. Compared to Figure 1.8, the marginal effects in Figure 1.9 are statistically insignificant. The randomization of the lead exposure channel dissolves the significance of $\overline{TEL}$ on blood lead, yet the remaining set of covariates are effectively unchanged with respect to the statistical significance of their coefficients. Though this strategy is imperfect, it is included as divergent validity evidence in support of the identifying assumptions underlying the empirical analysis presented above and in the chapters that follow.

1.5 Discussion and conclusion

Exploiting the spatial and temporal variation in gasoline lead concentrations following the Clean Air Act, we find considerable evidence supporting this flow source of environmental lead as a significant vector of lead exposure among children and adults. Following the work of Reyes (2007, 2015), we leverage the natural experiment created by the phase out of leaded gasoline to identify and estimate the causal effect of reductions in the flow of environmental lead from automobile emissions to individual blood lead levels at the national level. As average gasoline lead declined under the CAA, it produced quasi-random variations in lead emissions and in turn, re-

sulted in heterogeneous lead exposure risk across states during the phase out. Using annual, state averaged gasoline lead concentrations for the period 1976 to 1980, our results show a strong, positive relationship with individual blood lead levels among NHANESII survey respondents. Results from sensitivity and robustness tests show this relationship remains highly statistically significant, with the largest health benefits from the policy going to the youngest age groups. These findings emphasize the success of this national policy with respect to improving public health by stemming the flow of lead into the environment attributable to the consumption of leaded automobile gasoline.

While the CAA is responsible for the removal of lead from gasoline, the stock of legacy lead from this source, along with the flow of environmental lead from remaining source polluters continues to threaten public health. In the United States, flow sources of lead include point source toxic release inventory facilities and aviation gasoline. As of 2017, a total of 4,155 facilities reported on-site disposal or releases of lead to the Toxic Release Inventory (TRI) system of the EPA. Combined, these facilities disposed or released 10,894,687 pounds of lead on-site at their facilities. In addition, 3,809 facilities disposed or released just shy of 1 billion pounds (942,261,070) of lead compounds on-site at their facilities in 2017.

In contrast to the incidents reported to the TRI system, the dispersion of atmospheric lead extends far beyond the source. More than 50% of the current flow of atmospheric lead is from the deposition of leaded aviation gasoline used by piston-engine aircraft. About 160,000 piston-engine aircraft (PEA) are registered in the United States, constituting about 70% of the U.S. air fleet. These aircraft consume over 200 million gallons of avgas annually (Kessler, 2013), implying a flow of about a million pounds of lead each year.

Evidence of this can be found in the soil lead content across the U.S.. Soil samples drawn from a number of rural and urban counties in the mid 1970's and again in the early 2000's were found to have lead soil concentrations of 19.5 μg/kg and 21.11 μg/kg respectively (Clay et al., 2018). Soil lead as a cause for health concern has been documented in many empirical studies showing strong associations between neighborhood soil Pb, children's blood Pb, and learning or behavioral outcomes (Mielke and Reagan, 1998; Mielke et al., 2007; Johnson and Bretsch, 2002;

Zahran et al., 2011). For the approximately 16 million people (and 3 million children) that live within a kilometer of an airport that services piston-engine aircraft, the risk of lead exposure is serious. In fact, studies find that child blood lead levels: increase dose-responsively in proximity to airports, decline measurably among children sampled in the months after 9-11, resulting from the restriction of the flight behavior of PEA, increase dose-responsively in the flow of PEA traffic, and increase in the percent of prevailing wind days drifting in the direction of a child's residential location (Zahran et al., 2017b).

As national policies have eliminated the use of lead in paints and automobile gasoline, substantial reductions to the flow of lead into the lived environment have been realized. Yet, both stock and flow sources of exposure risk remain in the United States, presenting an ongoing health hazard that is both ubiquitous and indiscriminate.

Chapter 2

Childhood lead exposure, cognitive ability and risk preferences in adulthood

2.1 Introduction

Lead (Pb) is a neurotoxicant with developmentally harmful effects in children. Many studies link lead exposure to adverse cognitive, behavioral and physical health outcomes in children (Needleman and Bellinger, 1991; Dietrich et al., 2001; Canfield et al., 2003; Reyes, 2007; Jusko et al., 2008; Nigg et al., 2010). Despite recent national interest in lead exposure following preventable failures in water distribution systems in Flint, Michigan and Newark, New Jersey, lead pollution in the United States is generally regarded as a legacy problem. The enactment of various national policies – like the Clean Air Act which resulted in the removal of lead from automotive gasoline sold to consumers – effectively reduced the flow of lead (Pb) into the environment and caused blood lead levels (BLLs) in children to decline dramatically (Raymond et al., 2014).

While lead has been effectively banned in United States from historic uses – as a constituent in paint, plumbing, and automotive gasoline – scientists caution against the concept of lead as a strictly historic or rear-view problem. Lead-formulated aviation gasoline (avgas) used in piston engine aircraft remains an important source of new emissions. The flow of lead from avgas is about a million pounds per year (Kessler, 2013). While small compared to the amount consumed historically in automotive gasoline, the deposition of avgas remains a source of exposure risk to the estimated three million children residing within one kilometer of airport facilities that service piston engine aircraft (Zahran et al., 2017a).

An under-appreciated contemporary source of child lead exposure is lead-concentrated soils, primarily due to legacy deposition from lead-formulated automotive gasoline. Scientists have convincingly linked child BLLs to the accumulation of lead in residential soils. Contaminated soils enter the body through ingestion (involving hand-to-mouth behaviors) or inhalation of re-suspended

dust particles in summer months (Filippelli et al., 2005; Laidlaw et al., 2005, 2012; Zahran et al., 2010, 2011, 2013). Another exposure pathway for children in the United States is dust associated with deteriorating or haphazardly removed lead-based paint, covering the interior and exterior walls of older homes built before 1950 (when leaded paint was in widespread use).

Lead persists not only in the lived environment, but also in the human body. While the half-life of lead in the human bloodstream is about thirty days (Papanikolaou et al., 2005), the metal can persist in human tissue, the brain, and the skeletal system for many decades after an exposure event, causing measurable failures in cardiovascular and renal systems as well as early onset of neurodegenerative diseases in adulthood (Bellinger et al., 1987; Needleman and Bellinger, 1991; Zahran et al., 2017b). To a growing cadre of economists and epidemiologists, the persistence of the lead problem is evidenced in present-day intellectual and cognitive effects in adults exposed to lead in early childhood. Numerous studies have found that elevated BLLs in young children are associated with measurable reductions in intellectual ability (Bellinger, 2008, 2011, 2017; Clay et al., 2019).

More preliminary, but economically meaningful, evidence suggests that the intellectual impairments caused by early childhood lead exposure are lasting and possibly growing in age. In a recently published study in the Journal of the American Medical Association involving more than 500 adults observed repeatedly over four decades, Reuben et al. (2017) report that adults raised in the era of widespread leaded gasoline use experienced significant reductions in IQ and deficits in perceptual reasoning and working memory. The higher the blood lead level in childhood, the greater the loss in cognitive function and occupational status in adulthood.

A growing body of research also relates childhood lead exposure to abnormal behaviors in adulthood. Bellinger (2008) finds children with prior acute lead exposure exhibit abnormal behaviors such as irritability, impatience, and aggression which persist after BLLs return to pre-exposure levels. Economist Reyes (2015) has shown that persons exposed to lead in early life experience "an unfolding series of adverse behavioral outcomes: behavior problems as a child, pregnancy and aggression as a teen, and criminal behavior as a young adult." Far from being a legacy problem, lead

exposure appears to echo through the life-course, impairing the realization of health and human capital among its many victims.

Studies using neural-imaging find that adults exposed to lead as children have reduced gray matter in regions of the brain known to govern executive judgment and impulse control, cognitive factors that are implicated in the expression of preferences both revealed and stated (Cecil et al., 2008; Cecil, 2011). Behavioral scientists and economists have shown that judgment and impulse control meaningfully underwrite economic and social preferences and behaviors that involve risk and prospecting of uncertain futures (Thaler and Shefrin, 1981; Fudenberg and Levine, 2006; Khwaja et al., 2007). In this paper, we analyze whether lead exposure in childhood impacts cognitive ability and the presence of abnormal latent preferences toward risk and uncertainty in adulthood. To identify the causal relationship between BLLs in childhood and the realization of lower cognitive ability and abnormal risk preferences in adulthood, we exploit the phase out of lead from gasoline as an exogenous source of variation in lead emissions at a national scale.

Under the Clean Air Act (CAA), a timeline for the removal of lead from gasoline was implemented beginning in 1975. Specifically, this policy would target the anti-knock lead additive tetraethyl-lead (TEL), used to boost gasoline octane and improve engine performance (Needleman, 2000; Reyes, 2007). Over the following two decades, lead entering the environment from automobile emissions decreased precipitously. Though the policy was enacted at the national level, specific characteristics relating to the petroleum and automobile industries caused meaningful variation in lead emissions across states between 1975 and 1990. Since lead emissions resulting from the consumption of leaded gasoline was the most important source of lead exposure in children over this period, we use variation in gasoline lead concentrations to capture this quasi-random exposure channel.

This paper builds off existing literature that investigates the lasting health impacts of environmental externalities. Our analysis finds strong evidence supporting the effect of childhood lead exposure and later-in-life outcomes relating to both cognitive impairment and risk preferences. Estimating the causal relationship across a series of tests, we find that blood lead levels in childhood are a significant predictor of: 1) IQ loss, measured with standardized test scores; 2) increased

likelihood of low-IQ outcomes in exposure levels; and 3) increased abnormal risk response across a series of situations involving uncertain outcomes. To test the robustness of the identification strategy underlying our empirical analysis, we provide a divergent validity test in which the lead independent outcome, individual's dominant hand preference, is regressed on the childhood blood lead exposure channel. Across an ensemble of evidence, our results illustrate the significance and persistent effect of early in life lead exposure. These findings are a novel contribution to this area of research for the causal effect of lead exposure which we is modeled.

Analyses detailed in the following sections attempt to answer whether quasi-random exposure to lead in childhood produced significant cognitive and risk preference effects in adulthood. The remainder of the chapter is structured as follows: Section 2.2 summarizes pertinent literature on lead exposure as it relates to adverse health outcomes and abnormal psychology in adulthood. A discussion of the data used and empirical strategy used to identify the causal effects of lead exposure is found in Section 2.3, followed by results in Section 2.4. Finally, the chapter closes with a discussion and conclusion in Section 2.5.

2.2 Literature review

Unlike iron and magnesium, lead is not an essential element to the human body. While the symptoms of lead poisoning typically occur at high levels of lead exposure, any quantity entering the body along the extensive margin is effectively toxic (Bellinger and Bellinger, 2006). Both the Environmental Protection Agency and the Centers for Disease Control have stated that there is no known safe level of lead exposure (DHHS 2012; CDC 2012a, 2012b).

Although lead is a naturally occurring metal, legacy anthropogenic sources such as deteriorating lead-based paint in aging homes (Farfel et al., 2005b), as well as emissions from leaded gasoline (Needleman, 2000; Filippelli et al., 2005; Gould, 2009; Kessler, 2013) have posed more substantial risk to public health (Needleman and Bellinger, 1991; Filippelli et al., 2005; Abadin et al., 2007; Reyes, 2007). Exposure to these legacy sources of lead are typically in the form of lead dust. This dust can create a residue on foods prior to consumption (Abadin et al., 2007). Lead dust also mixes with topsoil, and researchers have found large, positive correlations between soil-

lead concentrations and blood lead levels (BLLs) in a given area (Reagan and Silbergeld, 1989; Johnson and Bretsch, 2002; Mielke et al., 2005; Laidlaw and Filippelli, 2008; Laidlaw et al., 2012; Filippelli and Laidlaw, 2010). This relationship between child BLLs and neighborhood soil lead accumulation is likely causal (Zahran et al., 2010; Clay et al., 2019).

In their review of existing literature comparing the two legacy sources of lead – paint versus soil - Mielke and Reagan (1998) note that "exposure to lead-contaminated soil, house lead dust, or street dust has consistently shown a positive correlation to blood lead and population blood lead levels." Inhalation of lead-laced dust has been shown to occur when particles are re-suspended during dry summer months or weather related events (Laidlaw et al., 2005, 2012; Filippelli et al., 2005; Zahran et al., 2013). Due to this phenomenon, the rate of uptake can vary significantly for children compared to adults, as children's higher rate of hand-to-mouth activities place them at greater risk of lead entering their bodies relative to other age groups (Mielke and Reagan, 1998; Nilsson, 2009).

The negative impacts of lead exposure on a child's cognitive development has long been recognized by the medical profession. Reports published by the EPA show that the estimated dose-response curve for lead exposure with respect to cognitive and behavioral effects in children are non-linear, with larger marginal effects at lower blood lead levels. These findings are echoed in a number of other studies which emphasize the marginal impact of increases in child BLLs along the intensive margins on IQ loss, school performance, and neuro-psychological impairments (Bellinger et al., 1987; Needleman and Bellinger, 1991; Lanphear et al., 2005; Bellinger, 2008). These findings have been replicated in other countries as well. In his analysis of lead exposure and cognitive development in Swedish children, Nilsson (2009) confirms the significance of the current reference threshold of 5 μg/dL, defined by the U.S. Center for Disease Control as a benchmark of concern (Needleman, 2004; Toscano and Guilarte, 2005; Bellinger and Bellinger, 2006).

Studies have shown that lead exposure in childhood impacts cognition and the realization of abnormal behaviors into adolescence. Aizer and Currie (2017) find that childhood lead exposure results in higher incidents of juvenile delinquency in adolescence. Reyes (2015) finds that childhood lead exposure is a significant predictor of risky behaviors in adolescence, including excessive

drinking and smoking and the risk of teen pregnancy. Recently, researchers have estimated the causal effects of childhood lead exposure on outcomes in adulthood, including fertility behavior and deleterious birth outcomes (Clay et al., 2018), violent crime (Nilsson, 2009; Curci and Masera, 2017; Mielke and Zahran, 2012), neurodegenerative diseases (Zahran et al., 2017a), and a series of abnormal behaviors in young adulthood (Graff Zivin and Neidell, 2013). Perhaps most compelling, Reuben et al. (2017) find that adult New Zealanders exposed to lead in childhood had measurable reductions in IQ and occupational status in midlife, with these negative effects of early childhood exposure amplifying over the life-course.

Providing a neurological basis for these observed effects of abnormal psychology and behavior, Cecil et al. (2008) find a strong, inverse dose response relationship between mean childhood blood lead levels and brain volume in adults. Using magnetic resonance imaging to measure brain matter in individuals born between 1979 to 1984, the authors find statistically significant losses in the structural volume of the brain, specifically in the prefrontal cortex that is considered the seat of decision making and impulse control in the brain. Moreover, Froehlich et al. (2007) show that lead impacts a set of genes controlling dopamine receptors that are associated with executive judgment, providing a plausible underlying biochemistry for the abnormal psychology of adults lead exposed as children.

Taken together, the research literature shows that childhood lead exposure significantly impairs the cognitive and socio-emotional development of children, with these impairments persisting through the life-course and manifesting in abnormal psychology and behavior. Insofar as exogenous exposure to lead in childhood compromises decision making and mood regulation, one might also expect to see cognitive effects and abnormalities in economically meaningful behavior involving the prospecting of risk and uncertainty in adulthood. In the next section, we describe our econometric strategy for evaluating cognitive effects and the presence of abnormal preferences in adults randomly exposed to lead in childhood.

2.3 Methods

2.3.1 Data and measurement

National Longitudinal Survey of Youth

The primary data set used in the analysis is from the National Longitudinal Survey of Youth (NLSY), sponsored and directed by the U.S. Bureau of Labor Statistics (BLS) and and managed by the Center for Human Resource Research (CHRR). The purpose of this survey is to track outcomes relating to transitions in and out of the labor force at various of stages of life (CHRR, 2019). Although the BLS has administered this survey to several cohorts, our analysis relies on data from the survey following individuals born between 1980-1984. Respondents from this cohort were first interviewed in 1997 (NLSY97), corresponding with the timing of the youngest respondents in the sample entering the labor force.

To create a nationally representative sample, the NLSY97 is constructed using a complex survey design.[9] Given the stated goal of the survey, the bulk of the data collected speak to employment outcomes, labor market conditions, and human capital investments. Complementing these data is a base set of demographic and socioeconomic variables on each respondent, as well as an inventory of family and household characteristics during the respondent's youth, physical and mental health assessments, and a complete geographic history of migration and places of residence.[10] In the analyses that follow, many control variables are used, including respondent's sex, racial status, age in years, and a set of indicator variables for household income to Federal Poverty Level (FPL) quartiles.[11] Table 2.1 reports descriptive statistics on demographic and socioeconomic variables for the 1980-1984 cohort, divided by cross-sectional and supplemental NLSY97 samples.

[9]To satisfy the survey design requirement and provide sufficient sample size for statistical analysis of minority groups, the NLSY97 cohort is comprised of two sample groups: a "cross-sectional" sample and an "oversample" of black and Hispanic or Latino respondents (CHRR, 2019).

[10]The geographic data is provided by the BLS for research purposes and is not available in the public data files to maintain respondent confidentiality. The non-public geographic data identifies each respondent's place of birth, migration history, and place of residency down to the county level for each survey round.

[11]The household income to poverty level data is generated by the BLS. Values are for the previous year and adjusted for household size (CHRR, 2019).

Table 2.1: Descriptive statistics, NLSY97

	Cross Sectional Sample		Supplemental Sample	
	Mean	Std. Dev.	Mean	Std. Dev.
Survey Round 1 (1997)				
Birth Year	1982	1.39	1982	1.39
Sex	51%	0.50	51%	0.50
White	74%	0.44	0%	0
Black	17%	0.37	75%	0.43
Hispanic	8%	0.27	25%	0.43
Teen Pregnancy	24%	0.43	40%	0.49
HS Graduate (Mother)	84%	0.37	71%	0.45
HS Graduate (Father)	83%	0.38	71%	0.45
Low Income Group	34%	0.47	67%	0.47
Middle Income Group	22%	0.42	25%	0.36
HH Income to Poverty Ratio	318	277.02	175	181.61
Survey Rounds 14,15 (2010, 2011)				
Age	28	1.45	28	1.45
Married	37%	0.48	21%	0.41
HS Graduate	82%	0.39	70%	0.46
College Graduate	29%	0.45	15%	0.36
HH Income to Poverty Ratio	370	335.83	230	273.37
Sample Size	3,838		998	

Notes: Labels with Mother (Father) are biological mother (father); Indicators are defined as: Sex=1 for male, Teen Pregnancy=1 if the respondent's biological mother was a teenager at respondent's time of birth; HS (College) Graduate are indicators for educational attainment, $1 =$ high school (4 year college) degree, 0 o.w.; Income groups are household income relative to Federal Poverty Level (FPL), defined as: Low=1 for < 2x FPL and Mid=1 for ratios $\leq$2x FPL and $<$3x FPL;

Risk preference and cognitive ability data

In addition to demographic and socioeconomic variables, the NLSY97 has data on survey questions pertaining to a respondent's self-reported appetite for risk as well as measures of intelligence based on results from standardized achievement tests. With respect to an individual's risk preferences, the NLSY97 has a set of questions measuring an individual's willingness to accept risk in various circumstances. Respondents are asked to rate their willingness to accept risk generally and across a series of contexts pertaining to a respondent's health, finances, romantic relationships, driving behavior, gambling and work life, among other domains of risk taking. The most general question, asked first in the NLSY97, reads:

> Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Rate yourself from 0 to 10, where 0 means "unwilling to take any risks" and 10 means "fully prepared to take risks."

The same response categories of 0 meaning "unwilling to take any risks" and 10 meaning "fully prepared to take risks" obtain for the more specific domains of risk-taking in health, finances, work and so on. The complete question set is detailed in Appendix Section A.

Taking direction from abnormal psychology, we calculate the degree to which an individual's response to each risk preference question deviates from the norm. Under this definition of abnormality, a respondent's appetite for risk is deemed abnormal if her answers are statistically deviant. Assuming the sample mean is representative of normal risk preferences, the distance between the mean and the stated preference of the respondent is indicative of abnormality.

Let $\boldsymbol{Z}$ be the set of survey questions describing a respondent's appetite for risk, then the stated preference for individual i is z_i for $z \in \boldsymbol{Z}$. Then, defining the sample mean for z as $\overline{z}$,[12] we calculate mean deviation among NLSY97 respondents as:

$$y_i^z = ln(|z_i - \overline{z}|) \tag{2.1}$$

[12]Sample means are calculated using inverse probability weights, adjusted for the NLSY97 survey design.

where y_i^z is the outcome of interest, measuring the relative deviation from central tendency for each risk question $z \in \boldsymbol{Z}$.

With respect to outcomes relating to general intelligence, we use exam scores for each respondent collected during round one of the survey (1997). The BLS administered a modified version of the Armed Services Vocational and Aptitude Battery (ASVAB) examination to all respondents in the NLSY97 sample. The values are reported as percentile scores mapping closely to IQ percentiles in similar tests, such as the SAT and ACT.

National Health and Exercise Survey blood lead data

Blood lead data are from the Second National Health and Exercise Survey (NHANESII), provided by the Centers for Disease Control (CDC). This nationally representative survey was intended to provide a detailed description of the overall health of the nation (NCHS, 1982). Approximately 20,000 individuals, aged 6 months to 74 years old, were surveyed between February 1976 and February 1980. Demographic, socioeconomic, employment and household characteristics were collected alongside full biochemistry and physiological analyses to deliver a comprehensive description of American health at the time.

A subset of respondents were selected for laboratory analysis of blood composition, which measured the presence of heavy metals known to be toxic, including lead. The selection criteria for laboratory analysis included all children between six months and six years old, along with half of the remaining sample aged 7 to 74 years old chosen at random (NCHS, 1982). In all, the sub-sample included blood lead data for 10,049 individuals reported in units of micro-grams-per-deciliter (μg/dL) of blood.

The NLSY97 data set does not include respondent blood lead data, therefore we must rely on the blood lead samples collected from the NHANESII survey respondents. To estimate BLLs for NLSY97 respondents we restrict to a set of regressors present in both NHANESII and NLSY97 data sets. Since the two surveys were conducted by different institutions, variation in definitions used for many common variables constrains the set of regressors which can be included in the model. Moreover, the NLSY97 survey incorporates geographic definitions used in the 1980 decennial census that was conducted after the NHANESII survey was fielded. The model selection

criteria for stage 1 was therefore based on best statistical fit of blood lead outcomes given the set of covariates available across both the NHANESII and NLSY97 data sets. Given this objective and constraint, the final model specification used for stage two is a log-log model of blood lead and gasoline lead relationship. Results of this strategy are presented in the next section, and a replication of Table 3 from Reyes's (2015) process for this cross-date set exercise can be found in Appendix Table C.1.

Gasoline lead data

To reconstruct a time series of gasoline lead (TEL) concentrations at the state level, we follow the methodology detailed in Reyes (2007, 2015). Utilizing petroleum industry records and government publications, we construct the variable of interest $\overline{TEL}$, as a weighted average of gasoline sales and TEL concentrations. To capture the year-over-year effect of the CAA on the flow of lead entering the environment from gasoline consumption, we produce a time-series of $\overline{TEL}$ for each state and year using both gasoline sales volumes and TEL concentrations for each grade of gasoline sold to consumers during the phase out.

To construct this metric, we first calculate total gasoline sales (Q_{jt}) for state j in year t, by summing the total sales of each grade sold. During the phase out, gasoline was sold to consumers in four grades: premium (p), regular (r), premium-unleaded (pu), and regular-unleaded (ru). Defining the set of gasoline grades as $G = \{p, r, pu, ru\}$ and letting total sales for each grade $g \in G$, be q^g_{jt}, total gasoline sales are:

$$Q_{jt} = \sum_{g \in G} q^g_{jt} \tag{2.2}$$

Taking Q_{jt} and q_{jt}, we calculate the share (s^g_{jt}) of gasoline grade g in total sales:

$$s^g_{jt} = \frac{q^g_{jt}}{Q_{jt}} \tag{2.3}$$

By construction, $\sum_{g \in G} s^g_{jt} = 1$ for each state j and year $t \in \{1976, ..., 1984\}$. Let tel^g_{jt} be the TEL concentration for gasoline grade g in state j and year t. Then, using s^g_{jt} from Equation 2.3 as

weights, we calculate $\overline{TEL}_{jt}$ as the sum of concentration times the share of total sales, for each grade of gasoline in a given state and year.

$$\overline{TEL}_{jt} = \sum_{g \in G} s^g_{jt} \times tel^g_{jt}$$

However, since TEL concentrations in both unleaded grades (tel^{pu}, tel^{ru}) are effectively zero, these two terms drop out of the calculation and the final equation simplifies to the following:

$$\overline{TEL}_{jt} = \ s^p_{jt} \times tel^p_{jt} + s^r_{jt} \times tel^r_{jt} \tag{2.4}$$

This calculation is performed for all states[13] and each year over the period of interest. A more detailed description of the data sources and this calculation are presented in Chapter 1, Section 1.3.1.

Recall from Chapter 1, to achieve the goal of phasing out leaded gasoline, the CAA set limits on average TEL concentrations across all grades produced at a given production facility (Needleman, 2000). As such, the policy created an incentive for producers to make the costly investments required to increase production of unleaded gasoline over the course of the phase out. Since the CAA directly targeted gasoline lead concentrations, impacts of the policy were immediate and distinct in the rate of reduction in TEL concentrations over time and across states. Take Figure 2.1 for example, which graphs the difference between national $\overline{TEL}$ and state level $\overline{TEL}$ for both the NHANESII years (1975-1980) and the NLSY97 years (1981-1985). The downward trend in $\overline{TEL}$ is evident in the deviation from the national average in the NLSY97 period versus the NHANESII period across states, as all gasoline lead concentrations eventually converge to zero. Furthermore, Figure 2.1 illustrates the extent to which the design and enforcement of the CAA induced quasi-random inter-state variations in $\overline{TEL}$ across both time periods.

2.3.2 Empirical model

To test the hypothesis that childhood lead exposure impacts cognitive ability and risk preferences in adulthood, we apply the methodology of Reyes (2015) of a split sample two-stage least

[13] Due to data limitations, Alaska, Washington D.C., and Hawaii are excluded from our analysis.

squares approach. This approach allows us to estimate the degree to which childhood blood lead levels explain cognitive outcomes in adulthood. A stylized representation of this relationship is shown in Equation 2.5.

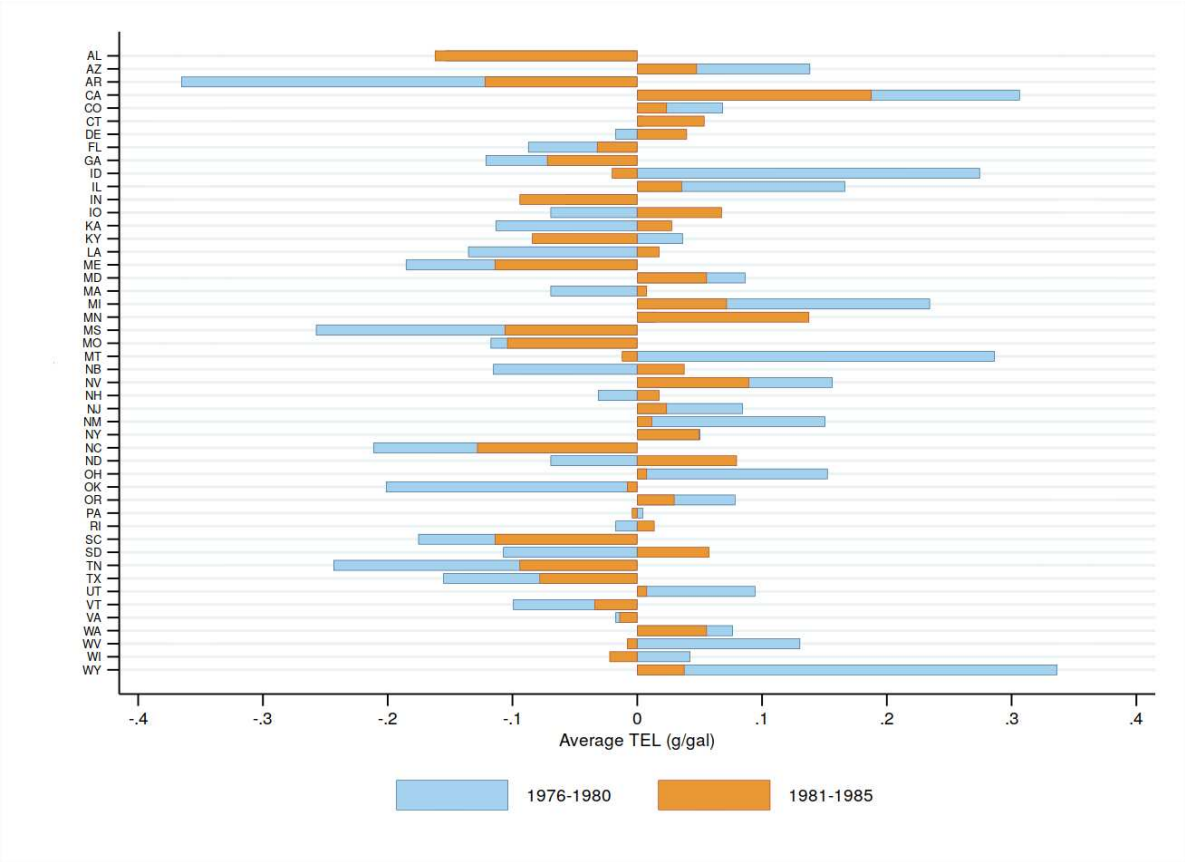


Figure 2.1: State average gasoline lead (TEL) relative to national average, 1976-1985

The dependent variable, in this case, is the degree to which an individual's response to risk deviates from the norm, defined in Equation 2.1. The predictor of interest is blood lead level (BLL), constituting our bio-marker proxying for childhood lead exposure. To estimate the causal impact of childhood lead exposure on later in life outcomes, we rely on the state of residence during childhood as the geographic feature linking variation in lead emissions to exposure. Recall, the design of the CAA resulted in quasi-random variation in lead entering the environment as a function of gasoline lead concentrations that varied across states over each year during the phase out. Letting the state of residence during childhood be s, then for each individual i, BLL_{is} is blood

lead averaged over ages 12 to 36 months of a child's life. The remaining explanatory variables and the dependent variable correspond to adulthood, which we define as period t, corresponding to the years in which the data underlying the outcome Y was collected.[14]

Defining Y_{ijt} as an outcome capturing individual i's IQ or appetite for risk in year t and state of residence j as a function of childhood lead exposure, the relationship of interest is expressed as:

$$Y_{ijt} = \alpha + \beta_1 BLL_{is} + \boldsymbol{X}'_{ijt}\boldsymbol{\beta} + \boldsymbol{\Gamma}_1 \boldsymbol{R}_j + \boldsymbol{\Gamma}_2 \boldsymbol{Z}_t + \epsilon_{ijt} \tag{2.5}$$

where, $\boldsymbol{X}_{ijt}$ is a vector of covariates controlling for demographic, socioeconomic, and heritable traits specific to the individual. Fixed effects for the state of residence and survey year are $\boldsymbol{R}_j$, and $\boldsymbol{Z}_t$ respectively.

The coefficient of interest in Equation 2.5 is β_1, capturing the expected marginal effect of a change in childhood lead exposure on a later in life outcome Y. Since BLL is an indicator of the underlying childhood lead exposure channel, a non-zero $\hat{\beta}_1$ which is statistically significant supports the hypothesis that childhood lead exposure can partly explain abnormal risk preferences in adulthood.

The first stage regression is described by Equation 2.6. Reductions in lead emissions following the regulations imposed by the CAA are captured in the variable $\overline{TEL}_{js}$, corresponding to the average gasoline lead concentration in state j and year s over the period in which NHANESII was conducted.[15] Then, using blood lead levels (BLL_{ijs}) as a measure of individual i's lead exposure, the affect of regulations on the lead exposure channel is modeled as:

$$BLL_{ijs} = \alpha + \gamma \overline{TEL}_{js} + \boldsymbol{Y}_{is}{}'\boldsymbol{\theta} + \epsilon_{ijs} \tag{2.6}$$

where the vector $\boldsymbol{Y}_{is}$ includes controls for age, sex, race, and household income.[16]

---

[14]For IQ outcomes, t is corresponds to round 1 (1997). For abnormal risk outcomes, t corresponds to rounds 14 and 15 (2010, 2011).

[15]The use of year subscript s is to re-enforce the time period across both surveys. NHANESII is conducted over the years 1976 to 1980, and all respondents in the NLSY97 sample are born between 1980 and 1984.

[16]Respondent's age is in years; sex (Female) and racial status (Black) are measured dichotomously, where Female=1 if non-male and Black=1 if black; indicator variables for household income are defined following Reyes

The coefficient estimates obtained from fitting Equation 2.6 with the NHANESII sample are then used to predict BLL's in the NLSY97 sample. Using the same set of regressors, blood lead is predicted for each respondent in the NLSY97 sample.

$$\widehat{BLL}_{ijs} = \widehat{\alpha} + \widehat{\gamma}\overline{TEL}_{js} + \boldsymbol{Y}_{is}{}'\widehat{\boldsymbol{\theta}} \tag{2.7}$$

Substituting $\widehat{BLL}_{ijs}$ from Equation 2.7 into Equation 2.5 produces the final specification for our analysis:

$$Y_{ijt} = \alpha + \beta_1\widehat{BLL}_{ijs} + \boldsymbol{X}_{ijt}{}'\boldsymbol{\beta} + \boldsymbol{\Gamma}_1\boldsymbol{R}_j + \boldsymbol{\Gamma}_2\boldsymbol{Z}_t + \epsilon_{ijt} \tag{2.8}$$

The aim of our approach is to estimate the BLL effects of lead exposure, which is widely recognized as a reliable measure of recent lead exposure risk across the academic literature.[17]

In defense of the identification strategy described in Section 2.2, our approach facilitates analysis at a national scale and over a period of time in which the risk of lead exposure varied as a direct result of an exogenous policy shock. The primary data sources are large, nationally representative, surveys, with approximately 2,400 observations of children with BLLs in the NHANESII data and approximately 4,500 respondents from NLSY97 satisfying the inclusion criteria to predict their BLL. Furthermore, respondents in the NLSY97 sample were born in the years immediately following the period in which the NHANESII was conducted. The NHANESII survey was conducted 1976 to 1980 and the cohort selected for the NLSY97 was born between 1980 to 1984, with the first interview taking place in 1997. Moreover, this time period coincides with the implementation and enforcement of the CAA, spanning the years in which the majority of TEL reductions took place, implying dramatic reductions in the risk of lead exposure attributable to automobile emissions. Given the NHANESII survey preceded the birth of NLSY97 respondents, coupled with

(2015) and derived from household income relative to Federal Poverty Line (FPL), where Low=1 for household income $< 2 \times$ FPL and Mid=1 for household income $\in [2 \times$ FPL, $3 \times$ FPL);

[17]To account for the short-term nature of this lead exposure metric, we apply a similar strategy to that of Reyes (2015) and estimate each respondent's blood lead for each month between ages one and three (age 12-36 months) for each respondent in the NLSY97 sample. we then take the two year centered average of estimated monthly blood lead to reflect the body burden associated with persistent exposure in early childhood.

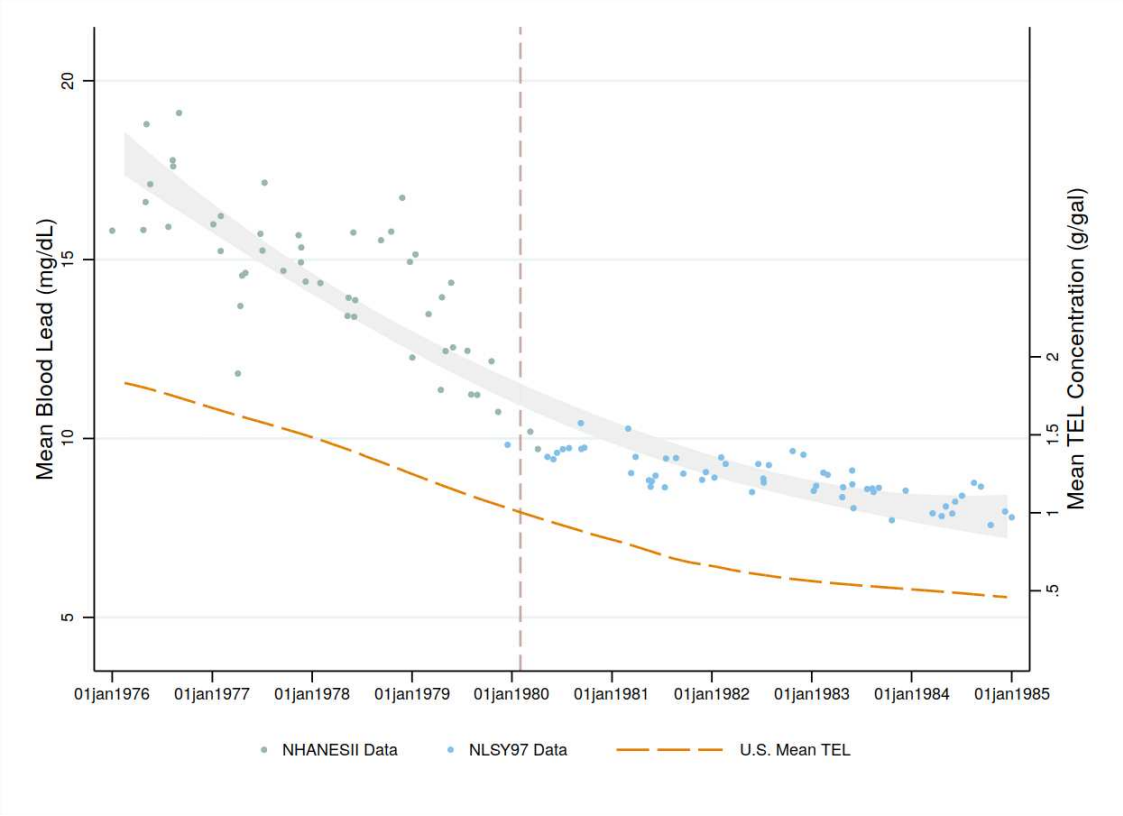


Figure 2.2: US average blood lead and gasoline lead (TEL) concentrations, 1976-1985

the year-over-year reductions in lead exposure from falling TEL concentrations, predicted BLLs among NLSY97 respondents should underestimate the true lead exposure risk during this period. This implies any negative effects of lead exposure on the response variables are likely underestimates using this empirical strategy. Take Figure 2.2 for example, which graphs the annual trends of national $\overline{TEL}$ and BLLs over the phase out. At the national level, average concentrations decreased steadily over a ten-year period beginning in 1976. The vertical line in Figure 2.2 separates the NHANESII data on the left, from the NLSY data on the right. The downward trend observed in $\overline{TEL}$ is mirrored by national average BLLs among NHANESII respondents, as well as the predicted values among NLSY97 respondents.

2.4 Results

2.4.1 Evidence of gasoline lead effects

Results from the first stage regressions in which Equation 2.6 is fit using a linear estimator and the NHANESII data are shown in Table 2.2. Observations are restricted to children aged six years and under at the date of examination.[18] The predictor of interest is average, annual state gasoline lead concentrations $\overline{TEL}$ from Equation 2.4. Control variables are included for respondent's sex ($Female$), racial status ($Black$), and Age in years. Indicator variables for household income are constructed following the definitions of Reyes (2015) and are household income relative to the Federal Poverty Level (FPL), where Low=1 for household income $< 2 \times$ FPL and Mid=1 for household income $\in [2 \times \text{FPL}, 3 \times \text{FPL})$. The dependent variable in Table 2.2 is BLL (μg/dL) or natural log of BLL where noted.

Across each of the seven specifications, $\overline{TEL}$ is a significant predictor of blood lead in children. Column 1 is the most sparse model, where child BLLs are regressed on $\overline{TEL}$ and month fixed effects. The estimated effect of a 1 g/gal reduction in $\overline{TEL}$ decreases BLL by 3.9 μg/dL (95% CI: 1.44, 6.50). The estimated marginal effect of gasoline lead concentration remains highly significant when state and year fixed effects are introduced in Column 2. The linear-linear specification in Column 2 highlights the impact of the geographic and temporal variation as the addition of year and state controls remove significance from the constant term, while the estimated effect of $\overline{TEL}$ is more than double that of the spare model in Column 1 (estimated marginal effect of 9.4). Column 3 is the fully saturated model reflecting Equation 2.6, with an estimated 8.1 μg/dL reduction in blood lead for a 1 g/gal decrease in $\overline{TEL}$ concentration (95% CI: 1.03, 15.14). As expected, age, racial status, and low income group are highly statistically significant correlates of child BLLs, and in the direction compatible with previous research (Campanella and Mielke, 2008). The results suggests that mean BLLs are an estimated 5 μg/dL higher for Blacks relative to Whites and Hispanics, and those in the Low income group (less than 2 times the Federal Poverty Level) have BLLs an estimated 1.9 μg/dL higher relative to those in the highest income group.

[18] This inclusion criteria is motivated by the sampling design of the NHANESII survey, in which all children in this age group were selected for laboratory analysis.

Table 2.2: Blood lead effect of gasoline lead (TEL) concentrations, linear regressions (NHANESII)

	(1) Blood Pb	(2) Blood Pb	(3) Blood Pb	(4) ln(Blood Pb)	(5) ln(Blood Pb)	(6) Blood Pb	(7) Blood Pb
Gasoline Lead (g/gal)	3.966***	9.390**	8.085**	0.630**	0.567**	7.887**	8.280**
	(1.257)	(4.354)	(3.506)	(0.278)	(0.236)	(3.505)	(3.774)
Age			-0.495***		-0.023***	-0.495***	-0.536***
			(0.140)		(0.007)	(0.141)	(0.140)
Female			-0.445*		-0.029*		-0.433*
			(0.234)		(0.015)		(0.246)
Black			5.151***		0.288***	5.149***	
			(0.502)		(0.021)	(0.496)	
Low			1.912***		0.110***	1.922***	2.465***
			(0.316)		(0.021)	(0.317)	(0.363)
Mid			-0.103		-0.012	-0.085	-0.318
			(0.425)		(0.026)	(0.428)	(0.397)
Constant	8.094***	-1.083	0.860	1.471**	1.547***	1.048	2.258
	(2.310)	(8.815)	(7.164)	(0.557)	(0.472)	(7.158)	(7.656)
Observations	2,430	2,430	2,430	2,430	2,430	2,430	2,430
R^2	0.086	0.192	0.301	0.260	0.368	0.301	0.243
State FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Ages	0-6	0-6	0-6	0-6	0-6	0-6	0-6

Notes: Blood lead is μg/dL; Robust state clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Gasoline lead is annual, state, averaged TEL (g/gal); Age is respondent's age in years when survey was conducted; Indicators for sex and racial status are defined as Female=1 and Black=1 if the respondent identifies as African American; Indicators for household income relative to Federal Poverty Level (FPL) are: Low=1 for < 2x FPL and Mid=1 for ratios $\leq$2x FPL to <3x FPL; State of residence, and Year and Month of survey fixed effects are included in regressions;

The remaining columns in Table 2.2 are included as sensitivity and robustness tests. Columns 4 and 5 substitute the dependent variable with the natural log of blood lead. In both cases, the results are largely unchanged with gasoline lead remaining statistically significant at the 95% confidence level. Similarly, coefficient estimates for Age, Black, and Low income remain highly statistically significant at the 99% confidence level. The fully saturated log-linear specification shown in Column 5 is the model used to predict BLLs for the NLSY97 respondents.

Lastly, sensitivity tests are included in Columns 6 and 7 using the linear-linear model from Column 3.[19] Removing the indicator for respondent sex in Column 6 reduces the coefficient estimate to a 7.8 μg/dL reduction in BLLs given a one unit decrease in gasoline lead, yet remains highly significant. Similarly, removal of the race indicator variable Black returns a slightly higher estimated marginal effect of gasoline lead on blood lead, suggesting a 1 g/gal reduction in gasoline lead reduces BLLs by 8.2 μg/dL and is significant at the 95% confidence level.

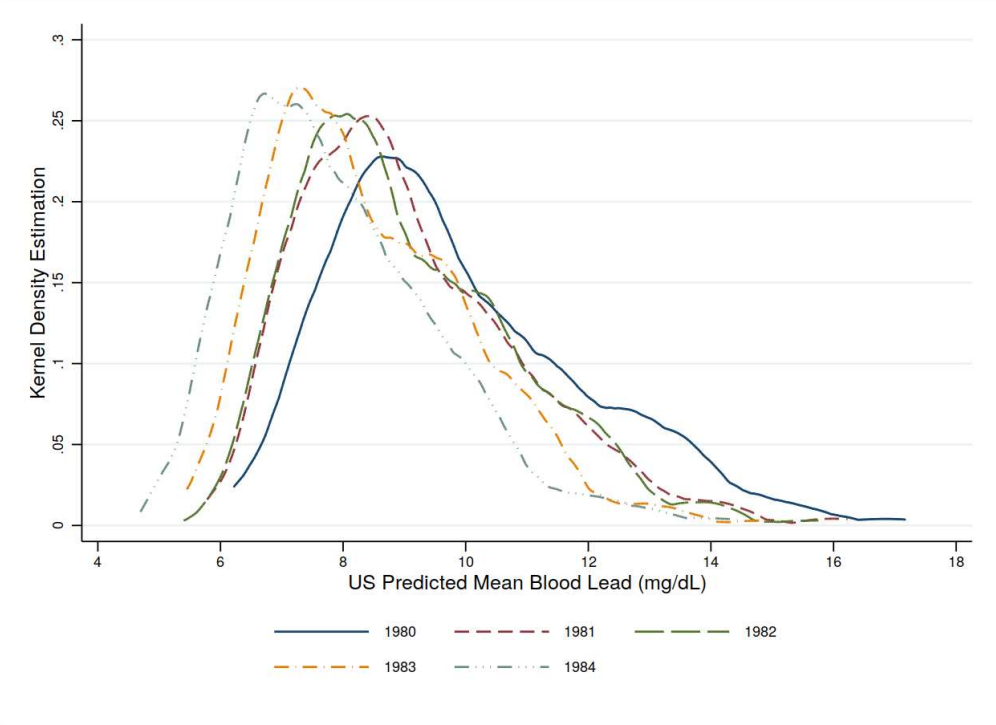


Figure 2.3: US average predicted blood lead kernel density

[19]We have chosen to include results of sensitivity tests using the linear-linear model over the log-linear model for ease of interpretation. Using the log-linear specification shown in Table 2.2, Columns 4 and 5 for this exercise returns results which are effectively identical.

Figure 2.3 plots kernel density estimates of BLLs by year of birth for the NLSY97 sample. Here, BLLs are the predicted values using the log-linear specification from Table 2.2, Column 5.The reductions in gasoline lead emissions during this period are evident in the behavior of uniform shifts in BLLs among the NLSY97 respondents. As $\overline{TEL}$ falls in time, BLLs follow a similar pattern as mean BLL decreases with each successive birth year. Similarly, Figure 2.3 shows the the year-over-year density distributions becoming more Leptokurtic as the right skew diminishes in mean BLLs across birth year cohorts. Descriptive statistics on NLSY97 predicted BLL are included in Table 2.3.

Table 2.3 reports descriptive statistics on NLSY97 cognitive outcomes and risk preference variables for the 1980-1984 cohort, as well as description of these outcomes across early childhood blood lead level quintiles. IQ is measured by respondents scores on the Armed Services Vocational Aptitude Battery (ASVAB) examination. All NLSY97 respondents were administered the examination in round one of the survey, and is used as a proxy for cognitive intelligence. The ASVAB has been shown to be highly correlated with other accepted standardized achievement tests, such as the SAT and ACT exams (Koenig et al., 2008). Of particular interest in Table 2.3 is the behavior of IQ percentile and Low-IQ across child blood lead quantiles. IQ percentile decreases dose-responsively across blood lead quantiles, from 57 (Q1) to 55 (Q2) to 50 (Q3) to 46 (Q4) and to 34 (Q5). Intuitively, the fraction of sampled adults scoring below the 25th percentile on the ASVAB exam increases dose-responsively in child blood lead quantiles.

Results in Table 2.4 extend the descriptive associations in Table 2.3, showing OLS models of adult percentile score on the ASVAB exam regressed on child blood lead and control variables. The predictor of interest is blood lead (μg/dL), which is the average predicted blood lead between 12 to 36 months old. Controls for respondent's sex (Sex), racial status ($Minority$) and fixed effects for survey sample group are included in all regressions.[20] State of residence and birth year cohort fixed effects are included where noted. Given prior literature showing a negative association childhood

[20]Indicators are Sex=1 for male and zero otherwise; $Minority$=1 for non-white racial status and zero otherwise; Sample group fixed effect is based on NLSY97 survey design and is an indicator with a value of 1 for respondents in the Cross Sectional Sample and 0 otherwise (e.g. respondent is classified as Supplemental Sample);

Table 2.3: Descriptive statistics - IQ and risk preferences

		Blood Lead Quintiles				
	Mean	Q1	Q2	Q3	Q4	Q5
Blood Lead (μg/dL)	8.88	6.53	7.51	8.27	9.18	11.36
Gasoline Lead (g/gal, ages: 0-36 m.)	0.50	0.43	0.48	0.50	0.49	0.55
IQ Percentile	47	57	55	50	46	34
Low-IQ ($< 25^{th}$ Percentile)	31%	17%	20%	26%	33%	51%
Left Handed (throwing)	8%	7%	7%	9%	7%	9%
Right Handed (writing)	89%	90%	91%	88%	89%	88%
Survey Questions (Rounds 14,15)						
General	5.58					
Driving	2.76					
Finance	3.89					
Work	4.64					
Health	2.91					
People	4.33					
Romance	4.42					
Life	5.04					
Gambling	5.37					
Sample Size	4,836	823	851	865	927	1,370

Notes: Blood lead is predicted blood lead (μg/dL); IQ Percentile is based on ASVAB examination scores; Low-IQ is an indicator for exam scores at or below the 25^{th} percentile; Survey questions are abnormal response values calculated as the absolute value of the sample mean less the quintile mean response;

Table 2.4: IQ effect of blood lead, linear regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Blood Lead (μg/dL)	-0.018***	-0.021***	-0.052***	-0.027***	-0.027***	-0.028***	-0.045***
	(0.005)	(0.006)	(0.005)	(0.005)	(0.006)	(0.005)	(0.005)
Sex	-0.023**	-0.022**	-0.021**	-0.027***	-0.016		-0.023**
	(0.009)	(0.009)	(0.009)	(0.009)	(0.009)		(0.009)
Minority	-0.173***	-0.167***	-0.081***	-0.092***	-0.054**	-0.089***	
	(0.023)	(0.025)	(0.019)	(0.019)	(0.022)	(0.019)	
Poverty Ratio (Q1 Reference Group)							
Poverty Ratio Q2				0.067***	0.013	0.065***	0.062***
				(0.017)	(0.019)	(0.017)	(0.017)
Poverty Ratio Q3				0.130***	0.030**	0.128***	0.122***
				(0.011)	(0.015)	(0.011)	(0.013)
Poverty Ratio Q4				0.194***	0.015	0.192***	0.190***
				(0.016)	(0.016)	(0.016)	(0.016)
Endowment					0.092***		
					(0.003)		
Constant	0.654***	0.697***	0.931***	0.624***	0.692***	0.626***	0.722***
	(0.040)	(0.051)	(0.044)	(0.045)	(0.053)	(0.045)	(0.053)
Observations	4,587	4,587	4,587	4,587	3,219	4,587	4,587
R^2	0.182	0.184	0.232	0.278	0.417	0.276	0.270
State FE	No	No	Yes	Yes	Yes	Yes	Yes
Cohort FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Oversample FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variable is percentile score on ASVAB examination; Robust state clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Blood lead is predicted mean blood lead in μg/dL; Poverty ratio quartiles are based on household income to poverty ratio, 1997 (survey round one); Endowment is the first principle component among mother's age at respondent's time of birth, and years of education for the respondent and biological mother and father; State of residence, over-sample group and cohort fixed effects included where noted;

lead exposure and cognitive IQ, the expectation for results in Table 2.4 is a negative coefficient estimate for blood lead.

In the spare model shown in Column 1, the coefficient estimate for blood lead suggests a 1 μg/dL increase in childhood blood lead reduces IQ by 1.8 percentile points (95% CI: -0.027, -0.008). Column 2 includes fixed effects for respondent's year of birth (Cohort FE) and Column 3 introduces state of residence fixed effects.[21] Controls for exogenous variation in access to and quality of education, are included in the remaining columns of Table 2.4 and analytical emphasis is placed on Column 4. A set of indicators for household income to Federal Poverty Level ratio, by quartiles are introduced in Column 4. With the introduction of socio-economic controls, the predicted marginal effect in Column 4 amplifies, with a unit increase in child BLL inducing a 2.7 percentile point reduction in IQ scores in young adulthood (95% CI: -0.036, -0.017). As expected, the coefficient estimates attached to poverty ratio quartiles ($Poverty\ Ratio$) are highly significant and positively correlated with IQ. Relative to respondents in the lowest income quartile (Q1), those in the second quartile score an estimated 6.7 percentile points higher on the ASVAB examination, all else equal. Similarly, those in the third and fourth quartiles score an estimated 13 and 19 percentile points higher than their counterparts in the lowest income quartile respectively. To place the estimated effect of blood lead on IQ into context, consider the difference in blood lead concentrations shown in Table 2.3 between those in the lowest (Q1) and highest (Q5) quintiles. Applying the coefficient estimate from Column 4 to the difference in average blood lead concentrations, we find an estimated differential of 13 IQ percentile points between individuals in the highest exposure group, Q5 (11.36 μg/dL) to those in Q1 (6.53 μg/dL). This is comparable to the estimated effect of socio-economic status between individuals in Q3 relative to their counterparts in the lowest socio-economic group Q1.

The fully saturated model shown in Column 5 adds a composite variable $Endowment$ to capture the factors potentially impacting educational outcomes and performance on standardized

[21] State of residence is defined as the state in which the respondent was living when the survey round was conducted. Data for outcomes relating to IQ were collected during round 1 (1997), and state fixed effects included in results of Table 2.4 and Table 2.5 are place of residence in 1997.

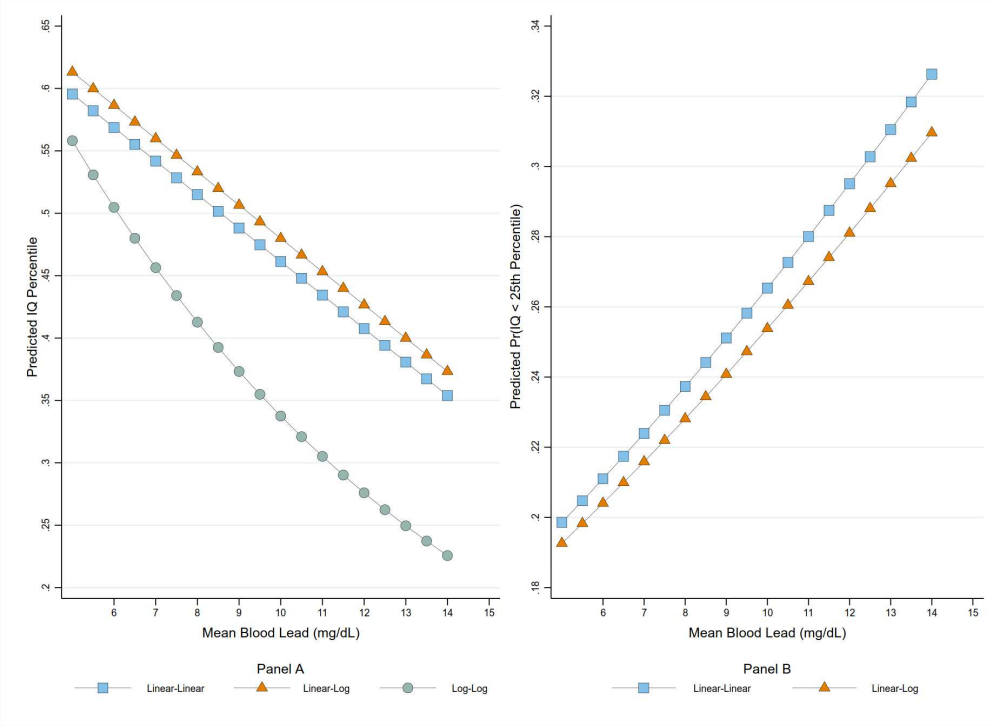


Figure 2.4: Marginal effect of blood lead on IQ outcomes

tests.[22] Again, blood lead remains highly significant with an estimated marginal effect of 2.7 percentile point reduction in IQ for a 1 μg/dL increase in BLL (95% CI: -0.04, -0.014). The sign of the coefficient estimate for $Endowment$ is in line with expectations, with IQ percentile score increasing in mother's age and the educational attainment of the respondents parents. A standard deviation increase in the endowment variable increases cognitive IQ by 9.2 percentage points. Finally, sensitivity tests are included in Columns 6 and 7. When controls for sex and racial status are removed, the results are largely unchanged and remain highly statistically significant in both cases. Results of Table 2.4 are illustrated in Figure 2.4, Panel A. Here, the estimated marginal effects are based on the linear-linear specification shown in Table 2.4, Column 4. The linear-log and and log-log variations have been included for comparison. Together, the marginal effects of blood lead on IQ shown in Panel A highlight the significant dose response relationship across all three functional forms.

[22]The variable $Endowment$ is the predicted first principal component across the following dimensions: age of the respondent's biological mother at the time of birth, educational attainment of the respondent, and educational attainment of both his biological mother and father.

Converting the continuous IQ outcome to an indicator for a low-IQ ASVAB score, we re-estimate the models used in Table 2.4 and present results in Table 2.5. Here, we define the outcome as low-IQ=1 for respondents with ASVAB scores in the at or below the 25^{th} percentile and zero otherwise. Controls are the same as those found in Table 2.4, including fixed effects for state of residence, NLSY97 sample group, and birth year.[23] The coefficient estimates for blood lead are expected to be strictly positive following the results in Table 2.4 where IQ is statistically significantly decreasing in blood lead exposure.

Again, emphasis is placed on the saturated model in Column 4 of Table 2.5. The spare model in Column 1 suggests a 1 μg/dL increase in BLL increases the probability of falling into the low-IQ group by an estimated 2.8 percentage points (95% CI: 0.015, 0.039), and an estimated 3.2 percentage points when cohort fixed effects are added in Column 2. The inclusion of state of birth fixed effects in Column 3 raises the marginal effect of BLL on the probability of a low-IQ outcome by an estimated 7.7 percentage points (95% CI: 0.059, 0.094). Household income to FPL quartiles are included in Column 4, and the estimated marginal effects on a low-IQ outcome, decrease in income relative to the lowest income group. BLL, sex and minority status remain highly statistically significant, with an estimated 4.3 percentage point increase in the probability of a low-IQ outcome per 1 μg/dL increase in BLL (95% CI: 0.025, 0.06). In Column 5, the previously described composite index variable $Endowment$ is introduced and the coefficient estimate is non-positive and highly significant. Given the inter-generational relationship of educational attainment, the coefficient estimate of -0.099 for endowment is within plausible range. The coefficient estimates for blood lead and minority remain statistically significant but are reduced to 0.039 and 0.088 respectively. Columns 6 and 7 of Table 2.5 are included as sensitivity tests based on the saturated model used in Column 4. As with Table 2.4, the exclusion of Sex in Column 6 does not result in a significant change on the estimated effect of BLL, nor does the exclusion of Minority status in Column 7. In both cases, the underlying lead exposure – IQ outcome channel remains highly statistically significant.

[23] We have chosen to use a linear probability model due to the need for time invariant geographic and sample group fixed effects.

Table 2.5: Low-IQ effect of blood lead, linear probability regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Blood Lead (μg/dL)	0.028***	0.032***	0.077***	0.043***	0.039***	0.046***	0.065***
	(0.006)	(0.006)	(0.009)	(0.009)	(0.011)	(0.009)	(0.009)
Sex	0.048***	0.046***	0.045***	0.052***	0.038***		0.048***
	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)		(0.012)
Minority	0.232***	0.224***	0.097***	0.114***	0.088***	0.108***	
	(0.033)	(0.034)	(0.030)	(0.025)	(0.030)	(0.025)	
Poverty Ratio (Q1 Reference Group)							
Poverty Ratio Q2				-0.113***	-0.053**	-0.108***	-0.107***
				(0.020)	(0.026)	(0.021)	(0.020)
Poverty Ratio Q3				-0.187***	-0.072***	-0.182***	-0.177***
				(0.023)	(0.024)	(0.023)	(0.025)
Poverty Ratio Q4				-0.242***	-0.042*	-0.238***	-0.237***
				(0.025)	(0.025)	(0.025)	(0.026)
Endowment					-0.099***		
					(0.004)		
Constant	0.031	-0.025	-0.415***	0.003	-0.066	-0.001	-0.119
	(0.055)	(0.061)	(0.091)	(0.097)	(0.107)	(0.099)	(0.109)
Observations	4,587	4,587	4,587	4,587	3,219	4,587	4,587
R^2	0.145	0.146	0.185	0.213	0.275	0.210	0.208
State FE	No	No	Yes	Yes	Yes	Yes	Yes
Cohort FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Oversample FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variable is low-IQ, where low-IQ=1 for ASVAB score $\leq 25^{th}$ percentile; Robust state clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Blood lead is predicted mean blood lead in μg/dL; Poverty ratio quartiles are based on household income to poverty ratio, 1997 (survey round one); Endowment is the first principle component among mother's age at respondent's time of birth, and years of education for the respondent and biological mother and father; State of residence, over-sample group, and cohort fixed effects included where noted;

Panel B in Figure 2.4, illustrates the dose response relationship estimated in Table 2.5, Column 4, with a linear-log variation included for comparison. Results of additional sensitivity and robustness tests of the blood lead – IQ channel are included in Appendix Table C.2.

Now that we have established a link between childhood lead exposure and cognitive intelligence in early adulthood, we turn attention to the possible link between child BLLs and abnormal risk preferences in various specific domains of risk.

Beginning with descriptive analysis, consider Figure 2.5 depicting the relationship between average childhood BLL and risk preference questions collected during rounds 14 and 15 of the NLSY. Along the x-axis are response values from the NLSY97 risk preference sample, plotted against respondent's BLLs in units of μg/dL on the y-axis. Sample means for each risk preference question are identified by the vertical reference lines.[24] In each case, the sample mean response is nearest to the minimum sample mean BLL. One can also see that average BLLs are substantially higher at both ends of the response distribution. For both the left tail (least willing to accept risk) and the right tail (completely willing to accept risk), mean BLLs increase sharply. Compatible with the expectations from the field of abnormal psychology, early childhood lead exposure increases the likelihood of risk preferences in adulthood deviating from central tendency.

Results modeling the effect of blood lead on abnormal risk preferences are presented in Table 2.6. Recall, the dependent variable is constructed using Equation 2.1 detailed in Section 2.3.1, constituting a measure of deviation from central tendency with respect to self reported risk preferences. Control variables match those of Table 2.4, though Table 2.6 has been truncated for clarity. Full results for the models in which blood lead is statistically significant are included in Appendix Table C.3.

Across a series of contexts, we find that childhood BLLs significantly increase deviation from normal willingness to assume risk in adulthood. Specifically, a 1 unit increase in child BLL, increases deviation from the norm by 2.1% for driving, 2.3% for personal finances, 2.9% for work,

[24]Modified survey sample weights have been applied to account for non-response as well as observations for which BLLs could not be assigned. See National Longitudinal Survey of Youth 1997 cohort for the NLSY's documentation on survey design and methodology for calculating custom weights (Accessible https://www.nlsinfo.org/content/cohorts/nlsy97/using-and-understanding-the-data/sample-weights-design-effects).

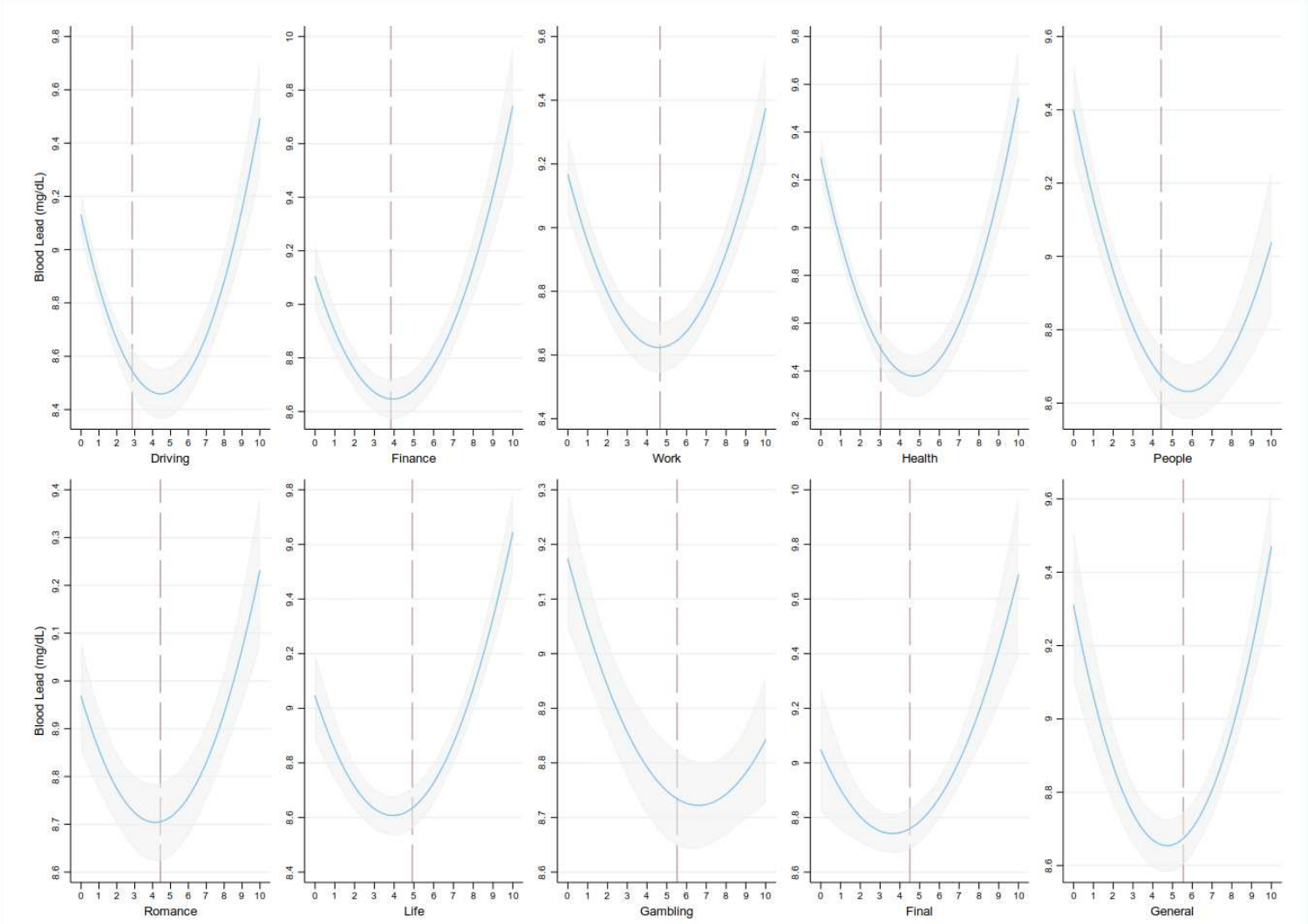


Figure 2.5: Abnormal risk preferences and blood lead

Table 2.6: Abnormal risk response (mean deviation) effect of blood lead, linear regressions

	(1) General	(2) Drive	(3) Finance	(4) Work	(5) Health	(6) People	(7) Romance	(8) Life	(9) Gamble	(10) Overall
Blood Lead (μg/dL)	0.007	0.021*	0.023***	0.029***	0.034*	0.014	0.017**	0.010	0.024**	0.017***
	(0.010)	(0.011)	(0.008)	(0.008)	(0.019)	(0.011)	(0.009)	(0.018)	(0.012)	(0.005)
Constant	0.449***	0.613***	0.291***	0.333***	0.144	0.594***	0.737***	0.114	0.736***	0.791***
	(0.104)	(0.120)	(0.107)	(0.107)	(0.237)	(0.122)	(0.108)	(0.211)	(0.126)	(0.056)
Observations	4,224	4,224	4,224	4,224	4,224	4,224	4,224	4,224	4,224	4,224
R^2	0.046	0.040	0.046	0.033	0.044	0.031	0.031	0.026	0.019	0.099
State FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variables are abnormal risk preferences (mean deviation), based on NLSY97 survey responses; Robust state clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Blood lead is predicted mean blood lead in μg/dL; Indicators controlling for for sex (Sex=1 if male), minority status (Minority=1, if non-white), and marital status (Married=1 if married when surveyed); Poverty ratio quartiles are based on household income to FPL poverty ratios; State of residence, age cohort, and over-sample survey group fixed effects included in all regressions;

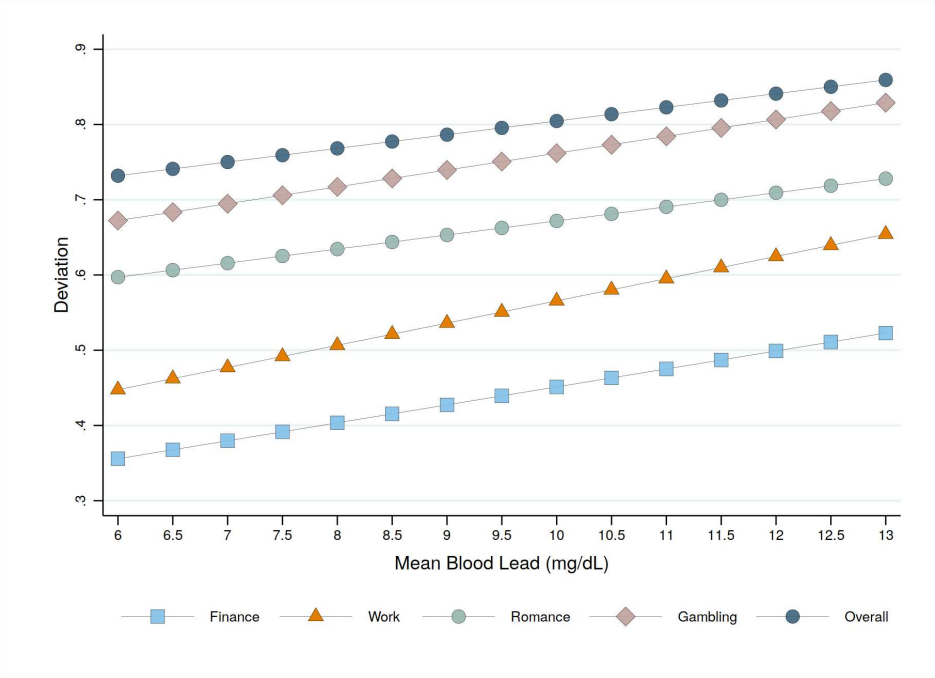


Figure 2.6: Marginal effect of blood lead on abnormal risk response

3.4% for personal health, 1.7% for romantic relationships, 2.4% for gambling, and 1.7% overall (representing the average across all variables).

Figure 2.6 illustrates the estimated marginal effect of blood lead across the response surface. Across the subset of outcomes with an estimated statistical significance of at least 90%, the estimated dose response is linear and near uniform across all risk preference outcomes.

2.4.2 Robustness and falsification

Robustness results are included in Table 2.7, which replace blood lead with average gasoline lead concentrations ($\overline{TEL}$). Borrowing the same set of regressors as those found in Table 2.4, the results are similar to those above. Across all four models, $\overline{TEL}$ has a statistically significant negative impact on IQ. At the 90% confidence level, a 1 g/gal increase in $\overline{TEL}$ reduces IQ by an estimated 11 to 15 percentile points. The controls for sex, minority status, and household income behave in a similar manner to those in Table 2.4.

As a falsification test, we model the effect of blood lead against an individual's dominant hand preference while writing and throwing. Borrowing the same notation in Equation 2.8, let s be the

Table 2.7: Robustness tests of IQ outcomes and average gasoline lead (TEL) concentrations

	(1) IQ Percentile	(2) IQ Percentile	(3) IQ Percentile	(4) Low-IQ	(5) Low-IQ	(6) Low-IQ
Gasoline Lead (g/gal)	-0.059**	-0.166**	-0.120*	0.099**	0.215**	0.155**
	(0.028)	(0.066)	(0.061)	(0.044)	(0.086)	(0.074)
Sex	-0.027***	-0.027***	-0.029***	0.054***	0.054***	0.057***
	(0.009)	(0.008)	(0.008)	(0.012)	(0.012)	(0.011)
Minority	-0.206***	-0.203***	-0.152***	0.283***	0.280***	0.211***
	(0.019)	(0.018)	(0.016)	(0.029)	(0.028)	(0.025)
Poverty Ratio (Q1 Reference Group)						
Poverty ratio Q2			0.093***			-0.151***
			(0.015)			(0.019)
Poverty ratio Q3			0.166***			-0.242***
			(0.011)			(0.023)
Poverty ratio Q4			0.228***			-0.292***
			(0.013)			(0.021)
Constant	0.536***	0.601***	0.453***	0.213***	0.141**	0.345***
	(0.028)	(0.053)	(0.048)	(0.040)	(0.067)	(0.059)
Observations	4,587	4,587	4,587	4,587	4,587	4,587
R^2	0.173	0.176	0.252	0.137	0.138	0.191
State FE	No	No	No	No	No	No
Cohort FE	No	Yes	Yes	No	Yes	Yes
Oversample FE	Yes	Yes	Yes	Yes	Yes	Yes
Age (months)	13-24	13-24	13-24	13-24	13-24	13-24

Notes: Dependent variables are percentile score on ASVAB examination and low IQ ($<25^{th}$ percentile) indicator; Robust state clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Gasoline lead concentrations (g/gal), calculated as the 12 month state average for each respondent, by state of and year of birth; Indicators are defined as: Sex=1 if male, Minority=1 for non-white race; Poverty ratio quartiles are based on household income to poverty ratio, 1997 (survey round one); State of residence, over-sample group, and cohort fixed effects included where noted;

state of residence during childhood and t denote the year in which survey round 1 was conducted (1997). Then for individual i, residing in state j we estimate effect of childhood lead exposure on dominant hand preference using the following linear probability model:

$$H_{ijt} = \alpha + \beta_1 \widehat{BLL}_{ijs} + \boldsymbol{X}_{ijt}'\boldsymbol{\beta} + \boldsymbol{\Gamma}_1 \boldsymbol{R}_j + \boldsymbol{\Gamma}_2 \boldsymbol{Z}_t + \epsilon_{ijt} \tag{2.9}$$

where, all terms carry from Equation 2.8 and the dependent variable H represents one of two dichotomous outcomes to be modeled: $Writing$ or $Throwing$. We define $Writing$=1 if individual i writes left-handed and zero otherwise, and $Throwing$=1 if i throws right-handed and zero otherwise.

There is no evidence in scientific literature linking early in life lead exposure to the development of a dominant hand in these activities. Therefore, expectations are that the predicted blood lead effect on dominant hand preferences are insignificant in all cases. The sample descriptive statistics in Table 2.3 illustrate this random process as the sample means for both $Writing$ and $Throwing$ outcomes mirroring the means across BLL quintiles.

Results of this strategy are presented in Table 2.8. The dependent variables are indicators for whether a respondent is left handed while writing ($Writing$=1) or right handed while throwing ($Throwing$=1). Fitting linear probability models, the estimates in Table 2.8 show that blood lead is not a statically significant predictor of dominant hand preference. These results hold across the five specifications which are based on those used in the primary analysis, and vary with respect to the choice of covariates and fixed effects included.

Figure 2.7 graphs the estimates of the fully saturated models shown in Table 2.8, Columns 2 and 5. A linear-log specification has been included for comparison. For both outcomes, left-handed throwing (Panel A) and right-handed writing (Panel B), the estimated marginal effect is statistically insignificant across the response surface.

Table 2.8: Dominant hand effect of blood lead falsification test, linear probability regressions

	(1) Writing	(2) Writing	(3) Throwing	(4) Throwing	(5) Throwing
Blood Lead (μg/dL)	-0.001	-0.002	-0.000	-0.002	-0.003
	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)
Sex	-0.025**	-0.025**	0.020**	0.020**	0.021**
	(0.012)	(0.012)	(0.008)	(0.008)	(0.008)
Minority	-0.017	-0.024*	0.018	0.021	0.015
	(0.011)	(0.012)	(0.013)	(0.014)	(0.013)
Age	-0.004	-0.004		0.004	0.006**
	(0.003)	(0.003)		(0.003)	(0.003)
Constant	1.008***	1.023***	0.074**	-0.030	-0.031
	(0.083)	(0.084)	(0.031)	(0.075)	(0.073)
Observations	4,362	4,362	4,555	4,555	4,361
R^2	0.016	0.017	0.011	0.012	0.016
State FE	Yes	Yes	Yes	Yes	Yes
Oversample FE	No	Yes	Yes	Yes	Yes
Poverty Ratio FE	Yes	Yes	No	No	Yes

Notes: Dependent variables are indicators for left-handed while writing (Writing=1) and right-handed while throwing (Throwing=1); Robust state clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Blood lead is predicted mean blood lead in μg/dL; Indicators for sex (Sex=1 if male) and minority status (Minority=1 if non-white); Age is respondent's age in years when surveyed; Fixed effects for income to poverty ratio quartiles, state of residence, and over-sample group included where noted;

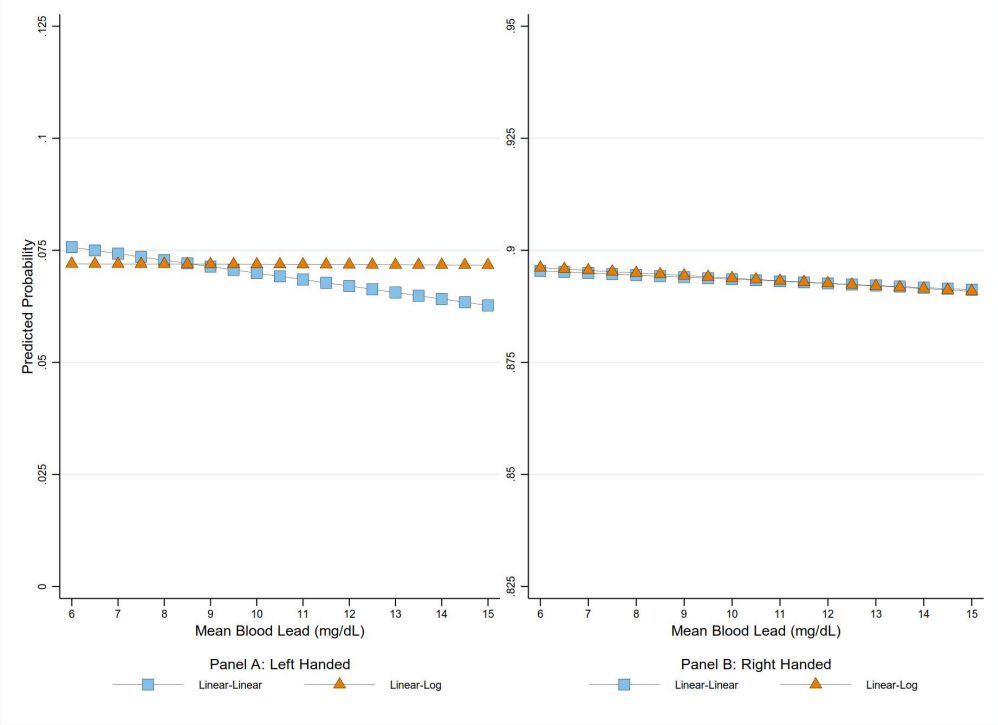


Figure 2.7: Falsification test of blood lead effect on dominant hand, predicted margins

2.5 Discussion and conclusion

Leveraging the spatial and temporal variation in lead exposure resulting from the differential phase-out of leaded gasoline, we find evidence in support of the lasting effects of early in life lead exposure on cognitive impairment and abnormal risk preferences in adulthood. Through an ensemble of evidence, we find that an increase in average childhood blood lead levels: 1) decreases IQ as measured by scores on the ASVAB examination; 2) increases the likelihood of a low-IQ (lowest quartile) ASVAB score; and 3) increases the extent to which responses to risk and uncertainty deviate from the norm. These findings are consistent across a number of alternative specifications of the treatment variable, blood lead levels (μg/dL), including log transformation and converted to a set of binary variables by quintiles. Sensitivity tests in which the treatment variable is substituted with average gasoline lead produces similar results. Furthermore, we lean against the identification strategy underlying the analysis with a divergent validity test in which the blood lead effect on individual's dominant hand preference is estimated.

The legacy of lead in the United States is complex and intertwined with public health. As concerns over the toxicity of lead have grown over time, policy makers have responded with a series of national policies aimed at minimizing exposure to lead across society. Our analysis of the phase out of lead in gasoline under the Clean Air Act adds to the body of research relating early in life lead exposure to later in life outcomes. Over a fifteen year period between 1975 to 1990, enforcement of the policy would significantly reduce the risk of lead exposure by stemming the flow of lead into the environment from automobile emissions, the primary source of environmental lead at the time. Despite the immense returns to public health from the policy, our findings suggest the legacy of this period in US history will persist into the future, both in the lives of individuals exposed as children and the ongoing risk posed by residual lead in soil and dust distributed across the country.

Chapter 3

Heavy metal and light babies: Maternal lead exposure and birth outcomes: Evidence from natural experiments in the phase-out of leaded gasoline

3.1 Introduction

The health and human capital costs of lead exposure in early childhood are lasting and substantial. Studies link elevated blood lead in childhood to physical health problems in adulthood, including hypertensive disorders and the malfunction of renal and cardiovascular systems. The intellectual and behavioral effects of lead exposure include poor academic achievement and impaired cognition, increased risk of attention-deficit and hyperactivity disorders, aggression and violence and ensnarement in the criminal justice system. Brain imaging studies find that adults exposed to lead as children have reduced gray matter in brain regions that govern judgment and mood regulation (Cecil et al., 2008; Cecil, 2011). Mood regulation is a socioemotional trait that economists have linked convincingly to long-term life outcomes (Cunha et al., 2010; Almond and Currie, 2011; Doyle et al., 2013). On the long reach of lead exposure, Reyes (2015) has linked elevated blood lead levels in childhood to "an unfolding series of adverse behavioral outcomes: behavior problems as a child, pregnancy and aggression as a teen, and criminal behavior as a young adult."

It might be tempting to assume that lead exposure is a rear-view problem, at least in the United States. Blood lead levels in children in the United States have fallen since the 1990s, coincident with a series of actions that banned lead from paint, plumbing, food cans and automotive gasoline. However, lead remains a constituent in aviation gasoline used by an estimated 160,000 piston-engine aircraft. These aircraft consume over 200 million gallons of avgas annually (Kessler, 2013), implying a flow into the environment of about a million pounds of lead. For the estimated 16 million people–and 3 million children–who live within a kilometer of airport facilities that service piston-engine aircraft, the atmospheric disposition of lead-formulated aviation gasoline remains a

potentially serious source of exposure risk. In fact, studies have shown that children proximate to airports with piston-engine aircraft traffic have elevated blood lead levels (Miranda et al., 2011; Zahran et al., 2017b).

Unlike other criteria pollutants, lead is a metal that persists in the lived environment, painted on interior and exterior walls of homes built prior to 1960 and accumulated in urban soils. Children in the United States are exposed to these legacy sources of lead as dust associated with deteriorating or haphazardly removed lead-based paint in renovation or demolition of homes (Farfel et al., 2003, 2005a; Rabito et al., 2007), and through the ingestion of contaminated soils (involving hand-to-mouth behaviors) or inhalation of lead-concentrated soils re-suspended in summer months (Filippelli et al., 2005; Laidlaw et al., 2005, 2012; Zahran et al., 2013). Studies find that children residing in neighborhoods with high soil lead concentration are at higher risk of elevated blood lead levels (Mielke et al., 2007; Zahran et al., 2010, 2011, 2013).

Child exposure to lead therefore remains a risk in the United States through the flow source of aviation gasoline, and through the legacy sources of paint, soil/dust and water. As noted, these sources typically enter a child's body through inhalation and ingestion. An underappreciated medium of child exposure to flow and legacy sources of lead is through maternal blood transmission in pregnancy. Transmission of lead to the fetus occurs via diffusion across the placental barrier over the course of a pregnancy. Child lead exposure and consequent development therefore has a fetal origin (Almond and Currie, 2011).

To evaluate the maternal blood transmission exposure pathway, we leverage the phase out of leaded gasoline following the passage of the Clean Air Act (CAA). The result of this policy was significant reduction in the concentrations of lead in leaded gasoline grades between 1975 and 1985, and eventual removal in 1990. Though the policy was implemented at the national level, enforcement was at the producer level, creating significant variation in lead emissions from automobile exhaust across states and over time. We leverage this variation in lead emissions as a quasi-random vector of lead exposure, as lead dust from automobile emissions was a significant source of lead exposure. Using sales of leaded gasoline along with gasoline lead concentrations

data during the initial period of the phase out as a lead exposure channel, we estimate the causal effect of maternal lead exposure on birth outcomes between 1976 to 1980.

Across an ensemble of tests, we find that an increase in maternal blood lead: 1) decreases infant birthweight; 2) increases the risk of low and very low birthweight; 3) shortens gestation length; 4) increases the risk of prematurity; and 5) increases the risk of a low APGAR score. These results are robust to various operations of our treatment variable, the relaxation of inclusion criteria, accounting for possible fertility selection effects, and various convergent validity tests. Moreover, we satisfied a divergent validity test by falsifying our causal channel of maternal blood lead with analyses of the lead-independent outcome of infant sex. Taken together, these results suggest considerable social and economic benefits from the phase-out of leaded gasoline in the United States through improved infant health. Calculating the economic benefits of the phase-out of leaded gasoline through the reduction of healthcare-related costs involved in treating low birthweight infants, alone, are in the tens of billions annually.

The rest of our manuscript is organized as follows. In Section 3.2 we discuss relevant scientific literature on the maternal transmission pathway and studies linking to deleterious birth outcomes to lead exposure. In Section 3.3, we detail elements of research design, including data collection efforts, variable operations and econometric strategies to identify the relationship between maternal lead exposure and birth outcomes. In Section 3.4 we present a series of results on the fetal effects of maternal lead exposure, including a series of robustness and falsification analyses. In Section 3.5, we conclude with a recapitulation of key findings, a narrow benefit analysis from the phase-out of leaded gasoline through the maternal blood lead exposure pathway, and discussion of the present challenges of limiting lead exposure in the United States and abroad.

3.2 Literature review

Lead typically enters the body through inhalation of lead dust or consumption of contaminated foods and liquids. Though the majority of lead is either exhaled or passed through the body in solid waste or urine, that which remains enters the circulatory system (Papanikolaou et al., 2005; Clay et al., 2018). Once lead enters the bloodstream it is transmitted throughout the body, accumulating

in soft tissue, bone, and blood (Needleman, 2004; Papanikolaou et al., 2005). The half life of lead in blood is approximately 35 days and four to six weeks for lead in soft tissue (Papanikolaou et al., 2005), though lead deposits in the brain has a half life of approximately two years (Lidsky and Schneider, 2003). Lead deposited in bone has an estimated half life of 20 to 30 years (Papanikolaou et al., 2005), and can re-enter the bloodstream during periods of high physiological stress such as pregnancy, lactation, and postmenopausal osteoporosis (Bellinger et al., 1987; Rothenberg et al., 1994; Needleman, 2004; Papanikolaou et al., 2005; Bellinger, 2011). Using isotopic measurements of blood lead during pregnancy, cross county analysis have found bone lead is re-mobilized at a higher rate during pregnancy, with skeletal lead attributable to approximately 30% of maternal blood lead levels during this period (Gulson et al., 1997a; Hertz-Picciotto et al., 2000; Harville et al., 2005).

Transmission of maternal blood lead to the fetus occurs via diffusion across the placental barrier over the course of the pregnancy and beginning at week twelve (Gershanik et al., 1974; Rothenberg et al., 1994; Harville et al., 2005). It is generally accepted that lead passes freely through the placental wall. Studies find that measured blood lead levels in a fetus are on average 70% of the mother's blood lead level (Gershanik et al., 1974; Cavalleri et al., 1978; Papanikolaou et al., 2005). Due to the consistent and significant correlation of this relationship, measuring umbilical cord lead levels relative to maternal blood lead is commonly used as a less invasive indicator of in-utero lead exposure in clinical studies (Rothenberg et al., 1994; Gulson et al., 1997a,b; Harville et al., 2005). Identification of the underlying biological and chemical determinants of this transmission channel has proven challenging, as many of the factors influencing maternal lead exposure, both prior to and during pregnancy, are exogenous (Bellinger et al., 1987; Bellinger, 2017; Schell et al., 2003). For example, studies which have focused on maternal behavior and dietary intake during pregnancy have returned mixed results (Gulson et al., 1999, 2000; Harville et al., 2005; Rothenberg et al., 1994). While the biochemistry remains open to debate, there is a consensus in the literature showing a highly positive correlation between maternal and infant blood lead.

Interdisciplinary studies have examined how lead exposure affects infant development during pregnancy as well as birth outcomes. Using the contamination of lead in drinking water of Flint

MI and surrounding areas, Grossman and Slutsky (2017) find a significant reduction in fertility rates and newborn health relative to similar Michigan cities following the switch to the Flint River and elevated lead levels in the water supply. Clay et al. (2018) find similar results using lead content from soil samples to show reduced fertility rates at the national scale. Other authors have identified the impact of vehicle emissions and environmental toxins on birth outcomes. Using the replacement of toll plazas with EZ-Pass systems, Currie and Walker (2011) found that improved air quality in neighboring areas reduced the incidence of low birthweight children by 8-11% and premature births by 7-9%.

Evidence supporting the consequences of lead exposure and environmental pollutants on birth outcomes continues to mount, though accounting for the costs is a challenge. Economists have shown that low birthweight babies fare worse in school performance and educational attainment, labor market outcomes, and worse health in adulthood (Currie and Hyson, 1999; Currie, 2009; Behrman and Rosenzweig, 2004; Almond et al., 2005; Almond and Currie, 2011; Black et al., 2007). In Almond and Currie (2011), the authors argue that the dynamic relationship between environmental conditions during pregnancy, exposing the unborn child to pollutants, resulting in deleterious epigenetic dynamics realized years later. In the next section we describe our research design to identify the relationship between maternal lead exposure and birth outcomes.

3.3 Methods

3.3.1 Data and measurement

Leaded gasoline data

Following a methodology advanced by Reyes (2007), we construct time series of the average concentration of Pb per gallon of gasoline sold (TEL) and total Pb emissions (PbE) at the state scale and over the period of interest. The data sources used to construct and series are detailed below.

Annual Sales of Leaded Gasoline

Annual gasoline sales for the years 1977-1980 are drawn from the "Yearly Report of Gasoline

Sales, by States," published by Ethyl Corporation. Ethyl was the primary supplier of the TEL additive used in gasoline from introduction in the 1920's until removal in 1990 and published annual marketing reports for the petroleum industry. We calculate total gasoline sales (Q_{jt}) for state j in year t by summing total sales for each grade sold.

During the phase-out period, gasoline was sold to consumers in four grades: premium (p), regular (r), premium-unleaded (pu), and regular-unleaded (ru).[25] From Ethyl reports we exploit information on total sales of each grade as well as shares of each grade in total sales. Defining the set of gasoline grades as $G = \{p, r, pu, ru\}$ and letting total sales for each grade $g \in G$ be q_{jt}^g, total gasoline sales are:

$$Q_{jt} = \sum_{g \in G} q_{jt}^g \tag{3.1}$$

Taking Q_{jt} and q_{jt}, we calculate the share (s_{jt}^g) of grade g in total sales as:

$$s_{jt}^g = \frac{q_{jt}^g}{Q_{jt}} \tag{3.2}$$

By construction, $\sum_{g \in G} s_{jt}^g = 1$ for each state j and year $t \in \{1977, ..., 1980\}$.

TEL Concentrations

Data on TEL concentrations for both regular and premium gasoline grades are from the "Petroleum Products Survey: Motor Gasolines, Summer" published annually by the National Institute for Petroleum Research (NIPER) under the Department of Energy. Samples of gasoline were taken each year from 17 districts across the United States. TEL concentrations were reported in grams per gallon (g/gal.) for both regular ("leaded non-premium") and premium ("leaded premium"). Using s^g from Equation 3.2 as weights, and letting tel^g denote the TEL concentration for grade g, $\overline{TEL}_{jt}$ is calculated as the sum of concentration tel_{jt}^g times the share s_{jt}^g of total sales across all gasoline grades $g \in G$, in a given state:

[25] Premium (p), regular (r), and regular-unleaded (ru) gasoline sales are observed over the entire observation period, whereas premium-unleaded (pu) was introduced in 1982.

$$\overline{TEL}_{jt} = \sum_{g \in G} s^g_{jt} \times tel^g_{jt} \tag{3.3}$$

Since TEL concentrations for both unleaded grades (tel^{pu}, tel^{ru}) are effectively zero, these terms drop from the calculation, simplifying to:

$$\overline{TEL}_{jt} = \ s^p_{jt} \times tel^p_{jt} + \ s^r_{jt} \times tel^r_{jt} \tag{3.4}$$

This calculation is rendered for all states and each year from 1977 to 1980 and constitutes the primary channel of maternal lead exposure throughout the study.

Pb Emissions

Finally, using both annual sales of leaded gasoline and TEL concentration data, we construct an annual time series of Pb emissions (PbE) at the state level over the period of interest. Total PbE in metric tons is calculated as the product of leaded gasoline sales (q) times average gasoline lead concentrations (tel) from NIPER reports. This is done for each state j and year t. Recall, total sales of leaded gasoline (millions of gallons) for leaded gasoline grade g is q^g_{jt}, and lead concentration per gallon of gasoline sold is tel^g_{jt}. Total Pb emissions (PbE) are then:

$$PbE_{jt} = \ q^p_{jt} \times tel^p_{jt} + \ q^r_{jt} \times tel^r_{jt} \tag{3.5}$$

Equation 3.5 produces PbE in metric tons, as total sales are in millions of gallons and tel concentrations are in g/gal. While $\overline{TEL}_{jt}$ is our primary lead exposure operation, we use PbE_{jt} to test the robustness of statistical claims linking maternal Pb exposure to deleterious infant health outcomes.

Natality files

Birth data (1977-1980) were obtained by National Bureau of Economic Research's inventory of natality files from the National Vital Statistics System of the National Center for Health Statistics. Birth data include unique information on the date of birth, spatial information on county of maternal residence and place of birth occurrence, as well as a host of birth outcomes and demographic

information on infants and birth mothers. We focus analysis on singleton births of plausible gestation age (greater than 20 weeks) whose delivery occurred in state with complete reporting of all births in the observation period of interest.

Our response variables include infant $birthweight$ (measured in grams), $low\ birthweight$ status $= 1$ if a child's birthweight is $< 2,500$g, $very\ low\ birthweight$ status $= 1$ if a child's birthweight is $< 1,500$g, $gestation\ length$ (measured in weeks), premature status $= 1$ if a child is born at < 37 weeks of gestation, and $low\ APGAR = 1$ if a child's APGAR score at 1 minute of life is < 7. Control variables include various maternal and infant characteristics associated with deleterious birth outcomes. Variables include maternal $education$, age, $race$, $interval\ since\ last\ birth$ and $child\ sex$.[26] All models also include fixed effects for county of maternal residence, year and month of birth.

3.3.2 Empirical model

Using natality data, we estimate the causal effects of maternal lead exposure on birth outcomes (B) for child i, born at time t in county j with the following model:

$$B_{ijt} = \alpha + \beta_1 BLL_{it} + \boldsymbol{X_{ijt}}'\boldsymbol{\beta} + \epsilon_{ijt} \tag{3.6}$$

where, $\boldsymbol{X}$ is vector of control variables, measuring maternal and infant characteristics (including maternal $education$, age, $race$, $interval\ since\ last\ birth$ and $child\ sex$) that affect birth outcomes of child i. The variable of theoretical interest is BLL_{it}, reflecting maternal blood lead concentration (μg/dL) of mother i in time t.

Because maternal blood lead information is not available in natality data, we estimate the blood lead concentration of mothers using a split sample, two-stage least squares approach that leverages National Health and Nutrition Exam Survey (NHANESII) data from the National Center for Health

[26]Maternal age is a categorical variable (1 = Under 20 years of age; 2 = 20-24 years; 3 = 25-29 years; 4 = 30-34 years; 5 = 35-39 years; and 6 = 40-49 years); maternal race is measured dichotomously (0 = white; 1 = non-white); maternal education is a categorical variable (1 = 0-8 years; 2 = 9-11 years; 3 = 12 years; 4 = 13-15 years; 5 = 16 year and over; 6 = Not stated); maternal interval since last live birth is measured categorically (2 = 1-11 months; 3 = 12-23 months; 4 = 24-35 months; 5 = 36-47 months; 6 = 48-71 months; 7 = over 72 months; 8 = Not stated); and infant sex is a binary variable (female $= 0$; male $= 1$).

Statistics. Specifically, we make use of NHANESII, a nationally representative sample of the US population that unfolded between 1976 and 1980, a period that is coincident with observed natality birth data. Of the 27,801 persons sampled in NHANESII, a sizable fraction underwent various tests of anthropometry, hematology, serology, and biochemistry. A total of 10,049 participants had their blood evaluated for Pb, measured in μg/dL units. Residential location of NHANESII participants (at the state scale) allows us to statistically link blood lead outcomes to exposure risk in females of reproductive age. Weighted least squares coefficient estimates from the analysis of NHANESII data inform blood lead estimates in our natality population (Solon et al., 2013).

Importantly, over the four-year period of interest (1977 to 1980), much of the historic phase out of leaded gasoline occurred. As shown in Figure 3.1, Panel A, all states experienced reductions in gasoline-related Pb emissions of at least 40% over this period. The decrease in Pb emissions ranged from 65% in Michigan to 40% in Montana. Reductions in Pb emissions were driven largely by reductions in the TEL concentration of gasoline sold. As shown in Panel B, the reduction in total emissions is driven largely by the reduction in the concentration of TEL per unit of gasoline sold not the quantity of gasoline sold. In terms of the change in concentration, reductions ranged from 62.0% in Washington to 37.9% in Montana.

Since the rate of lead entering the environment during this period declined precipitously and exogenously across states, our estimates of the exposure channel in the first stage are defensibly representative of the population over the same period. Therefore, the first stage regression equation of blood lead level (BLL) on TEL, using the NHANESII survey is summarized as follows:

$$BLL_{it} = \alpha + \gamma \overline{TEL}_{jt} + \boldsymbol{Y_{it}}'\boldsymbol{\theta} + \epsilon_{it} \tag{3.7}$$

where, $\boldsymbol{Y}$ is a vector of controls for the respondent's age, sex, race, and educational attainment. The variable $\overline{TEL}$ is the measured concentration of lead per gallon gasoline in state i in year t defined by Equation 3.4. Using the coefficient estimates obtained from Equation 3.7 and the same set of regressors, we then predict blood lead levels for each mother in natality files. The predicted blood lead values are described in Equation 3.8.

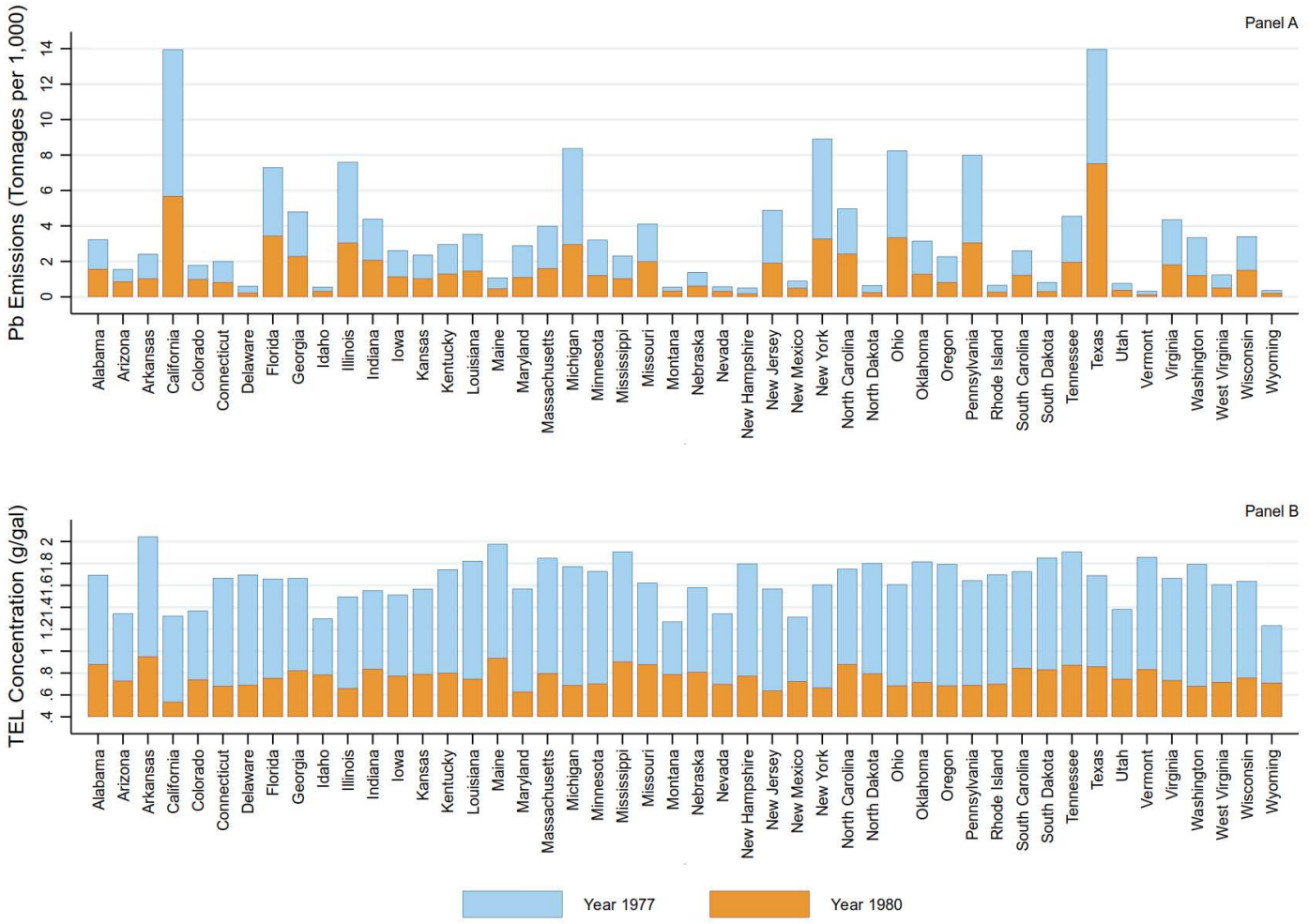


Figure 3.1: Pb emissions (tonnages per 1,000) and average TEL concentration (g/gal) exposure risk by state in time

$$\widehat{BLL}_{it} = \widehat{\alpha} + \widehat{\gamma}\overline{TEL}_{jt} + \boldsymbol{Y_{it}}'\widehat{\boldsymbol{\theta}} \tag{3.8}$$

Substituting estimated blood lead from Equation 3.8 into equation produces the final model to be used in analyses:

$$B_{ijt} = \alpha + \beta_1\widehat{BLL}_{ijt} + \boldsymbol{X}'_{ijt}\boldsymbol{\beta} + \boldsymbol{\Gamma}_1\boldsymbol{Z}_t + \boldsymbol{\Gamma}_2\boldsymbol{M}_t + \boldsymbol{\Gamma}_3\boldsymbol{C}_j + \epsilon_{ijt} \tag{3.9}$$

where, again, $\boldsymbol{X}$ is vector of control variables, measuring maternal and infant characteristics (including include maternal $education$, age, $race$, $interval\ since\ last\ birth$ and $child\ sex$) that affect birth outcomes specific to child i. $\boldsymbol{Z_t}$ is the year of birth, measured as a series of dummy variables with 1977 constituting our reference year, $\boldsymbol{M_t}$ is the month of birth, again measured as a series of dummy variables with January as our reference month, and $\boldsymbol{C_j}$ is the county of maternal residence j at the time of birth. The coefficient of analytic interest is β_1, reflecting the effect of maternal blood lead level BLL_{it}, on various birth outcomes. The expectation is that β_1 is negative for outcomes measured continuously like birthweight (grams) and gestation length (weeks), and positive for deleterious binary outcomes like low birthweight, very low birthweight, prematurity, and low APGAR score.

3.4 Results

3.4.1 Evidence of lead effects

Table 3.1 presents descriptive statistics on birth outcomes by quantiles of estimated maternal blood lead. Birth events are limited to singletons with plausible gestation length (> 20 weeks), occurring in states with full reporting over the observation period of 1977 to 1980. In Column 1 we observe unconditional mean birthweight of infants born to mothers grouped by blood lead level. Mean birthweight decreases incrementally in going from Quantile 1 (Q1) to Quantile 5 (Q5). As compared to infants born to Q5 mothers, Q1 infants weight 151 grams or 4.64% more at birth. As shown in Column 2, the same pattern of increasing risk by maternal blood lead obtains for low birthweight status. About 1 in 12 infants born to Q5 mothers arrive at 2,500 grams or less, whereas

Table 3.1: Descriptive statistics on birth outcomes by quantiles of estimated maternal blood (μg/dl)

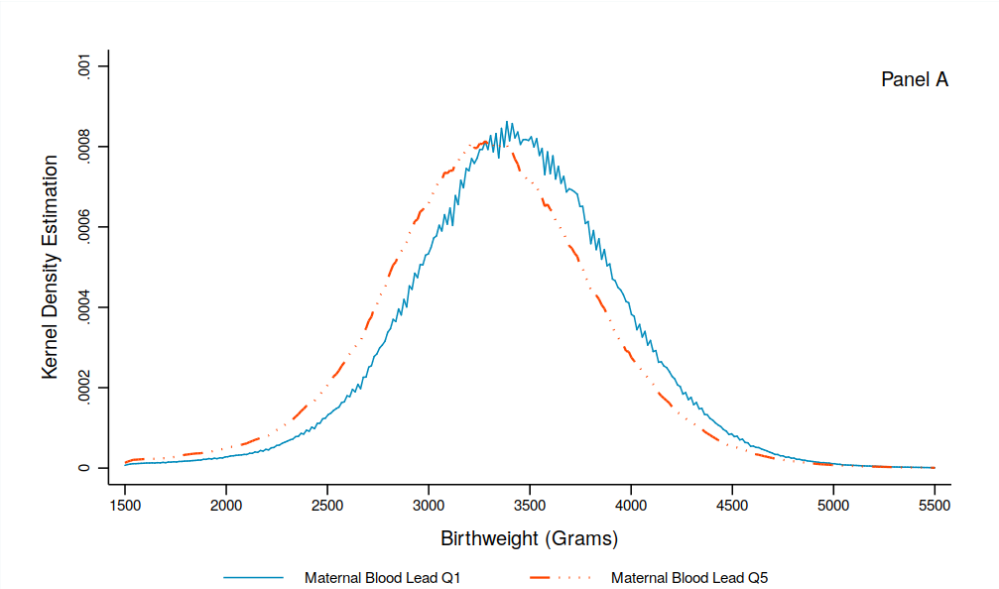
Maternal Blood Lead (μg/dL)	Birthweight (Grams)	Low Birthweight (<2,500g)	V. Low Birthweight (<1,500g)	Gestation Length (Weeks)	Premature (<37 Weeks)	Low APGAR (<7, 1 min)	Low APGAR (<7, 5 min)
Maternal Blood Lead Q1	3407.826	0.0473882	0.0066778	39.73966	0.0709061	0.0840116	0.0156776
	(555.77)	(0.212)	(0.082)	(2.572)	(0.257)	(0.277)	(0.124)
Maternal Blood Lead Q2	3399.087	0.0492124	0.0067449	39.71019	0.0735085	0.0863954	0.016342
	(557.86)	(0.216)	(0.082)	(2.591)	(0.261)	(0.281)	(0.127)
Maternal Blood Lead Q3	3387.731	0.051437	0.0073819	39.67169	0.0752023	0.0856951	0.0172108
	(562.97)	(0.221)	(0.086)	(2.610)	(0.264)	(0.280)	(0.130)
Maternal Blood Lead Q4	3359.678	0.0569863	0.0084298	39.56352	0.0825275	0.0838764	0.0179081
	(568.39)	(0.232)	(0.091)	(2.669)	(0.275)	(0.277)	(0.133)
Maternal Blood Lead Q5	3256.604	0.0823191	0.0132513	39.20303	0.116137	0.1003102	0.0267118
	(595.47)	(0.275)	(0.114)	(2.970)	(0.320)	(0.300)	(0.161)
Total	3,366.086	0.0566103	0.0083296	39.59164	0.0825822	0.0867286	0.0177008
	(569.80)	(0.231)	(0.091)	(2.684)	(0.275)	(0.281)	(0.132)
	6,332,747	6,332,747	6,332,747	6,332,747	6,332,747	3,858,278	3,870,901

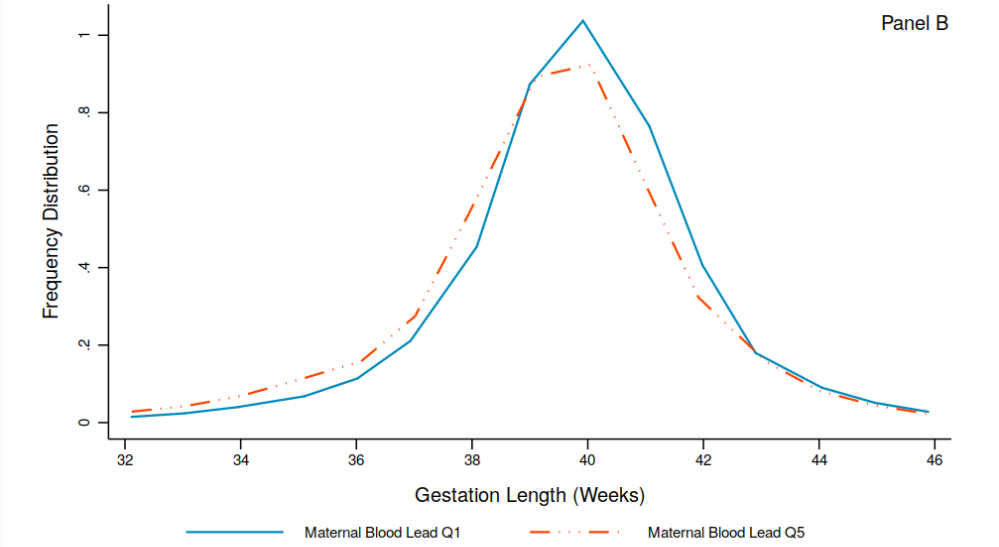
Notes: Birth events are limited to singletons with plausible gestation length (>20 weeks), occurring in states with full reporting over the observation period of 1977 to 1980; Mean and standard deviation of blood lead (μg/dL) of mothers in: Q1 (9.63, 0.50), Q2 (10.72, 0.24), Q3 (11.46, 0.20), Q4 (12.21, 0.24), and Q5 (13.70, 0.92).

about 1 in 20 children born to Q1 mothers are at risk of low birth weight. With respect to gestation length in Column 4, on average, Q1 infants age 3.8 days longer in womb than Q5 infants (39.74 weeks vs 39.20 weeks). With the exception of a low APGAR score (< 7) at 1 minute of life, all birth outcomes worsen dose-responsively in the estimated blood lead of mothers.

Figure 3.2, Panel A presents a kernel density plot of infant birthweight (grams) and Panel B displays a histogram of infant gestation length (weeks) by Q5 versus Q1 blood lead mothers. The negative effect of maternal blood lead appears to operate more or less uniformly across the unconditional distributions of infant birthweight and gestation length, shifting each distribution toward the origin.

Table 3.2 reports the birthweight effects of maternal blood lead, involving various combinations of time and county fixed effects, a suite of control variables, and transformations of both response and treatment variables. Interpretive emphasis is on fully saturated models in Columns 3, 6 and 9. In Column 3 we find that, other things held equal, a 1 microgram per deciliter increase in maternal blood lead reduces infant birthweight by 19.83 grams (95% CI: -24.30 to -15.36). In Column 6 we find that the birthweight elasticity of maternal blood lead is 0.108, with a 1 percent increase in maternal blood lead decreasing infant birthweight by a tenth of one percent (95% CI: -0.124 to -0.093). In Column 9, maternal blood lead is divided into five quantiles. As compared to infants born to reference group mothers in Q1, newborns to Q5 mothers are 46.78 grams (95% CI: -60.89 to -32.67) lighter. The mean blood lead concentration differential of mothers in Q5 versus Q1 is 4.1 micrograms per deciliter. Estimated coefficients of the negative birthweight effect of maternal blood lead increase incrementally in quantiles, going from -11.49 (Q2) to -23.29 (Q3) to -34.24 (Q4) and to -46.78 (Q5), indicative of a linear dose-response relationship between infant birthweight and maternal blood lead level. Appendix Table C.4 shows birthweight effects involving the substitution of our treatment variable of maternal blood lead for $\overline{TEL}$ concentration (g/gal), Appendix Table C.5 substitutes our exposure variable as PbE, and Appendix Table C.6 shows birthweight effects where analyses are extended to livebirth events in all states, regardless of the completeness of reporting. Results behave similarly across treatment variable substitutions and sample extensions.





Notes: Birth events are limited to singletons with plausible gestation length ($>$ 20 weeks), occurring in states with full reporting over the observation period of 1977 to 1980.

Figure 3.2: Kernel density plot of birthweight (grams) and histogram of gestation length (weeks) by q5 versus q1 maternal blood lead

Table 3.2: Birthweight (grams) effect of maternal blood lead

	(1) BW	(2) BW	(3) BW	(4) BW	(5) ln BW	(6) ln BW	(7) BW	(8) BW	(9) BW
Maternal Blood Lead (μg/dL)	-37.248***	-80.266***	-19.833***						
	(2.794)	(4.588)	(2.280)						
ln Maternal Blood Lead (μg/dL)				-278.953***	-0.334***	-0.108***			
				(24.306)	(0.016)	(0.008)			
Reference (Maternal Blood Lead Q1)									
Maternal Blood Lead Q2							-8.739***	-68.514***	-11.487***
							(1.942)	(3.314)	(2.614)
Maternal Blood Lead Q3							-20.095***	-117.161***	-23.288***
							(2.922)	(5.285)	(4.281)
Maternal Blood Lead Q4							-48.148***	-166.598***	-34.139***
							(5.908)	(7.027)	(6.437)
Maternal Blood Lead Q5							-151.222***	-273.149***	-46.781***
							(9.712)	(13.262)	(7.194)
Constant	3,791.346***	4,402.041***	3,392.100***	3,850.772***	8.959***	8.297***	3,407.826***	3,572.108***	3,180.017***
	(28.179)	(60.323)	(34.407)	(63.589)	(0.042)	(0.021)	(4.043)	(10.243)	(16.492)
Observations	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047
R^2	0.009	0.030	0.069	0.069	0.027	0.058	0.009	0.026	0.069
County FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Month FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	Yes	No	Yes	No	No	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); Control variables include: maternal age ($1 =$ Under 20 years of age; $2 = 20$-24 years; $3 = 25$-29 years; $4 = 30$-34 years; $5 = 35$-39 years; and $6 = 40$-49 years); maternal race ($0 =$ white; $1 =$ non-white); maternal education ($1 = 0$-8 years; $2 = 9$-11 years; $3 = 12$ years; $4 = 13$-15 years; $5 = 16$ year and over; $6 =$ Not stated); interval since last live birth ($2 =$1-11 months; $3 =$12-23 months; $4 =$24-35 months; $5 =$36-47 months; $6 =$48-71 months; $7 =$over 72 months; $8 =$Not stated); and infant sex (female $= 0$, male $= 1$).

Table 3.3 reports coefficients of the gestation length effects of maternal blood lead concentration. As with Table 3.2, interpretive emphasis is on fully saturated models in Columns 3, 6 and 9. In Column 3, and other things held equal, we observe that a 1 μg/dL increase in maternal blood lead decreases gestation length by 0.087 weeks (95% CI: -0.101 to -0.072) or six-tenths of a day. Column 6 shows that the elasticity of gestation length vis-á-vis maternal blood lead is about one-third the size of the birthweight elasticity (in Table 3.2), with a one percent increase in maternal blood lead inducing a shortening of gestation length of -0.035 percent (95% CI: -0.040 to -0.029). In Column 9, we find that the gestation length effect of maternal blood lead decreases progressively from one maternal blood lead quantile to the next, going from -0.058 (Q2) to -0.106 (Q3) to -0.146 (Q4) to -0.208 (Q5). On average, the gestational age of infants born to Q5 mothers is 1.5 days shorter than children born to Q1 mothers.

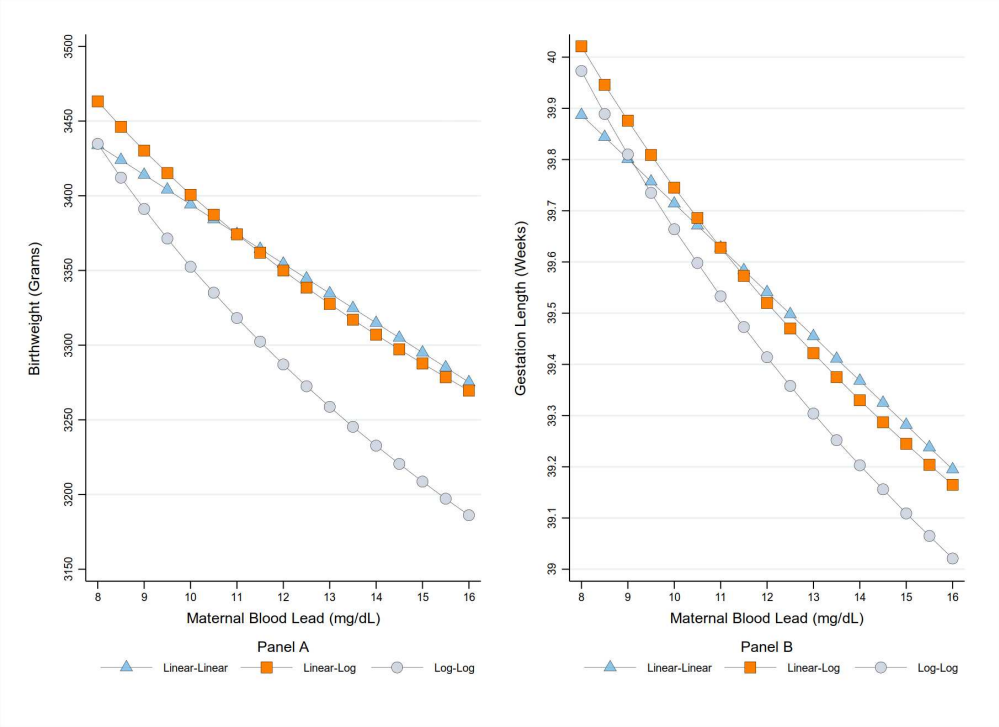
Taken together, Tables 3.2 and 3.3 allow one to roughly adjudicate between lead exposure mechanisms of birthweight reduction as arising from either intrauterine growth retardation or gestation length shortening. First, birthweight increases by about 90 grams per week of gestation. Second, as shown in Column 3, Table 3.3, gestation length decreases by six-tenths of day for a 1 μg/dL increase in maternal blood lead, implying -7.7 grams of expected birthweight from gestational shortening (given expected gains in birthweight per week of gestation). Given that the estimated birthweight effect of a 1 μg/dL increase in maternal blood lead of -19.8 grams, about 39% of estimated birthweight loss is from gestation shortening, and 61% of loss is from growth retardation.

Figure 3.3 summarizes results in Tables 3.2 and 3.3 graphically, depicting predicted birthweight (grams) and gestation length (weeks) across a range of maternal blood lead concentrations (8 to 16 μg/dL) and three different operations of response and treatment variables. Other covariates fixed at their means, the linear-linear (blue triangles), linear-log (orange squares), and log-log (gray circles) models of birthweight indicate that predicted birthweight decreases from 3,434 to 3,275 grams, 3,463 to 3,279 grams, and 3,434 to 3,186 grams, respectively, in going from 1^{st} (8 μg/dL) to 99^{th} percentile (16 μg/dL) in maternal blood lead concentration. While the log-log model results imply a steeper dose-response, the 95% confidence intervals of the three specifications depicted overlap

Table 3.3: Gestation length (weeks) effect of maternal blood lead

	(1) GL	(2) GL	(3) GL	(4) GL	(5) ln GL	(6) ln GL	(7) GL	(8) GL	(9) GL
Maternal Blood Lead (μg/dL)	-0.132***	-0.314***	-0.087***						
	(0.010)	(0.013)	(0.007)						
ln Maternal Blood Lead (μg/dL)				-1.234***	-0.106***	-0.035***			
				(0.101)	(0.003)	(0.003)			
Reference (Maternal Blood Lead Q1)									
Maternal Blood Lead Q2							-0.029***	-0.310***	-0.058***
							(0.006)	(0.009)	(0.010)
Maternal Blood Lead Q3							-0.068***	-0.516***	-0.106***
							(0.010)	(0.014)	(0.014)
Maternal Blood Lead Q4							-0.176***	-0.716***	-0.146***
							(0.018)	(0.019)	(0.021)
Maternal Blood Lead Q5							-0.537***	-1.113***	-0.208***
							(0.034)	(0.028)	(0.026)
Constant	41.095***	43.732***	40.614***	42.658***	3.949***	3.760***	39.740***	40.536***	39.689***
	(0.096)	(0.172)	(0.197)	(0.301)	(0.009)	(0.008)	(0.014)	(0.024)	(0.181)
Observations	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047
R^2	0.005	0.019	0.027	0.027	0.019	0.028	0.005	0.016	0.027
County FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Month FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	Yes	No	Yes	No	No	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); Control variables include: maternal age ($1 =$ Under 20 years of age; $2 =$ 20-24 years; $3 =$ 25-29 years; $4 =$ 30-34 years; $5 =$ 35-39 years; and $6 =$ 40-49 years); maternal race ($0 =$ white; $1 =$ non-white); maternal education ($1 =$ 0-8 years; $2 =$ 9-11 years; $3 =$ 12 years; $4 =$ 13-15 years; $5 =$ 16 year and over; $6 =$ Not stated); interval since last live birth ($2 =$1-11 months; $3 =$12-23 months; $4 =$24-35 months; $5 =$36-47 months; $6 =$48-71 months; $7 =$over 72 months; $8 =$Not stated); and infant sex (female $= 0$, male $= 1$).



Notes: Models results depicted are fully saturated with control variables, year, month, and county fixed effects; All covariates are fixed at sample means; The linear-linear models are from Column 3, linear-log models are from Column 4, and log-log models are from Column 6 in Tables 3.2 (for Birthweight) and 3.3 (Gestation Length);

Figure 3.3: Predicted birthweight (grams) and gestation length (weeks) by maternal blood lead

each other. As estimated mean (and median) blood levels of mothers decreased from about 12.5 to 10 μg/dL over the four-year observation period, depicted models imply that infant birthweights were pushed upward by 40 to 80 grams by the phase-out of leaded gasoline.

Health economists have produced a considerable body of research estimating the birth weight effects of social policies, including the Food Stamp Program (Almond and Currie, 2011) and the Women, Infants, and Children (WIC) (Hoynes et al., 2011). Almond and Currie (2011), for instance, find that the Food Stamp Program increased infant birth weight by 15-20 grams for whites and 13-42 grams for African Americans. Hoynes et al. (2011), report average birthweight increased approximately 29 grams, a 10 percent increase, among WIC participants. Our findings are comparable with the effects of the Food Stamp Program and WIC on birthweight when one con-

siders the CAA reduced lead emissions nationwide, benefiting all mothers rather than the subset of mothers with the most to gain.

Table 3.4 reports coefficients from linear probability models estimating the effect of maternal blood lead on risk of low birthweight ($< 2,500$ grams). We focus interpretation on fully saturated models in Columns 3 and 7. Beginning with Column 3, we find that a 1 μg/dL increase in maternal blood lead increases the probability of a mother giving birth to a low birthweight infant increases by nine-tenths of one percentage point (95% CI: 0.008, 0.011), or about 18%.[27] Results from a logistic regression model (not shown) supports this result, indicating that a unit increase in maternal blood lead increases the risk of a low birthweight outcome by a multiplicative factor of 1.065 (95% CI: 1.039, 1.092). In Column 7, maternal blood lead is divided into quantiles. The probability of a low birthweight outcome increases progressively in maternal blood lead category. Other things held equal, the probability of a low birthweight event for an infant born to a Q5 mother is 1.6 percentage points (or 32%) higher than infants born to Q1 mothers (0.050 versus 0.066).

In Table 3.5 we summarize linear probability models of very low birthweight risk ($< 1,500$ grams). As with other birth outcomes, we find that the risk of a very low birthweight increases substantially in maternal blood lead. Again, we focus interpretation on models reported in Columns 3 and 7 containing the full set of controls. In Column 3, we find that a 1 μg/dL increase in maternal blood lead increases the probability of a mother giving birth to a very low birthweight infant increases by two-tenths of one percentage point (95% CI: 0.002, 0.003). In Column 7, maternal blood lead is divided into quantiles. The probability of a very low birthweight outcome increases incrementally in maternal blood lead category. Other things held equal, the probability of a very low birthweight event for an infant born to a Q5 mother is about 45.2% higher than infants born to Q1 infants (0.010 vs 0.007).

Table 3.6 reports coefficients of the prematurity effects of maternal blood lead concentration. Again, interpretive emphasis is on fully saturated models in Columns 3 and 7. In Column 3, and other things held equal, we observe that a 1 μg/dL increase in maternal blood lead increases the

[27] The low birthweight effect size of maternal lead exposure is approximately equal to the low birthweight risk advantage of male over female infants (-0.010, 95% CI: -0.011, -0.009).

Table 3.4: Low birthweight ($< 2,500$ grams) effect of maternal blood lead

	(1) LBW	(2) LBW	(3) LBW	(4) LBW	(5) LBW	(6) LBW	(7) LBW
Maternal Blood Lead (μg/dL)	0.009***	0.020***	0.009***				
	(0.001)	(0.001)	(0.001)				
ln Maternal Blood Lead (μg/dL)				0.243***	0.118***		
				(0.011)	(0.012)		
Reference (Maternal Blood Lead Q1)							
Maternal Blood Lead Q2						0.017***	0.004***
						(0.001)	(0.001)
Maternal Blood Lead Q3						0.028***	0.007***
						(0.001)	(0.001)
Maternal Blood Lead Q4						0.040***	0.010***
						(0.002)	(0.001)
Maternal Blood Lead Q5						0.067***	0.016***
						(0.004)	(0.002)
Constant	-0.042***	-0.201***	-0.008	-0.565***	-0.190***	0.005**	0.098***
	(0.008)	(0.012)	(0.013)	(0.029)	(0.030)	(0.003)	(0.007)
Observations	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047
R^2	0.003	0.010	0.018	0.010	0.018	0.009	0.018
County FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Yes	No	Yes	No	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); Control variables include: maternal age ($1 =$ Under 20 years of age; $2 =$ 20-24 years; $3 =$ 25-29 years; $4 =$ 30-34 years; $5 =$ 35-39 years; and $6 =$ 40-49 years); maternal race ($0 =$ white; $1 =$ non-white); maternal education ($1 =$ 0-8 years; $2 =$ 9-11 years; $3 =$ 12 years; $4 =$ 13-15 years; $5 =$ 16 year and over; $6 =$ Not stated); interval since last live birth ($2 =$1-11 months; $3 =$12-23 months; $4 =$24-35 months; $5 =$36-47 months; $6 =$48-71 months; $7 =$over 72 months; $8 =$Not stated); and infant sex (female $= 0$, male $= 1$).

Table 3.5: Very low birthweight ($< 1,500$ grams) effect of maternal blood lead

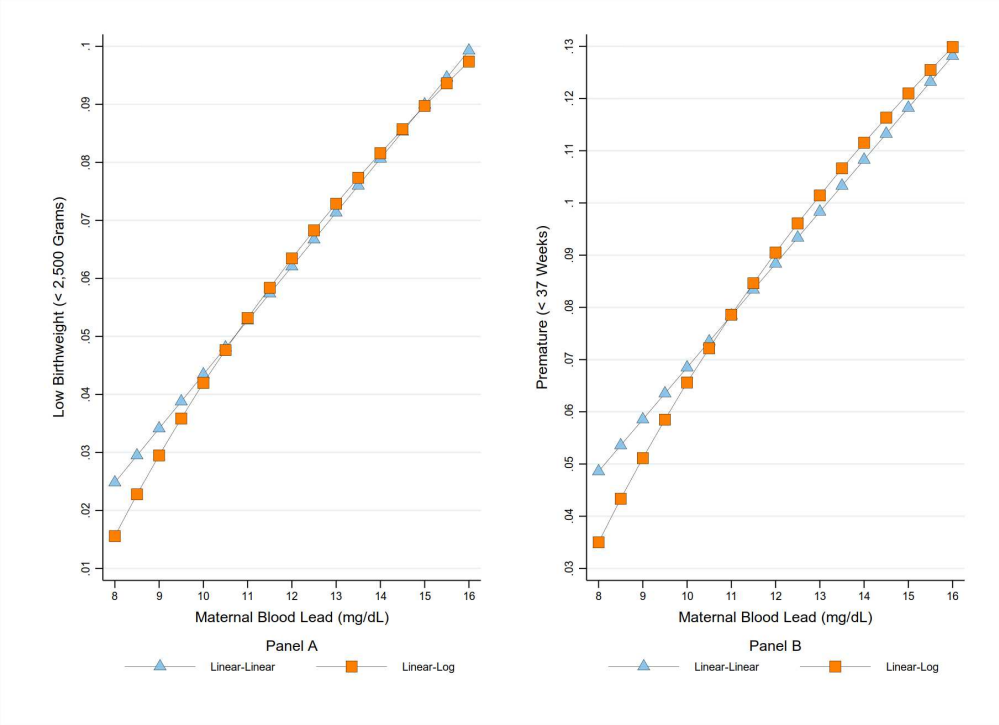
	(1) VLBW	(2) VLBW	(3) VLBW	(4) VLBW	(5) VLBW	(6) VLBW	(7) VLBW
Maternal Blood Lead (μg/dL)	0.002***	0.004***	0.002***				
	(0.000)	(0.000)	(0.000)				
ln Maternal Blood Lead (μg/dL)				0.046***	0.027***		
				(0.002)	(0.003)		
Reference (Maternal Blood Lead Q1)							
Maternal Blood Lead Q2						0.003***	0.001***
						(0.000)	(0.000)
Maternal Blood Lead Q3						0.005***	0.001***
						(0.000)	(0.000)
Maternal Blood Lead Q4						0.008***	0.002***
						(0.000)	(0.000)
Maternal Blood Lead Q5						0.013***	0.003***
						(0.001)	(0.001)
Constant	-0.011***	-0.042***	0.004	-0.111***	-0.038***	-0.002***	0.028***
	(0.002)	(0.002)	(0.006)	(0.006)	(0.010)	(0.001)	(0.005)
Observations	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047
R^2	0.001	0.003	0.006	0.003	0.006	0.003	0.006
County FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Yes	No	Yes	No	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); control variables include maternal age ($1 =$ Under 20 years of age; $2 =$ 20-24 years; $3 =$ 25-29 years; $4 =$ 30-34 years; $5 =$ 35-39 years; and $6 =$ 40-49 years); maternal race ($0 =$ white; $1 =$ non-white); maternal education ($1 =$ 0-8 years; $2 =$ 9-11 years; $3 =$ 12 years; $4 =$ 13-15 years; $5 =$ 16 year and over; $6 =$ Not stated); interval since last live birth ($2 =$1-11 months; $3 =$12-23 months; $4 =$24-35 months; $5 =$36-47 months; $6 =$48-71 months; $7 =$over 72 months; $8 =$Not stated); and infant sex (female $= 0$, male $= 1$)

Table 3.6: Prematurity (< 37 weeks) effect of maternal blood lead

	(1) Prem	(2) Prem	(3) Prem	(4) Prem	(5) Prem	(6) Prem	(7) Prem
Maternal Blood Lead (μg/dL)	0.011***	0.028***	0.010***				
	(0.001)	(0.001)	(0.001)				
ln Maternal Blood Lead (μg/dL)				0.340***	0.137***		
				(0.011)	(0.014)		
Reference (Maternal Blood Lead Q1)							
Maternal Blood Lead Q2						0.026***	0.006***
						(0.001)	(0.001)
Maternal Blood Lead Q3						0.042***	0.010***
						(0.001)	(0.001)
Maternal Blood Lead Q4						0.059***	0.014***
						(0.002)	(0.002)
Maternal Blood Lead Q5						0.096***	0.021***
						(0.003)	(0.003)
Constant	-0.042***	-0.275***	-0.003	-0.787***	-0.225***	0.009***	0.106***
	(0.008)	(0.014)	(0.014)	(0.029)	(0.037)	(0.003)	(0.009)
Observations	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047
R^2	0.003	0.013	0.024	0.013	0.024	0.011	0.024
County FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Yes	No	Yes	No	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); control variables include maternal age ($1 =$ Under 20 years of age; $2 =$ 20-24 years; $3 =$ 25-29 years; $4 =$ 30-34 years; $5 =$ 35-39 years; and $6 =$ 40-49 years); maternal race ($0 =$ white; $1 =$ non-white); maternal education ($1 =$ 0-8 years; $2 =$ 9-11 years; $3 =$ 12 years; $4 =$ 13-15 years; $5 =$ 16 year and over; $6 =$ Not stated); interval since last live birth ($2 =$1-11 months; $3 =$12-23 months; $4 =$24-35 months; $5 =$36-47 months; $6 =$48-71 months; $7 =$over 72 months; $8 =$Not stated); and infant sex (female $= 0$, male $= 1$)



Notes: Models results depicted are fully saturated with control variables, year, month, and county fixed effects; All covariates are fixed at sample means; The linear-linear models are from Column 3, and linear-log models are from Column 4 in Tables 3.4 (for Low Birthweight) and 3.6 (Prematurity);

Figure 3.4: Predicted low birthweight ($< 2,500$ grams) and prematurity (< 37 weeks) by maternal blood lead

probability of a shortened gestation (< 37 weeks) length by 1 percentage point (95% CI: .008, .012). In Column 7, we find that the risk of premature birth increases as maternal blood lead increases from one maternal blood lead quantile to the next, going from 0.006 (Q2) to 0.105 (Q3) to 0.142 (Q4) to 0.21 (Q5). Other things held equal, infants born to Q5 mothers are 28.96% more likely than children born to Q1 mothers to be born premature (0.094 vs 0.073).

Next, in Figure 3.4 we graphically summarize results from Table 3.4 and Table 3.6, showing predicted probabilities of low birthweight ($< 2,500$ Grams, Panel A) and prematurity (< 37 Weeks, Panel B) by maternal blood lead concentration. In both panels, predictions from linear-linear and linear-log models are displayed. As shown, the risk of low birthweight and prematurity increases unmistakably in the maternal blood lead concentration of the birth mother. In going from

8 to 16 μg/dL, under the linear-linear model, the risk of low birthweight quadruples (0.099 vs 0.025) while the risk of prematurity increases from 3.5 to 13.0 events per 100 live births.

Finally, in Table 3.7 we summarize linear probability models of risk of an infant presenting with low APGAR score (< 7) in the first minute of life. The APGAR score is a clinical measure of an infant's breathing effort, heart rate, muscle tone, reflexes and skin color. An APGAR score of 7 or greater is a sign that a newborn is in good health. As with other birth outcomes, we find that the risk of a low APGAR score increases significantly in maternal blood lead. In Column 3 showing results with the full set of controls, we find that a 1 μg/dL increase in maternal blood lead increases the probability of a mother giving birth to low APGAR score infant by one percentage point (95% CI: .008, .012). In Column 7, maternal blood lead is divided into quantiles. The probability of a newborn presenting with a low APGAR score increases incrementally in maternal blood lead category. Other things held equal, the probability of a very low birthweight event for an infant born to a Q5 mother is about 20.2% higher than infants born to Q1 infants (0.097 vs 0.081).

Appendix Table C.8 shows results for linear probability models of risk of a newborn presenting with low APGAR score (< 7) at the five-minute mark of life. Results behave similarly, with the risk of an abnormal APGAR score at the five-minute mark increasing dose-responsively in maternal blood lead concentration.

One limitation of the above analysis, regarding the temporal feature of the data, is the potential for the blood lead effect to be overstated due to confounding factors known to negatively affect birth outcomes. For example, coincidental changes in atmospheric pollutants known to reduce birthweight and occurring during the observation period would bias upward the estimated maternal blood lead effect on this birth outcome. To address this potential source of criticism we examine data from the NHANESII sample measuring blood Carboxyhemoglobin (COHb) concentrations, a bio-marker for carbon monoxide (CO) inhalation. Studies have found ambient CO concentrations to negatively impact birth outcomes, including reductions in birthweight with the prevalence of low-birthweight occurrences increasing in atmospheric CO concentrations (Ritz and Yu, 1999; Maisonet et al., 2001; Salam et al., 2005). Data from respondents in the sample show that blood COHb concentrations increase year-over-year during our 1976-1980 study period. In-

Table 3.7: APGAR score (< 7) 1-minute effect of maternal blood lead

	(1) AGAR	(2) AGAR	(3) AGAR	(4) AGAR	(5) AGAR	(6) AGAR	(7) AGAR
Maternal Blood Lead (μg/dL)	0.003***	0.013***	0.010***				
	(0.001)	(0.001)	(0.001)				
ln Maternal Blood Lead (μg/dL)				0.154***	0.119***		
				(0.007)	(0.013)		
Reference (Maternal Blood Lead Q1)							
Maternal Blood Lead Q2						0.011***	0.004***
						(0.001)	(0.001)
Maternal Blood Lead Q3						0.017***	0.007***
						(0.001)	(0.001)
Maternal Blood Lead Q4						0.024***	0.009***
						(0.002)	(0.002)
Maternal Blood Lead Q5						0.044***	0.016***
						(0.003)	(0.003)
Constant	0.055***	-0.072***	-0.007	-0.297***	-0.183***	0.064***	0.100***
	(0.010)	(0.007)	(0.011)	(0.019)	(0.034)	(0.002)	(0.003)
Observations	3,852,811	3,852,811	3,852,811	3,852,811	3,852,811	3,852,811	3,852,811
R^2	0.000	0.016	0.021	0.016	0.021	0.016	0.021
County FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	No	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	Yes	No	Yes	No	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); control variables include maternal age (1 = Under 20 years of age; 2 = 20-24 years; 3 = 25-29 years; 4 = 30-34 years; 5 = 35-39 years; and 6 = 40-49 years); maternal race (0 = white; 1 = non-white); maternal education (1 = 0-8 years; 2 = 9-11 years; 3 = 12 years; 4 = 13-15 years; 5 = 16 year and over; 6 = Not stated); interval since last live birth (2 =1-11 months; 3 =12-23 months; 4 =24-35 months; 5 =36-47 months; 6 =48-71 months; 7 =over 72 months; 8 =Not stated); and infant sex (female $= 0$, male $= 1$)

sofar as the inhalation of ambient carbon monoxide is associated with gestational outcomes like birth weight and risk of low birth weight, our estimates showing an increase in birth weight as a function of gasoline formulation (that caused a decrease in atmospheric lead) can be interpreted as conservative.

While it is statistically unlikely that the consistency of results reported above are governed by chance alone, in the next section we nonetheless perform a series of falsification and robustness tests. We begin with a falsification exercise involving child sex, an outcome plausibly unrelated to maternal lead exposure. After that, we examine whether possible compositional effects from the economic recession in 1980 meaningfully impact parameter estimates of maternal blood lead. We end with a test that exploits maternal migration during pregnancy, showing that migration decisions are spatially independent of the concentration in TEL (g/gal) sold to consumers.

3.4.2 Robustness and falsification

With respect to falsification of our maternal blood lead channel, we model an outcome variable that is unrelated to maternal lead exposure, namely infant sex. We estimate the effect of maternal lead exposure on sex identity (S) of child i, born at time t in county j with the following linear probability model:

$$S_{ijt} = \alpha + \beta_1 \widehat{BLL}_{ijt} + \boldsymbol{X}'_{ijt}\boldsymbol{\beta} + \boldsymbol{\Gamma}_1 \boldsymbol{Z}_t + \boldsymbol{\Gamma}_2 \boldsymbol{M}_t + \boldsymbol{\Gamma}_3 \boldsymbol{C}_j + \epsilon_{ijt} \tag{3.10}$$

where, all terms carry from Equation 3.9, and our outcome variable S is equal to 1 if the newborn is male and 0 if female. Given that human sex is determined by chromosomal assignment at fertilization, followed by gonadal differentiation and secretion of masculinizing hormones, processes that operate independently of lead exposure, our statistical expectation of male sex determination from maternal blood lead is chance indistinguishable or β_1 indistinguishable from zero.

Table 3.8 and Figure 3.5 show results from this channel falsification exercise. As expected, in Table 3.8 we see that whatever the operationalization of maternal blood lead (original units, log transformed or categorical), the male sex outcome is independent of fetal lead exposure. All maternal blood lead coefficients are effectively zero. Figure 3.5 graphs predicted probabilities

Table 3.8: Falsification test of male sex effect of maternal blood lead

	(1) Male	(2) Male	(3) Male
Maternal Blood Lead (μg/dL)	0.000		
	(0.000)		
ln(Maternal Blood Lead) (μg/dL)		-0.000	
		(0.006)	
Reference (Maternal Blood Lead Q1)			
Maternal Blood Lead Q2			-0.001
			(0.001)
Maternal Blood Lead Q3			-0.001
			(0.001)
Maternal Blood Lead Q4			-0.000
			(0.001)
Maternal Blood Lead Q5			-0.000
			(0.002)
Constant	0.539***	0.541***	0.541***
	(0.022)	(0.027)	(0.022)
Observations	6,325,047	6,325,047	6,325,047
R^2	0.000	0.000	0.000
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Month FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); Control variables include: maternal age ($1 =$ Under 20 years of age; $2 =$ 20-24 years; $3 =$ 25-29 years; $4 =$ 30-34 years; $5 =$ 35-39 years; and $6 =$ 40-49 years); maternal race ($0 =$ white; $1 =$ non-white); maternal education ($1 =$ 0-8 years; $2 =$ 9-11 years; $3 = 12$ years; $4 =$ 13-15 years; $5 = 16$ year and over; $6 =$ Not stated); interval since last live birth ($2 =$1-11 months; $3 =$12-23 months; $4 =$24-35 months; $5 =$36-47 months; $6 =$48-71 months; $7 =$over 72 months; $8 =$Not stated); and infant sex (female $= 0$, male $= 1$).

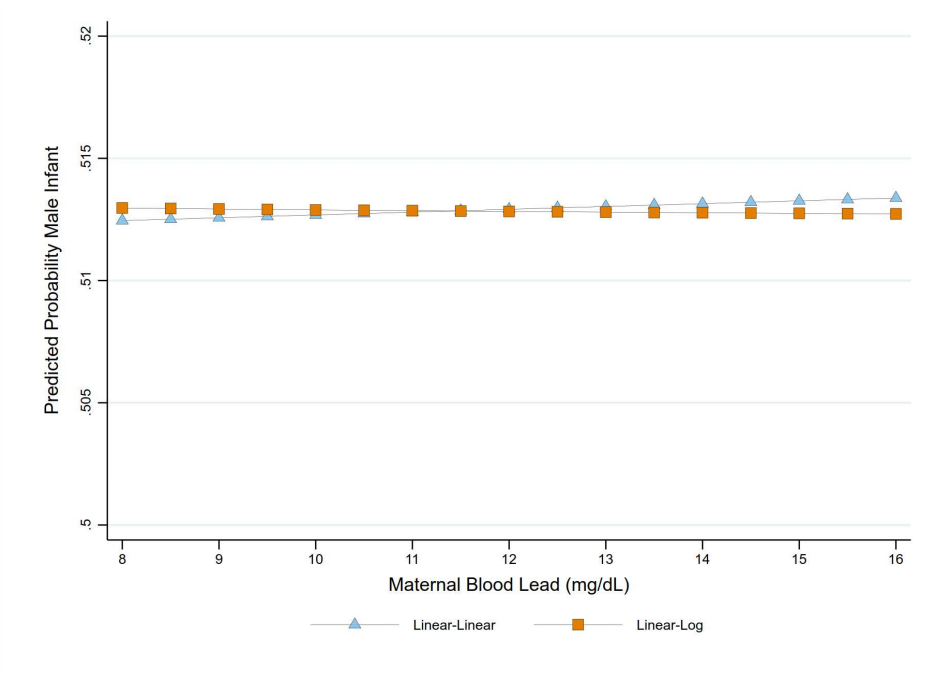


Figure 3.5: Predicted male sex of newborn by maternal blood lead

of male sex from linear (Column 1, Table 3.8) and natural log (Column 2, Table 3.8) models across various levels of maternal blood lead (μg/dL), confirming our expectation[28] of statistical independence.

Next, we examine whether compositional effects from the economic recession in 1980 meaningfully impact parameter estimates of maternal blood lead. Recessions have known income and opportunity cost effects on reproductive choice, functioning on net to increase the average human capital and socioeconomic status of birth mothers. Between 1977-1979 (pre-recession) and 1980 (recession), we observe decreases in the percentage of birth mothers with less than a high school degree, percentage non-white mothers, and the percentage of teen mothers, as well as increases in the percentage of birth mothers with a college degree. While previous models have year and month fixed effects, given compositional change in birth mothers toward lower risk of lead exposure, previous estimates on maternal blood lead may be understated. Appendix Section B details the logic.

[28]While results behave as expected, there is a possible indirect theoretical link between that maternal blood lead and male sex. Supposing that maternal lead exposure increases the risk of fetal mortality, and given that the risk of fetal demise is higher in male versus female unborn, primary sex ratios can depart from standard expectation. This indirect pathway implies a negative association between maternal blood lead and probability of male infant sex. To our knowledge, no study has convincingly linked lead exposure to differential demise by fetal sex.

We estimate the following linear probability model, where R_t denotes the two birth outcomes of infant birth weight and gestational age:

$$B_{ijt} = \alpha + \beta_1 \widehat{BLL}_{ijt} + \beta_2 R_t + \delta\left(\widehat{BLL}_{it} \times R_t\right) \boldsymbol{X}'_{ijt}\boldsymbol{\beta} + \boldsymbol{\Gamma}_1 \boldsymbol{M}_t + \boldsymbol{\Gamma}_2 \boldsymbol{C}_j + \epsilon_{ijt} \quad (3.11)$$

where, again, terms carry from Equation 3.9 with the exception of R_t that is equal to 1 if the birth event occurred in 1980, and zero otherwise. Of theoretical interest is the δ parameter, capturing the interaction effect of maternal blood lead and economic recession of 1980.

Table 3.9 shows results for linear probability models of the infant birthweight and gestational length effects of maternal blood lead interacted with the 1980 economic recession. As before, interpretive emphasis is on fully saturated linear-linear models in columns 3 (birth weight) and 7 (gestation length). With respect to birthweight, we find that the negative effect of maternal blood lead worsened in the 1980 recession, evident in the negative sign on the interaction term ($\hat{\beta}$ = -6.28). The same is true of gestation length, with the negative effect of maternal blood lead amplifying in the recession ($\hat{\beta}$ = -0.019). The combined effects of the maternal blood lead and interaction with recession indicator are approximately equal to main effects reported in Tables 3.2 and 3.3, corresponding to birthweight and gestation length, respectively. Figure 3.6 traces the differential response in birth outcomes to maternal blood lead dosage. Other things held equal, and focusing on linear-linear model results, in Panel A showing predicted birthweight in grams we find that infants born in the economic recession to mothers with 16 μg/dL, for instance, weighted on average about 52.96 grams less than the same children before the recession (3,255.63g vs 3,308.59g). Similarly, in Panel B showing predicted gestation length in weeks we find that infants born in the economic recession to mothers with 16 μg/dL, for instance, realized pregnancy lengths that were 1.35 days shorter than like children before the recession (39.23 vs 39.43 weeks).

All together, these results on the amplification effect of the economic recession are indicative of co-variation of maternal blood lead and unmeasured demographic characteristics of mothers that are correlated with negative birth outcomes.

Table 3.9: Birthweight (grams) and gestation length effects of maternal blood lead by 1980 recession

	(1) BW	(2) BW	(3) BW	(4) BW	(5) GL	(6) GL	(7) GL	(8) GL
Maternal Blood Lead (μg/dL)	-48.598***	-65.279***	-13.570***		-0.184***	-0.248***	-0.042***	
	(3.344)	(3.504)	(1.479)		(0.012)	(0.010)	(0.006)	
Recession (1980)	76.417***	148.203***	47.480***	68.618***	0.274***	0.543***	0.113**	0.190*
	(12.237)	(10.862)	(8.282)	(19.300)	(0.048)	(0.053)	(0.047)	(0.109)
Maternal Blood Lead $\times$ Recession	-15.168***	-25.604***	-6.277***		-0.061***	-0.100***	-0.019***	
	(1.132)	(0.974)	(0.920)		(0.004)	(0.005)	(0.005)	
ln Maternal Blood Lead (μg/dL)				-177.065***				-0.570***
				(18.106)				(0.076)
ln Maternal Blood Lead $\times$ Recession				-38.591***				-0.125**
				(8.817)				(0.049)
Constant	3,943.391***	4,176.144***	3,314.661***	3,593.206***	41.790***	42.708***	40.056***	40.981***
	(36.969)	(43.642)	(25.665)	(48.956)	(0.127)	(0.128)	(0.195)	(0.262)
Observations	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047
R^2	0.012	0.029	0.069	0.069	0.008	0.017	0.027	0.027
County FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Month FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Controls	No	No	Yes	Yes	No	No	Yes	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); Control variables include: maternal age ($1 =$ Under 20 years of age; $2 =$ 20-24 years; $3 =$ 25-29 years; $4 =$ 30-34 years; $5 =$ 35-39 years; and $6 =$ 40-49 years); maternal race ($0 =$ white; $1 =$ non-white); maternal education ($1 =$ 0-8 years; $2 =$ 9-11 years; $3 =$ 12 years; $4 =$ 13-15 years; $5 =$ 16 year and over; $6 =$ Not stated); interval since last live birth ($2 =$1-11 months; $3 =$12-23 months; $4 =$24-35 months; $5 =$36-47 months; $6 =$48-71 months; $7 =$over 72 months; $8 =$Not stated); and infant sex (female $= 0$, male $= 1$).

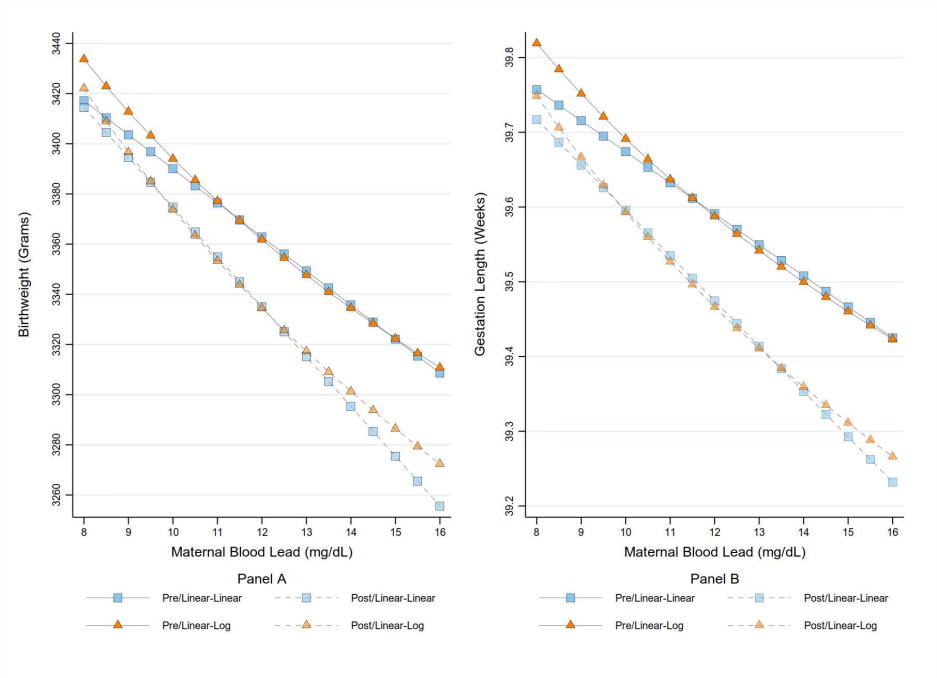
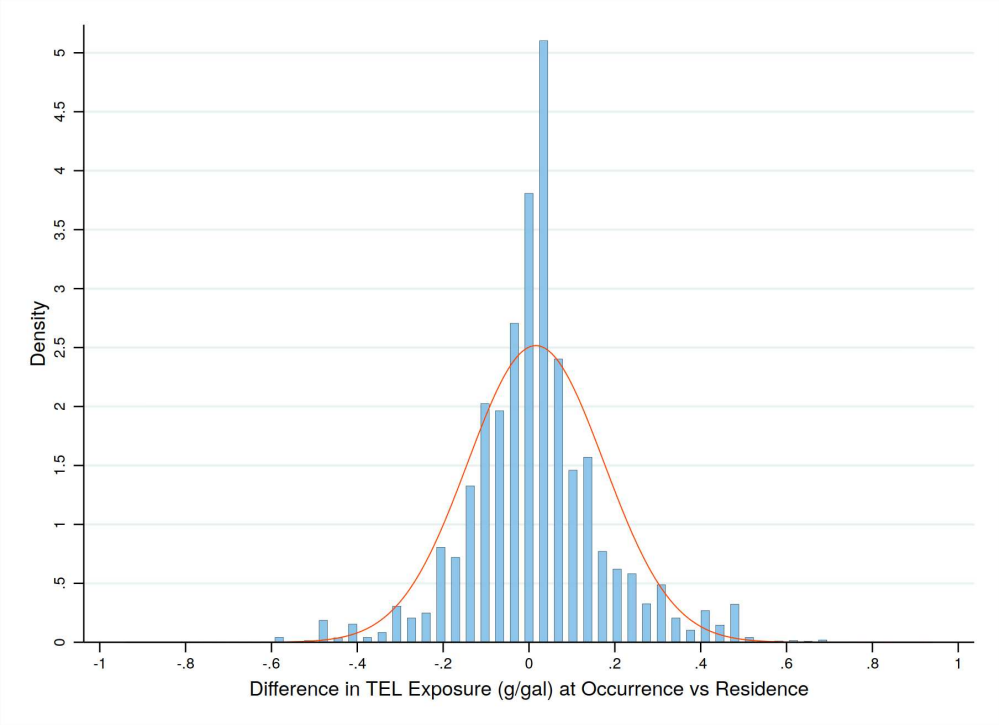


Figure 3.6: Predicted birthweight and gestation length by maternal blood lead and 1980 recession

Of the 6,325,047 birth events satisfying our inclusion criteria, just shy of 184,000 had a reported state of residence different from the state of birth occurrence. Because leaded-gasoline was phased-out differently by states in time, the movement of these mothers is potentially exploitable as a natural experiment in TEL exposure. Toward this end, we present evidence showing that: 1) $\Delta TEL = TEL_{occurrence} - TEL_{residence}$ is a standard normal variable; and 2) mothers moving from higher to a lower TEL environment are statistically similar to mothers moving from a lower to a higher TEL environment.

Taking the TEL (g/gal) difference at the place of birth occurrence and place of residence, we derive whether a mother moved up or down in TEL exposure risk to birth her child and the precise change in TEL exposure. Although the distance moved from place of residence to place of birth occurrence was highly positively skewed, as moving mothers were substantially more likely to travel shorter than longer distances, the distribution of ΔTEL had near zero skew ($S = 0.227$). Figure 3.7 is a histogram of migrating mothers on state of residence (TEL g/gal) versus state of



Notes: All other covariates fixed at means.

Figure 3.7: Histogram of migrating mothers on state of residence (TEL g/gal) versus state of birth occurrence (TEL g/gal)

birth occurrence (TEL g/gal). The distribution of ΔTEL is modestly leptokurtic ($Kurtosis = 5.744$), with $\mu = 0.006$, $\sigma = 0.149$.

Second, on the observed risk factors of deleterious birth outcomes, up- and down-moving mothers are statistically similar with respect to the change in TEL exposure risk. Consider Table 3.10 comparing up-moving and down-moving mothers on the probability of up-moving in TEL exposure risk. Limiting to moving mothers only, we regressed up-moving status ($1 =$ up-moving mother; $0 =$ down-moving mother) on relevant covariates of birth outcomes. If movement-caused TEL exposure risk is a pseudo-randomly assigned, then we ought to see no meaningful differences between up-moving and down-moving mothers on the determinants of birth outcomes with respect to up-moving in TEL exposure risk. All coefficients on covariates of interests in Table 3.10 are chance indistinguishable, supporting the notion that maternal up versus down movement in TEL exposure risk is a quasi-random process. Mothers made migration decisions independent of the concentration in TEL (g/gal) sold to consumers.

Table 3.10: Linear probability model coefficients of maternal up-movement in TEL exposure among moving mothers

	$\Delta\, TEL > 0$		$\Delta\, TEL > 0$
Reference = January		Reference = Maternal Age < 20 years	
February	-0.005	20-24 years	-0.002
	(0.004)		(0.004)
March	-0.004	25-29 Years	-0.003
	(0.004)		(0.008)
April	-0.005	30-34 years	-0.007
	(0.006)		(0.010)
May	0.001	35-39 years	-0.004
	(0.005)		(0.012)
June	-0.001	40-49 years	-0.019
	(0.004)		(0.011)
July	-0.005	Reference = Birth Interval (No Previous)	
	(0.004)		
August	-0.003	1-11 months	-0.019
	(0.004)		(0.012)
September	-0.004	12-23 months	0.003
	(0.004)		(0.005)
October	-0.005	24-35 months	0.006
	(0.004)		(0.005)
November	-0.002	36-47 months	0.003
	(0.004)		(0.006)
December	-0.004	48-71 months	0.001
	(0.004)		(0.006)
Reference = 1977		> 72 months	0.010
			(0.009)
Year 1978	-0.033	Not stated	-0.017
	(0.099)		(0.020)
Year 1979	0.044	Reference = Female	
	(0.081)		
Year 1980	0.014	Male infant	0.001
	(0.081)		(0.002)
Reference = Education (Not Applicable)		Reference = Non-Minority	
0-8 years	-0.210	Minority	0.049
	(0.314)		(0.032)
9-11 years	-0.195		
	(0.313)	Constant	0.742**
12 years	-0.192		(0.304)
	(0.313)		
13-15 years	-0.200	Observations	183,605
	(0.313)	R^2	0.532
16 years	-0.205	County FE	Yes
	(0.314)		
Not Stated	-0.211		
	(0.314)		

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks).

Insofar as observable and unobservable maternal and infant characteristics are correlated (i.e., the proportionality assumption), and assuming that the proportion of time spent in the state of birth occurrence versus the state of residence is statistically independent of whether a mother moves up or down in TEL exposure, then we ought to see a worsening in birth outcomes as ΔTEL increases. The effect of TEL exposure is estimated with the following model:

$$B_{ijt} = \alpha + \beta_1 \Delta TEL_{ijt} + \beta_2 MV_{it} + \boldsymbol{X}'_{ijt}\boldsymbol{\beta} + \boldsymbol{\Gamma}_1 \boldsymbol{Z}_t + \boldsymbol{\Gamma}_2 \boldsymbol{M}_t + \boldsymbol{\Gamma}_3 \boldsymbol{C}_j + \epsilon_{ijt} \qquad (3.12)$$

where, all terms from Equation 3.9 carry, with the exception of ΔTEL_{ijt} measured as $TEL_{occurrence} - TEL_{residence}$, and MV_{it} which an indicator variable equal to 1 if the mother birthed her infant in a state different from her state of residence. The coefficient of analytic interest is β_1 reflecting the effect of the change in TEL exposure on various birth outcomes. The expectation is that β_1 is negative for birthweight (grams) and gestation length (weeks), and positive for dichotomous outcomes like low birthweight, very low birthweight, prematurity, and low APGAR score.

Table 3.11 reports results. Other things equal, in Column 1 we find that a 1 g/gal increase in ΔTEL exposure decreases infant birthweight by just shy of 60 grams (95% CI: -99.21, -20.41). In Column 2, we find that gestation length decreases by 0.278 weeks (or about 2 days) for a unit increase in ΔTEL. The probability of low birthweight ($< 2,500$ grams), very low birthweight ($< 1,500$ grams), and premature birth (< 37 weeks) increase 2.6 percentage points, eight-tenths of one percentage points, and 3.1 percentage points, respectively, as ΔTEL increases 1 g/gal. Also noteworthy is that moving mothers, as compared to non-moving mothers, deliver infants that are lighter, that are more likely to be of low and very low birthweight status, and are more likely to be born prematurely.

Finally, Figure 3.8 summarizes results in Table 3.11, showing predicted birthweight and gestation length in Panel A, and predicted low and very low birthweight status in Panel B for unit changes in ΔTEL. For all predictions, other model covariates are fixed at sample means. Fo-

Table 3.11: Birth outcome effects of maternal TEL exposure through migration

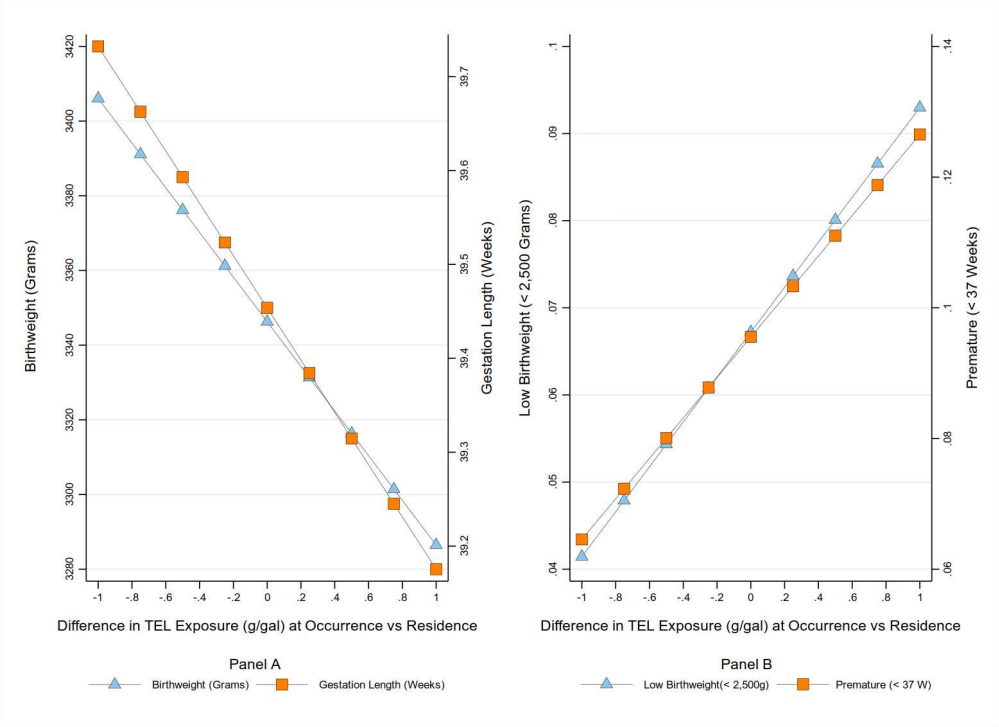
	(1) Birthweight(g)	(2) Gestation(w)	(3) LBW (<2,500g)	(4) VLBW (1,500g)	(5) Premature(<37w)
$TEL\ \Delta$ (g/gal)	-59.810***	-0.278***	0.026***	0.008**	0.031***
	(20.091)	(0.095)	(0.010)	(0.004)	(0.009)
Mover	-20.417***	-0.142***	0.011***	0.006***	0.013***
	(4.698)	(0.022)	(0.002)	(0.001)	(0.002)
Constant	3,144.834***	39.538***	0.108***	0.030***	0.121***
	(16.723)	(0.198)	(0.007)	(0.005)	(0.011)
Observations	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047
R^2	0.069	0.027	0.018	0.006	0.024
County FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); control variables include maternal age ($1 =$ Under 20 years of age; $2 =$ 20-24 years; $3 =$ 25-29 years; $4 =$ 30-34 years; $5 =$ 35-39 years; and $6 =$ 40-49 years); maternal race ($0 =$ white; $1 =$ non-white); maternal education ($1 =$ 0-8 years; $2 =$ 9-11 years; $3 =$ 12 years; $4 =$ 13-15 years; $5 =$ 16 year and over; $6 =$ Not stated); interval since last live birth ($2 =$1-11 months; $3 =$12-23 months; $4 =$24-35 months; $5 =$36-47 months; $6 =$48-71 months; $7 =$over 72 months; $8 =$Not stated); and infant sex (female $= 0$, male $= 1$)

cusing on Panel B, a 1 g/gal increase in ΔTEL from zero increases the risk of low birthweight by 38.8% (0.067 versus 0.093 probability) and the risk of very low birthweight by 57.1% (0.014 versus 0.022 probability).

3.5 Discussion and conclusion

Exploiting natural experiments in the spatial and temporal distribution of maternal exposure arising from the differential phase-out of leaded gasoline across states, and the independent migration behavior of mothers during pregnancy, we find considerable evidence that fetal exposure to lead through the maternal blood lead pathway significantly compromises infant health. Across an ensemble of tests, we find that an increase in maternal blood lead: 1) decreases infant birthweight; 2) increases the risk of low and very low birthweight; 3) shortens gestation length; 4) increases the risk of prematurity; and 5) increases the risk of a low APGAR score. These results are robust to various operations of our treatment variable – whether measured in original units (μg/dL),



Notes: All other covariates fixed at means.

Figure 3.8: Predicted birthweight (grams), gestation length (weeks), low birthweight status ($< 2,500$ grams), and prematurity (< 37 weeks) by difference of state of residence (TEL g/gal) versus state of birth occurrence (TEL g/gal) of migrating mothers

log transformed, or divided in quantiles – and the relaxation of inclusion criteria to birth events in all states irrespective of the universality of reporting, accounting for possible fertility selection effects through the 1980 economic recession, and the substitution of our treatment variable of maternal blood lead for concentration (g/gal) exposure. Moreover, we satisfied a divergent validity test by falsifying our causal channel of maternal blood lead with analyses of the lead-independent outcome of infant sex.

Taken together, these results suggest considerable social and economic benefits from the phase-out of leaded gasoline in the United States through improved infant health. To gauge these gains, first consider the sizable ex-ante estimated social and economic benefits of the phase-out of lead in paint. According to Gould (2009), each dollar invested in lead paint hazard control delivered a return of $17-221 and hundreds of billions in net savings. These benefits were realized through reductions in health care costs, special education costs, and costs of psychological disorders, as

well as by increased lifetime earnings of children through significant cognitive gains. Because leaded gasoline was a far more important determinant of lead exposure in the United States in the second half of the 20^{th} Century (see Laidlaw and Filippelli (2008)), the phase-out as a constituent in automotive gasoline likely delivered benefits on par or in excess of what was had from lead paint hazard control efforts.

In Table 3.12 we calculate the economic benefits of the phase-out of leaded gasoline through the reduction of healthcare-related costs involved in treating low birthweight infants. Importantly, healthcare-related costs of low birthweight constitute a fraction[29] of the total costs of maternal and infant lead exposure over the life-course. To capture our slice of costs, we first estimate the conditional probability $P(LBW)$ of a child being born at less than 2,500 grams as well as the complement probability of normal birthweight P in 1977, 1980, and 1987, given a maternal blood lead (μg/dL) level and leveraging the coefficient on maternal blood lead vis-á-vis low birthweight risk in Table 3.4, Column 3.

Three scenarios (S) are presented, involving assumptions about the healthcare-related costs (C) of normal birthweight (NBW), low birthweight (LBW) at birth (Lewit et al., 1995; Almond et al., 2005) and costs of a LBW infant till age 15 (Lewit et al., 1995). To get the expected healthcare-related costs per infant (EC) we calculate: $EC = P(NBW) \times C_{NBW} + P(LBW) \times C_{LBW}$. Across the three scenarios, the healthcare-related costs of a NBW infant is assumed to be \$3,432 (Lewit et al., 1995). In the first scenario, healthcare-related costs at birth of LBW is assumed to be \$27,095 (Lewit et al., 1995), in the second \$25,750 (Almond et al., 2005), and in the third scenario healthcare-related costs of a LBW infant till age 15 is \$89,975 (Lewit et al., 1995). All figures are in 1999 USD. Assuming 4 million births annually and an expected cost of per infant (EC) of \$5,017 (S1), \$4,927 (S2) and \$9,230 (S3), total annual healthcare-related costs of live-births are \$21.1, \$19.7, and \$36.9 billion in 1977, respectively. By 1980, following from a rapid reduction in the concentration of TEL in gasoline nationwide, total annual healthcare-related costs of live-

[29] In Gould's (2009) assessment of the benefits of lead paint hazard control, health care costs are between 6 and 20% of the total net savings (involving other dimensions of life earnings, tax revenue, special education, attention deficit-hyperactivity disorder, and the direct costs of crime). Also, studies find that maternal exposure to lead increases the risk of hypertensive disorders like eclampsia (Zahran et al., 2014).

Table 3.12: Estimated savings from low birthweight from the phase-out of leaded gasoline

Year	Maternal Blood Lead (μg/dL)	P(LBW)	P(NBW)	Expected Cost	Δ Cost from 1977	% Cost from 1977	Annual Savings ($ Billions)[‡‡]
1977	12.56	0.067	0.933	$5,017			
1980	10.15	0.043	0.957	$4,450	-$568	-11.32%	$2.27
1987	8.500[*]	0.029[‡]	0.971	$4,118	-$899	-17.92%	$3.60
Costs at Birth $27,095[†]							
1977	12.56	0.067	0.933	$4,927			
1980	10.15	0.043	0.957	$4,392	-$536	-10.87%	$2.14
1987	8.500[*]	0.029[‡]	0.971	$4,079	-$848	-17.21%	$3.39
Cost at Birth $25,750[††]							
1977	12.56	0.067	0.933	$9,230			
1980	10.15	0.043	0.957	$7,153	-$2,077	-22.50%	$8.31
1987	8.500[*]	0.029[‡]	0.971	$5,942	-$3,289	-35.63%	$13.16
Costs <15 years $89,975[†]							

Notes: Healthcare-related costs of a non-low birth weight (NLBW) infant is assumed to be $3,432 and from Lewit et al. (1995); [†]Healthcare-related costs of low birthweight (LBW) till age 15 and at birth are from Lewit et al. (1995); [††]Healthcare-related costs of LBW at birth are from Almond et al. (2005); Mean maternal blood lead is limited to mothers birthing singletons with plausible gestational age from states with complete reporting; [*]Estimated blood lead of mothers in 1987 is derived from the statistical relationship of blood lead and TEL concentration, anchored on the observed TEL concentration in 1987 of 0.10 g/gal; The estimated fraction of low birthweight infants in 1987 leverages the coefficient on maternal blood lead vis-á-vis low birthweight risk in Table 3.4, Column 3; To get the expected healthcare-related costs per infant (EC) we calculate: $EC = P(NWB) \times C_{NBW} + P(LBW) \times C_{LBW}$; [‡‡]To get total savings, we multiply EC by an assumed 4 million births per year in the United States;

births decreased by an estimated $2.1 (at birth) to $8.3 billion (by age 15). By 1987, when leaded gasoline was largely phased-out, annual savings reach $3.4 (at birth) to $13.2 billion (by age 15). In 2019 USD, total healthcare-related savings from 1987 to the present exceed $150 billion (at birth) to $600 billion by age 15.

The foregoing benefit analysis adds to the measured gains amassed from lead hazard control policies, detailing yet another channel by which lead destroys human potential and societal welfare. While the elimination of leaded gasoline substantially reduced the flow of lead into the lived environment, both stock and flow sources of exposure risk remain in the United States and abroad. These sources present an ongoing risk to infant health.

Bibliography

Henry Abadin, Annette Ashizawa, Fernando Llados, and Yee-Wan Stevens. Toxicological profile for lead. 2007.

Anna Aizer and Janet Currie. Lead and juvenile delinquency: New evidence from linked birth, school and juvenile detention records. Technical report, National Bureau of Economic Research, 2017.

Douglas Almond and Janet Currie. Killing me softly: The fetal origins hypothesis. *Journal of Economic Perspectives*, 25(3):153–72, 2011.

Douglas Almond, Kenneth Y Chay, and David S Lee. The costs of low birth weight. *The Quarterly Journal of Economics*, 120(3):1031–1083, 2005.

Jere R Behrman and Mark R Rosenzweig. Returns to birthweight. *Review of Economics and Statistics*, 86(2):586–601, 2004.

David Bellinger, Alan Leviton, Christine Waternaux, Herbert Needleman, and Michael Rabinowitz. Longitudinal analyses of prenatal and postnatal lead exposure and early cognitive development. *New England Journal of Medicine*, 316(17):1037–1043, 1987.

David C Bellinger. Very low lead exposures and children's neurodevelopment. *Current Opinion in Pediatrics*, 20(2):172–177, 2008.

David C Bellinger. The protean toxicities of lead: New chapters in a familiar story. *International Journal of Environmental Research and Public Health*, 8(7):2593–2628, 2011.

David C Bellinger. Childhood lead exposure and adult outcomes. *Journal of the American Medical Association*, 317(12):1219–1220, 2017.

David C Bellinger and Andrew M Bellinger. Childhood lead poisoning: The torturous path from science to policy. *The Journal of Clinical Investigation*, 116(4):853–857, 2006.

Sandra E Black, Paul J Devereux, and Kjell G Salvanes. From the cradle to the labor market? The effect of birth weight on adult outcomes. *The Quarterly Journal of Economics*, 122(1):409–439, 2007.

Richard Campanella and Howard W Mielke. Human geography of New Orleans' high-lead geochemical setting. *Environmental Geochemistry and Health*, 30(6):531–540, 2008.

Richard L Canfield, Charles R Henderson Jr, Deborah A Cory-Slechta, Christopher Cox, Todd A Jusko, and Bruce P Lanphear. Intellectual impairment in children with blood lead concentrations below 10 μg per deciliter. *New England Journal of Medicine*, 348(16):1517–1526, 2003.

Alessandro Cavalleri, Claudio Minoia, Luigi Pozzoli, Franco Polatti, and Pier Francesco Bolis. Lead in red blood cells and in plasma of pregnant women and their offspring. *Environmental Research*, 17(3):403–408, 1978.

CDC. Childhood lead poisoning prevention: Blood lead levels in children. Technical report, Centers for Disease Control and Prevention, 7 2019. URL https://www.cdc.gov/nceh/lead/prevention/blood-lead-levels.htm. Web Accessed December 14, 2019.

Kim M Cecil. Effects of early low-level lead exposure on human brain structure, organization and functions. *Journal of Developmental Origins of Health and Disease*, 2(1):17–24, 2011.

Kim M Cecil, Christopher J Brubaker, Caleb M Adler, Kim N Dietrich, Mekibib Altaye, John C Egelhoff, Stephanie Wessel, Ilayaraja Elangovan, Richard Hornung, Kelly Jarvis, et al. Decreased brain volume in adults with childhood lead exposure. *PLoS Medicine*, 5(5):e112, 2008.

Karen Clay, Margarita Portnykh, and Edson Severnini. Toxic truth: Lead and fertility. Technical report, National Bureau of Economic Research, 2018.

Karen Clay, 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 and Human Biology*, 33:181–192, 2019.

Flavio Cunha, James J Heckman, and Susanne M Schennach. Estimating the technology of cognitive and noncognitive skill formation. *Econometrica*, 78(3):883–931, 2010.

Federico Curci and Federico Masera. *Flight from urban blight: Lead poisoning, crime and suburbanization*. PhD thesis, 2017.

Janet Currie. Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature*, 47(1):87–122, 2009.

Janet Currie and Rosemary Hyson. Is the impact of health shocks cushioned by socioeconomic status? The case of low birthweight. *American Economic Review*, 89(2):245–250, 1999.

Janet Currie and Reed Walker. Traffic congestion and infant health: Evidence from E-ZPass. *American Economic Journal: Applied Economics*, 3(1):65–90, 2011.

Kim N Dietrich, Ris M Douglas, Paul A Succop, Omer G Berger, and Robert L Bornschein. Early exposure to lead and juvenile delinquency. *Neurotoxicology and Teratology*, 23(6):511–518, 2001.

Orla Doyle, Colm Harmon, James J Heckman, Caitriona Logue, and Seong Moon. Measuring investment in human capital formation: An experimental analysis of early life outcomes. Technical report, National Bureau of Economic Research, 2013.

Mark R Farfel, Anna O Orlova, Peter SJ Lees, Charles Rohde, Peter J Ashley, and J Julian Chisolm Jr. A study of urban housing demolitions as sources of lead in ambient dust: Demolition practices and exterior dust fall. *Environmental Health Perspectives*, 111(9):1228–1234, 2003.

Mark R Farfel, Anna O Orlova, Rufus L Chaney, Peter SJ Lees, Charles Rohde, and Peter J Ashley. Biosolids compost amendment for reducing soil lead hazards: A pilot study of Orgro® amendment and grass seeding in urban yards. *Science of the Total Environment*, 340(1-3):81–95, 2005a.

Mark R Farfel, Anna O Orlova, Peter SJ Lees, Charles Rohde, Peter J Ashley, and J Julian Chisolm Jr. A study of urban housing demolition as a source of lead in ambient dust on sidewalks, streets, and alleys. *Environmental Research*, 99(2):204–213, 2005b.

Gabriel M Filippelli and Mark AS Laidlaw. The elephant in the playground: Confronting lead-contaminated soils as an important source of lead burdens to urban populations. *Perspectives in Biology and Medicine*, 53(1):31–45, 2010.

Gabriel M Filippelli, Mark AS Laidlaw, Jennifer C Latimer, and Robyn Raftis. Urban lead poisoning and medical geology: An unfinished story. *GSA Today*, 15(1):4–11, 2005.

National Center for Health Statistics. National Health and Nutrition Examination Survey, 1976-1980: Hematology and biochemistry. Technical report, Hyattsville, MD: Public Health Service, 1982.

Tanya E Froehlich, Bruce P Lanphear, Kim N Dietrich, Deborah A Cory-Slechta, Ning Wang, and Robert S Kahn. Interactive effects of a DRD4 polymorphism, lead, and sex on executive functions in children. *Biological Psychiatry*, 62(3):243–249, 2007.

Drew Fudenberg and David K Levine. A dual-self model of impulse control. *American Economic Review*, 96(5):1449–1476, 2006.

Juan J Gershanik, G Gary Brooks, and Joseph A Little. Blood lead values in pregnant women and their offspring. *American Journal of Obstetrics and Gynecology*, 119(4):508–511, 1974.

Elise Gould. Childhood lead poisoning: Conservative estimates of the social and economic benefits of lead hazard control. *Environmental Health Perspectives*, 117(7):1162–1167, 2009.

Joshua Graff Zivin and Matthew Neidell. Environment, health, and human capital. *Journal of Economic Literature*, 51(3):689–730, 2013.

Daniel Grossman and David JG Slutsky. The effect of an increase in lead in the water system on fertility and birth outcomes: The case of Flint, Michigan. 2017.

BL Gulson, CW Jameson, KR Mahaffey, KJ Mizon, MJ Korsch, and G Vimpani. Pregnancy increases mobilization of lead from maternal skeleton. *Journal of Laboratory and Clinical Medicine*, 130(1):51–62, 1997a.

BL Gulson, KJ Mizon, JM Palmer, MJ Korsch, Nicole Patison, CW Jameson, and JB Donnelly. Urinary lead isotopes during pregnancy and postpartum indicate no preferential partitioning of endogenous lead into plasma. *Journal of Laboratory and Clinical Medicine*, 136(3):236–242, 2000.

Brian L Gulson, Kathryn R Mahaffey, M Vidal, CW Jameson, Alistair J Law, Karen J Mizon, AJ Smith, and Michael J Korsch. Dietary lead intakes for mother/child pairs and relevance to Pharmacokinetic models. *Environmental Health Perspectives*, 105(12):1334–1342, 1997b.

Brian L Gulson, Kathryn R Mahaffey, C William Jameson, Nicole Patison, Alistair J Law, Karen J Mizon, Michael J Korsch, and David Pederson. Impact of diet on lead in blood and urine in female adults and relevance to mobilization of lead from bone stores. *Environmental Health Perspectives*, 107(4):257–263, 1999.

EW Harville, Irva Hertz-Picciotto, M Schramm, M Watt-Morse, K Chantala, J Osterloh, PJ Parsons, and W Rogan. Factors influencing the difference between maternal and cord blood lead. *Occupational and Environmental Medicine*, 62(4):263–269, 2005.

Irva Hertz-Picciotto, Margaret Schramm, Margaret Watt-Morse, Kim Chantala, John Anderson, and John Osterloh. Patterns and determinants of blood lead during pregnancy. *American Journal of Epidemiology*, 152(9):829–837, 2000.

Hilary Hoynes, Marianne Page, and Ann Huff Stevens. Can targeted transfers improve birth outcomes?: Evidence from the introduction of the WIC program. *Journal of Public Economics*, 95 (7-8):813–827, 2011.

David L Johnson and Jennifer K Bretsch. Soil lead and children's blood lead levels in Syracuse, NY, USA. *Environmental Geochemistry and Health*, 24(4):375–385, 2002.

Todd A Jusko, Charles R Henderson Jr, Bruce P Lanphear, Deborah A Cory-Slechta, Patrick J Parsons, and Richard L Canfield. Blood lead concentrations< 10 μg/dl and child intelligence at 6 years of age. *Environmental health perspectives*, 116(2):243–248, 2008.

Rebecca Kessler. Sunset for leaded aviation gasoline? *Environmental Health Perspectives*, 121 (2):A54–A57, 2013.

Ahmed Khwaja, Dan Silverman, and Frank Sloan. Time preference, time discounting, and smoking decisions. *Journal of Health Economics*, 26(5):927–949, 2007.

Katherine A Koenig, Meredith C Frey, and Douglas K Detterman. ACT and general cognitive ability. *Intelligence*, 36(2):153–160, 2008.

Mark AS Laidlaw and Gabriel M Filippelli. Re-suspension of urban soils as a persistent source of lead poisoning in children: A review and new directions. *Applied Geochemistry*, 23(8):2021–2039, 2008.

Mark AS Laidlaw, Howard W Mielke, Gabriel M Filippelli, David L Johnson, and Christopher R Gonzales. Seasonality and children's blood lead levels: Developing a predictive model using climatic variables and blood lead data from Indianapolis, Indiana, Syracuse, New York, and New Orleans, Louisiana (USA). *Environmental Health Perspectives*, 113(6):793–800, 2005.

Mark AS Laidlaw, Sammy Zahran, Howard W Mielke, Mark P Taylor, and Gabriel M Filippelli. 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, 2012.

Bruce P Lanphear, Richard Hornung, Jane Khoury, Kimberly Yolton, Peter Baghurst, David C Bellinger, Richard L Canfield, Kim N Dietrich, Robert Bornschein, Tom Greene, et al. Low-level environmental lead exposure and children's intellectual function: An international pooled analysis. *Environmental Health Perspectives*, 113(7):894, 2005.

Eugene M Lewit, Linda Schuurmann Baker, Hope Corman, and Patricia H Shiono. The direct cost of low birth weight. *The Future of Children*, pages 35–56, 1995.

Theodore I Lidsky and Jay S Schneider. Lead neurotoxicity in children: Basic mechanisms and clinical correlates. *Brain*, 126(1):5–19, 2003.

Mildred Maisonet, Timothy J Bush, Adolfo Correa, and JJ Jaakkola. Relation between ambient air pollution and low birth weight in the Northeastern United States. *Environmental Health Perspectives*, 109(3):351–356, 2001.

Howard Mielke, Christopher Gonzales, Eric Powell, and Paul Mielke. Changes of multiple metal accumulation (MMA) in New Orleans soil: Preliminary evaluation of differences between survey I (1992) and survey II (2000). *International Journal of Environmental Research and Public Health*, 2(2):308–313, 2005.

Howard W Mielke and Patrick L Reagan. Soil is an important pathway of human lead exposure. *Environmental Health Perspectives*, 106(1):217, 1998.

Howard W Mielke and Sammy Zahran. The urban rise and fall of air lead (Pb) and the latent surge and retreat of societal violence. *Environment International*, 43:48–55, 2012.

Howard W Mielke, Christopher R Gonzales, M Kelley Smith, and Paul W Mielke. The urban environment and children's health: Soils as an integrator of lead, zinc, and cadmium in New Orleans, Louisiana, USA. *Environmental Research*, 81(2):117–129, 1999.

Howard W Mielke, Chris R Gonzales, Eric Powell, Morten Jartun, and Paul W Mielke Jr. 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, 2007.

Marie Lynn Miranda, Rebecca Anthopolos, and Douglas Hastings. A geospatial analysis of the effects of aviation gasoline on childhood blood lead levels. *Environmental Health Perspectives*, 119(10):1513–1516, 2011.

Herbert Needleman. Lead poisoning. *Annual Review of Medicine*, 55:209–222, 2004.

Herbert L Needleman. The removal of lead from gasoline: Historical and personal reflections. *Environmental Research*, 84(1):20–35, 2000.

Herbert L Needleman and David Bellinger. The health effects of low level exposure to lead. *Annual Review of Public Health*, 12(1):111–140, 1991.

Rick Nevin. How lead exposure relates to temporal changes in IQ, violent crime, and unwed pregnancy. *Environmental Research*, 83(1):1–22, 2000.

Joel T Nigg, Molly Nikolas, G Mark Knottnerus, Kevin Cavanagh, and Karen Friderici. Confirmation and extension of association of blood lead with attention-deficit/hyperactivity disorder (ADHD) and ADHD symptom domains at population-typical exposure levels. *Journal of Child Psychology and Psychiatry*, 51(1):58–65, 2010.

J Peter Nilsson. The long-term effects of early childhood lead exposure: Evidence from the phase-out of leaded gasoline. *Institute for Labour Market Policy Evaluation (IFAU) Working Paper*, 2009.

Bureau of Labor Statistics. Sample weights and design effects. https://www.nlsinfo.org/content/cohorts/nlsy97/using-and-understanding-the-data/sample-weights-design-effects. Accessed: 2019-12-27.

Bureau of Labor Statistics. National Longitudinal Survey of Youth 1997 cohort, 1997-2017 (rounds 1-18). Technical report, U.S. Department of Labor, Columbus, OH: The Ohio State University, 2019. Produced and distributed by the Center for Human Resource Research (CHRR).

Nikolas C Papanikolaou, Eleftheria G Hatzidaki, Stamatis Belivanis, George N Tzanakakis, and Aristidis M Tsatsakis. Lead toxicity update. A brief review. *Medical Science Monitor*, 11(10): RA329–RA336, 2005.

Felicia A Rabito, Shahed Iqbal, Charles Franklin Shorter, Patricia Osman, PE Philips, E Langlois, and LE White. The association between demolition activity and children's blood lead levels. *Environmental Research*, 103(3):345–351, 2007.

J Raymond, W Wheeler, and MJ Brown. Lead screening and prevalence of blood lead levels in children aged 1-2 years–Child Blood Lead Surveillance System, United States, 2002-2010 and National Health and Nutrition Examination Survey, United States, 1999-2010. *Morbidity and Mortality Weekly Report Supplements*, 63(2):36–42, 2014.

PL Reagan and EK Silbergeld. Establishing a health based standard for lead in residential soils. *Trace Substances in Environmental Health*, 23:199–238, 1989.

Aaron Reuben, Avshalom Caspi, Daniel W Belsky, Jonathan Broadbent, Honalee Harrington, Karen Sugden, Renate M Houts, Sandhya Ramrakha, Richie Poulton, and Terrie E Moffitt. Association of childhood blood lead levels with cognitive function and socioeconomic status at age 38 years and with IQ change and socioeconomic mobility between childhood and adulthood. *Journal of the American Medical Association*, 317(12):1244–1251, 2017.

Jessica Wolpaw Reyes. Environmental policy as social policy? The impact of childhood lead exposure on crime. *The B.E. Journal of Economic Analysis and Policy*, 7(1):51, 2007.

Jessica Wolpaw Reyes. Lead exposure and behavior: Effects on antisocial and risky behavior among children and adolescents. *Economic Inquiry*, 53(3):1580–1605, 2015.

Beate Ritz and Fei Yu. The effect of ambient carbon monoxide on low birth weight among children born in southern california between 1989 and 1993. *Environmental Health Perspectives*, 107(1): 17–25, 1999.

Stephen J Rothenberg, Samuel Karchmer, Lourdes Schnaas, Estela Perroni, Francisco Zea, and J Fernandez Alba. Changes in serial blood lead levels during pregnancy. *Environmental Health Perspectives*, 102(10):876–880, 1994.

Muhammad T Salam, Joshua Millstein, Yu-Fen Li, Frederick W Lurmann, Helene G Margolis, and Frank D Gilliland. Birth outcomes and prenatal exposure to ozone, carbon monoxide, and particulate matter: Results from the Children's Health Study. *Environmental Health Perspectives*, 113(11):1638–1644, 2005.

Lawrence M Schell, Melinda Denham, Alice D Stark, Marta Gomez, Julia Ravenscroft, Patrick J Parsons, Aida Aydermir, and Renee Samelson. Maternal blood lead concentration, diet during pregnancy, and anthropometry predict neonatal blood lead in a socioeconomically disadvantaged population. *Environmental Health Perspectives*, 111(2):195–200, 2003.

Ella Mae Shelton. Motor gasolines, Summer 1978. Technical report, Department of Energy, Bartlesville, OK (USA). Bartlesville Energy Technology Center, 1979.

Gary Solon, Steven J. Haider, and Jeffrey Wooldridge. What are we weighting for? NBER Working Papers 18859, National Bureau of Economic Research, Inc, February 2013. URL https://ideas.repec.org/p/nbr/nberwo/18859.html.

Richard H Thaler and Hersh M Shefrin. An economic theory of self-control. *Journal of Political Economy*, 89(2):392–406, 1981.

Christopher D Toscano and Tomás R Guilarte. Lead neurotoxicity: From exposure to molecular effects. *Brain Research Reviews*, 49(3):529–554, 2005.

Sammy Zahran, Howard W Mielke, Christopher R Gonzales, Eric T Powell, and Stephan Weiler. New Orleans before and after Hurricanes Katrina/Rita: A quasi-experiment of the association between soil lead and children's blood lead. *Environmental Science and Technology*, 44(12): 4433–4440, 2010.

Sammy Zahran, Lori Peek, Jeffrey G Snodgrass, Stephan Weiler, and Lynn Hempel. Economics of disaster risk, social vulnerability, and mental health resilience. *Risk Analysis*, 31(7):1107–1119, 2011.

Sammy Zahran, Mark AS Laidlaw, Shawn P McElmurry, Gabriel M Filippelli, and Mark Taylor. Linking source and effect: Resuspended soil lead, air lead, and children's blood lead levels in Detroit, Michigan. *Environmental Science and Technology*, 47(6):2839–2845, 2013.

Sammy Zahran, Ian M Breunig, Bruce G Link, Jeffrey G Snodgrass, and Stephan Weiler. A quasi-experimental analysis of maternal altitude exposure and infant birth weight. *American Journal of Public Health*, 104(S1):S166–S174, 2014.

Sammy Zahran, Terrence Iverson, Shawn P McElmurry, and Stephan Weiler. The effect of leaded aviation gasoline on blood lead in children. *Journal of the Association of Environmental and Resource Economists*, 4(2):575–610, 2017a.

Sammy Zahran, Shawn P McElmurry, and Richard C Sadler. Four phases of the Flint Water Crisis: Evidence from blood lead levels in children. *Environmental Research*, 157:160–172, 2017b.

Appendix A

Risk assessment question set, NLSY97

To characterize individual's appetite for risk, respondents were given a set of nine questions, and asked to rate their willingness to take risks in a given context. The lead in question reads:

> Are you generally a person who is fully prepared to take risks or do you try to avoid taking risks? Rate yourself from 0 to 10, where 0 means "unwilling to take any risks" and 10 means "fully prepared to take risks."

After responding to the previous question, respondents are read the following statement:

> People can behave differently in different situations. How would you rate your willingness to take risks in the following areas? For each situation, rate yourself from 0 to 10, where 0 means "unwilling to take any risks" and 10 means "fully prepared to take risks."

then asked to answer the following sequence of risk assessment questions, in order:

1. "...While driving?"
2. "...In financial matters?"
3. "...In your occupation?"
4. "...With your health?"
5. "...In your faith in other people?"
6. "...In your romantic relationships?"
7. "...In making major life changes?"
8. "...In placing a bet where you have a 50-50 chance of winning $20 and a 50-50 chance of losing $10?"

Appendix B

Recession effect on maternal blood lead

To elucidate how an economic recession may implicate the estimated effect of maternal blood lead on infant birthweight through selection of birth mothers, let $E(BW|Pb, s)$ be the expected birthweight (BW) on an infant given maternal blood lead (Pb) and socioeconomic status (s). This is a physiological relationship which won't be affected by an embargo or recession.

Letting $f_0(s)$ be the probability density function (pdf) of the socio-economic factors (SEF) of women who give birth before the event, and $f_1(s)$ the pdf of SEF women who give birth after the event.

Assume that $s \in [0, S]$. We expect higher SEF women to be more likely to give birth after an event. So, $f_1(\bullet)$ will stochastically dominate $f_0(\bullet)$ (i.e. $F_0(s) \geq F_1(s) \; \forall \; s$, and $F_j(\bullet) \; \forall \; j \in \{1, 2\}$ is the cumulative distribution function (cdf) of the pdf's).

The BW-Pb relationship in the period before is $BW_0(Pb) = \int_0^S E(BW \mid Pb, s) \cdot f_0(s) \, ds$

The BW-Pb relationship in the period after is $BW_1(Pb) = \int_0^S E(BW \mid Pb, s) \cdot f_1(s) \, ds$

So, the difference between the two relationships is:

$$\Big[BW_1(Pb) \; - \; BW_0(Pb) \Big] \; = \; \int_0^S E(BW \mid Pb, s) \cdot f_1(s) \, ds \; - \; \int_0^S E(BW \mid Pb, s) \cdot f_0(s) \, ds \qquad \text{(B.1)}$$

Focusing on the first component in the right-hand side of Equation B.1 and using Integration by Parts, we get:

$$\int_0^S E(BW \mid Pb, s) \cdot f_1(s)\, ds \;\; =$$

$$\Big[\; E(BW \mid Pb, S) \cdot F_1(S) - \; E(BW \mid Pb, 0) \cdot F_1(0) \;\Big]$$

$$- \quad \int_0^S \left(\frac{\partial E(BW \mid Pb, s)}{\partial s} \right) \cdot F_1(s)\, ds$$

Since $F_1(S) = 1$ and $F_1(0) = 0$, it becomes:

$$\Big\{\; E(BW \mid Pb, S) - \; \int_0^S \left(\frac{\partial E(BW \mid Pb, s)}{\partial s} \right) \cdot F_1(s)\, ds \;\Big\} \tag{B.2}$$

Similarly, the second component of in the right-hand side of Equation B.1 is:

$$\Big\{\; E(BW \mid Pb, S) - \; \int_0^S \left(\frac{\partial E(BW \mid Pb, s)}{\partial s} \right) \cdot F_0(s)\, ds \;\Big\} \tag{B.3}$$

Combining Equation B.2 and Equation B.3, we get:

$$\left\{ E(BW \mid Pb, S) \; - \; \int_0^S \left(\frac{\partial E(BW \mid Pb, s)}{\partial s} \right) \; \cdot F_1(s)\, ds \right\} \qquad -$$

$$\left\{ E(BW \mid Pb, S) \; - \; \int_0^S \left(\frac{\partial E(BW \mid Pb, s)}{\partial s} \right) \; \cdot F_0(s)\, ds \right\}$$

$$= \int_0^S \left(\frac{E(BW \mid Pb, s)}{s} \right) \cdot F_0(s)\, ds \; - \; \int_0^S \left(\frac{\partial E(BW \mid Pb, s)}{\partial s} \right) \cdot F_1(s)\, ds$$

$$= \int_0^S \left(\frac{\partial E(BW \mid Pb, s)}{\partial s} \right) \cdot \; \Big[F_0(s) \; - F_1(s) \Big]\, ds$$

Substituting back into Equation B.1, returns the final BW-Pb relationship of interest:

$$\Big[\, BW_1(Pb) \;-\; BW_0(Pb) \Big] \;=\; \int_0^S \left(\frac{\partial E(BW \mid Pb, s)}{\partial s} \right) \cdot \Big[F_0(s) \;-\; F_1(s) \Big]\, ds \qquad \textbf{(B.4)}$$

Because we expect socio-economic factors to have a positive impact on birthweight, $\frac{\partial\, E(BW \mid Pb, s)}{\partial\, s} > 0$. Furthermore, if $f_1(s)$ stochastically dominates $f_0(s)$, then $\Big[F_0(s) \;-\; F_1(s) \Big]\, ds > 0$. Thus, we would expect the BW-Pb relationship to be higher.

Appendix C

Results and figures

Table C.1: Blood lead effect of average gasoline lead concentrations, Reyes (2015) replication

	(1)	(2)	(3)	(4)	(5)	(6)
Gasoline Lead (g/gal)	4.17***	6.73***	4.39*	5.73**	4.56*	8.67*
	(1.285)	(1.186)	(2.438)	(2.211)	(2.565)	(4.651)
Age	-0.27**	-0.27**	-0.29**	-0.34***	-0.30**	-0.38
	(0.124)	(0.128)	(0.123)	(0.126)	(0.125)	(0.445)
Female	-0.53**	-0.46*	-0.44	-0.52*	-0.48*	-0.27
	(0.256)	(0.261)	(0.264)	(0.287)	(0.256)	(0.355)
Black	6.84***	6.58***	6.41***		7.08***	6.38***
	(0.833)	(0.647)	(0.602)		(0.638)	(0.994)
Low	1.82***	1.77***	1.78***	2.70***		1.95***
	(0.414)	(0.370)	(0.373)	(0.487)		(0.663)
Mid	-0.41	-0.44	-0.47	-0.73*		-0.53
	(0.458)	(0.398)	(0.393)	(0.403)		(0.647)
Constant	8.26***	1.06	8.64*	7.63*	9.15*	2.50
	(2.365)	(2.171)	(4.795)	(4.482)	(4.946)	(9.718)
Observations	2,192	2,192	2,192	2,192	2,192	833
R^2	0.216	0.285	0.295	0.218	0.280	0.386
State FE	No	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes	Yes
Ages	0-6	0-6	0-6	0-6	0-6	2-4

Notes: Results are a replication of Reyes (2015), Table 3; Dependent variable is blood lead level (BLL) in μg/dL; Robust state clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Gasoline lead is $\overline{TEL}$ concentrations (g/gal); Age is respondent's age (years) when the survey was conducted; Indicators for sex and racial status are defined as Female=1, and Black=1 if the respondent identifies as African American; Indicators for household income relative to Federal Poverty Level (FPL) are: Low=1 for < 2x FPL and Mid=1 for ratios $\leq$2x FPL and <3x FPL; State of residence, and Year and Month of survey fixed effects are included as noted; Regression inclusion criteria denoted by Age group;

Table C.2: IQ and low-IQ effect of blood lead by quintiles, linear regressions

	(1) IQ Percentile	(2) IQ Percentile	(3) IQ Percentile	(4) Low-IQ	(5) Low-IQ	(6) Low-IQ
Blood Lead (μg/dL, Q1 Reference Group)						
Blood Lead Q2	-0.061***	-0.060***	-0.031**	0.093***	0.091***	0.050*
	(0.013)	(0.013)	(0.015)	(0.023)	(0.024)	(0.026)
Blood Lead Q3	-0.103***	-0.101***	-0.038**	0.151***	0.149***	0.064**
	(0.013)	(0.013)	(0.017)	(0.024)	(0.024)	(0.029)
Blood Lead Q4	-0.143***	-0.141***	-0.062***	0.223***	0.221***	0.113***
	(0.019)	(0.019)	(0.019)	(0.033)	(0.033)	(0.037)
Blood Lead Q5	-0.213***	-0.210***	-0.097***	0.319***	0.317***	0.162***
	(0.023)	(0.022)	(0.026)	(0.043)	(0.043)	(0.046)
Poverty Ratio (Q1 Reference Group)						
Poverty Ratio Q2			0.073***			-0.122***
			(0.016)			(0.019)
Poverty Ratio Q3			0.139***			-0.200***
			(0.012)			(0.021)
Poverty Ratio Q4			0.203***			-0.256***
			(0.016)			(0.024)
Constant	0.599***	0.604***	0.446***	0.073	0.069	0.288***
	(0.024)	(0.024)	(0.022)	(0.050)	(0.049)	(0.054)
Observations	4,587	4,587	4,587	4,587	4,587	4,587
R^2	0.224	0.225	0.276	0.178	0.178	0.211
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Oversample FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variables are percentile score on ASVAB examination and low-IQ ($<25^{th}$ percentile) indicator. Robust state clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Blood lead quintiles are NLSY sample weighted and using predicted mean blood lead in μg/dL. Poverty ratio quartiles are based on household income to poverty ratio, 1997 (survey round one). State of residence and cohort fixed effects are included in regressions. Controls for sex, minority status, and prior ASVAB excluded for clarity.

Table C.3: Abnormal risk response effects of blood lead, statistically significant models and coefficients

	(1) Drive	(2) Finance	(3) Work	(4) Health	(5) Romance	(6) Gamble
Blood Lead (μg/dL)	0.021*	0.023***	0.029***	0.034*	0.017**	0.024**
	(0.011)	(0.008)	(0.008)	(0.019)	(0.009)	(0.012)
Sex	0.047	0.124***	0.047	-0.081	0.073**	0.054*
	(0.032)	(0.028)	(0.030)	(0.049)	(0.032)	(0.031)
Minority	0.140***	0.114**	0.073*	0.272***	0.035	-0.021
	(0.050)	(0.044)	(0.043)	(0.068)	(0.036)	(0.046)
Married	-0.047	-0.022	0.038	-0.020	0.134***	0.025
	(0.030)	(0.028)	(0.034)	(0.057)	(0.033)	(0.033)
Poverty Ratio (Q1 Reference Group)						
Poverty Ratio Q2	-0.052	-0.116***	-0.067*	-0.115	-0.065*	-0.123**
	(0.039)	(0.038)	(0.036)	(0.069)	(0.035)	(0.052)
Poverty Ratio Q3	-0.116**	-0.139***	-0.042	-0.102	-0.003	-0.035
	(0.045)	(0.028)	(0.043)	(0.078)	(0.033)	(0.042)
Poverty Ratio Q4	-0.121**	-0.139***	-0.134***	-0.357***	-0.108***	-0.048
	(0.050)	(0.040)	(0.048)	(0.092)	(0.037)	(0.043)
Constant	0.613***	0.291***	0.333***	0.144	0.737***	0.736***
	(0.120)	(0.107)	(0.107)	(0.237)	(0.108)	(0.126)
Observations	4,224	4,224	4,224	4,224	4,224	4,224
R-squared	0.040	0.046	0.033	0.044	0.031	0.019
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Cohort FE	Yes	Yes	Yes	Yes	Yes	Yes
Oversample FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Dependent variables are abnormal risk preferences (mean deviation), based on NLSY97 survey responses; Robust state clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; Blood lead is predicted mean blood lead in μg/dL; Indicators controlling for for sex (Sex=1 if male), minority status (Minority=1, if non-white), and marital status (Married=1 if married when surveyed); Poverty ratio quartiles are based on household income to FPL poverty ratios; State of residence, age cohort, and over-sample survey group fixed effects included in all regressions;

Table C.4: Birthweight (grams) effect of TEL

	(1) BW	(2) BW	(3) ln(BW)	(4) BW	(5) ln(BW)
TEL (g/gal)	-11.506***				
	(1.278)				
ln(TEL) (g/gal)		-13.394***	-0.004***		
		(1.356)	(0.000)		
Reference TEL (g/gal) Q1					
TEL (g/gal) Q2				-8.461***	-0.003***
				(1.112)	(0.000)
TEL (g/gal) Q3				-10.022***	-0.003***
				(1.147)	(0.000)
TEL (g/gal) Q4				-10.230***	-0.003***
				(1.177)	(0.000)
TEL (g/gal) Q5				-13.663***	-0.004***
				(1.416)	(0.000)
Constant	3,159.524***	3,147.277***	8.023***	3,152.386***	8.024***
	(16.481)	(16.451)	(0.007)	(16.582)	(0.007)
Observations	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047
R^2	0.069	0.069	0.057	0.069	0.057
County FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); control variables include maternal age (1 = Under 20 years of age; 2 = 20-24 years; 3 = 25-29 years; 4 = 30-34 years; 5 = 35-39 years; and 6 = 40-49 years); maternal race (0 = white; 1 = non-white); maternal education (1 = 0-8 years; 2 = 9-11 years; 3 = 12 years; 4 = 13-15 years; 5 = 16 year and over; 6 = Not stated); interval since last live birth (2 =1-11 months; 3 =12-23 months; 4 =24-35 months; 5 =36-47 months; 6 =48-71 months; 7 =over 72 months; 8 =Not stated); and infant sex (female $= 0$, male $= 1$).

Table C.5: Birthweight (grams) effect of Pb gasoline emissions

	(1) BW	(2) BW	(3) ln(BW)	(4) BW	(5) ln(BW)
Pb Emissions	-3.055				
	(0.274)				
ln(Pb Emissions)		-12.613***	-0.004***		
		(1.197)	(0.000)		
Reference Pb Emissions Q1					
Pb Emissions Q2				-7.543***	-0.002***
				(1.579)	(0.001)
Pb Emissions Q3				-8.527***	-0.002***
				(2.228)	(0.001)
Pb Emissions Q4				-16.019***	-0.005***
				(2.185)	(0.001)
Pb Emissions Q5				-21.585***	-0.006***
				(2.375)	(0.001)
Constant	3,152.221***	3,246.911***	8.051***	3,151.932***	8.024***
	(16.664)	(19.558)	(0.008)	(16.418)	(0.007)
Observations	6,325,047	6,325,047	6,325,047	6,325,047	6,325,047
R^2	0.069	0.069	0.057	0.069	0.057
County FE	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); control variables include maternal age ($1 =$ Under 20 years of age; $2 =$ 20-24 years; $3 =$ 25-29 years; $4 =$ 30-34 years; $5 =$ 35-39 years; and $6 =$ 40-49 years); maternal race ($0 =$ white; $1 =$ non-white); maternal education ($1 =$ 0-8 years; $2 =$ 9-11 years; $3 =$ 12 years; $4 =$ 13-15 years; $5 =$ 16 year and over; $6 =$ Not stated); interval since last live birth ($2 =$1-11 months; $3 =$12-23 months; $4 =$24-35 months; $5 =$36-47 months; $6 =$48-71 months; $7 =$over 72 months; $8 =$Not stated); and infant sex (female $= 0$, male $= 1$); Pb emissions are mesured in tonnages/1000.

Table C.6: Birthweight (grams) effect of maternal blood lead, all states

	(1) BW	(2) BW	(3) BW	(4) BW	(5) ln(BW)	(6) ln(BW)	(7) BW	(8) BW	(9) BW
Maternal Blood Lead (μg/dL)	-36.638***	-79.675***	-22.192***						
	(2.311)	(3.511)	(2.072)						
ln(Maternal Blood Lead) (μg/dL)				-304.168***	-0.331***	-0.120***			
				(20.615)	(0.013)	(0.006)			
Reference (Maternal Blood Lead Q1)									
Maternal Blood Lead Q2							-7.856***	-67.581***	-13.687***
							(1.537)	(2.756)	(1.913)
Maternal Blood Lead Q3							-19.875***	-112.779***	-25.662***
							(2.159)	(4.438)	(3.019)
Maternal Blood Lead Q4							-42.157***	-157.387***	-36.636***
							(4.548)	(5.880)	(4.415)
Maternal Blood Lead Q5							-150.129***	-269.674***	-52.293***
							(7.909)	(10.592)	(5.088)
Constant	3,787.969***	4,402.354***	3,514.070***	4,005.595***	8.955***	8.368***	3,409.278***	3,573.915***	3,273.031***
	(23.388)	(46.377)	(28.719)	(53.568)	(0.033)	(0.017)	(3.147)	(8.416)	(5.865)
Observations	8,568,194	8,568,194	8,568,194	8,568,194	8,568,194	8,568,194	8,568,194	8,568,194	8,568,194
R^2	0.009	0.029	0.068	0.068	0.026	0.057	0.009	0.025	0.068
County FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Month FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	Yes	No	Yes	No	No	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); control variables include maternal age ($1 =$ Under 20 years of age; $2 = 20$-24 years; $3 = 25$-29 years; $4 = 30$-34 years; $5 = 35$-39 years; and $6 = 40$-49 years); maternal race ($0 =$ white; $1 =$ non-white); maternal education ($1 = 0$-8 years; $2 = 9$-11 years; $3 = 12$ years; $4 = 13$-15 years; $5 = 16$ year and over; $6 =$ Not stated); interval since last live birth ($2 =$1-11 months; $3 =$12-23 months; $4 =$24-35 months; $5 =$36-47 months; $6 =$48-71 months; $7 =$over 72 months; $8 =$Not stated); and infant sex (female $= 0$, male $= 1$).

Table C.7: Gestational length (weeks) effect of maternal blood lead, all states

	(1) GL	(2) GL	(3) GL	(4) GL	(5) ln(GL)	(6) ln(GL)	(7) GL	(8) GL	(9) GL
Maternal Blood Lead (μg/dL)	-0.130***	-0.313***	-0.097***						
	(0.008)	(0.010)	(0.007)						
ln(Maternal Blood Lead) (μg/dL)				-1.338***	-0.105***	-0.038***			
				(0.077)	(0.003)	(0.002)			
Reference (Maternal Blood Lead Q1)									
Maternal Blood Lead Q2							-0.025***	-0.300***	-0.064***
							(0.005)	(0.008)	(0.007)
Maternal Blood Lead Q3							-0.063***	-0.491***	-0.111***
							(0.007)	(0.012)	(0.011)
Maternal Blood Lead Q4							-0.154***	-0.677***	-0.150***
							(0.014)	(0.016)	(0.015)
Maternal Blood Lead Q5							-0.534***	-1.093***	-0.226***
							(0.028)	(0.024)	(0.019)
Constant	41.088***	43.730***	40.870***	43.040***	3.947***	3.773***	39.738***	40.530***	39.808***
	(0.078)	(0.132)	(0.095)	(0.201)	(0.007)	(0.006)	(0.011)	(0.020)	(0.027)
Observations	8,568,194	8,568,194	8,568,194	8,568,194	8,568,194	8,568,194	8,568,194	8,568,194	8,568,194
R^2	0.005	0.018	0.027	0.027	0.018	0.027	0.005	0.016	0.026
County FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Year FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Month FE	No	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	Yes	No	Yes	No	No	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); control variables include maternal age (1 = Under 20 years of age; 2 = 20-24 years; 3 = 25-29 years; 4 = 30-34 years; 5 = 35-39 years; and 6 = 40-49 years); maternal race (0 = white; 1 = non-white); maternal education (1 = 0-8 years; 2 = 9-11 years; 3 = 12 years; 4 = 13-15 years; 5 = 16 year and over; 6 = Not stated); interval since last live birth (2 =1-11 months; 3 =12-23 months; 4 =24-35 months; 5 =36-47 months; 6 =48-71 months; 7 =over 72 months; 8 =Not stated); and infant sex (female = 0, male = 1).

Table C.8: APGAR score (< 7) 5-minute effect of maternal blood lead

	(1) APGAR	(2) APGAR	(3) APGAR	(4) APGAR	(5) APGAR	(6) APGAR	(7) APGAR
Maternal Blood Lead (μg/dL)	0.002***	0.006***	0.003***				
	(0.000)	(0.000)	(0.000)				
ln(Maternal Blood Lead) (μg/dL)				0.041***			
				(0.004)			
Reference (Maternal Blood Lead Q1)							
Maternal Blood Lead Q2					0.001***	0.004***	0.001***
					(0.000)	(0.000)	(0.000)
Maternal Blood Lead Q3					0.001***	0.008***	0.002***
					(0.000)	(0.000)	(0.000)
Maternal Blood Lead Q4					0.002***	0.010***	0.003***
					(0.000)	(0.001)	(0.001)
Maternal Blood Lead Q5					0.010***	0.019***	0.006***
					(0.001)	(0.001)	(0.001)
Constant	-0.006***	-0.049***	-0.017***	-0.076***	0.016***	0.009***	0.021***
	(0.002)	(0.003)	(0.004)	(0.010)	(0.000)	(0.001)	(0.001)
Observations	5,270,952	5,270,952	5,270,952	5,270,952	5,270,952	5,270,952	5,270,952
R^2	0.00	0.04	0.06	0.06	0.01	0.04	0.06
County FE	No	Yes	Yes	Yes	No	Yes	Yes
Year FE	No	Yes	Yes	Yes	No	Yes	Yes
Month FE	No	Yes	Yes	Yes	No	Yes	Yes
Controls	No	No	Yes	Yes	No	No	Yes

Notes: Robust county clustered errors in parentheses *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$; All models limit to singletons with plausible gestational length (> 20 weeks); control variables include maternal age (1 = Under 20 years of age; 2 = 20-24 years; 3 = 25-29 years; 4 = 30-34 years; 5 = 35-39 years; and 6 = 40-49 years); maternal race (0 = white; 1 = non-white); maternal education (1 = 0-8 years; 2 = 9-11 years; 3 = 12 years; 4 = 13-15 years; 5 = 16 year and over; 6 = Not stated); interval since last live birth (2 =1-11 months; 3 =12-23 months; 4 =24-35 months; 5 =36-47 months; 6 =48-71 months; 7 =over 72 months; 8 =Not stated); and infant sex (female $= 0$, male $= 1$)

TABLE 4. - MOTOR GASOLINE SURVEY, SUMMER 1979
AVERAGE DATA FOR BRANDS IN EACH DISTRICT--CONTINUED

REGULAR GASOLINE

DISTRICT NO. AND NAME	NO. OF BRANDS	SAMPLES	GR., ASTM D287 API	SULF., ASTM D1266 WT %	GUM ASTM D381 MG	BENZENE ASTM D3606 VOL. %	LEAD ASTM D526 G/GAL	OCTANE NUMBER RES. ASTM D2699	OCTANE NUMBER MOT. ASTM D2700	OCTANE NUMBER R+M --- 2	RVP, ASTM D323 LB	20 V/L ASTM D439 F	DISTILLATION, ASTM D86 TEMPERATURE, F (CORRECTED TO 760 MM HG) IBP	5	10	20	30	50	70	90	95	EP	RES %	LOSS %
1 NORTHEAST	11	24	61.6	0.045	2	1.50	2.41	93.0	86.5	89.7	10.5	129	86	102	114	136	158	204	265	341	368	407	0.8	1.4
2 MID-ATLANTIC COAST	17	130	61.2	.034	1	.87	2.21	93.0	86.3	89.7	10.2	131	88	103	115	138	160	210	268	345	378	417	1.0	1.6
3 SOUTHEAST	18	104	61.5	.050	1	1.45	2.50	93.1	86.2	89.7	9.7	133	89	105	118	139	160	210	269	346	380	415	.9	1.5
4 APPALACHIAN	14	71	61.0	.026	1	.79	1.74	93.1	86.0	89.6	10.4	130	87	102	116	137	160	209	264	344	376	421	.8	2.2
5 MICHIGAN	13	41	61.3	.058	2	1.00	1.73	93.1	85.7	89.4	10.9	128	86	101	115	136	158	207	267	353	392	434	1.3	1.7
6 NORTH ILLINOIS	11	29	61.2	-	1	-	1.88	92.9	85.9	89.4	10.9	128	85	102	116	136	157	204	262	349	390	428	1.5	1.4
7 CENTRAL MISSISSIPPI	14	33	61.4	.050	2	1.18	2.41	93.2	86.0	89.6	10.4	130	88	105	117	137	159	205	259	335	369	413	.8	1.8
8 LOWER MISSISSIPPI	15	45	61.5	.070	0	1.60	2.27	92.8	86.3	89.6	9.5	134	90	106	118	138	158	204	257	334	367	411	1.2	1.6
9 NORTH PLAINS	11	12	61.9	.019	1	.83	1.93	92.3	85.3	88.8	10.5	126	72	-	114	130	146	189	250	340	-	410	.5	2.5
10 CENTRAL PLAINS	12	45	61.2	.039	2	1.17	2.11	92.3	85.5	89.0	9.1	136	90	107	120	141	161	204	257	332	366	409	.8	1.3
11 SOUTH PLAINS	23	64	61.2	.027	1	1.00	2.05	92.7	85.5	89.1	9.5	135	91	107	119	139	160	208	264	340	373	412	.8	1.4
12 SOUTH TEXAS	12	38	61.0	.030	1	1.10	2.19	93.1	86.2	89.7	9.0	136	91	107	119	138	159	207	269	348	379	415	.9	1.0
13 SOUTH MOUNTAIN STATES	19	155	60.9	.044	1	1.27	1.59	92.1	84.9	88.5	8.3	143	95	112	127	148	169	213	263	337	371	414	.9	1.5
14 NORTH MOUNTAIN STATES	12	66	60.8	.044	1	.61	1.02	91.5	84.2	87.9	9.3	137	95	108	122	144	166	211	259	338	375	412	1.0	2.5
15 PACIFIC NORTHWEST	7	23	59.7	.025	2	1.25	1.66	92.6	85.7	89.2	9.9	132	87	102	114	136	160	209	268	347	388	416	1.1	2.1
16 NORTH CALIFORNIA	9	35	56.8	.019	2	1.44	.92	93.9	85.2	89.6	8.1	147	96	116	131	154	177	223	274	343	372	415	1.0	1.3
17 SOUTH CALIFORNIA	10	40	57.5	.046	1	1.02	1.16	93.4	85.1	89.3	8.2	143	96	116	128	146	166	215	273	343	374	415	1.0	1.3
AVERAGE			60.7	.039	1	1.13	1.87	92.8	85.7	89.3	9.7	134	89	106	119	140	161	208	264	342	376	416	1.0	1.7
MINIMUM			56.8	.019	0	.61	.92	91.5	84.2	87.9	8.1	126												
MAXIMUM			61.9	.070	2	1.60	2.50	93.9	86.5	89.7	10.9	147												
SAMPLES	955																							

Figure C.1: Motor Gasoline Survey, summer 1979: Regular grade

TABLE 4. - MOTOR GASOLINE SURVEY, SUMMER 1979
AVERAGE DATA FOR BRANDS IN EACH DISTRICT

UNLEADED GASOLINE

DISTRICT NO. AND NAME	NO. OF BRANDS	SAMPLES	GR., ASTM D287 API	SULF., ASTM D1266 WT %	GUM, ASTM D381 MG	BENZENE ASTM D3606 VOL. %	OCTANE NUMBER RES., ASTM D2699	OCTANE NUMBER MOT., ASTM D2700	OCTANE NUMBER R+M --- 2	RVP, ASTM D323 LB	20V/L ASTM D439 F	DISTILLATION, ASTM D86 TEMPERATURE, F (CORRECTED TO 760 MM HG) IBP	PERCENT EVAPORATED 5	10	20	30	50	70	90	95	EP	RES %	LOSS %
1 NORTHEAST	15	32	58.1	0.037	2	1.00	93.8	84.3	89.1	10.5	130	85	99	112	133	157	220	271	332	358	400	0.8	1.4
2 MID-ATLANTIC COAST	21	169	58.5	.023	1	.84	93.5	84.2	88.9	10.2	132	87	102	115	139	165	222	270	334	364	408	1.0	1.6
3 SOUTHEAST	21	98	58.5	.033	1	1.03	93.2	84.3	88.8	9.6	134	88	104	116	138	163	219	270	335	365	403	.8	1.3
4 APPALACHIAN	17	81	59.3	.017	2	1.02	92.8	84.0	88.4	10.6	132	86	101	116	142	170	224	274	341	371	422	.9	2.1
5 MICHIGAN	17	52	59.6	.036	1	1.02	92.9	84.4	88.7	10.9	131	85	99	117	144	172	223	272	344	380	425	1.3	2.4
6 NORTH ILLINOIS	13	36	59.4	-	1	-	93.2	84.5	88.9	10.9	131	83	99	115	142	171	224	272	342	379	421	1.5	1.8
7 CENTRAL MISSISSIPPI	17	36	58.7	.041	2	1.17	92.9	83.8	88.4	10.4	133	87	104	117	143	170	220	269	338	371	412	.9	1.7
8 LOWER MISSISSIPPI	17	49	57.4	.050	1	1.00	93.5	84.2	88.9	9.5	137	88	101	119	144	171	223	271	336	363	406	.9	1.5
9 NORTH PLAINS	12	14	61.0	.015	1	.75	92.8	84.2	88.5	9.2	143	89	-	136	154	191	231	275	373	-	434	.5	4.5
10 CENTRAL PLAINS	14	55	60.1	.036	1	1.10	92.9	84.2	88.5	9.2	137	91	105	119	142	168	214	252	328	362	409	.8	1.6
11 SOUTH PLAINS	25	66	60.6	.014	1	1.54	92.2	84.1	88.2	9.5	137	89	106	121	147	174	220	260	329	363	406	.9	1.3
12 SOUTH TEXAS	14	46	58.4	.025	1	1.00	93.0	83.9	88.5	9.4	137	91	106	119	141	166	224	268	326	356	396	1.0	1.1
13 SOUTH MOUNTAIN STATES	20	118	59.3	.025	1	1.24	91.4	83.1	87.3	8.7	142	91	110	125	149	174	222	266	333	368	412	1.0	1.4
14 NORTH MOUNTAIN STATES	13	38	61.5	.038	1	.19	91.2	82.9	87.0	9.6	135	90	105	118	140	166	212	253	325	364	404	.9	2.6
15 PACIFIC NORTHWEST	8	26	57.3	.012	7	1.26	93.1	84.5	88.8	10.3	136	88	104	118	149	180	229	269	330	364	405	1.4	2.8
16 NORTH CALIFORNIA	11	37	54.9	.011	2	1.62	94.1	84.4	89.3	8.0	148	97	110	133	158	183	228	272	333	361	410	.9	1.4
17 SOUTH CALIFORNIA	13	48	55.1	.027	1	1.39	94.3	84.3	89.3	8.4	146	95	116	130	155	181	230	274	331	360	411	.9	1.2
AVERAGE			58.7	.028	2	1.07	93.0	84.1	88.6	9.7	137	89	104	120	145	172	223	268	336	366	411	1.0	1.9
MINIMUM			54.9	.011	1	.19	91.2	82.9	87.0	8.0	130												
MAXIMUM			61.5	.050	7	1.62	94.3	84.5	89.3	10.9	148												
SAMPLES 1,001																							

Figure C.2: Motor Gasoline Survey, summer 1979: Unleaded grade