Has LRRP Moved The Needle? Examining the Effectiveness of EPA's 2010 Lead Renovation, Repair & Painting Rule

A thesis submitted by

Jeffrey F. Higgins

In partial fulfillment of the requirements for the degree of

Masters of Science
In
Economics

Tufts University
August 2020
Advisor: Ujjayant Chakravorty

Abstract

Though cases of severe lead poisoning in children have fallen dramatically in the U.S. since the mid-20th century, the contemporary medical literature has established that lead is still toxic at low levels of exposure (Lanphear 2007; NTP 2012; ATSDR 2019). In April 2010, EPA implemented the Lead Renovation, Repair, and Painting Rule (LRRP) (EPA 2011; EPA 2019), which requires renovators disturbing lead-based paint in residential homes and child-occupied facilities to attain certification and use best-practice cleaning and containment methods. I exploit time series variation in lead poisoning outcomes to estimate the efficacy of LRRP, employing regression discontinuity design with time as the running variable. I find LRRP had a small impact, reducing mild cases of lead poisoning (blood-lead level>10 μg/dL) by 0.87% and more severe cases of lead poisoning (blood-lead level>25 μg/dL) by 0.06%. However, these results appear to be carried by states without robust state-level lead programs. I evaluate heterogeneity in the treatment effect, comparing the effect of LRRP in states with and without a state-level lead abatement program. I find there were 1.16% more mild cases and 0.12% more severe cases in states with a lead abatement program compared to states without one.

Acknowledgements

I would like to thank my advisor, Ujjayant Chakravorty, for his patience and advice. I would also like to thank Jeffrey Zabel and Gilbert Metcalf for their advice at my defense and in research seminars. I would also like to thank my friends and family who supported me along the way, as well as my partner, Amelia Cohen, for her love and support.

Table of Contents

List of Tables

List of Figures

Hence lazy jaundice with her saffron face;
Palsy, with shaking head and tott'ring knees.
And bloated dropsy, the staunch sot's disease;
Consumption, pale, with keen but hollow eye,
And sharpened feature, shew'd that death was nigh.
The feeble offspring curse their crazy sires,
And, tainted from his birth, the youth expires.

- Description of lead poisoning by an anonymous Roman hermit, translated by Humelbergius Secundus, 1829 (Lewis 1985)

1. Introduction

The long-run health effects of synthetic materials and industrial metals are seldom explored or tested before their introduction into civil environments. The introduction of such chemicals in the age of industrialization could aptly be characterized as an inadvertent, yet far-reaching grand experiment in public health. Few toxins illuminate the long term consequences of this experiment as much as lead. In the U.S., the cost of lead's proliferation in air, water, and infrastructure continues to be paid by firms through increased regulatory burden and by children, who suffer from a set of adverse health outcomes that range from permanent brain damage and IQ loss to premature death.

Though anecdotal evidence of lead's toxicity reaches back to antiquity, lead's beneficial properties—such as its malleability, resistance to corrosion, good insulation, and low cost—resulted in expansive use of the chemical in the 20th century (EPA 2019). According to public health researchers Markowitz and Rosner (2013), by the 1920's "virtually every item a toddler touched" contained lead. By the mid-20th century, lead use metastasized to tragic proportions; it is likely millions of children experienced lead poisoning's maladaptive effects, while tens of thousands died (Markowitz and Rosner 2013, Chisolm 2001).

The onset of regulations, which include banning lead from paint and phasing it out of gasoline, facilitated a substantial decline in lead poisoning rates beginning in the 1970's (Markowitz and Rosner 2013; Lanphear 2007). From a certain perspective, this reduction amounts to a tremendous achievement in public health. However, as lead poising rates declined, the public health literature uncovered an array of dangers associated with smaller levels of exposure (Lanphear 2007; NTP 2012; ATSDR 2019). The CDC now acknowledges that no level of lead is safe, demarcating the 97.5 percentile of children's blood-lead levels as threshold of

acceptable exposure (CDC 2017). This dissonance lead public health researcher Bruce Lanphear to label the declines in lead poisoning a “pyrrhic victory” (Lanphear 2007).

The U.S. EPA formally banned lead from paint in 1978. Therefore, a substantial percentage of the U.S. building stock still contains leaded-paint. Lead paint from these homes represents a potent threat. To illustrate, consider that despite the national headlines covering the lead water crisis in Flint, Michigan, data released by Michigan’s Department of Health and Human Services suggest that one zip code in Grand Rapids saw more cases of lead poisoning attributed to lead paint than all seven zip codes in Flint before, during, and after Flint’s crisis (Michigan Department of Health and Human Services (MDHHS) 2016; Schmidt 2018).

Renovations to older buildings risk disturbing this paint, generating paint dust that can contaminate the air as well as surrounding indoor and outdoor surfaces (EPA 2011; EPA 2019). In effort to mitigate this hazard, EPA enacted the Lead Renovation, Repair, and Painting Rule (LRRP), which took effect on April 22, 2010. The rule requires that renovators of older buildings undergo a training and certification process and utilize a set of best-practice cleaning and containment guidelines. In this paper, I utilize a state-by-year panel collected by the U.S. Centers for Disease Control (CDC) to evaluate whether the rule reduced the percent of children screened who have an elevated blood-lead level (EBLL). To evaluate LRRP’s impact, I exploit time-series variation in lead poisoning outcomes, building regression discontinuity model with time as the running variable and the onset of LRRP as the cut point. I find LRRP had a small impact, reducing mild cases of lead poisoning (Blood-lead level>10 µg/dL) by 0.87% and more severe cases of lead poisoning (Blood-lead level>25 µg/dL) by 0.06%. I also find a larger effect of LRRP in states without a lead abatement program. There was a 1.16% increase in mild cases in

states with a lead abatement program and a 0.12% increase in severe cases (see Table 5) when compared to states without one.

The remainder of this paper adheres to the following structure: Section 2 provides background information on the history of lead exposure, health outcomes linked with lead poisoning, the current regulatory picture, and regression discontinuity design; Section 3 reviews the current economic literature regarding lead poisoning; Section 4 describes the data used in this study; Section 5 explains the paper's empirical approach and explores limitations to causal inference; and Section 6 presents the model's results and discusses them.

2. Background

2.1 Lead Exposure Past and Present

Historically, the polity used lead in many products, but they primarily added it to paint as a pigment and to water pipes, fixtures, and solder (Dickman 2017). Industry mixed lead into gasoline in the 1920's to improve octane ratings (Dickman 2017). Today, however, individual's exposure to lead occurs primarily through paint chips and paint dust from residual paint still covering older homes, through water from old plumbing fixtures, and through soil from past car exhaust that entered the ground and never degraded. While exposure to airborne lead decreased significantly after the phase-out of leaded gasoline, airborne exposure remains possible in connection to emissions from lead smelters, battery manufacturing plants, and solid waste incinerators (EPA 2019). Lead hazards are substantially more dangerous to young children and fetuses, who still have developing nervous systems.

While lead exposure from water presents a unique danger due to lead's high absorption rate in water, according to the President's Task Force on Environmental Health Risks and Safety Risks to Children (2016), a downward trend in exposure from air, water, and food implies that,

for many children, the most likely exposure pathway is chips and dust from lead paint. Lead poisoning has been connected to Pica Disorder. In the mid-20th century, Pica was thought to be the primary source of lead exposure (Markowitz and Rosner 2013). However, typically children are exposed to lead dust or soil though typical hand-to-mouth activities in their own homes (EPA 2019). Children may even discover an affinity for the taste of lead paint (Fee 1990). Renovation activities where lead based paint is disturbed increases the risk of exposure if it increases the concentration of lead in the dust and soil of the renovated home (EPA 2019).

Lead paint is still present in a substantial portion of the U.S. housing stock, with paint manufactured prior to 1950 comprised by weight of up to 50% lead (Reissman et al. 2001). Cox et al. (2015) estimated that approximately 23.2 million homes contain lead-based paint hazards, 3.6 million of which house one or more children under six years of age. Figure 2 and Figure 3 in Section 5.1 present the change in the percent of children with an elevated blood-lead level (EBLL)—defined alternatively as above 10 μg/dL or 25 μg/dL—over the timeframe of the analysis.

Based on data from the National Health and Nutrition Examination Survey from release periods 1999–2002 through 2007–2010, there were an estimated 535,000 cases of lead poisoning (Wheeler and Brown 2013) [1]. U.S. Department of Housing and Urban Development (HUD) estimates that 15.5 million housing units have levels of dust-lead greater than EPA's standard (HUD 2012). Figure 1 presents the percentage of the U.S. Housing Stock that was built before 1980 (lead was banned in 1978). Any house built prior to 1978 may contain lead paint, but does not necessarily.

[1]NHANES estimates were extrapolated to the population based on 2010 U.S. Census figures.

Figure 1: Percent of U.S. Housing Stock Built Before 1980

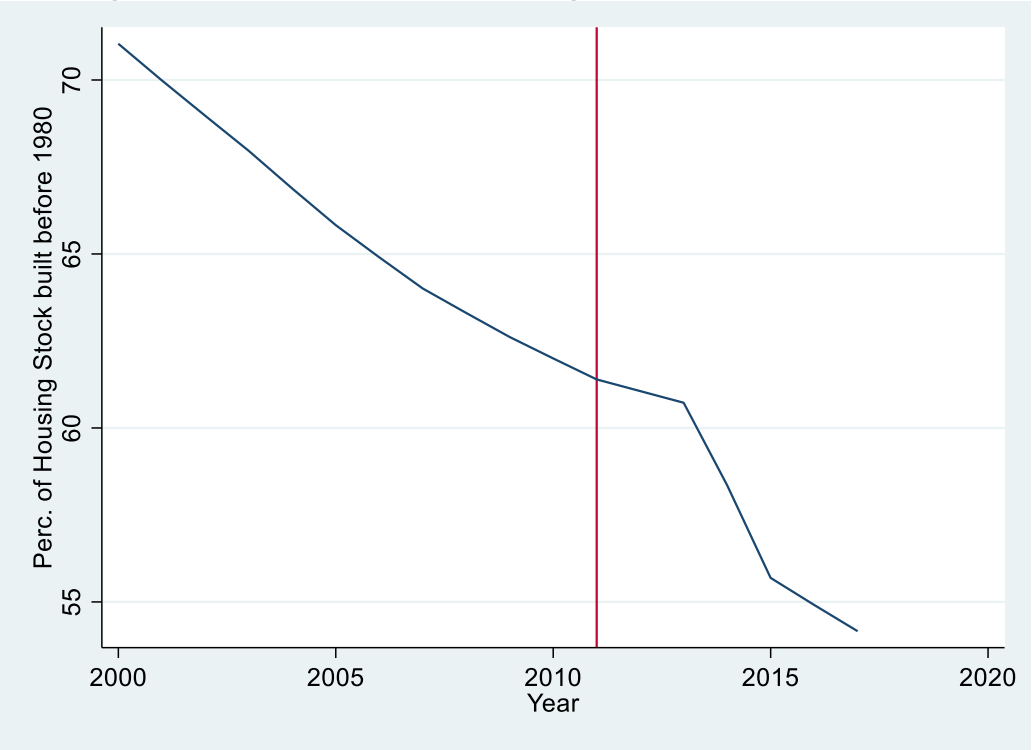


Figure 1 presents the percentage of the U.S. Housing Stock that was built before 1980 by year between 2000 and 2017. The U.S. prohibited lead-based paint in 1978; any house built prior to 1978 may contain lead paint, but does not necessarily.

2.2 Health Outcomes of Lead Exposure

At the highest levels of exposure, acute lead toxicity prompts symptoms such as abdominal pain, colic, vomiting, constipation, peripheral neuropathy, cerebral edema, and encephalopathy—leading to seizures, coma, and death (ATSDR 2019). However, as blood-lead levels in children began falling in the second half of the 20th century, awareness in the medical community of the maladaptive health effects common at even levels of exposure below 10 µg/dL grew (Lanphear 2007; NTP 2012; ATSDR 2019). CDC currently attests that "no safe" blood-lead level in children exists. As such, the CDC classifies blood-lead in children as elevated when it falls above the 97.5th percentile for the U.S. population (CDC 2017).

At low lead levels, the health literature on lead has catalogued neurological, renal, cardiovascular, hematological, immunological, reproductive, and developmental effects.

However, at low levels, neurological effects of lead have been best substantiated; these effects include reduction of cognitive function (learning and memory), altered behavior and mood (attention, hyperactivity, impulsivity, irritably, delinquency), and altered neuromotor and neurosensory function (visual-motor integration, dexterity, postural sway, changes in hearing and visual thresholds) (ATSDR 2019).

According to a 2012 National Toxicology Program Monograph, there is sufficient evidence of decreased cognitive performance in children with blood-lead levels below 5 μg/dL, measured by lower academic achievement, lower IQ, and decreased performance in other specific cognitive measures (NTP 2012). Neurological effects of lead poisoning are irreversible and treatment can only prevent further accumulation of lead (Rogan et al. 2001). Lanphear et al. (2005) estimated that, on average, an increase in blood-lead levels would lead to an IQ loss of 3.9 points (95% CI, 2.4–5.3), an increase of 10 to 20 μg/dL would lead to an incremental loss of 1.9 points (95% CI, 1.2–2.6), and an increase from 20 to 30 μg/dL would lead to an incremental loss of 1.1 points (95% CI, 0.7–1.5).

2.3 Regulatory Background

Amid mounting scientific evidence of lead's toxicity and proliferation across markets, along with an impassioned public health campaign, governments at all levels in the U.S. began regulating lead in the second half of the 20th century (Markowitz and Rosner 2013). In the late 1950's, some paint manufacturers voluntarily reduced the content of lead in paint to 1%. In February of 1978, the Consumer Product Safety Commission (CPSC) took formal action to eliminate lead paint. The U.S. also began phasing lead out of gasoline in the early 1970's; the EPA issued the first reduction standards in 1973, which would apply to 1975 car models (Newell and Rogers 2003; EPA 1996). By the early 1980's, lead in gasoline declined by 80% (Newell

and Rogers 2003). EPA completed the phase out in 1996, when they officially banned lead as an additive in gasoline for road vehicles.

Despite the large reduction of new lead due to the aforementioned regulations, lead hazards remain in the pre-1978 housing stock, in soil, and in water. Several U.S. federal government agencies share responsibility for eliminating lead hazards, including the EPA, HUD, and CPSC, as well as the Department of Agriculture, (USDA), the Department of Health and Human Services (HHS), the Department of Justice (DOJ), the Department of Transportation (DOT), the Department of Labor (DOL), and the Department of Education (ED) (EPA 2019).

EPA Lead-Based Paint Program

In 1992, the U.S. Congress passed Title X of the Housing and Community Development Act, also known as Residential Lead-Based Paint Hazard Reduction Act. Under Title X, Congress made the elimination of lead-based paint hazards a national goal, to be attained by mobilizing resources, building programs, and educating the public. Under Title X, EPA was tasked with establishing a training and certification program to address exposure to lead-paint from renovation and remodeling projects; ensuring disclosure of potential lead hazards from sellers to purchases, landlords to tenants, and renovators to occupiers in pre-1978 housing; identifying de minimis standard levels of lead in paint, dust, and soil; and providing information to the public regarding lead hazards (EPA 2019).

In an effort to fulfill these requirements, EPA promulgated the Lead Renovation, Repair, and Painting Program (LRRP) in 2008. It took effect in April 2010 and requires firms that perform renovation, repair, and painting work in residential homes and child occupied facilities (e.g. daycare centers, pediatric outpatient clinics) meet EPA certification requirements and use

lead-safe work practices when disturbing lead-based paint (U.S. Environmental Protection Agency (EPA) 2011). Specifically, renovators are required to:

1) Distribute EPA's lead pamphlet;
2) Ensure their firm is certified by the LRRP Program and that all workers are trained; and
3) Follow lead-safe work practices.[2]

Other Federal Regulations

HUD maintains the Lead Safe Housing Rule, originally promulgated in 1999. The rule only applies to housing sold by the federal government and housing that receives federal government financial assistance, but it grants HUD the regulatory authority to perform inspections, risk assessments, abatement, paint stabilization, or interim controls under certain circumstances (HUD 2012; EPA 2019).

The CDC supplies state, city, and county lead programs with grants and cooperative agreement funding for purposes of primary prevention, case management and screening, surveillance, strategic partnerships, and program evaluation (CDC 2019). The CDC also maintains a national surveillance system for children that records all reported blood-lead levels by state, city, and county lead programs. This thesis relies on data collected from this effort. In addition, as previously discussed, the CDC makes recommendations regarding the blood-lead level threshold for program intervention.

[2] EPA's Small Entity Compliance Guide contains a full list of lead-safe work practices. The mandated practices generally involve work area containment, work area cleaning, and the prohibition of specific paint removal practices that generate large amounts of lead-dust (EPA 2011).

State Regulations

Forty three states, including the District of Columbia, have at least one state-level policy regarding lead (NCLS 2017). There is considerable heterogeneity among these regulations; they cover lead testing, standards, penalties, abatement requirements, employee protections, and training for renovators, inspectors, and abatement professionals. Table 1 provides information on state-level testing policies, abatement policies, and renovation policies.

2.4 Regression Discontinuity Design Background

To estimate the effect of the LRRP rule, this paper adopts a regression discontinuity framework. While I relate the details of my specific empirical strategy in Section 5, this section serves to provide a general background on regression discontinuity design.

Fundamentally, regression discontinuity relies upon exploiting precise and arbitrary rules (Angrist and Pischke 2009). It involves a cutoff value of some continuous variable that acts as a threshold between the treatment and control groups. For instance, the seminal regression discontinuity example, Thistlewaite and Campbell (1960), estimates the effect of receiving a national merit scholarship on graduate school attendance rates. Under a simple OLS framework, this regression would be confounded by unobservables (i.e. “aptitude”). In their model, Thistlewaite and Campbell exploit the fact that eligibility for a national merit scholarship depends on receiving a PSAT score above a pre-determined cutoff score. They set PSAT scores as the running variable, while the threshold is the scholarship cutoff test score. The rationale behind RD design is that the students groups just above and just below the score threshold are similar enough that confounders are independent of the outcome variable within the neighborhood of scores near the threshold. In other words, if the students who scored just above

and just below the scholarship threshold on their PSAT's have the same aptitude as one another, a comparison between the mean outcomes of the two groups is unbiased.

Analytically, treatment status is a deterministic and discontinuous function of running variable x_i, such that:

$$D_i = \begin{cases} 1 \; if \; X_i \geq X_c \\ 0 \; if \; X_i < X_c \end{cases} \qquad \text{Equation 1}$$

where x_c represents the predetermined cutoff value of x_i. Incorporating D_i into a regression format, we obtain:

$$\begin{gathered} Y_i = \alpha + \beta X_i + \rho D_i + \varepsilon_i \\ Y_i = f(X_i) + \rho D_i + \varepsilon_i \end{gathered} \qquad \text{Equation 2}$$

Where Y_i represents the outcome variable and X_i represents the running variable. In the first equation, the functional form describing the relationship between Y_i and X_i is linear, while it is generalized in the second equation. Correctly specifying this functional form is particularly critical RD design. This becomes clear when presenting the model under the potential outcomes framework[3]:

$$\begin{gathered} E[Y_{0i}|X_i] = f(X_i) \\ E[Y_{1i}|X_i] = f(X_i) + \rho \end{gathered} \qquad \text{Equation 3}$$

Typically, under the potential outcomes framework, $E[Y_{0i}|X_i]$ signifies expected value of the outcome variable conditional upon X_i for the i^{th} unit, if the i^{th} unit not receiving the treatment. Conversely, $E[Y_{1i}|X_i]$ typically represents the same were the i^{th} unit to receive the treatment. The treatment effect ρ is the measure of the discontinuity in the running variable at the threshold:

$$\rho = \lim_{X_i \to X_c} E[Y_{1i}|X_i = X_c] - \lim_{X_c \leftarrow X_i} [Y_{0i}|X_i = X_c] \qquad \text{Equation 4}$$

[3] See (Rubin 1974; Rubin 1977)

Looking at Equation 3, if the functional form of $f(X_i)$ is improperly represented, there would be clear bias in ρ.

Lee and Lemieux (2010) contend that to more accurately capture the relationship between the running and outcome variable, the empiricist should employ a generalized model that includes two separate trend functions split at the cutoff x_c. We first allow the functional forms of the respective trend functions to be polynomial function of order *p*, such that:

$$E[Y_{0i}|X_i] = \alpha + \beta_{01}\tilde{X}_i + \cdots + \beta_{0p}\tilde{X}_i^p$$

$$E[Y_{1i}|X_i] = \alpha + \beta_{11}\tilde{X}_i + \cdots + \beta_{1p}\tilde{X}_i^p$$

These trend functions plug into the below regression, obtained by combining Equations X and Y:

$$E[Y_i|X_i] = E[Y_{0i}|X_i] + (E[Y_{1i}|X_i] - E[Y_{0i}|X_i])D_i$$

This yields the following model:

$$Y_i = \alpha + \beta_{01}\tilde{X}_i + \cdots + \beta_{0p}\tilde{X}_i^p + \rho D_i + \beta_1^*\tilde{X}_i + \cdots + \beta_p^*\tilde{X}_i^p + \varepsilon_i$$

Where $\beta_1^* = \beta_{11} - \beta_{01}$ and $\beta_p^* = \beta_{1p} - \beta_{0p}$. The treatment effect at the threshold X_c is ρ and the treatment effect at some other value $X_i - X_c = c > 0$ is equal to $\rho + \beta_1^* c + \cdots + \beta_p^* c^p$. Note also that the running variable threshold should be normalized to 0, so that $\tilde{X}_i = X_i - X_c$. This ensures that the coefficient ρ is the treatment effect at X_c when the time trends are specified through an interaction with D_i.

Angrist and Pischke (2009) emphasize that the validity of the RD estimate depends on whether the polynomial model accurately depicts $E[Y_{0i}|X_i]$. Otherwise, an apparent jump due to the treatment may instead be unaccounted for nonlinearity. To reduce the likelihood of this mistake, they suggest evaluating a non-parametric model—limiting the dataset to values near the threshold. However, this becomes difficult if there are not many observations close to the threshold. The choice to widen the bandwidth of data included in the model results in a tradeoff

between the statistical efficiency of the regression and the probability of bias from miss-specifying the controlling function.

Gelman and Imbens (2014) denote additional concerns, cautioning the empiricist against using high-order polynomial functions to control for the running variable. They argue that high-order polynomial functions may assign large weights to values away from the threshold, yield estimates highly sensitive to the polynomial function's order, and yield confidence intervals that inflate the likelihood of Type I error. Gelman and Imbens (2014) suggest using local linear or local quadratic specifications. In addition to limiting the sample of X around c, another strategy is to user a non-parametric smoothing function based on a kernel weighting scheme, which give more weight to observations closer to the threshold.[4] Hahn et al. (2001) developed local linear non-parametric regressions to this effect.

There are two key assumptions underpinning the RD framework. The first, as has been discussed, is that the chosen functional form meant to control for the running variable is properly assessed. The second is the continuity assumption—that the population average potential outcomes, $E[Y_{0i}|X_i]$ and $E[Y_{1i}|X_i]$, are continuous at the cutoff X_c. In other words, in a possible world where no units received the treatment, $E[Y_{0i}|X_i]$ would be a continuous function at the cutoff X_c, and vice versa if all units were to receive the treatment. This assumption would be violated if there is a discontinuity in another variable at the threshold or if there is sorting behavior (i.e. anticipatory effects from agents adjusting in advance of the treatment).

There is debate about the validity of causal claims based on regression discontinuity models. There is question as to how closely regression discontinuity design hedges to the

[4] Hahn et al. (2001) point out using a nonparametric rectangular kernel results in bias from systematically overestimating the magnitude of the discontinuity at the threshold. See Figure 38 in Hanh et al. 2001. Cunningham (2018) suggests using a triangular kernel instead.

randomized control trials, the experimental gold standard, versus other quasi-experimental frameworks. Whether the former or the latter is true depends as much on the specific RD model in question as anything else. Jacob et al. (2012) note two common lenses through which to interpret RD models. The first is based on the premise that any differences are random between observations on either side of a threshold, but within the neighborhood of the threshold; and if that is the case, any differences in mean between the two groups can be attributed to the threshold. The second is to focus on measuring the magnitude and direction of the discontinuity at the threshold. In a given experimental scenario, the appropriateness of each respective viewpoint depends on the quality of the data close to the threshold. If there is a large cross-section of data close to the threshold and sorting behavior is not possible, then it is easier to attribute any difference to the threshold. Hausman and Rapson (2017) contend that focusing on the discontinuity at the threshold is more appropriate in circumstances when time is the forcing variable.

3. Literature Review

Though researchers in the field of public health have written on lead extensively, in economics coverage has been more limited. Broadly speaking, the economic literature on lead falls into two categories: 1) analyzing economic and health outcomes associated with lead exposure and 2) analyzing economic outcomes associated with policy efforts to alleviate lead poisoning. The latter group, analysis of policy, focuses either on evaluating the unintended consequences of policy or on evaluating whether a policy accomplished its objective of abating lead poisoning and adverse economic outcomes associated with lead poisoning. This review focuses on recent lead economics papers in both categories in order to characterize where the literature currently stands.

The economic literature links lead exposure to a variety of detrimental economic outcomes. As such, it is reasonable to claim that any reduction in EBLL due to LRRP has resulted in a commensurate reduction in adverse outcomes. For example, Aizer et al. (2018) characterize the relationship between preschool blood-lead levels and third grade test scores in Rhode Island. They find a one-unit decrease in average blood-lead levels reduces the probability of being substantially below proficient in reading by 0.96 percentage points and in math by 0.79 percentage points. Clay (2018) examines the relationship connecting lead exposure from air and soil with fertility rates. They find reductions in airborne lead between 1978 and 1988 increased fertility rates and that higher lead in topsoil in the 2000's decreased fertility rates. Reyes (2007) and Reyes (2015a) examine the relationship connecting lead poisoning to crime and to risky and antisocial behavior among adolescents. Reyes finds that reduced childhood lead exposure in the late 1970's and early 1980's was responsible for a substantial reduction in violent crime in the 1990's.

I focus in more detail on papers that evaluate lead policy. As is true in other modes of environment and health policy, policy intended to target lead poisoning can have unintended and sometimes adverse effects on associated industries and markets. Billings (2017) estimate the benefits of lead-paint remediation on housing prices in Charlotte, NC. They combine data from LeadSafe Charlotte with the area's County Assessor Parcel Records. They utilize a difference in difference model that exploits variation between homes in the LeadSafe Charlotte Program where inspectors found there not to be a lead hazard with homes where lead was remediated. They find a reduction in residential turnover and that each dollar spent on lead remediation generated $2.60 in benefits. Gazze (2019) examined price and allocation effects on state-level lead abatement mandates using a hedonic pricing model. She finds that the mandates decrease

prices of old homes by 7.1% (this figure includes homes that both did and did not have an abatement renovation) and that families with children became 14.6% less likely to live in old homes.

Hyunhoe Bae has authored two papers evaluating the disclosure rule introduced in 1996 as part of the previously discussed Title X. The rule mandates sellers in the housing market disclose known lead-based paint hazards to buyers. Bae (2016) posits that the disclosure rule might result in a decline in the house prices of older home from reduced demand and therefore an increased occupancy of older homes by low-income residents. Bae finds that the policy did not lower prices of old homes. Bae (2012) examines the impact of the disclosure rule on home buyers and seller's behavior. Bae finds that the disclosure rule increased the probability of homebuyer lead testing and decreased the probability of the existence of peeling paint in old homes. Bae also found that the disclosure rule did not result in a substantial switch from old houses to new houses in any socioeconomic status groups, but that the policy reduced the number of older homes occupied by young children.

One of the fundamental questions driving of which this paper seeks to contribute is in regard to the relative efficacy of different types of lead policies. Two papers, and Gazze (2016), Reyes (2015b) contribute to this question through their respective evaluations of the impact of state-level abatement policies in Massachusetts and in the United States, respectively, on lead poisoning and economic and health outcomes associated with it. Another paper, Billings (2018) estimates the long-term impacts of early-life policy-based interventions (such interventions may include lead remediation, nutritional assessment, medical evaluation, developmental surveillance, and public assistance referrals) on educational performance and antisocial behavior. They use individual level data in Charlotte, NC and employ an OLS regression with covariates to

estimate the relationship between receiving an intervention and adverse outcomes. They find an intervention is associated with a 0.184 standard deviation decrease in antisocial behavior and a 0.117 increase in primary and middle school educational performance.

In 1971, Massachusetts implemented a lead program that requires the removal or covering of lead paint hazards in homes with children under 72 months and mandates the prevention, screening, diagnosis, and treatment of lead poisoning. Reyes (2015b) uses data from the Massachusetts Department of Public Health and the Massachusetts Department of Elementary and Secondary Education to evaluate the relationship between lead poisoning and test scores in Massachusetts and obtain insight into the effectiveness of Massachusetts policy. Reyes uses a fixed effects regression model, conditioned on covariates, on a panel dataset of cohorts of children who were third and fourth graders in the 2000's.[5] Reyes also develops a differences in differences framework that treats communities that experienced big declines in lead in the 1990s and uses as a control communities that experienced small declines. Reyes finds that the policy lowered the rate of children scoring unsatisfactory on standardized tests by 1–2 percentage points.

Gazze (2016) analyzes the effect of state-level abatement mandates on blood-lead levels and several other economic and health outcomes, including child disability, educational attainment, and infant health and fertility. Similarly to this thesis, she also utilizes the CDC National Lead Surveillance Data. Gazze provides evidence for the exogeneity of abatement mandate implementation, and uses a difference in differences approach that compares states with and without an abatement program. Gazze finds that abatement mandates reduce the EBLL rates

[5] Did you know: As a fourth grader in 2000 in the Massachusetts Public School System, I was probably an unwitting participant in this study?

by 29%. Gazze also finds that the mandates decrease the rate of enrollment in special education by 8.1%.

4. Data

To evaluate potential impacts of LRRP, I collected data from four publicly available sources across three organizations: the Centers for Disease Control's National Childhood Blood-lead Surveillance Data, the National Centers for Education Statistics DataLab, and the U.S. Census' Current Population Survey and American Housing Survey.

4.1 CDC's National Surveillance Data

The CDC's National Childhood Blood-lead Surveillance Dataset provides yearly state-level data on blood-lead levels in children under 72 months (CDC 2019). The data provides counts of the number of children in each state whose blood-lead falls within a given range. Table 2 and Table 3 provide summary statistics of each specific range in the dataset. The dataset also provides the total population of children under 72 months in each state as well as the number of children screened in CDC programs in each state per annum.

The CDC's national dataset integrates test results collected from state and local health departments. The CDC provides funding for state and local lead testing programs. In U.S. fiscal year 2014, CDC awarded $11 million to 29 States, DC, and 5 cities (Dickman 2017). State and local programs supported by these cooperative agreements must provide blood-lead testing data to the CDC, while other programs can provide data on a voluntary basis (CDC 2019). The CDC receives approximately 4 million blood-lead test results annually, which in turn are subjected to a cleaning and deduplication process to ensure only one test per child is counted in a given year.

The CDC recommends that states develop statewide blood-lead screening plans based on local data and conditions; it is the CDC's position that not all children are at risk for lead

exposure and thus all children do not need to be tested (CDC 2019). Some states that report to the CDC test children universally while other states maintain a targeted testing policy. There are also other states where testing is only recommended. Column 7 in Table 1 indicates which of these applies on a state-by-state basis. There is likely upwards bias in the percent of children with a high blood-lead level as the percent of the population tested shrinks and the testing becomes more targeted. In addition, children who receive Medicaid assistance are required to receive a blood-lead test at 12 months, 24 months, and between 36 and 72 months of age, yet many enrolled children do not receive the required screening (NCQA 2019).

In the summer of 2019, CDC reposted the national surveillance data to include lead testing results from the years 2016 and 2017. However, CDC did not repost data from 1997 to 2011—only data from 2012 to 2017 is publically available. Furthermore, there are minor inconsistencies between the two datasets in the years that overlap (2012-2015). I use the more recent data for those years. In some circumstances, data in the years 2012-2015 was reported in the earlier dataset but was not reported in the update. In these cases, I use the figures from the old dataset to substitute for the missing values.

4.2 Covariate Data

National Center for Education Statistics

I gathered data on the number of public school students in each state eligible for either a free or a reduced price per year from the National Center for Education Statistic's ELSI (Elementary and Secondary Information System) Table Generator (NCES 2019). Public school data from the ELSI generator draws from the Department of Education's Common Core of Data (CCD), NCES's primary database on public elementary and secondary education in the United

States. CCD surveys are conducted annually, collecting data from all public elementary and secondary schools as well as all local and state education agencies (NCES 2019).

Current Population Survey

The Current Population Survey (CPS) represents a combined effort from the U.S. Census Bureau and the Bureau of Labor Statistics (BLS). Conducted monthly, it is the primary source of labor statistics for the population of the United States (U.S. Census 2019b). The analysis collected data on the percent of the population universe[6] whose income is below the poverty line.

American Housing Survey

The American Housing Survey (AHS) is a nationally-representative survey of housing data conducted biennially by the Department of Housing and Urban Development and the U.S. Census Bureau (U.S. Census 2019a). It collects data on a variety of housing and occupant characteristics, including housing size, condition, financing, and occupant demographics, and consists of approximately 115,000 housing units. I collected national-level data on the percent of housing units built before 1980.[7] I computed linear midpoints to fill in missing data in the years AHS was not conducted.

5. Empirical Strategy and Results

5.1 Primary Identification Strategy

I built a regression discontinuity model to assess whether LRRP reduced lead poisoning in American children under 6 years old. The principle outcome variable is the percentage of children with an elevated blood-lead level (EBLL) in a given state-year pairing. Mirroring EPA's current and previous thresholds, I regress on two outcome variables. In the first, I define an

[6] The poverty universe does not include children under the age of 15 who are not related to a reference person within the household by way of birth, marriage or adoption (for example, foster children) (U.S. Census 2019b).
[7] AHS collects state-level data, but U.S. Census does not make the state indicator variable publically available.

elevated blood-lead level as above 10 μg/dL, and in the second as above 25 μg/dL. The running variable is time measured in years. The threshold is the year 2011, which represents the point in time at which LRRP took effect.[8] Following Hausman and Rapson (2017), as well as others[9] who have utilized a regression discontinuity in time framework, I adopt the following basic model:

$$Y_{st} = f_1(T) * 1(T \geq T_{2011}) + f_0(T) * 1(T < T_{2011}) + \tau T_{2011} + X'_{st}\beta + \gamma_s + \varepsilon_{st}$$

Equation 5

Where Y_{st} is the rate of children with an elevated blood level in state *s* and time *t*; $f_1(T) * 1(T \geq T_{2011})$ and $f_0(T) * 1(T < T_{2011})$ are two different time trend functions representing the time before and the time after LRRP took effect; T_{2011} is the variable of interest, a binary variable that indicates whether LRRP is in effect; X'_{st} is a vector of covariates; and γ_s and ε_{st} represent state fixed-effects and the error term, respectively.

There are two principle characteristics of the data that shaped model selection. First, as evidenced by Figure 2 and Figure 3, there is a downward nonlinear trend in EBLL rates overall across the time period of the dataset. Any empirical strategy that successfully evaluates the effect of LRRP needs to decompose any effect that could be attributed to LRRP from the variables—observed and unobserved—that are causing the downward trend in EBLL percentage. Second, EPA implemented the treatment—the onset of LRRP, uniformly and simultaneously at a national level. This means that there is not sufficient variation in the data I can exploit to identify the treatment effect cross-sectionally, ruling out differences-in-differences. If there were cross-sectionally variation in the treatment, downward trend would not be problematic provided that the treatment and control groups followed parallel downward trends prior to the onset of the

[8] LRRP took effect on April 22, 2010, meaning that a child with an elevated blood-lead level measured in 2010 may have been exposed before or after LRRP took effect. Because the data is annual, it is not know which. However, because there is likely some delay between when a child is exposed and when their blood is tested, I use 2011 as the discontinuity threshold.

[9] Other examples include Auffhammer and Kellogg (2011); Davis and Kahn (2010); Lang and Siler (2013).

treatment. The fact that there is not geographic heterogeneity in the implementation of LRRP also rules out a regression discontinuity based on borders.

Figure 2: Downward Trend in EBLL>10 µg/dL

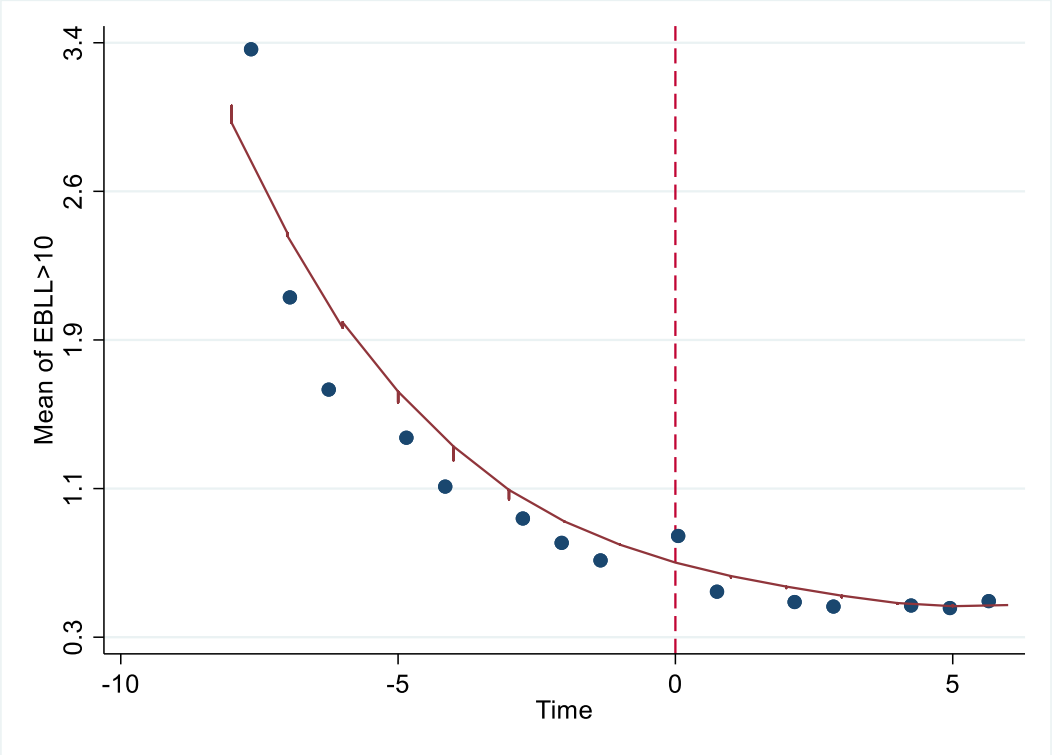


Figure 2 presents the change in the percentage of children with an elevated blood-lead level (EBLL), where the EBLL threshold is 10 µg/dL. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The graphed function was computed using locally weighted scatterplot smoothing.

Figure 3: Downward Trend in EBLL>25 µg/dL

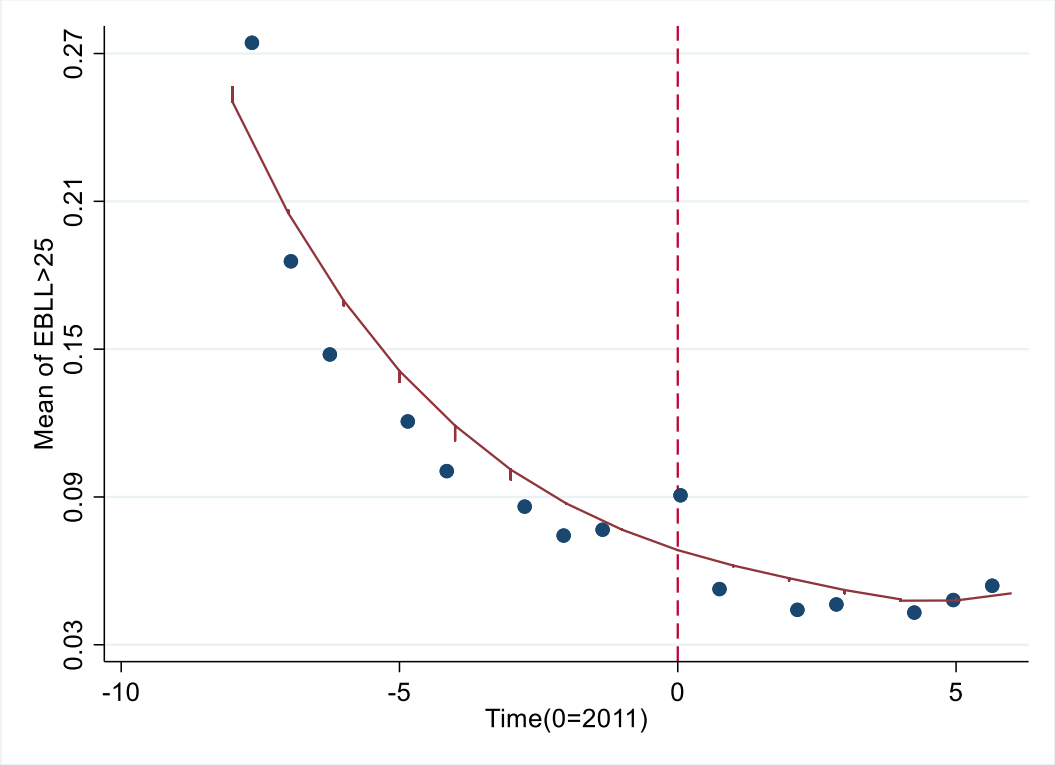


Figure 3 presents the change in the percentage of children with an elevated blood-lead level (EBLL), where the EBLL threshold is 25 µg/dL. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The graphed function was computed using locally weighted scatterplot smoothing.

5.2 Continuity Assumption

Deploying time as the forcing variable begets a number of empirical challenges. Some of them are common to customary cross-sectional regression discontinuity identification strategies while others common to with other empirical frameworks that rely on identification via time-series variation (Hausman and Rapson 2017). The first of these challenges is the continuity assumption. As detailed in Section 2.4, the continuity assumption states that the population average potential outcomes are continuous at the threshold. In other words, in a possible world where no units received the treatment, the conditional expectation function would be a continuous at the cutoff, and vice versa if all units were to receive the treatment. This assumption

may be violated for two reasons: if there is sorting behavior or if there is discontinuity in an unobserved variable at the threshold.

Sorting and Anticipatory Effects

Hausman and Rapson (2017) argue that sorting behavior is especially concerning in RDiT settings and can cause unwanted anticipatory effects. In policy scenarios, the effective date is often chosen by policymakers and announced in advance. This was certainly the case for LRRP, which was promulgated in 2008 but did not take effect until 2010. A substantial part of LRRP is the training certification requirement. If a contractor were to finish training prior to the onset of the rule, it is well within the realm of possibility they implemented the best practices required by LRRP before the onset of the rule. Cross-sectional regression discontinuities rely on the McCrary (2008) density test to evaluate whether there is sorting behavior. However, in RDiT the McCrary test is not an option because the density of the running variable (time) is uniform. Uncontrolled sorting behavior would reduce the magnitude of the discontinuity. In the case of LRRP, if contractors are following the rule's requirements on projects prior to the legal onset of the rule enough to result in a measurable difference in EBLL outcomes, it would compel me to understate the effect of LRRP. In order to control for a possible sorting, I test my model for dynamic completeness. I regress each of the outcome variables on time and add various numbers of leads and lags of LRRP and the outcome variable to the model. I find one significant lead when of LRRP is significant when EBLL>25 is the outcome variable. I thus include a one-period lead in my model when EBLL>25 is the outcome variable.

Covariate Selection

An RD analysis still meets the continuity assumption if other variables are impacting the outcome variable, as long as they do not change discontinuously at the threshold. If an RD model

is optimally designed—there are ample observations in the neighborhood of an exogenous threshold—that fact alone provides convincing evidence the model meets the conditional independence assumption. In this case, covariates may be included to reduce the standard error, but they are not required for the model to yield an unbiased estimator. Hausman and Rapson (2017) argue that in RDiT cases, there are nearly always discontinuous covariates at the threshold. Hausman and Rapson (2017) observe that most applications of RDiT fall in the field of environmental and energy economics, citing a hypothetical example of a regression where the outcome variable is power plant emissions. In these cases, simply comparing emissions the day before and day after a policy was implemented would suffer from omitted variables bias because it is unlikely emissions progress smoothly over the course of a given week. After controlling discontinuous observables (perhaps by including day-of-week or weekend effects in this example), the empiricist may then need to increase the bandwidth of the regression to attain enough statistical precision to minimize the possibility of Type I error. RDiT therefore requires any potential confounders to either be controlled for explicitly or be sufficiently captured in the time trend. This opens up the RDiT specification to omitted variables bias.

Hausman and Rapson (2017) note that RDiT studies often have "little or no" cross-sectional variation. Practitioners are forced to either increase the frequency of observations or expand the overall bandwidth of the sample. Distinct from those examples, as noted in Section 4.1, I have 28 cross-sectional observations per time period in my panel. However, my dataset lacks clarity around the threshold because the frequency of the data is yearly. LRRP took effect precisely on April 22, 2010. For any observation in 2010, whether a healthcare provided took a given elevated blood-lead measurement before or after the policy implementation date is obscured. Due to the persistence of lead in the blood, it is also a strong possibility that a sizable

portion of blood-lead measurements taken after the policy implementation date arose from renovation projects not yet subject to LRRP. To account for this source of bias, I include 2010 outcome observations in my control group and widen the bandwidth of my dataset to 6 years before and after LRRP implementation. The fact that the frequency of the data is so large implies a greater possibility for other discontinuities that affect the outcome. As such, I include covariates in my model.

Population EBLL's—both observed and unobserved—directly depend on children's exposure to lead. Since lead exposure is strongly correlated with poverty, the first covariate I investigate is the rate of students in a given state who receive a free or reduced price lunch (data from National Center for Education Statistics). This is my preferred poverty metric because it specifically targets children (though there is limited overlap in age since most children enrolled in school are older than 72 months). Figure 8 and Figure 9 graph changes in free and reduced price lunch over time. Due to the discontinuous increase in free or reduced price lunch rate evident from Figure 8, I include it in my model. I also examine the poverty rate (data from Current Population Survey) and find a potential small discontinuous increase (see Figure 10 and Figure 11). Therefore I also include it in my model. I next look at housing vintage (data from American Housing Survey), defined as the percent of the housing stock build before 1980. Lead was banned from paint at the federal level in 1978, this variable approximates for changes to the distribution of lead in the current housing stock. Housing vintage data was provided in decade ranges, therefore introducing a 1978 cutoff point was not possible. In Figure 12 and Figure 13, I find no clear discontinuity and therefore do not include it in my model. Lastly, I look at lead poisoning screening rates (data from Center for Disease Control) and find a discontinuous decrease. I therefore also include in my model (see Figure 14 and Figure 15).

I also include other covariates in my model. Based on my test for dynamic completeness, I also include two dependent variable lags in my EBLL>10 model and one dependent variable lag and one independent variable lead in my EBLL>25 model. Additionally, to investigate the possibility of a lagged treatment effect, I include one independent variable lag. To test for heterogeneity in the treatment effect, I also add an indicator for the presence of a state-level lead abatement program. Lastly, in order to control for missingness in the rate of free and reduced price lunch variable, I include a binary covariate indicating when it is missing.

5.3 Dataset Limitations

In addition to the imprecision around the policy date threshold I acknowledged in the previous sub-section, there are other empirical issues born from the particularities of the dataset. As detailed in Section 4.1, data collection of blood-lead is heterogeneous. State and local health departments collect data for their own localities, based both on stipulations outlined in their respective lead programs and cultural unobservables, then report that data to the CDC. The distinction between states with targeted and universal testing policies is particularly concerning. States that practice targeted testing report an artificially high rate of children with an elevated blood-lead level. To control for this source of bias, I include the screening rates as a covariate in my model. There are also likely other state-specific differences that might affect blood-lead levels such as housing and zoning policy, as well as economic indicators such as poverty and inequality.

Though many of the differences between states are likely continuous across the LRRP threshold, I am cognizant of the extensive heterogeneity between states as I make national treatment estimates. As such, I include state fixed effects in my model. Following Auffhammer and Kellogg (2011), Lang and Siler (2013), and Davis and Kahn (2010), I interact the state fixed

effects with the time trends, yielding discontinuity estimates for each state. I also include a version that weighs states based on the screening rates of children in each included state. See Section 5.7 for further discussion of the weighted fixed effects regression.

A second dataset-related concern is missingness, which is widespread in the CDC's Lead National Surveillance Dataset. Between the 50 U.S. States and the District of Columbia, 10 do not report to the CDC at all. Another 18 states have missing or incomplete data[10] for at least one year. Table 1 indicates which states were excluded from the analysis based on missingness or incompleteness. States missing an additional three years or more of during the timeframe of the collected data were not included. Given the sporadic nature of randomness across the analysis, it is most likely that missingness is independent of lead outcomes; therefore, states with two or fewer years missing are left as such and the panel was run unbalanced. I also assume incompleteness, as demarcated by the CDC, is random and not a result of intentionally targeted testing policies. The prevalence of incomplete reporting provides another reason to include screening rates as a covariate.

Measurement error is also of concern. Aizer et al. (2018) highlight how measurement error affects analyses based on blood-lead testing. Blood-lead is an imperfect measurement because lead will eventually leave the blood and pass into bones, hair, and other body organs. Additionally, health professionals often utilize capillary measurements (or finger pricks) because they are inexpensive and less painful. The CDC itself acknowledges that measurement error from capillary tests is relatively high, particularly at BLLs under 10 μg/dL (Aizer et al. 2018). Aizer et al. (2018) use data from the Rhode Island Health Department to instrument for measurement error by using an average blood-lead test result for each child. Without access to multiple test

[10] The extent of incompleteness for each recorded incident was not provided by the CDC.

results per child, I could not implement this instrument. In line with Aizer et al. 2018, I presume measurement error is random, which increases the variance of my dependent variable, weakening my results.

5.4 Functional Form Specification

As discussed in Section 2.4, the validity of an RD treatment estimate pivots upon whether the functional form of the running variable accurate characterizes the relationship between the running and outcome variable. Lee and Lemieux (2010) note that graphical presentation of an RD design is not only helpful and informative to the reader, but can help identify the proper functional form to specify. As such, I present a graph of EBLL rates over the timeframe on the analysis, with two time trends on each side of the threshold, mirroring the split trends in the RD regression.

Figure 4: Quadratic Fit of EBLL>10 with Discontinuity

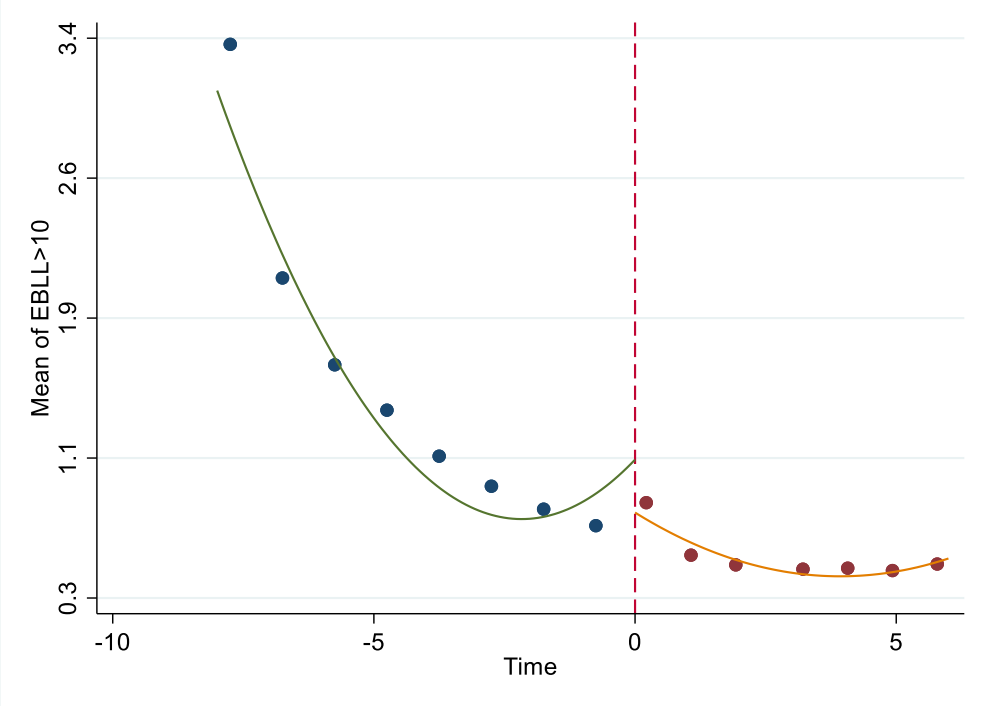


Figure 4 graphs the percentage of children with an elevated blood-lead level (EBLL), where the EBLL threshold is 10 µg/dL. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The two time trend functions represent a quadratic fit, distinct on each side of the threshold date.

Figure 5: Quadratic Fit of EBLL>25 with Discontinuity

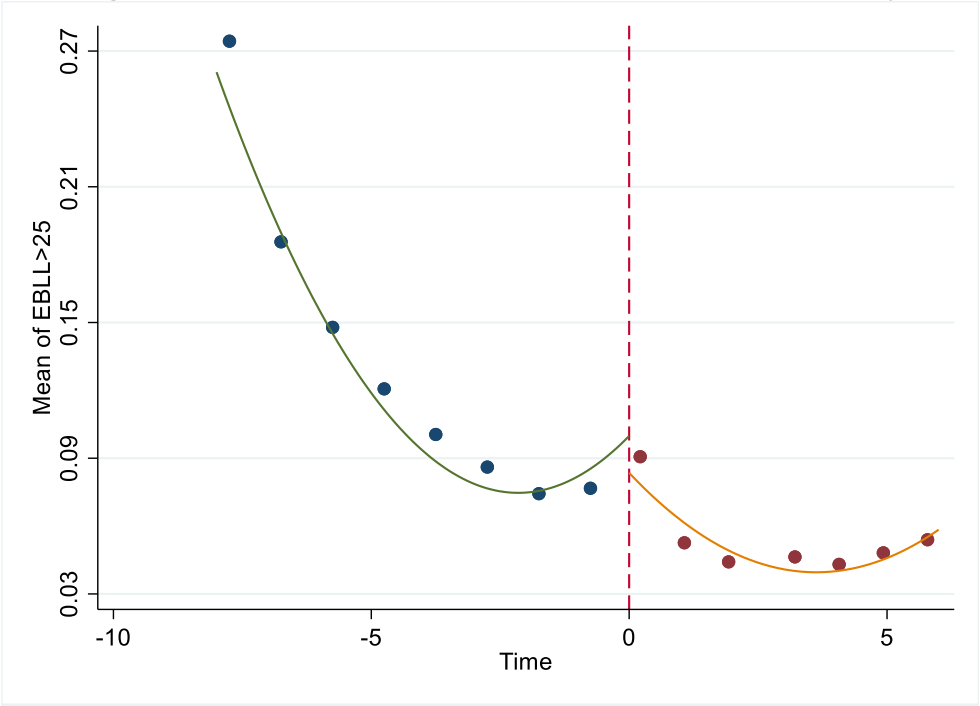


Figure 5 graphs the percentage of children with an elevated blood-lead level (EBLL), where the EBLL threshold is 25 µg/dL. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The two time trend functions represent a quadratic fit, distinct on each side of the threshold date.

Comparing these graphs to Figure 6 and This figure graphs the percentage of children with an elevated blood-lead level (EBLL), where the EBLL threshold is 10 µg/dL. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The time trend function represent a quadratic fit.

Figure 7 in the appendix, there is not strong visual evidence of a discontinuity between the year 2010 and 2011, when RRP took effect. However, there is visual evidence of a discontinuity between 2011 and 2012. Though the graphs above do not control for covariates, they somewhat reflect my empirical results, which finds small but statistically significant discontinuities between 2010 and 2011 (see Table 4) and between 2011 and 2012 (see Table 7). To evaluate the possibility the discontinuity between 2011 and 2012 results from a delayed effect of LRRP, I include a lagged independent variable in my model.

Based on these graphs, the EBLL rates time series is clearly nonlinear and downward trending, ruling out a linear specification. Following Gelman and Imbens (2014) (see Section 2.4), I do not present a trend with higher-order polynomial functions as my main result. Table 9

presents the treatment effect coefficients, standard errors, and p-values when I specify the model with higher polynomial orders, ranging from 3 to 5. My primary regression uses a quadratic polynomial function localized to a bandwidth of six periods before and after the threshold. Regressions with a smaller bandwidth included caused unwanted exclusions due to multicollinearity. In Table 8, I present variations of my regression with more localized specifications, with bandwidths ranging from 3 to 6. Results are consistent between regressions that include 4, 5, or 6 years of data.

Also of note is that, given the imprecision around the threshold cataloged in the *Continuity Assumption* section (Section 5.2), any specification of my model that is extremely localized is likely to be biased. This is also an issue with respect to the weighted nonparametric triangular kernel functional form, which is common in the literature, but places more weight on observations closer to the threshold. Furthermore, the kernel weighting scheme is not compatible with unit fixed effects in Stata. I include a regression with a nonparametric triangular kernel and no state fixed effects as a robustness check (see Table 10).

5.5 Other Empirical Concerns

In this section I address a series of other empirical concerns. This list includes time-varying treatment effects, serial correlation, clustering, heteroskedasticity, and exogeneity of assignment.

Time-Varying Treatment Effect

While the implementation of the treatment for all cross sectional unit occurs at the same time, it is possible that the impact of the treatment varies heterogeneously over time. Hausman and Rapson (2017) note that a time-varying treatment effect violates the assumption that the functional form of the running variable is properly specified. Under typical RD specification, the

empiricist assumes the treatment effect evolves smoothly through the window of the analysis. If the treatment effect tapers off prior to the conclusion of the window of the analysis, this would result in an overfitted model. The more observations there are away from the threshold, the higher the likelihood of overfitting. In addition to the discontinuity between 2011 and 2012, I also test for the possibility of continuities in other years as a robustness check. The results of this test are discussed further in the Results Section.

Serial Correlation

The temporal dimension of my panel introduces dynamic considerations that present numerous risks to inference. Serial correlation, or autocorrelation, is correlation between the dependent variable and lags of itself. This relationship is then caught by the error term, introducing bias into a dynamic model. The persistent nature of lead in the blood, combined with the fact that children with an elevated blood measurement in the previous period are more likely to have one the next if no one accounts for the source of their exposure, make the presence of serial correlation highly plausible. As such, I employ the Wooldridge test for autocorrelation in panel data and confirm the presence of serial correlation. As discussed in Section 5.2, I test for dynamic completeness in each outcome variable. To control for serial correlation, I include two periods of dependent variable lags of EBLL>10 and a one period lag for EBLL>25.

Clustering

Clustering is another risk associated with dynamics in panel data. Angrist and Pischke (2009) note that in models with state and year fixed effects, the error term represents idiosyncratic variations in potential outcomes across people, states, and time. They characterize the error terms as the sum of state-year shocks. Unobserved similarities between people within a state and across time begets clustering, which also introduces unwanted correlations into the

error term. Angrist and Pischke (2009) suggest building models that analyze multiple time period, many states, or both to mitigate the risk derived from such shocks. They also suggest clustering standard errors one level higher than the data. In a state-year example, akin to my dataset, they suggest clustering standard errors by state instead of by state and year. In my model, I also cluster standard errors at the state level.

Heteroskedasticity

Unbiased inference also depends on the assumption of homoskedasticity, that the variance of the conditional means estimated at each value of the x variable are constant across all values of the x variable. I test this assumption using modified Wald test for groupwise heteroskedasticity in a fixed effect regression model. I find that the assumption of homoskedasticity is violated. The clustered standard errors are also robust to account for heteroskedasticity.

Exogeneity of Treatment

Cunningham (2018) notes that the validity of regression discontinuity design does not require the assignment rule to be arbitrary, only that it be precise and free of manipulation. This standard could be compromised by uncontrolled sorting behavior (see Section 5.2), which would result in an understated treatment effect. It could also be threatened if there is not a clear chain of causation from the variable of interest to the outcome variable—in other words, the outcome variable determines the independent variable. This is especially a concern in public policy analyses, where a government entity promulgates some regulatory solution in response to a perceptible problem. Yet, it is unlikely that EPA enacted Lead Renovation, Repair, and Painting, which was passed in 2008 and took effect in April 2010, in direct response to the EBLL rates. As stated in Section 2.3, EPA enacted LRRP in fulfillment of its congressional mandate derived

from Title X. Congress passed Title X in 1992, nearly two decades prior to LRRP's enactment. The implementation of LRRP, in 2010 specifically, is therefore more likely a function of exogenous factors such as the EPA's pre-established rulemaking process than of EBLL trends. To test for reverse causality, Angrist and Pischke (2009) suggest utilizing a model specification designed in the spirit of Granger Causality. Referred to as the parallel trends test in differences-in-differences parlance, I add leads of the treatment to the regression to evaluate pre-treatment effects. Angrist and Pischke (2009) argue that if there is no reverse causality, that treatment causes the outcome, and not the inverse, then dummies for future policy changes will not matter. Results indicate that there are not statistically significant anticipatory effects. The results of this regression are displayed in the top panel of Table 12.

5.6 Heterogeneous Treatment Effect

In my model, I also evaluate whether the effect from LRRP's enactment itself differs between states that do and do not have a state-level lead abatement program. On one hand, states with a state-level lead abatement program are likely to have a more robust regulatory infrastructure around lead in place and a generally higher awareness of lead hazards, both in the overall population and among educators, healthcare, and social workers. Those states are likely to more fully adopt and execute LRRP practices, occupants are more likely to be aware of the risks to their children lead paint poses, and public officials are more likely enforce environmental regulations. One the other hand, because states with a lead abatement program have to some extent mitigated the preponderance of lead hazards in their localities, LRRP may have a larger effect in states without a program.

As highlighted in Section 2.3, 43 states plus the District of Columbia have at least one lead regulation (NCLS 2017). These rule range from establishing testing requirements to

occupational protections to direct bans. Many of these regulations are historic (e.g. years-old measures that ban incorporating lead into articles which have long since been lead-free). To evaluate treatment effect heterogeneity, I interact an indicator variable of whether a given state has a lead abatement program in a given year with the treatment variable. Lead abatement programs are a set of programs which loosely consist of blood-lead testing for a subset of the population, followed by a mandated investigation of the source of exposure if a test reveals an EBLL, from which various parties take action to mitigate the hazard. Gazze (2016) offers evidence that abatement regulations were exogenous at the time of their enactment.

Exogeneity of Heterogeneous Subgroups

To evaluate whether the presence of state-level abatement regulations alters the impact of LRRP, there must be proof that the two group's outcomes were on parallel trends prior to the introduction of LRRP. If this is not the case, it implies that the presence of a state-level abatement regulation is endogenous—it is not independent of measured outcomes. There are compelling reasons to suspect that the presence of a state-level abatement regulation is endogenous at the time of treatment. As covered in the literature review, Gazze (2016) finds that state-level abatement regulations mandates reduce the EBLL rates by 29%. All of the state-level abatement regulations Gazze analyzes were in place at the time LRRP took effect.

In effort to credibly argue for independence of state-level abatement regulations, I offer evidence corroborating the parallel trends assumption. First I conduct t-tests assessing the likelihood that the differential between the two subgroups group is not equivalent to zero for each year leading up to the onset of LRRP. In each year, I fail to reject the null—indicating that there is not a significant difference between the two subgroups in outcome prior to the treatment (see Table 11). I also conduct a parallel trends test common in the differences-in-differences

literature. I run a version of my model that includes additional leads of the interaction variable assigning the heterogeneous subgroups and the treatment (see Table 12, bottom panel). Results are discussed in Section 6.

5.7 Weighted Fixed Effects Regression

In effort to access the effect of LRRP more accurately, as an alternative specification I introduce a weighted fixed effects regression based on the screening rates of children in each included state. In a panel regression that contains unit fixed effects, the unit fixed effects themselves allow the intercept of the model to vary from unit to unit, allowing the empiricist to evaluate the treatment effect "within" each unit before assessing an overall effect of the treatment across units. As discussed in Section 5.3, different states have different populations and different testing policies. Overall, states who test more children should play a greater role in determining the overall cross-unit treatment effect. As such, I introduce a weighted fixed effects regression based on the screening rates of children in each included state.

These weights should not be understood as propensity scores or inverse probability weights—their function is not to more accurately align the observed sample with the overall population, but to influence the calculation of the average treatment effect such that it more heavily considers states that have tested more children. Weights for each state are homogenous over time. I utilize analytic weights in Stata both because the within-state effects used to calculate the average treatment effects are themselves means and because the analytic weights function adjusts standard errors such that larger weights denote more accurately measured observations. The weights were determined according to following function:

$$\frac{\sum_{t=2005}^{t=2017} children\_tested_{st}}{\sum_{s=1}^{s=j} \sum_{t=2005}^{t=2017} children\_tested_{st}}$$

The function represents the share of the total number of children screened over the lifespan of the panel in a given state above the share of the total number of children screened for all states. Subscripts s & t represent state and year. Results of the weighted regression are presented in Table 6 and discussed in Section 6. To compute accurate standard errors in the weighted regressions, I was not able to cluster standard errors in these regressions. These results may therefore be biased due to unwanted correlation.

6. Results and Discussion

There is some evidence LRRP reduced elevated blood-lead levels in children under 6 years old. My results show a 0.87% reduction in the short term of mild cases of lead poisoning and a 0.06% reduction in the short term of severe cases (see Table 4). LRRP was considerably more impactful in states without a lead abatement program; there was a 1.16% increase in mild cases in states with a lead abatement program and a 0.12% increase in severe cases (see Table 5) when compared to states without one. However, these results are not necessarily robust to empirical concerns outlined in Section 5. As such, I do not contend that my model captures a causal link between LRRP and EBLL. In the remainder of this section, I discuss my results in their empirical context and conclude by discussing implications for future policy and research.

Returning to Figure 4 and Figure 5, there is not compelling graphical evidence of a downward discontinuity during the year LRRP took effect. However, there is compelling graphical evidence of a discontinuity one year after. This may be a result of a lagged effect of the onset of LRRP. To test this notion, I included a one-period independent variable lag of LRRP. The lag was not significant. I also ran a placebo regression—the same in all respects as my primary model except it assumes the treatment took effect in 2012 instead of 2011. The results of this model show a statistically significant decrease in EBLL rates (-0.37% at EBLL>10 µg/dL

and -0.06% at EBLL>25 µg/dL). As a further robustness check, I ran placebo regressions that assumes LRRP took effect in 2009 and 2013—years where I did not expect there to be a discontinuity based on the graph. I found that the results were not statistically different than zero, implying consistency between the model and the graphs. The results of my placebo regressions are presented in Table 7.

One possible explanation for the mismatch between my model results and the graphical evidence is that the small but significant decrease in EBLL levels evidenced by my model manifests only once I control for other discontinuous variables. Another possibility is that there is some omitted variable my model is not capturing. There is considerable possibility of omitted variables bias in the model for other reason too. Returning to the discussion in Section 5.2, investigations into four potential covariates I collected data on (as seen in Figure 8 through Figure 15) show that there is likely a discontinuity in three of the four covariates, raising the possibility of discontinuities in unobserved covariates. As also discussed in Section 5.2, the fact that the time series data is yearly leaves a large gap in time in which some shock to an unobserved covariate could cause a discontinuity. Omitted variables bias is also a possibility because I use a more global model to avoid collinearity. I can therefore only invoke the conditional independence assumption if I assume the covariates and running variable function capture all unobserved heterogeneity affecting lead poisoning. As a robustness check, I include the results of the LRRP coefficient at different bandwidths (Table 8). The fact that results are similar between my model and models with a smaller bandwidth may partially mitigate this concern, but the coefficients of the models with a smaller bandwidth may not be reliable due to considerable multicollinearity.

My results are also not particularly robust to other functional form specifications. While Gelman and Imbens (2014) warn about the use of higher-order polynomials in regression discontinuity design, I ran a series of higher order polynomial models as a sensitivity check. In general, results are not robust to higher order polynomials. In Table 9, I include the coefficient, standard deviation, and p-value of the treatment effect estimated from higher order specifications. I also ran a series of regressions with a nonparametric triangular weighting scheme and find that my results are also not robust to those specifications. I present those results in Table 10. As discussed in Section 5.2, there is an increased change of bias in the kernel regressions because they more highly weigh observations close to the LRRP threshold, where my dataset lacks clarity. The kernel specifications might also be biased because they are not compatible with state fixed effects.

In Section 5.7, I highlighted concern regarding the "between" units estimation in the fixed effects framework. Results could be misleading if the model considers states with a large and small number of children screened equally as it estimates the overall effect of LRRP. I run a weighted fixed effects regression to counteract this concern. I find a statistically significant 0.90% decrease in the percent of children screened with EBLL>10. Results for EBLL>25 were not significant. I display the full results in Table 6.

To achieve an unbiased estimation of the effect of LRRP, the onset of LRRP itself must be exogenous. Similarly, to achieve an unbiased estimation of the heterogeneous treatment effect of LRRP comparing states with and without a lead abatement program, the state-level lead programs themselves must be exogenous with respect to lead poisoning outcomes. Given the regulatory history of LRRP, presented in Section 2.3 and again in Section 5.5, I think it is unlikely LRRP is endogenous. However to test this assumption I run a version of my model with

two additional LRRP pre-trends on the log of EBLL. These pre-trends, presented in the top panel of Table 12, do not yield significant results, providing additional evidence of LRRP's exogeneity. It is more likely that the state-level lead programs are endogenous, though I conduct a similar test to assess. I find that the pre-trends are also insignificant. In Figure 18 through Figure 21, I presents plots of the leads and lags from Table 12.

My results, if true, provide some evidence that LRRP is mildly effective, particularly in states that do not have a lead abatement program in place. But, due to the concerns discussed above, I refrain from making causal claims based on my results. However, there are alternative specification options that might yield a better understanding of both the effect of LRRP and the role of renovations in contributing to lead poisoning. Future researchers interested in evaluating the effect of LRRP on EBLL's or associated health or economic outcomes could build more compelling models with access to better data. For example, a researcher with full access to NHANES data could adopt a similar regression discontinuity in time approach. Public iterations of NHANES are released in two year cycles to protect the identity of participants. To evaluate LRRP using and RDiT framework using publically available data, the researcher would have to identify a treatment effect by comparing data from the 2007-2008 release period to data from the 2011-2012 release period. Data in the respective release periods would range from 3 to 5 years apart, and yielding what would surely be a biased and imprecise model. Another specification option would be to take advantage of private access data from the American Housing Survey, which has geo-tagged locations of renovations. If combined with a state's microdata on lead testing, one could identify not only whether proximity to a renovation project impacts the likelihood of having an elevated blood-lead level, but also whether that measure changed before and after the introduction of LRRP. Researchers interested in lead can also engage with other

aspects of the economic literature on lead. There is still a poor understanding of which types of lead policies most effectively reduce EBLL rates and while imposing minimal costs on associated industries.

In conclusion, lead is a toxic chemical that, crucially and cruelly, most negatively impacts children. It is a public health predicament caused by a lack of foresight from policy makers and industry leaders in previous generations to take action to prevent an obvious negative externality. This thesis provides limited evidence that retrospective action taken by the EPA has at least mitigated a small portion of the problem.

7. References

Agency for Toxic Substances and Disease Registry (ATSDR). (2019). Toxicological Profile for Lead.

Aizer, A., Currie, J., Simon, P. and Vivier, P. (2018). Lead Exposure and Racial Disparities in Test Scores. *American Economic Journal: Applied Economics, 10*(1), 307–341.

Angrist, J., and Pischke, J. (2009). *Mostly Harmless Econometrics: An Empiricist's Guide*. Princeton, NJ: Princeton University Press.

Auffhammer, M., and Kellogg, R. (2011). Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality. *American Economic Review, 101*(6), 2687-2722.

Bae, H. (2012). Reducing Environmental Risks by Information Disclosure: Evidence in Residential Lead Paint Disclosure Rule. *Journal of Policy Analysis and Management, 31*(2), 404-431.

Bae, H. (2016). The Impact of the Residential Lead Paint Disclosure Rule on House Prices: Findings in the American Housing Survey. *Journal of Housing and the Built Environment, 31*(1), 19-30.

Billings, S., and Schnepel, K. (2017). The Value of a Healthy Home: Lead Paint Remediation and Housing Values. *Journal of Public Economics, 153*, 69-81.

Billings, S., and Schnepel, K. (2018). Life after Lead: Effecs of Early Childhood Intervention on Lead Exposure. *American Economic Journal: Applied Economics, 10*(3), 315-344.

Centers for Disease Control and Prevention (CDC). (2017). What Do Parents Need to Know to Protect Their Children? Retrieved from https://www.cdc.gov/nceh/lead/acclpp/blood_lead_levels.htm

Chisolm, J. J. (2001). Evolution of the Management and Prevention of Childhood Lead Poisoning: Dependence of Advances in Public Health on Technological Advances in the Determination of Lead and Related Biochemical Indicators of Toxicity. *Environmental Research, 86*, 111-121.

Clay, K., Portnykh, M., and Severnini, E. (2018). *Toxic Truth: Lead and Fertility*. IZA Institute of Labor Economics. Retrieved from http://ftp.iza.org/dp11541.pdf

Cox, D. C., O'Haver, R., Salatino, B., Holmes, D., Pinzer, E. A., & Warren Friedman PhD, C. (2015). Prevalence of Lead Hazards and Soil Arsenic in US Housing. *Journal of environmental health, 78*(5), 22.

Cunningham, S. (2018). *Causal Inference: The Mixtape (V.1.7)*. tufte-latex.googlecode.com. Retrieved from https://www.scunning.com/cunningham_mixtape.pdf

Davis, L. W., and Kahn, M. E. (2010). International Trade in Used Vehicles: The Environmental Consequences of NAFTA. *American Economic Journal: Economic Policy, 2010*(4), 58-82.

Dickman, J. (2017). *Children at Risk: Gaps in State Lead Screening Policies*. Safer Chemicals, Healthy Families.

Fee, E. (1990). Public Health in Practice: An Early Confrontation with the "Silent Epidemic" of Childhood Lead Paint Poisoning. *Journal of the History of Medicine and Allied Sciences, 45*, 579-606.

Gazze, L. A. (2016). *Lead Policies, Lead Poisoning, and Government Spending*. Retrieved from http://home.uchicago.edu/lgazze/webfiles/LeadPolicies.pdf

Gazze, L. A. (2019). *The Price and Allocation Effects of Targeted Mandates: Evidence from Lead Hazards*. Retrieved from http://home.uchicago.edu/~lgazze/webfiles/LGazze_LeadHousing.pdf

Gelman, A., and Imbens, G. (2014). Why High-Order Polynomials Should Not be Used in Regression Discontinuity Designs. *National Bureau of Economic Research, 20405*.

Hahn, J., Todd, P., and Van der Klaauw, W. (2001). Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design. *Econometrica, 69*(1), 201-209.

Hausman, C., and Rapson, D. S. (2017). Regression Discontinuity in Time: Considerations for Empirical Applications. *National Bureau of Economic Research, 23602*.

Jacob, R., Zhu, P., Somers, M., and Bloom, H. (2012). A Practical Guide to Regression Discontinuity. *MDRC Working Paper*.

Lang, C., and Siler, M. (2013). Engineering Estimates versus Impact Evaluation of Energy Efficiency Projects: Regression Discontinuity Evidence from a Case Study. *Journal of Environmental Economics and Management, 32*, 360-370.

Lanphear, B. P. (2007). The Conquest of Lead Poisoning: A Pyrrhic Victory. *Environmental Health Perspectives, 15*(10), A484–A485.

Lanphear et al. (2005). Low-Level Environmental Lead Exposure and Children's Intellectual Function: An International Pooled Analysis. *Environmental Health Perspectives, 113*(7), 894–899.

Lee, D. S., and Lemieux, T. (2010). Regression Discontinuity Design in Economics. *Journal of Economic Literature, 48*, 281-355.

Lewis, J. (Producer). (1985). Lead Poisoning: A Historical Perspective. *EPA Journal*. Retrieved from https://archive.epa.gov/epa/aboutepa/lead-poisoning-historical-perspective.html

Markowitz and Rosner. (2013). *Lead Wars: The Politics of Science and the Fate of America's Children*. Berkeley and Los Angeles, California: University of California Press.

McCrary, J. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics, 142*, 698-714.

Michigan Department of Health and Human Services (MDHHS). (2016). *Data Report on Childhood Lead Testing and Elevated Levels: Michigan*. Retrieved from: https://www.michigan.gov/documents/lead/2016_CLPPP_Annual_Report_5-1-18_621989_7.pdf

National Center for Education Statistics (NCES). (2019). *ElSi Table Generator*. Retrieved from: https://nces.ed.gov/ccd/elsi/tableGenerator.aspx

National Committee for Quality Assurance (NCQA). (2019). Lead Screening in Children (LSC). Retrieved from https://www.ncqa.org/hedis/measures/lead-screening-in-children/

National Conference of State Legislatures (NCLS). (2017). Lead Hazards Project. Retrieved from http://www.ncsl.org/research/environment-and-natural-resources/lead-hazards-project.aspx

National Toxicology Program (NTP). (2012). *NTP Monograph. Health Effects of Low-Level Lead*. Retrieved from: https://ntp.niehs.nih.gov/ntp/ohat/lead/final/monographhealtheffectslowlevellead_newisstn_508.pdf

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

President's Task Force on Environmental Health Risks and Safety Risks to Children. (2016). *Key Federal Programs to Reduce Childhood Lead Exposures and Eliminate Associated*

Health Impacts. Retrieved from: https://ptfceh.niehs.nih.gov/features/assets/files/key_federal_programs_to_reduce_childhood_lead_exposures_and_eliminate_associated_health_impactspresidents_508.pdf

Reissman et al. (2001). Use of Geographic Information System Technology to aid Health Department Decision Making about Childhood Lead Poisoning Prevention Activities. *Environmental Health Perspectives, 109*(1), 89-94.

Reyes, J. W. (2007). Environmental Policy as Social Policy? The Impact of Childhood Lead Exposure on Crime. *BE Journal of Economic Analysis & Policy, 7*(1).

Reyes, J. W. (2015a). Lead Exposure and Behavior: Effects on Antisocial and Risky Behavior Among Children and Adolescents. *Economic Inquiry, 53*(3), 1580-1605.

Reyes, J. W. (2015b). Lead Policy and Academic Performance: Insights From Massachusetts. *Harvard Educational Review, 85*(1), 75-108.

Rogan et al. (2001). The Effect of Chelation Therapy with Succimer on Neuropsychological Development in Children Exposed to Lead. *New England Journal of Medicine, 344*(19), 1421-1426.

Rubin, D. B. (1974). Estimating the Causal Effects of Treatments in Randomized and Non-Randomized Studies. *Journal of Educational Psychology, 66*(688-701).

Rubin, D. B. (1977). Assignment to a Treatment Group on the Basis of a Covariate. *Journal of Educational Statistics, 2*(1-26).

Schmidt, C. (2018). America's Misguided War on Childhood Lead Exposures. *Undark*. Retrieved from https://undark.org/2018/03/21/lead-testing-child-blood-levels/

Thistlewaite, D., and Campbell, D. (1960). Regression-Discontinuity Analysis: An Alternative to the Ex-Post Facto Experiment. *Journal of Educational Psychology, 51*, 309-317.

U.S. Census. (2019a). *American Housing Survey (AHS)*. Retrieved from: https://www.census.gov/programs-surveys/ahs.html

U.S. Census. (2019b). *Current Population Survey (CPS)*. Retrieved from: https://www.census.gov/programs-surveys/cps.html

U.S. Centers for Disease Control (CDC). (2019). CDC National Childhood Blood Lead Surveillance Data. Retrieved from https://www.cdc.gov/nceh/lead/data/national.htm

U.S. Department of Housing and Urban Development (HUD). (2012). *Guidelines for the Evaluation and Control of Lead-Based Paint Hazards in Housing*. Retrieved from: https://www.hud.gov/program_offices/healthy_homes/lbp/hudguidelines

U.S. Environmental Protection Agency (EPA). (1996). EPA Takes Final Step in Phaseout of Leaded Gasoline. Retrieved from https://archive.epa.gov/epa/aboutepa/epa-takes-final-step-phaseout-leaded-gasoline.html

U.S. Environmental Protection Agency (EPA). (2011). Small Entity Compliance Guide to Renovate Right: EPA's Lead-Based Paint Renovation, Repair, and Painting Program: Office of Chemical Safety and Pollution Prevention.

U.S. Environmental Protection Agency (EPA). (2019). *Economic Analysis of the Final Rule to Revise the TSCA Dust-Lead Hazard Standards*. Retrieved from https://www.regulations.gov/document?D=EPA-HQ-OPPT-2018-0166-0573.

8. Appendix Tables and Figures

Table 1 (Appendix): States Reference Table

State	Reported Data to CDC National Surveillance Program	Data is Fully or Partially Missing or Incomplete	Data Included (Based on Criteria)	Subgroup Assignment	Abatement Policy	Required Testing Policy	Renovation Policy	Number of Other Policies
Alabama	Yes		Yes	Group 1 (No Abatement)		Recommended		2
Alaska		Yes				Recommended		
Arizona	Yes		Yes	Group 1 (No Abatement)		Recommended		1
Arkansas		Yes						2
California	Yes	Yes				Targeted	Yes	7
Colorado	Yes	Yes				Recommended		2
Connecticut	Yes		Yes	Group 2 (Abatement)	All Housing	Universal	Yes	5
Delaware	Yes		Yes	Group 1 (No Abatement)		Universal	Yes	2
DC	Yes	Yes	Yes	Group 2 (Abatement)	All Housing	Universal		2
Florida	Yes	Yes				Recommended		1
Georgia	Yes	Yes			Rentals Only	Recommended	Yes	1
Hawaii		Yes				Recommended		2
Idaho		Yes				Recommended		
Illinois	Yes		Yes	Group 2 (Abatement)	All Housing	Targeted		2
Indiana	Yes		Yes	Group 1 (No Abatement)		Recommended		3
Iowa	Yes	Yes				Universal		1
Kansas	Yes	Yes				Recommended		2
Kentucky	Yes		Yes	Group 2 (Abatement)	All Housing	Recommended		1
Louisiana	Yes	Yes	Yes	Group 2 (Abatement)	All Housing	Universal		2
Maine	Yes	Yes	Yes	Group 2 (Abatement)	All Housing	Targeted		2
Maryland	Yes		Yes	Group 2 (Abatement)	Rentals Only	Universal		3
Massachusetts	Yes		Yes	Group 2 (Abatement)	All Housing	Universal		1
Michigan	Yes		Yes	Group 2 (Abatement)	Rentals Only	Targeted	Yes	2

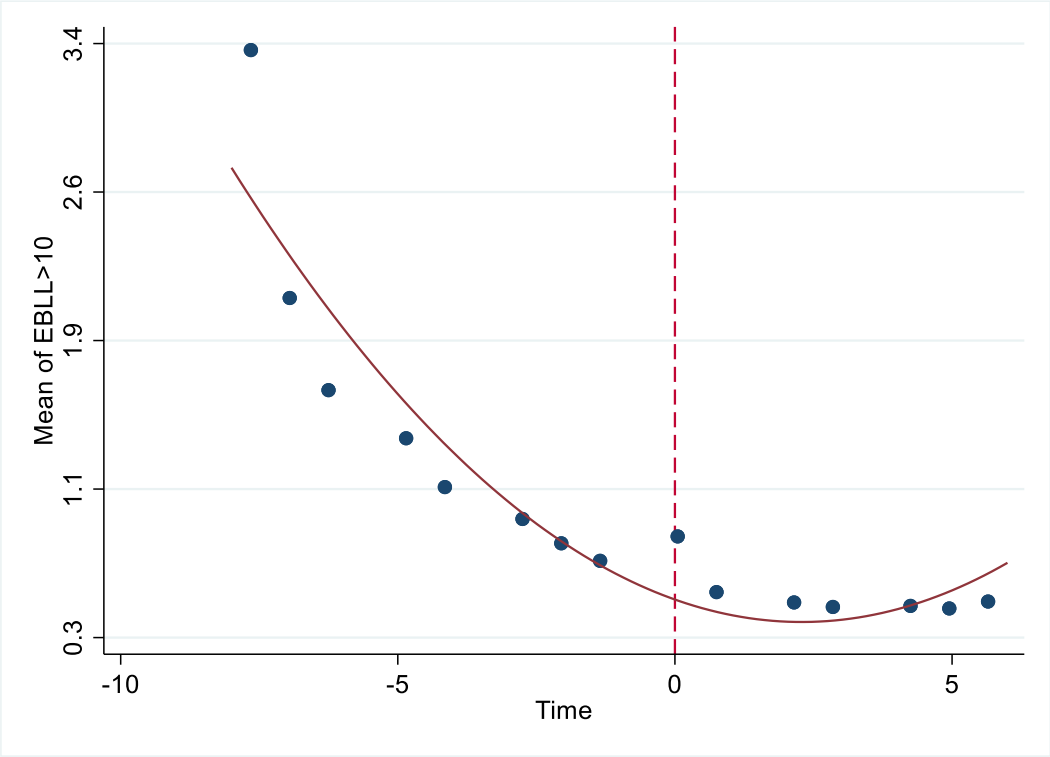
State	Reported Data to CDC National Surveillance Program	Data is Fully or Partially Missing or Incomplete	Data Included (Based on Criteria)	Subgroup Assignment	Abatement Policy	Required Testing Policy	Renovation Policy	Number of Other Policies
Minnesota	Yes		Yes	Group 2 (Abatement)	All Housing	Recommended	Yes	2
Mississippi	Yes		Yes	Group 1 (No Abatement)		Recommended	Yes	1
Missouri	Yes	Yes	Yes	Group 2 (Abatement)	All Housing	Targeted		1
Montana		Yes						
Nebraska		Yes				Recommended		1
Nevada	Yes	Yes				Recommended		
New Hampshire	Yes		Yes	Group 2 (Abatement)	Rentals Only	Recommended		1
New Jersey	Yes		Yes	Group 2 (Abatement)	All Housing	Universal		2
New Mexico	Yes	Yes				Recommended		
New York	Yes	Yes	Yes	Group 1 (No Abatement)		Universal		1
North Carolina	Yes		Yes	Group 2 (Abatement)	All Housing	Recommended		2
North Dakota		Yes						1
Ohio	Yes		Yes	Group 2 (Abatement)	All Housing	Targeted		1
Oklahoma	Yes		Yes	Group 1 (No Abatement)		Recommended		3
Oregon	Yes		Yes	Group 1 (No Abatement)		Recommended	Yes	1
Pennsylvania	Yes		Yes	Group 1 (No Abatement)		Recommended	Yes	1
Rhode Island	Yes		Yes	Group 2 (Abatement)	Rentals Only	Universal		2
South Carolina		Yes			All Housing	Recommended		1
South Dakota	Yes	Yes						
Tennessee	Yes	Yes	Yes	Group 1 (No Abatement)		Recommended		1
Texas	Yes	Yes				Recommended	Yes	2
Utah		Yes				Recommended	Yes	1
Vermont	Yes		Yes	Group 2 (Abatement)	Rentals Only	Universal		1
Virginia	Yes	Yes				Targeted		5
Washington	Yes	Yes	Yes	Group 1 (No Abatement)		Recommended	Yes	1

State	Reported Data to CDC National Surveillance Program	Data is Fully or Partially Missing or Incomplete	Data Included (Based on Criteria)	Subgroup Assignment	Abatement Policy	Required Testing Policy	Renovation Policy	Number of Other Policies
West Virginia	Yes		Yes	Group 1 (No Abatement)		Targeted	Yes	1
Wisconsin	Yes		Yes	Group 1 (No Abatement)		Recommended		1
Wyoming		Yes						

This table presents information both on the relative thoroughness of state's reporting to the CDC and on state's lead policies. The first column indicates whether a given state reports lead testing data to the CDC. The second column indicates whether a given state's data is fully or partially incomplete. The third column indicates whether a given state was included in the analysis. States with three or more years missing over the timeframe of the analysis were not included. The fourth column indicates whether a given state has a lead abatement program and hence which subgroup it is part of. The fifth column indicates whether a given state has a lead abatement program. The sixth column indicates whether in a given state blood-lead testing is universal (all children), targeted (only at-risk children), or only recommended. The seventh column indicates whether a given state has a lead policy concerning renovations. The eighth column indicates the total number of lead policies.

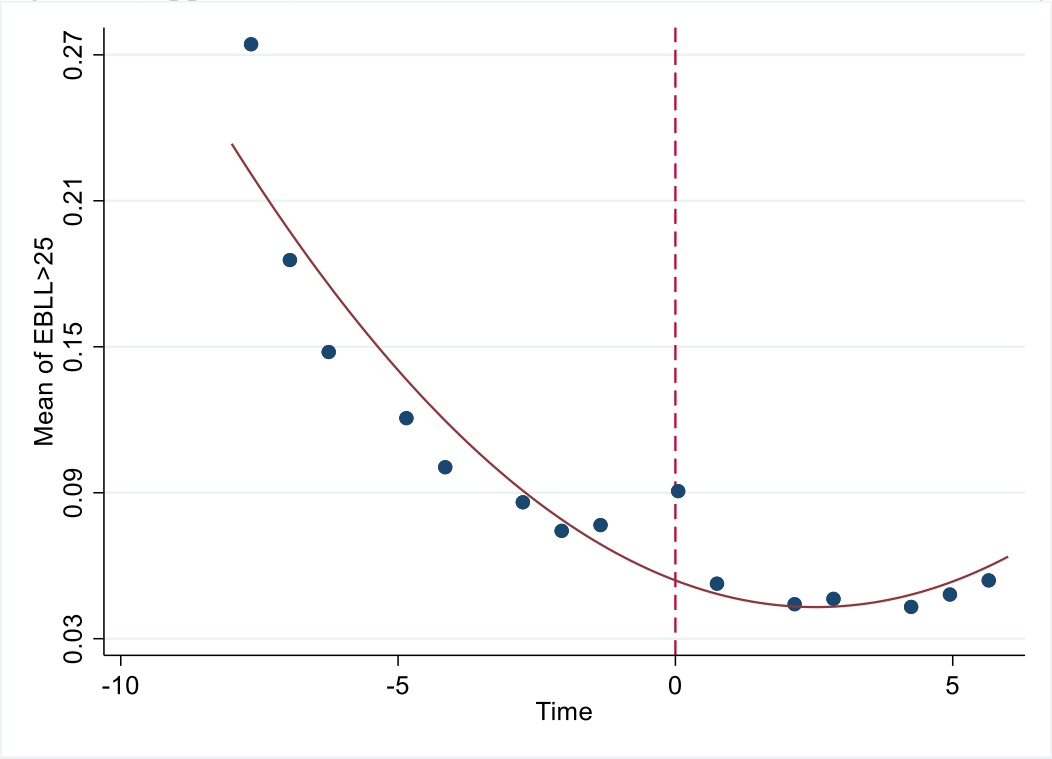
Sources: CDC 2019, Dickman 2017, NCLS 2017

Figure 6 (Appendix): Quadratic Fit of EBLL>10 without Discontinuity



This figure graphs the percentage of children with an elevated blood-lead level (EBLL), where the EBLL threshold is 10 µg/dL. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The time trend function represent a quadratic fit.

Figure 7 (Appendix): Quadratic Fit of EBLL>25 without Discontinuity



This figure graphs the percentage of children with an elevated blood-lead level (EBLL), where the EBLL threshold is 25 µg/dL. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The time trend function represent a quadratic fit.

Table 2 (Appendix): Summary Statistics of Control Group

Control Group	N	mean	sd	min	max
Blood-lead Variables					
Children Under 72 Mo.	178	472,077	404,289	38,156	2,104,904
Count Children Tested	176	87,356	103,782	4831	557,022
Percent EBLL > 10*	176	1.08%	0.82%	0.15%	4.98%
Percent EBLL > 25*	176	0.10%	0.08%	0.01%	0.53%
Count EBLL > 10	176	1,042	1,478	25	7,051
Count EBLL > 25	176	100	146	3	771
Count 10 < EBLL < 14	176	639	908	14	4,425
Count 15 < EBLL < 19	176	213	302	5	1,496
Count 20 < EBLL < 24	176	90	127	1	645
Count 25 < EBLL < 24	176	87	126	2	651
Count EBLL > 45	176	13	20	0	120
Covariates					
percent_tested*	176	18.53%	10.73%	0.96%	51.12%
percpoverty*	180	12.94%	3.58%	5.43%	23.11%
perclunch*	172	40.61%	11.76%	16.52%	70.03%
vintage	180	63.78%	1.31%	62.00%	65.83%

*These variables are included in the preferred model

Table 3 (Appendix): Summary Statistics of Treatment Group

Treatment Group	N	mean	sd	min	max
Blood-lead Variables					
Children Under 72 Mo.	176	457,866	379,384	36,857	2,051,411
Count Children Tested	175	84,561	85,750	1865	557,697
Percent EBLL > 10*	175	0.54%	0.55%	0.07%	6.43%
Percent EBLL > 25*	175	0.05%	0.07%	0.00%	0.75%
Count EBLL > 10	175	466	614	14	3,411
Count EBLL > 25	178	49	69	0	327
Count 10 < EBLL < 14	175	276	364	6	2,088
Count 15 < EBLL < 19	178	95	127	0	682
Count 20 < EBLL < 24	178	41	58	0	314
Count 25 < EBLL < 24	178	43	58	0	286
Count EBLL > 45	178	6	11	0	56
Covariates					
percent_tested*	175	19.18%	10.77%	0.50%	48.96%
percpoverty*	180	13.88%	3.61%	6.38%	23.05%
perclunch*	177	48.07%	11.54%	25.12%	99.17%
vintage	180	58.69%	2.60%	54.91%	61.40%

*These variables are included in the preferred model

Table 4 (Appendix): Model Results

VARIABLES	EBLL>10	EBLL>25
Treatment Effect		
LRRP	-0.872***	-0.0578***
	(0.148)	(0.013)
Dep. Var. Lags and Leads		
Dep Var Lag 1	-0.382***	-0.300**
	(0.091)	(0.142)
Dep Var Lag 2	0.0378**	
	(0.018)	
LRRP Lag 1	0.028	0.00365
	(0.047)	(0.0117)
LRRP Lead 1		0.0118
Covariates		
Abatement Regulation	-0.0833	-0.0452**
	(0.122)	(0.018)
Lead Screening Rate	-0.00394	-0.00102
	(0.008)	(0.001)
Poverty Rate	0.000575	-0.000327
	(0.009)	(0.002)
Percent Free or Reduc. Lunch	0.00699	0.000438
	(0.007)	(0.001)
Lunch Missing Indicator	0.296	0.0161
	(0.331)	(0.041)
State-Specific Effects		
Arizona	0.668***	0.0035
	(0.257)	(0.009)
Connecticut	1.112***	0.0653***
	(0.205)	(0.021)
Delaware	0.569***	-0.0137*
	(0.112)	(0.008)
District of Columbia	0.112	0.011
	(0.211)	(0.212)
Illinois	0.681***	0.0648***
	(0.130)	(0.011)

Indiana	0.845***	0.0585***
	(0.124)	(0.016)
Kentucky	0.680***	-0.159***
	(0.161)	(0.029)
Louisiana	6.592***	0.687***
	(0.197)	(0.026)
Maine	0.809***	0.0982***
	(0.158)	(0.016)
Maryland	0.930***	0.0650***
	(0.152)	(0.018)
Massachusetts	0.889***	0.0686***
	(0.172)	(0.016)
Michigan	0.731***	0.0283
	(0.190)	(0.026)
Minnesota	0.914***	0.0445***
	(0.157)	(0.014)
Mississippi	0.868***	0.0595**
	(0.171)	(0.026)
Missouri	0.275*	-0.0257
	(0.167)	(0.021)
New Hampshire	0.912***	0.0987***
	(0.154)	(0.020)
New Jersey	1.320***	0.0721***
	(0.179)	(0.021)
New York	0.741***	0.0403
	(0.216)	(0.038)
North Carolina	0.866***	0.0620***
	(0.123)	(0.013)
Ohio	1.193***	0.0775***
	(0.128)	(0.010)
Oklahoma	1.037***	0.0808***
	(0.165)	(0.019)
Oregon	0.898***	0.0374**
	(0.150)	(0.017)
Pennsylvania	1.267***	0.0988***
	(0.196)	(0.018)
Rhode Island	0.304***	0.114***
	(0.082)	(0.028)
Tennessee	-1.182	0.544**
	(1.499)	(0.233)
Vermont	1.161***	0.124***

	(0.176)	(0.024)
Washington	0.956***	0.100***
	(0.156)	(0.036)
West Virginia	0.773***	0.114***
	(0.127)	(0.021)
Wisconsin	0.799***	0.0610***
	(0.165)	(0.016)
Constant	1.331***	0.106*
	(0.443)	(0.055)
Observations	364	346
Number of stateid	30	30

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 5 (Appendix): Heterogeneous Treatment Results

VARIABLES	EBLL>10	EBLL>25
Treatment Effect		
LRRP	-0.872***	-0.0578***
	-0.148	-0.0126
Dep. Var. Lags and Leads		
Dep Var Lag 1	-0.382***	-0.300**
	(0.091)	(0.142)
Dep Var Lag 2	0.0378**	
	(0.018)	
LRRP Lag 1	-0.0282	0.00365
	(0.047)	-0.0117
LRRP Lead 1		0.0118
		(0.015)
Heterogeneous Effects		
Abatement Regulation	-0.0833	-0.0452**
	(0.122)	(0.018)
Subgroup Indicator	1.161***	0.124***
	(0.176)	(0.024)
Covariates		
Lead Screening Rate	-0.00394	-0.00102
	(0.008)	(0.001)
Poverty Rate	0.000575	-0.000327
	(0.009)	(0.002)
Percent Free or Reduc. Lunch	0.00699	0.000438
	(0.007)	(0.001)
Lunch Missing Indicator	0.296	0.0161
	(0.331)	(0.041)
Constant	1.331***	0.106*
	(0.443)	(0.055)
Observations	364	346
Number of stateid	30	30

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table 6 (Appendix): Weighted Fixed Effects Results

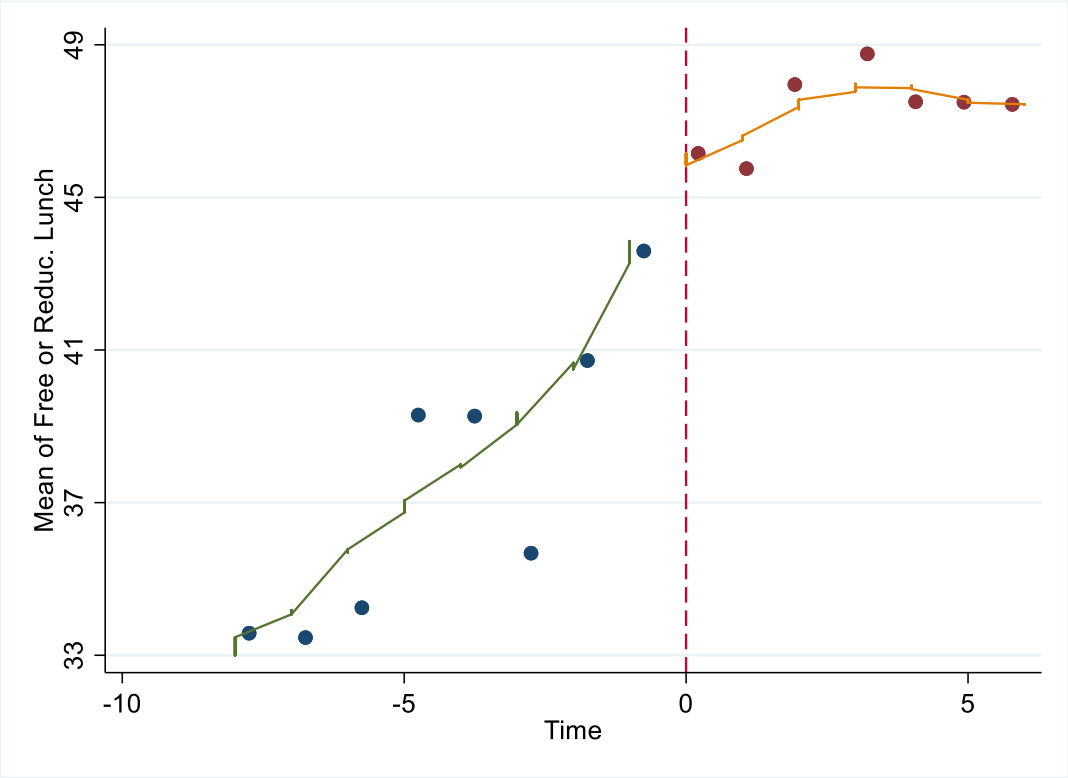
VARIABLES	EBLL>10	EBLL>25
Treatment Effect		
LRRP	-0.898**	-0.0609
	(0.352)	(0.055)
Dep. Var. Lags and Leads		
Dep Var Lag 1	-0.372***	-0.292***
	(0.067)	(0.062)
Dep Var Lag 2	0.0271**	
	(0.013)	
LRRP Lag 1	-0.0243	0.00803
	(0.037)	(0.007)
LRRP Lead 1		0.00325
		(0.007)
Covariates		
Abatement Regulation	-0.184	-0.0463
	(0.181)	(0.028)
Lead Screening Rate	0.00256	-0.000731
	(0.003)	(0.000)
Poverty Rate	-0.00131	0.000289
	(0.007)	(0.001)
Percent Free or Reduc. Lunch	0.0110***	0.000491
	(0.003)	(0.001)
Lunch Missing Indicator	0.513***	0.0283
	(0.125)	(0.020)
State-Specific Effects		
Arizona	0.603	0.00686
	(0.477)	(0.069)
Connecticut	1.182***	0.0647
	(0.409)	(0.064)
Delaware	0.576	-0.016
	(0.611)	(0.097)
District of Columbia	0.0895	0.0109
	(2.451)	(0.277)
Illinois	0.717*	0.0654
	(0.372)	(0.058)

Indiana	0.876**	0.0594
	(0.429)	(0.068)
Kentucky	0.702	-0.156*
	(0.546)	(0.086)
Louisiana	6.735***	0.694***
	(0.459)	(0.073)
Maine	0.817	0.0942
	(0.601)	(0.095)
Maryland	0.966**	0.066
	(0.393)	(0.062)
Massachusetts	0.941**	0.0674
	(0.371)	(0.058)
Michigan	0.858**	0.0315
	(0.390)	(0.061)
Minnesota	0.951**	0.044
	(0.401)	(0.063)
Mississippi	0.911**	0.0622
	(0.454)	(0.071)
Missouri	0.327	-0.027
	(0.398)	(0.063)
New Hampshire	0.946*	0.0959
	(0.553)	(0.088)
New Jersey	1.413***	0.0761
	(0.386)	(0.060)
New York	0.774**	0.0534
	(0.385)	(0.060)
North Carolina	0.891**	0.0614
	(0.377)	(0.059)
Ohio	1.160***	0.0714
	(0.373)	(0.059)
Oklahoma	1.080**	0.0822
	(0.473)	(0.074)
Oregon	0.925	0.0373
	(0.630)	(0.099)
Pennsylvania	1.293***	0.0971
	(0.393)	(0.060)
Rhode Island	0.316	0.112
	(0.452)	(0.073)
Tennessee	-1.943	0.602
	(3.614)	(0.573)
Vermont	1.196*	0.12

	(0.664)	(0.105)
Washington	0.987	0.0966
	(0.694)	(0.107)
West Virginia	0.804	0.113
	(0.589)	(0.093)
Wisconsin	0.833**	0.0572
	(0.396)	(0.062)
Constant	0.508	0.136
	(79528.000)	(29732.000)
Observations	364	346
R-Squared	0.988	0.97
Number of stateid	30	30

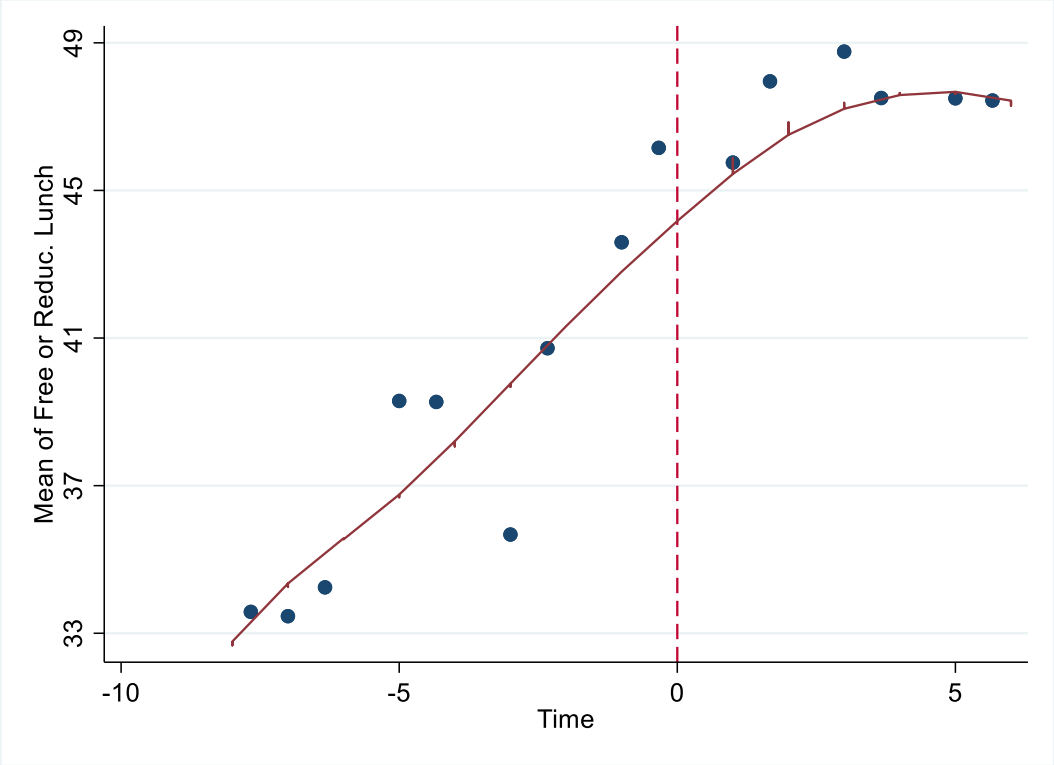
Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 8 (Appendix): Free or Reduced Price Lunch Rate with Discontinuity



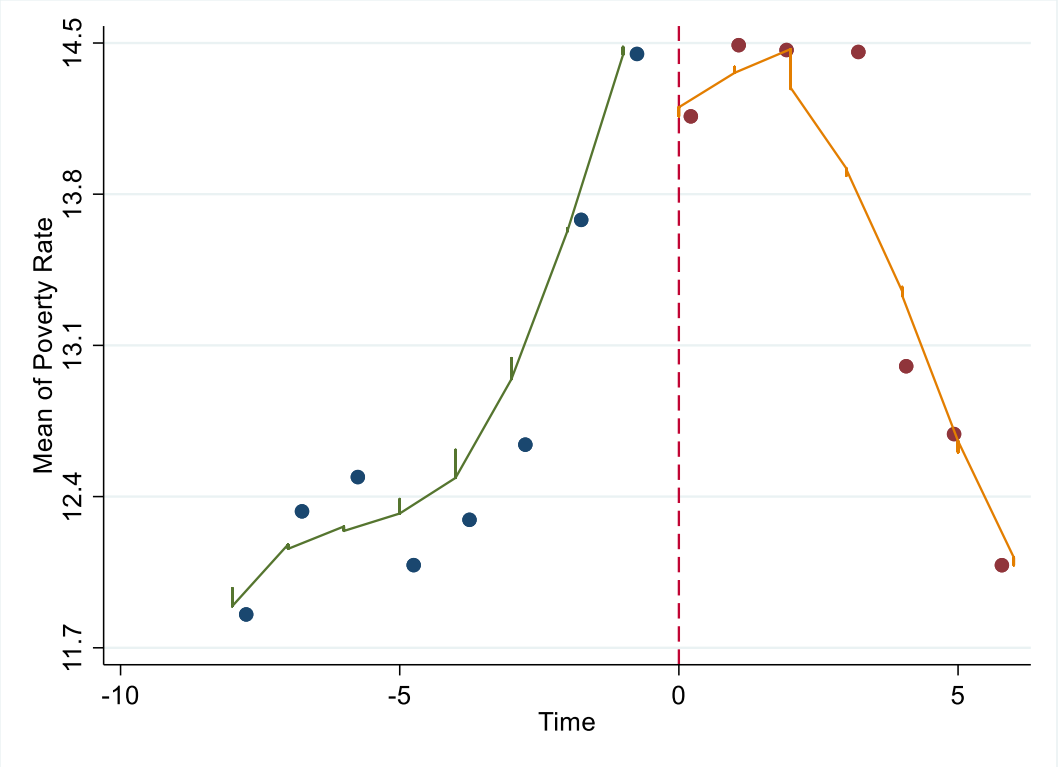
This figure presents the change in percent of children with a free or reduced price lunch. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The graphed functions were computed using locally weighted scatterplot smoothing.

Figure 9 (Appendix): Free or Reduced Price Lunch Rate without Discontinuity



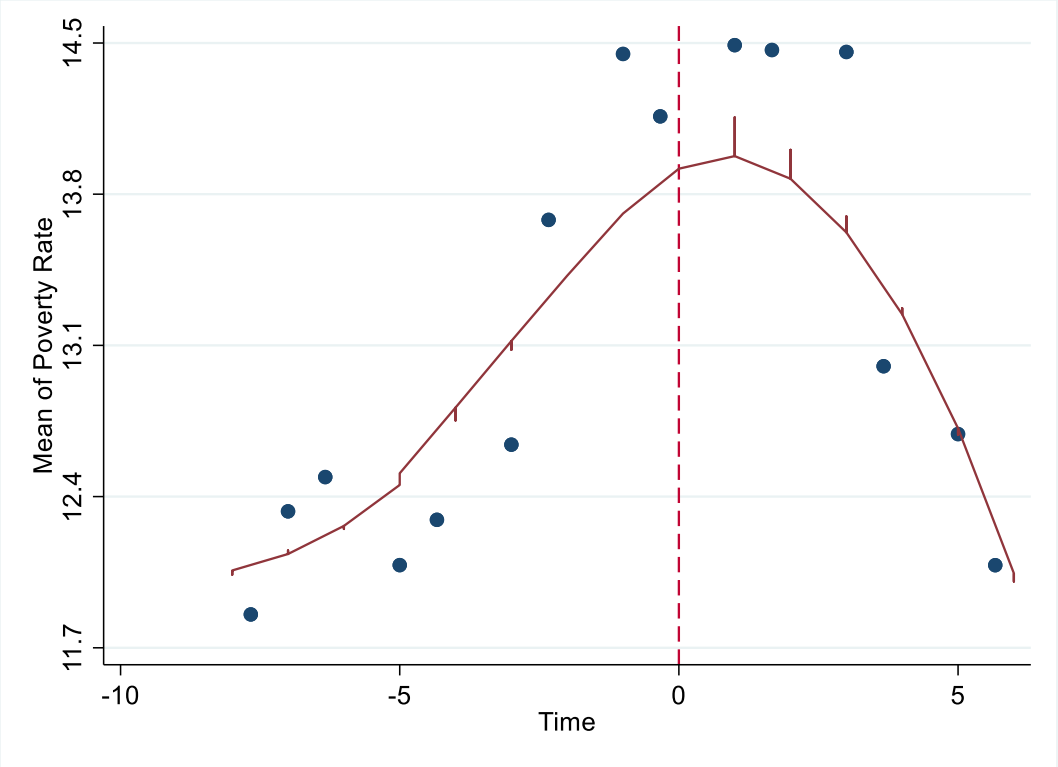
This figure presents the change in percent of children with a free or reduced price lunch. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The graphed function was computed using locally weighted scatterplot smoothing.

Figure 10 (Appendix): Poverty Rate without Discontinuity



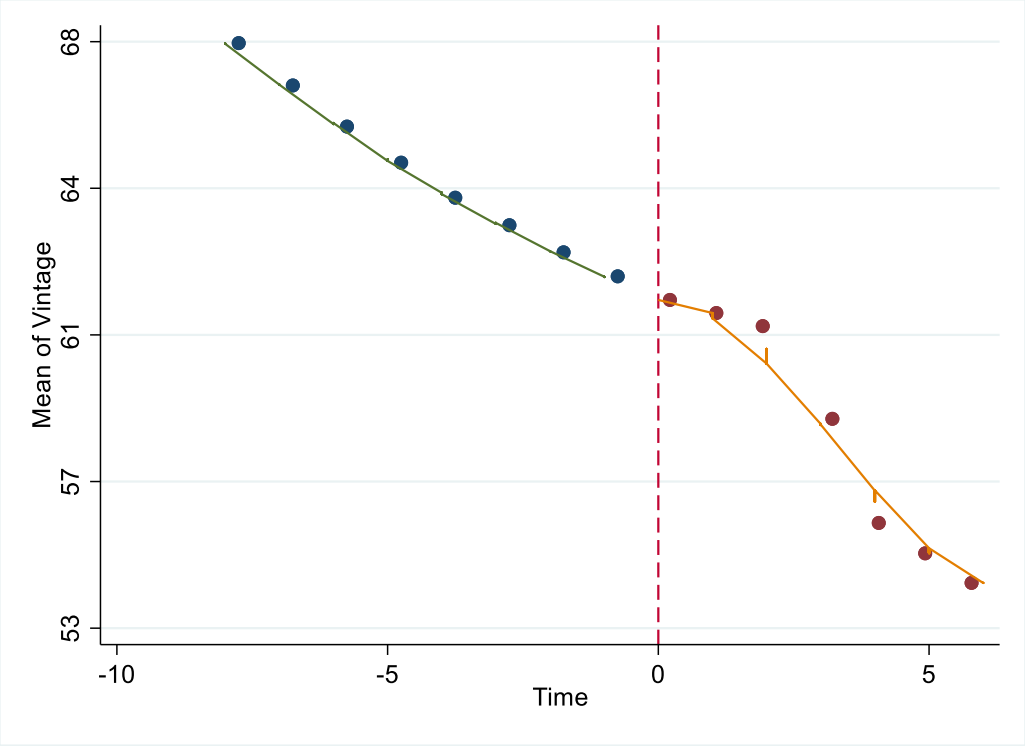
This figure presents the change in U.S poverty rate. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The graphed functions were computed using locally weighted scatterplot smoothing.

Figure 11 (Appendix): Poverty Rate without Discontinuity



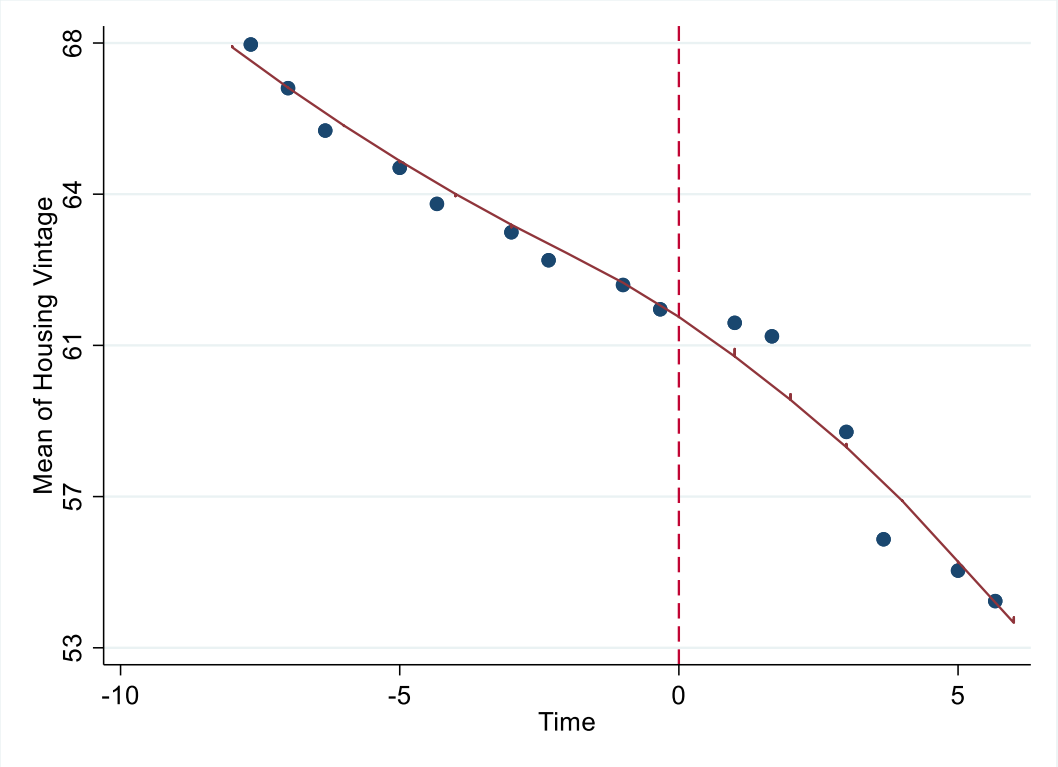
This figure presents the change in U.S poverty rate. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The graphed function was computed using locally weighted scatterplot smoothing.

Figure 12 (Appendix): Housing Vintage Trend with Discontinuity



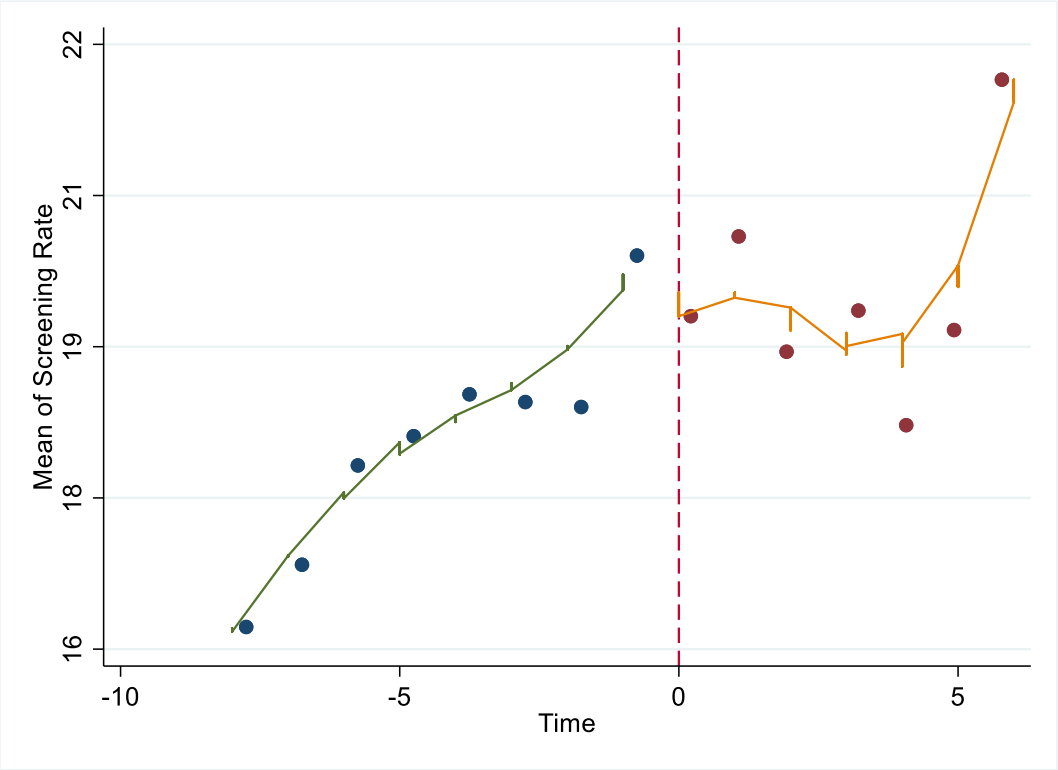
This figure presents the change in housing vintage, defined as the percent of U.S. housing that was built prior to 1980. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The graphed functions were computed using locally weighted scatterplot smoothing.

Figure 13 (Appendix): Housing Vintage Trend without Discontinuity



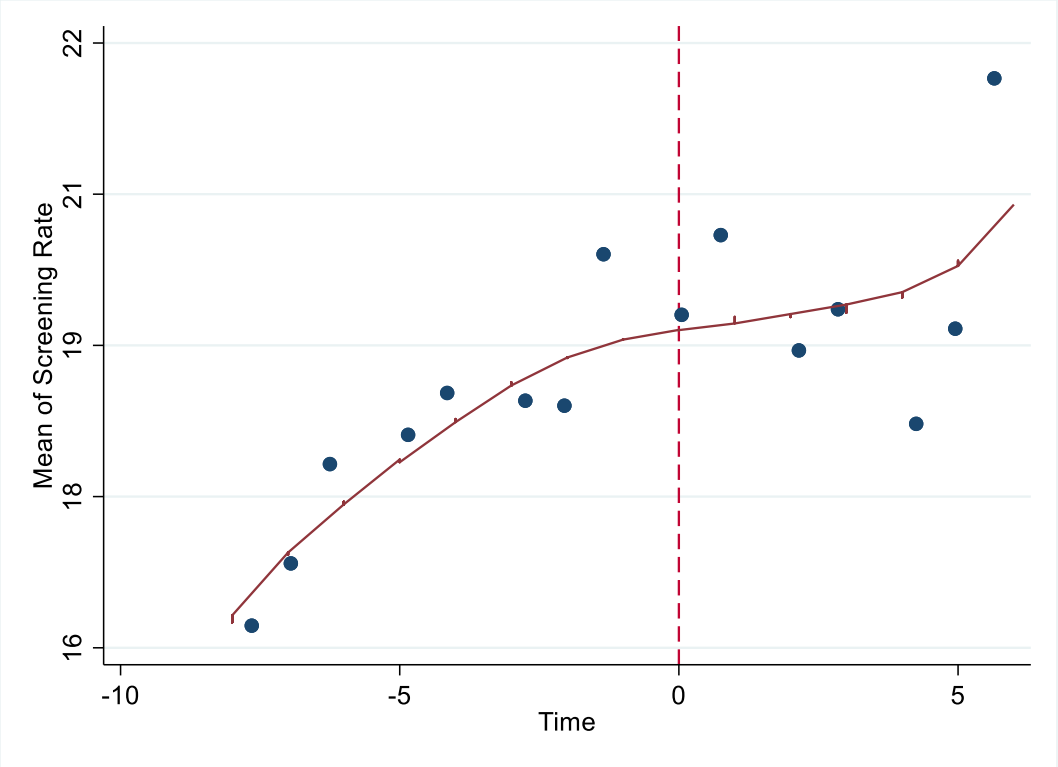
This figure presents the change in housing vintage, defined as the percent of U.S. housing that was built prior to 1980. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The graphed function was computed using locally weighted scatterplot smoothing.

Figure 14 (Appendix): Screening Rate Trend with Discontinuity



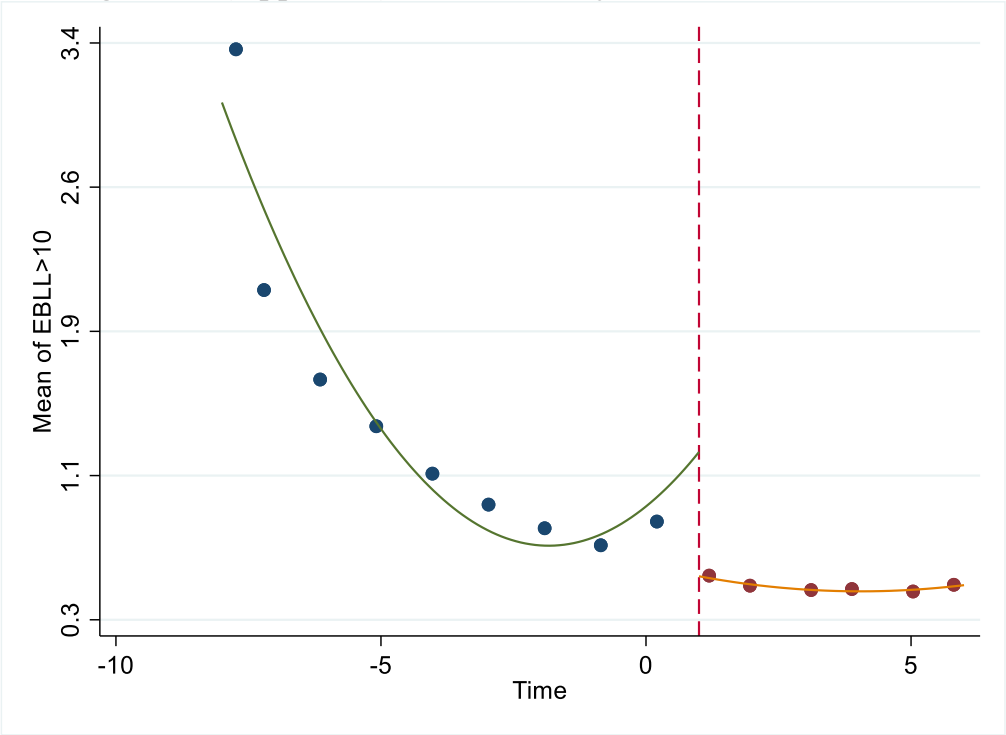
This figure presents the change in the percent of the child population who were tested for elevated blood-lead. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The graphed functions were computed using locally weighted scatterplot smoothing.

Figure 15 (Appendix): Screening Rate Trend without Discontinuity



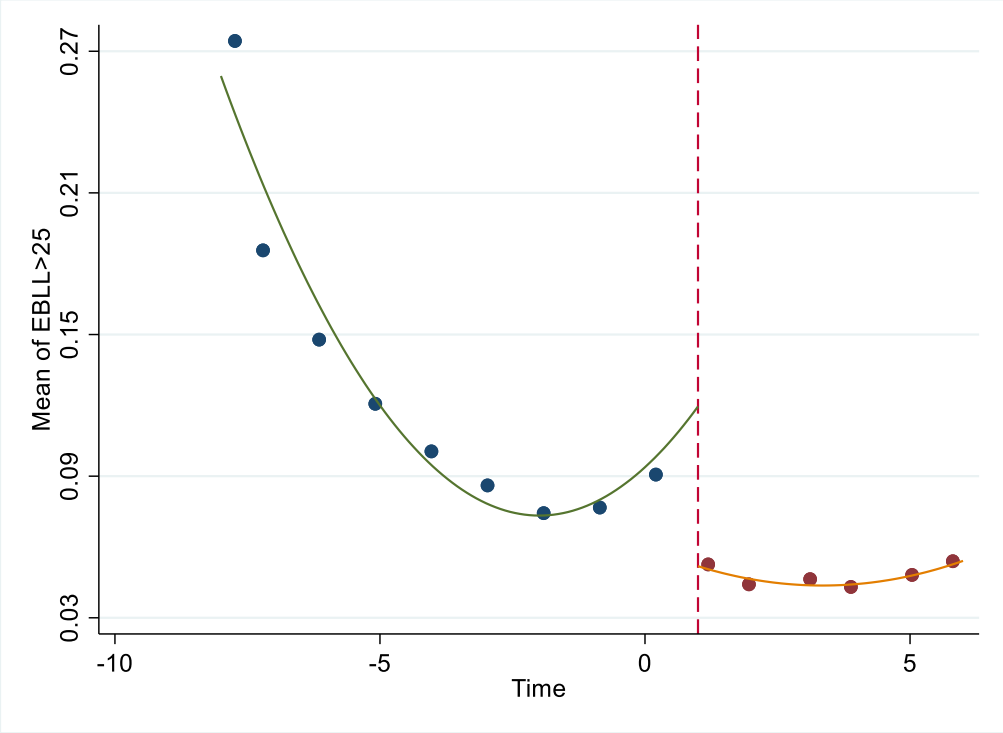
This figure presents the change in the percent of the child population who were tested for elevated blood-lead. Time is centered on the onset of LRRP, so T=0 represents the year 2011. The graphed function was computed using locally weighted scatterplot smoothing.

Figure 16 (Appendix): Discontinuity in 2012 of EBLL>10



This figure graphs the percentage of children with an elevated blood-lead level (EBLL), where the EBLL threshold is 10 µg/dL. Time is centered on the *year after* the onset of LRRP, so T=0 represents the year 2012. The two time trend functions represent a quadratic fit, distinct on each side of the threshold point.

Figure 17 (Appendix): Discontinuity in 2012 of EBLL>25



This figure graphs the percentage of children with an elevated blood-lead level (EBLL), where the EBLL threshold is 25 µg/dL. Time is centered on the *year after* the onset of LRRP, so T=0 represents the year 2012. The two time trend functions represent a quadratic fit, distinct on each side of the threshold point.

Table 7 (Appendix): Other Date Placebo Regression Treatment Effects

Year	Outcome Variable	Coefficient	Robust Standard Error	P-Value
2009	EBLL>10	0.097	0.195	0.620
2009	EBLL>10	-0.020	0.053	0.704
2013	EBLL>25	0.013	0.025	0.607
2013	EBLL>25	-0.321	0.215	0.135
2012	EBLL>10	-0.374	0.226	0.099
2012	EBLL>25	-0.063	0.021	0.002

This table presents the coefficients of LRRP in placebo regressions that represent the results in the counterfactual scenarios of LRRP being implemented in 2009, 2013, and 2012.

Table 8 (Appendix): Treatment Effects at Different Bandwidths

Bandwidth	Outcome Variable	Coefficient	Standard Error	P-Value
3*	EBLL>10	-0.087	0.484	0.858
3*	EBLL>25	0.007	0.004	0.044
4*	EBLL>10	-0.865	0.176	0.000
4*	EBLL>25	-0.050	0.026	0.057
5*	EBLL>10	-0.910	0.147	0.000
5*	EBLL>25	-0.057	0.012	0.000
6	EBLL>10	-0.929	0.107	0.000
6	EBLL>25	-0.066	0.010	0.000

*Some variables were omitted due to multicollinearity.

This table presents the coefficients of LRRP in regressions where the bandwidth (years included in the dataset when running the regression) was varied. The bandwidth represents the length of time both before and after the rule (e.g. a bandwidth of 3 implies 6 years of data was considered).

Table 9 (Appendix): Treatment Effects with Different Polynomial Function Orders

Polynomial Order	Outcome Variable	Coefficient	Standard Error	P-Value
3*	EBLL>10	-0.088	0.0652	0.177
3*	EBLL>25	-0.088	0.065	0.177
4*	EBLL>10	0.723	0.4512	0.109
4*	EBLL>25	0.000	0.236	1.000
5*	EBLL>10	-0.680	6.7135	0.919
5	EBLL>25	Insufficient Observations		

*Some variables were omitted due to multicollinearity

This table presents the coefficients of LRRP in regressions where the order of the time trend function polynomials were varied.

Table 10 (Appendix): Treatment Effects from Nonparametric Triangular Kernel Regression

Bandwidth	Outcome Variable	Coefficient	Standard Error	P-Value
4	EBLL>10	0.177	0.433	0.682
4	EBLL>25	-0.003	0.050	0.951
5	EBLL>10	0.166	0.320	0.604
5	EBLL>25	0.002	0.037	0.951
6	EBLL>10	0.133	0.274	0.628
6	EBLL>25	0.004	0.032	0.894

This table presents the coefficients of LRRP in regressions where time trends each take the functional form of a nonparametric triangular kernel. The bandwidth (years included in the dataset when running the regression) of the nonparametric regressions were also varied. The bandwidth represents the length of time both before and after the rule (e.g. a bandwidth of 3 implies 6 years of data was considered).

Table 11 (Appendix): Two-Sample T-Tests of Subgroup Outcomes

Year	EBLL>10	EBLL>25
2010	0.4336	0.3985
2009	0.7495	0.9843
2008	0.3858	0.8605
2007	0.3895	0.6683
2006	0.3842	0.5692

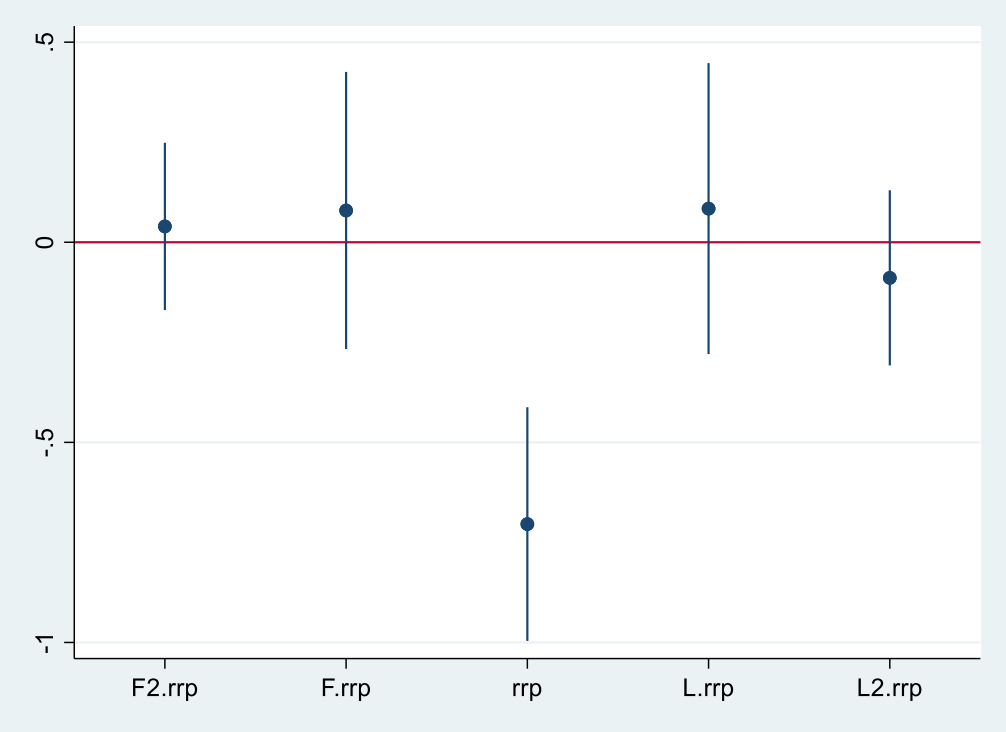
Two-Sample T-Tests were run that compare whether the EBLL rates in each heterogeneous subgroup are significantly different from each other in the years preceding the onset of LRRP. The figures in the body of the table represent p-values of these hypothesis tests.

Table 12 (Appendix): Parallel Trends Test Regressions Lags and Leads Results

VARIABLES	Log EBLL>10	Log EBLL>25
LRRP Lags and Leads		
LRRP	-0.704***	-0.773
	(0.149)	(0.514)
LRRP Lag 1	0.084	0.139
	(0.19)	(0.40)
LRRP Lag 2	-0.089	-0.011
	(0.112)	(0.293)
LRRP Lead 1	0.079	0.144
	(0.177)	(0.428)
LRRP Lead 2	0.040	-0.072
	(0.107)	(0.266)
Subgroup Interaction Lags and Leads		
Interaction	1.320***	2.113**
	(0.228)	(0.836)
Interaction Lag 1	0.004	0.115
	(0.216)	(0.583)
Interaction Lag 2	-0.097	0.080
	(0.133)	(0.424)
Interaction Lead 1	0.145	0.364
	(0.188)	(0.798)
Interaction Lead 2	0.028	0.078
	(0.090)	(0.582)
Observations	314	291
Number of stateid	30	30

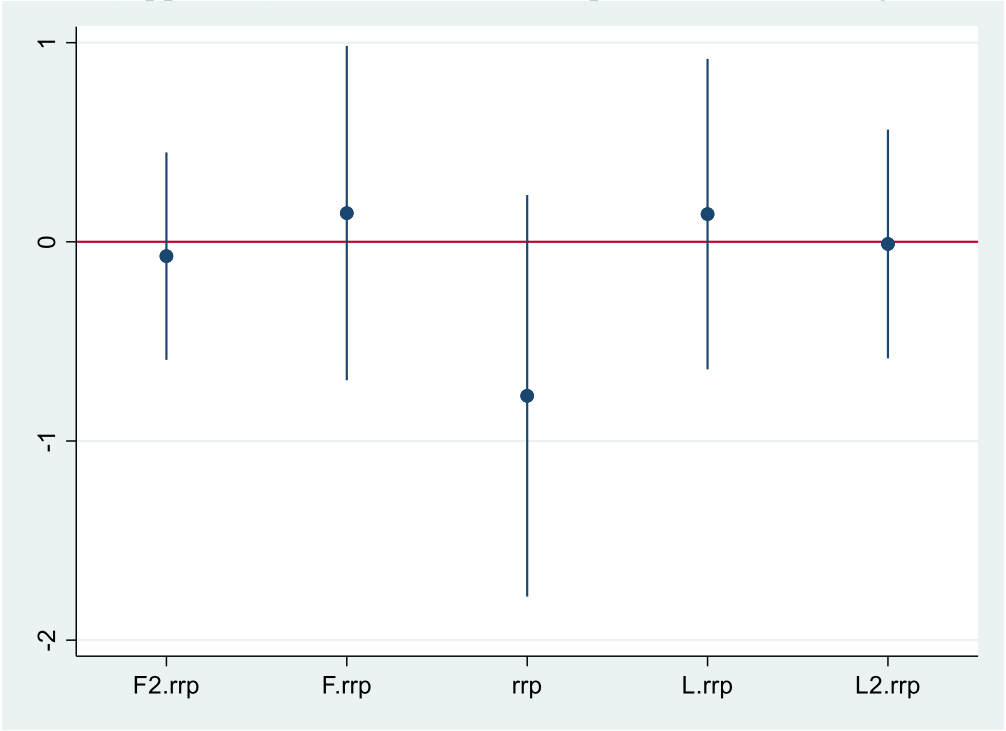
Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Figure 18 (Appendix): Parallel Trends Graph of LRRP for Log EBLL>10



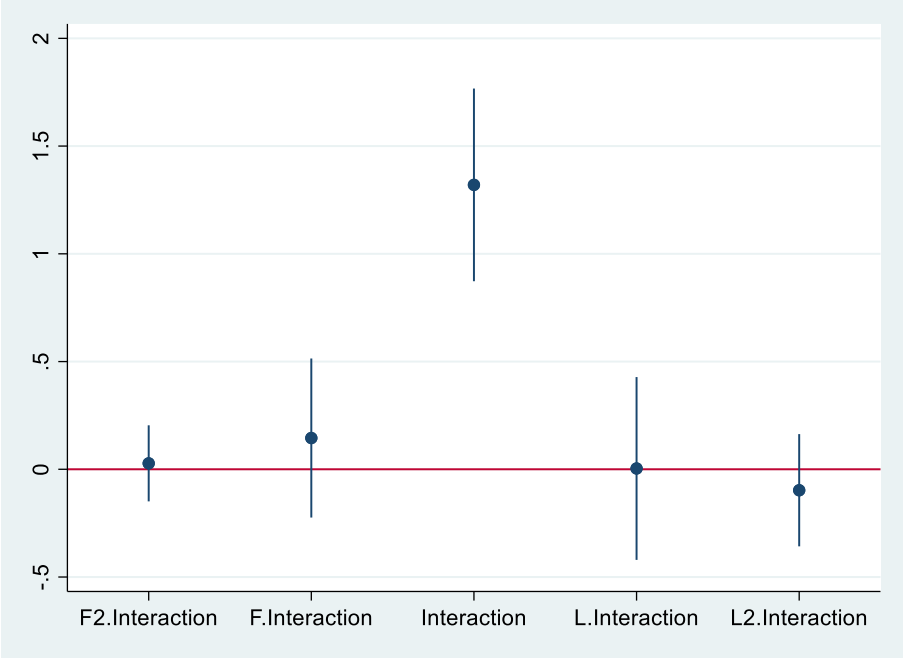
This figure plots lead and lag coefficients of LRRP when regressed on an elevated blood-lead level rate of 10 µg/dL.

Figure 19 (Appendix): Parallel Trends Graph of LRRP for Log EBLL>25



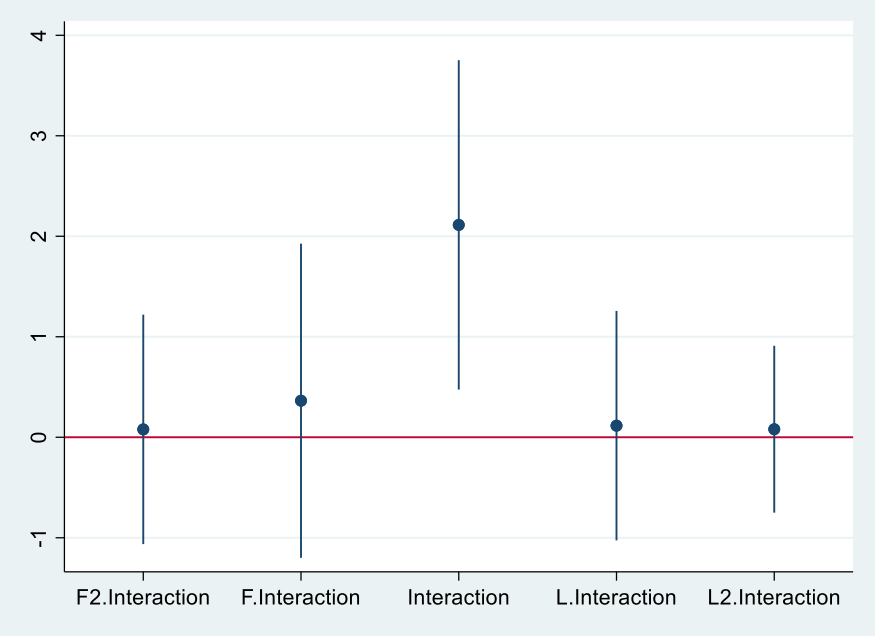
This figure plots lead and lag coefficients of LRRP when regressed on an elevated blood-lead level rate of 25 µg/dL.

Figure 20 (Appendix): Parallel Trends Graph of Subgroup Interaction for Log EBLL>10



This figure plots lead and lag coefficients of the subgroup interaction when regressed on an elevated blood-lead level rate of 10 µg/dL.

Figure 21 (Appendix): Parallel Trends Graph of Subgroup Interaction for Log EBLL>25



This figure plots lead and lag coefficients of subgroup interaction when regressed on an elevated blood-lead level rate of 25 µg/dL.