

C A G E

# **Hassles and environmental health screenings: Evidence from lead tests in Illinois**

CAGE working paper no. 509

September 2020

Ludovica Gazze

# Hassles and Environmental Health Screenings: Evidence from Lead Tests in Illinois

Ludovica Gazze\*

August 24, 2020

## Abstract

I study the determinants of childhood lead screening using all Illinois birth records (2001-2014), matched to lead testing records and geocoded housing age data. Housing age measures lead risk, as older houses disproportionally have lead paint. Changes in providers' availability, inferred from testing data, provide variation in non-monetary costs of testing. Travel costs reduce screening among low- and high-risk households alike. Thus, self-selection based on travel costs does not appear to improve targeting, even though high-risk households are willing to pay \$29-389 more than low-risk households for screening. Screening incentives would be cost-effective for reasonable values of lead poisoning externalities.

JEL CODES: I1, Q52.

---

\*Department of Economics, University of Warwick. Email: [Ludovica.Gazze@warwick.ac.uk](mailto:Ludovica.Gazze@warwick.ac.uk). Marcella Alsan, Alex Bartik, Fiona Burlig, Thomas Covert, Golvine de Rochambeau, Catie Hausman, Michael Kofoed, Michael Greenstone, Rebecca Meyerson, Roland Rathelot, Nick Sanders, Emilia Simeonova, Tommaso Sonno, Dan Waldinger, Mirko Wiederholt and seminar participants at EEA, Collegio Carlo Alberto, University of Chicago, JHU, Indiana University, Sciences Po, University of Barcelona, University of Padua, University of Warwick, APPAM, ASHEcon, H2D2, and the 4th Marco Fanno Alumni Workshop provided helpful comments. The Illinois Department of Public Health provided feedback, yet the conclusions, opinions, and recommendations in this paper are not necessarily theirs. The Joyce Foundation provided generous support. Bridget Pals and Xiyue Song provided excellent research assistance.

# 1 Introduction

Sources of lead exposure are still pervasive in US homes despite evidence that early childhood poisoning is associated with reduced IQ (Ferrie et al. 2015) and educational attainment (Aizer et al. 2018, Grönqvist et al. forthcoming, Reyes 2015a), and an increased risk of criminal activity (Aizer & Currie 2019, Feigenbaum & Muller 2016, Reyes 2015b, 2007). Two thirds of the Illinois housing stock, almost 3.6 million homes, was built prior to the residential lead paint ban in 1978 and may have lead paint.<sup>1</sup> Yet, poisoning rates are relatively low: at current levels, 2.2 percent of Illinois children born in 2014 had lead poisoning (Figure 1: Panel A).<sup>2</sup> Ensuring each Illinois home is lead-safe for children by stripping or painting over the lead paint could cost up to \$37.9 billion.<sup>3</sup> Thus, it is key to target clean-ups to high-value homes in ways similar to the targeting of energy efficiency programs (Boomhower & Davis 2014, Allcott & Greenstone 2017).

To identify homes requiring clean-up, lead poisoning prevention programs in the US rely on early childhood health screenings that reveal lead exposure. Because small children are not systematically in school, this approach hinges on families travelling to their doctor’s office for lead screening when the children are about 9 and 24 months old. This sort of barrier to policy uptake is known as a *hassle* or *ordeal*, and hassles may explain why lead screening rates are lower than 60 percent even in areas where the State of Illinois requires universal screening (Figure 1: Panel B).

This paper investigates the impact of ordeals, that is higher screening costs, on lead screening and poisoning prevention. Do these ordeals improve targeting efficiency, or do they hinder detection and remediation of lead hazards? When only program recipients know their private value of receiving a program, ordeals may reduce inclusion errors. That is, recipients who do not need it may select out of the program to avoid these ordeals (Nichols & Zeckhauser 1982), lowering health care costs. Households may know the state of the paint coat or have off-the-record property inspection results. However, households with high potential benefits may also face higher costs per

<sup>1</sup>Source: American Community Survey (2017).

<sup>2</sup>During my sample period, the Illinois Department of Public Health (IDPH) referred children to services if they had a blood lead level of  $10\mu\text{g}/\text{dL}$  or higher. In 2019, IDPH lowered the threshold to  $5\mu\text{g}/\text{dL}$  following Centers for Disease Control and Prevention guidelines that recognize no safe level of lead exposure.

<sup>3</sup>Source: Author’s calculation based on data from the Department of Housing and Urban Development.

ordeal ([Alatas et al. 2016](#)), for example because they do not have a car. Then, ordeals may increase exclusion errors: poisoned children may forego screening, leading to high private and social costs.

To study the effect of ordeals on lead poisoning prevention, I link geocoded birth records for the universe of over 2 million children born in Illinois between 2001 and 2014 to blood lead screening records and housing age information from assessor files. Screening data include information on realized poisoning risk for screened children, and housing age measures ex-ante observable risk for both screened and unscreened children. First, I estimate the elasticity of screening with respect to travel costs in terms of distance to health care providers. To assuage concerns of endogeneity in households' location relative to providers, I exploit providers' openings and closings and compare children born in the same location in different years who face different sets of providers.<sup>4</sup> The key identifying assumption is that openings and closings of medical doctor offices are orthogonal to trends in lead screening. Second, I study how travel costs affect which households select into screening, in terms of both ex-ante observable and ex-post realized risk. The identifying assumption to study selection is that, while children may obtain other services when they get screening, households with a high or low risk of lead poisoning value these additional services similarly.

First, being 15 minutes farther away from a lead-screening provider (two-way) decreases the likelihood of screening by 9 percent, on average. Second, I find no evidence that households who get screened despite facing higher costs have higher observable or unobservable exposure risk. In other words, I find no evidence that ordeals improve targeting efficiency. Third, proximity to providers improves detection of lead poisoning, but it does not increase take-up of remediation funding. Thus, removing barriers to screening may not lead to increased remediations, perhaps due to partial compliance with abatement regulations or limited awareness of remediation funding. Moreover, proximity to high-quality providers, as measured either by screening outcomes or medical school attended, increases screening more than proximity to low-quality providers, suggesting both travel costs and providers' discretion affect screening take-up.

---

<sup>4</sup>A growing literature leverages closures of health care providers, such as abortion clinics, Social Security Administration field offices, and bank branches to estimate the effect of travel costs on take-up of different programs ([Deshpande & Li 2019](#), [Nguyen 2019](#), [Lu & Slusky 2016, 2017](#), [Lindo et al. 2019](#), [Venator & Fletcher 2019](#)).

Households' revealed preference for screening relative to travel costs allow me to estimate the social value of the existing lead screening policy and counterfactual prevention policies. I use travel costs in the logit framework to estimate the willingness-to-pay (WTP) for screening of households with different lead exposure risk. I simulate the impact of four screening policies: travel subsidies, pay-for-performance incentives for providers, an increase in screening locations, and universal screening for children in old homes. I estimate that the average household in the most at-risk homes has a WTP for screening of \$6.14, \$29-389 higher than the average low-risk household, consistent with the low incidence of lead poisoning, behavioral hazards (Baicker et al. 2015, Chandra et al. 2019), or non-standard preferences (Kőszegi 2003, Oster et al. 2013). All counterfactual screening policies I examine result in modest benefits for the marginal households. Yet, these policies may be cost-effective when accounting for reductions in lead exposure externalities, consistent with the large impacts of programs targeting disadvantaged children found by Hendren & Sprung-Keyser (2020). By contrast, increasing remediations does not appear to be cost-effective.

This paper contributes to a robust body of evidence that identifies travel costs as an important determinant of take-up of social benefits, including childcare subsidies, disability insurance, small business loans, and health care services (Currie 2006, Rossin-Slater 2013, Herbst & Tekin 2012, Deshpande & Li 2019, Nguyen 2019, Lu & Slusky 2016, 2017, Einav et al. 2016, Lindo et al. 2019, Venator & Fletcher 2019). In the US, information barriers, scheduling challenges, and transportation costs appear to contribute to vaccine delays among disadvantaged families (Brito et al. 1991, Carpenter & Lawler 2019). In India, small financial incentives appear more cost-effective at increasing immunization take-up than improving supply (Banerjee et al. 2010). I use travel costs to elicit households' willingness-to-pay for information about their exposure risks, related to a large environmental economics literature that uses travel costs to estimate the recreational value of environmental amenities (Kuwayama & Olmstead 2015). My paper shows that travel costs decrease detection of lead hazards, potentially imposing a large externality on society.

This paper also contributes to a large literature studies the targeting efficiency of welfare programs (Hanna & Olken 2018). Hoffmann (2018) finds that poor Indian households are very elastic

with respect to non-monetary prices, such as travel costs. I find no evidence that high-risk households differentially select into screening at higher distances, suggesting that households at high risk for lead exposure in the US may disproportionately dislike travel hassles, too. My findings suggest that travel costs may have worse targeting properties than bureaucratic ordeals, which have been shown to improve targeting efficiency in the US (Kleven & Kopczuk 2011, Finkelstein & Notowidigdo 2019, Einav et al. forthcoming).

Section 2 provides institutional background and models how travel distance may affect targeting efficiency. Section 3 describes the data I use in this paper. Sections 4 and 5 analyze screening take-up and the costs and benefits of different lead poisoning prevention policies.

## 2 Background and Theoretical Framework

First, this section provides background on lead screening. Second, it builds on the classical work of Nichols & Zeckhauser (1982) and its extension by Alatas et al. (2016) to show how travel costs affect selection into screening. Third, it discusses the role of lead poisoning externalities in the planner's screening decision.

### 2.1 Lead Screening Background

Lead exposure becomes more likely when children start crawling and walking in contaminated homes and may ingest or inhale lead particles. Thus, federal guidelines mandate that all children on Medicaid must be screened for lead poisoning at ages one and two.<sup>5</sup> In addition, Illinois requires screening for all children living in high-risk zip codes, defined by housing age and demographic characteristics. Even in these high-risk zip codes, which include the whole city of Chicago, less than two thirds of children born between 2001 and 2014 were screened (Figure 1: Panel B). This low compliance with the screening guidelines raises the question of what barriers hinder screening.

---

<sup>5</sup>The effects of lead exposure are worst in small children. Therefore, in the remainder of my paper I focus on screening and exposure by age two. The findings and conclusions carry through in the larger sample.

I hypothesize that travel costs to the doctor constitute a barrier to screening since house visits do not include blood lead screening in Illinois. For comparison, around 80 percent of Illinois children eligible for Family Case Management had three or more well-child visits in the period 2005-2010, even though the American Academy of Pediatrics recommends six visits during the first year of life, to occur at one, two, four, six, nine, and twelve months of age (IDHS, 2010). As the likelihood that a child has a well-child visit decreases with age (Caldwell & Berdahl 2013), missed well-child visits after 9 months of age may explain the gap between well-child visits and lead screening rates. Section 4.4 investigates the alternative explanation that providers' discretion drives the low compliance with the screening guidelines. In particular, a household may have limited choice of providers due to insurance network and providers' capacity constraints, leading to a correlation between provider's quality and household's lead exposure risk. For reference, a survey by the American Academy of Pediatrics reports that 20 percent of pediatricians do not accept patients on Medicaid in Illinois, compared to 15 nationwide (AAP, 2012). My empirical findings support the hypothesis that both travel costs and providers' quality affect screening rates.

## 2.2 The Household's Screening Take-Up Decision

I model screening as an insurance mechanism, with benefits accruing from assignment to case management aimed at reducing damages if a child is found to be lead-poisoned.<sup>6</sup> This modelling choice ignores potential screening benefits accruing from negative tests, such as learning that a home is lead-safe: given I find underscreening, allowing for such benefits would only increase the gap. Parents' perceived screening benefits depend on several factors, including information about exposure risk, degree of risk aversion, degree of altruism towards the child,<sup>7</sup> beliefs about treatment costs and feasibility (which may correlate with home-ownership) as well as recovery

---

<sup>6</sup>Case management occurs mostly at home and includes nutritional education and information about reducing exposure in the home, a home inspection, and referral to lead remediation services, which are generally subsidized for low-income households. Billings & Schnepel (2018) show that such case management fully reverses lead poisoning damages in a sample of North Carolina children.

<sup>7</sup>The evidence on how much parents value reductions in their children's health risk relative to reductions in their own risk is mixed (see for example, Gerking & Dickie 2013, Gerking et al. 2014)

probability,<sup>8</sup> and additional benefits from visiting the doctor, such as having a physical examination or an immunization shot.<sup>9</sup> My model does not require assumptions on these parameters; the revealed-preference approach in Section 5 allows me to compare willingness-to-pay (WTP) estimates to estimates of screening benefits computed for different parameter values.

Household  $i$  perceives benefit  $b_i$  from screening their child. The cost of screening child  $i$ ,  $c_i$ , is a function of the nominal screening price,  $p$ , and the opportunity cost in terms of the parents' wage,  $w_i$  and travel time,  $t_i$ , which is proportional to distance from a healthcare provider,  $d_i$ . I abstract from heterogeneity in  $p$  for simplicity, although the cost of a blood lead test in Illinois varies with a child's insurance coverage.<sup>10</sup> Then, child  $i$  is screened if and only if

$$b_i \geq c_i = w_i t_i + p. \quad (1)$$

Because  $t_i \propto d_i$ , this inequality yields a cutoff  $\bar{d}_i$  above which a child is not screened:

$$\bar{d}_i = \frac{b_i - p}{w_i}. \quad (2)$$

I assume that benefits increase with risk, that is  $b'(r_i) > 0$ : the higher the potential exposure, the more likely that screening leads to crucial intervention. Then, riskier children have a higher willingness-to-travel for screening: as in the classic ordeals model (Nichols & Zeckhauser 1982), the cutoff increases with risk:

$$\frac{\partial \bar{d}_i}{\partial r} = \frac{\partial b_i}{\partial r} \frac{1}{w_i} \geq 0. \quad (3)$$

Figure 2 illustrates how risk affects the relationship between screening and distance. High-risk households are less sensitive to distance than low-risk households: their screening rates decline

<sup>8</sup>Myerson et al. 2018 show that increasing treatment access increases screening, evidence of an “ostrich effect”, a term coined by Galai & Sade (2006).

<sup>9</sup>Not observing these additional services does not bias the selection analysis if benefits from these additional services are not correlated with screening benefits after controlling for observables.

<sup>10</sup>While lead screening is fully covered for children enrolled in Medicaid or All Kids, nominal prices range between \$0-43 for uninsured or private insurance (<http://www.leadsafeillinois.org/uploads/documents/LeadSafeILDirectory061.pdf>, accessed in June 2019), with an average venous test costing \$31 (Kaplowitz et al. 2012). I discuss how this variation in prices affects my estimates of households' WTP for screening in Section 5.1.



less sharply with distance (left panel). Therefore, the share of screened children that is high-risk increases with distance (right panel).

However, the model's predictions become ambiguous if we consider travel mode, following [Alatas et al. \(2016\)](#). Assume that family assets  $a_i$  are negatively correlated with risk,  $a'(r_i) < 0$ , and that travel time is negatively correlated with assets. For example, travelling by car is faster than walking or using public transit:  $t_i(a_i, d_i) \propto \frac{d_i}{a_i}$ . Then,

$$\bar{d}_i \propto a_i \frac{b_i - p}{w_i}, \quad (4)$$

$$\frac{\partial \bar{d}_i}{\partial r_i} \propto \underbrace{\frac{\partial a_i}{\partial r} \frac{b_i - p}{w_i}}_{<0} + a_i \underbrace{\frac{\partial b_i}{\partial r} \frac{1}{w_i}}_{>0} \stackrel{?}{\leq} 0. \quad (5)$$

In a model with assets, cutoffs may either increase or decrease in risk. While the second term in equation (5) is still positive, the first term is negative: riskier households face higher travel times conditional on distance. Thus, the effect of reducing distance to providers on the average riskiness of screened children is an empirical question. In Section 4.2, I exploit providers' openings and closings to answer this question.

### 2.3 The Planner's Problem

The socially optimal level and targeting of screening may not coincide with the individual optimum due to externalities. First, lead-poisoned children negatively affect their classroom peers ([Gazze et al. 2020](#)) and are more likely to engage in risky and criminal behavior ([Aizer & Currie 2019](#), [Feigenbaum & Muller 2016](#), [Reyes 2015b, 2007](#)). Second, detecting lead hazards may prevent exposure of future residents.

I model the social benefits of screening a child as the sum of three components.<sup>11</sup> First, I consider the private benefit,  $b_i - c_i$ . Second, I add the averted externality  $i$  would have imposed on society if they had not been screened,  $e_i$ . Third, I add the discounted value of the avoided

---

<sup>11</sup>Here, I abstract from the medical sector costs of increasing screening.

externalities from preventing exposure among children  $j \in J$  who will live in  $i$ 's building in the future.<sup>12</sup> Summing over the set of screened children  $S$ , this yields

$$B = \sum_{i \in S} ( \underbrace{b_i - c_i}_{\text{Private Value}} + \underbrace{e_i}_{\text{Externality}} + \underbrace{\delta \sum_j e_j * \text{Lives in } i\text{'s building}_j}_{\text{Prevention Value}} ). \quad (6)$$

Thus, some households with low private benefits may have a high social value of screening if they have a large externality or prevention value.

The planner cannot optimally target screening without knowing  $e_i$  and  $e_j$ . Housing age may proxy for exposure risk at each home. Then, targeting screening based on observable risk may improve upon self-selection on private benefits. I estimate both the average prevention value of screening (Section 4.3) and the societal benefits of different screening policies (Section 5.2).

### 3 Data

My analysis requires data on children's screening outcomes, travel costs, lead exposure risk, and lead remediations. First, I link birth records to blood lead test data to construct children's screening histories. Second, I geocode children's addresses at birth and lead-screening providers' addresses to measure the distance a child has to travel to get screening. Third, I link these individual-level data to address-level housing age and remediation data to construct unique measures of exposure risk and remediation activity at birth addresses. Appendix Table A.1 provides child-level summary statistics for the variables and sample of children included in the analysis.

#### 3.1 Childhood Lead Screening Measures

The Illinois Department of Public Health (IDPH) provided birth and death certificates for almost 4.5 million children born in Illinois between 1991 and 2016. These records include each child's name and birth date, allowing me to link these data to the universe of 5.4 million blood lead tests

<sup>12</sup> $e_j$  will depend on the riskiness of each building, and may be zero.

performed in Illinois between 1997 and 2016, with a match rate of 86 percent (Appendix Figure A.1). Matched and unmatched tests have similar observable characteristics (Appendix Table A.2). Because lead test records are incomplete prior to 2001, I limit my analysis to children born after 2000. I also limit the analysis to children born before 2015 to ensure I observe each child's outcome by age two. I classify non-deceased children not linked to any tests as not screened. Appendix Tables A.3 and A.4 show the number of tests and unique children in my original sample, and the number remaining after each data cleaning and linkage step.

IDPH collects children's blood lead records from physicians and laboratories. These records include test date, blood lead level (BLL), test type (capillary or venous), provider and laboratory identifiers, and Medicaid status (albeit incomplete). Capillary tests are prone to false positives. Thus, capillary tests that show elevated blood lead levels (EBLLs), defined as blood lead levels above 9 micrograms per deciliter of blood ( $\mu\text{g}/\text{dL}$ ), need to be confirmed by another test. For each child, I keep the highest venous test when available, or the highest confirmed capillary test when available. My sample includes 70,000 confirmed EBLLs from over 22,000 children (Appendix Table A.5). Some laboratories have minimum reporting limits, meaning BLLs are bottom-censored; I correct for these limits to obtain correct population estimates of lead exposure.<sup>13</sup>

Birth records also include family characteristics, such as mother's marital status, age, education, and race, as well as child's address at birth. I geocode these addresses to link the blood lead data to housing age information (see Section 3.3) and Census block group median income from the 2015 American Communities Survey. After geocoding, I obtain a sample of over 2 million children and over 2.9 million tests linked to these children. I use birth address rather than address at testing time because I only observe subsequent addresses conditional on a child being screened for lead. Appendix Table A.6 shows that even if a third of households in my sample move within a two-year period, most households remain in homes and zip codes with the same exposure risk.

---

<sup>13</sup>I determine the cutoff for each laboratory based on the distribution of test results for that laboratory by both test type and year. Appendix Figure A.2 shows an example of a laboratory with a very apparent cutoff at  $5\mu\text{g}/\text{dL}$ . Some laboratories have a thin left tail of test results below the estimated cutoff: I reassign those test results to the cutoff value. For each cutoff-year-type cell, I use laboratories without cutoffs to compute the average BLL for tests below that cutoff and I reassign all test results at the cutoff to this average value.

## 3.2 Provider Access Measures

IDPH collects the name and address of providers who perform lead tests. Screening providers can be individuals, small groups of doctors, or hospitals. I code a provider as entering or exiting the sample the first or last year that I observe them ordering tests, respectively. On average, 4.5 percent of providers enter each year and 4.8 percent exit, with similar figures when excluding providers performing fewer than ten tests per year (4.1 and 4.4 percent, respectively).<sup>14</sup> These numbers are roughly consistent with statistics on physicians' churning: for example, 9-13 percent of physicians report plans to retire within three years, and about 2.5 percent new students graduate from medical school each year (Walker et al. 2016, Young et al. 2015). Appendix Table A.7 shows that 24 percent of providers in my sample are individuals, and that providers who enter or exit throughout my sample are generally similar to the average provider.<sup>15</sup>

To construct a measure of travel costs, I calculate the distance “as the crow flies” between a child’s birth residence and the closest provider open during the child’s birth year.<sup>16</sup> The median child has a provider within 1.2 kilometers (Appendix Figure A.6). Over 90 percent of screened children did not get tested at their closest provider (Appendix Figure A.7), likely due to preference for continued care after a move (Raval & Rosenbaum 2018) or insurance network constraints. Section 4.1 investigates the relationship between distance to closest provider and distance travelled.

The impact of nearby providers may depend on the quality of the available providers. For example, Vivier et al. (2001) document heterogeneous screening rates across providers in Rhode Island. I use the 2019 USNews ranking of the medical school the provider attended as one measure of quality (Schnell & Currie 2018). I obtain medical school attended by linking providers to the 2019 Medicare Physician Compare File (MPCF) through name, address, and practice name.<sup>17,18</sup>

<sup>14</sup>The median provider performs 11 tests in a year (Appendix Figure A.3).

<sup>15</sup>Appendix Figure A.4 displays how providers' locations change from the beginning to the end of my sample, while Appendix Figure A.5 shows the distribution of providers across neighborhoods and years.

<sup>16</sup>For computational reasons, to identify closest providers I use a search algorithm that conditions on the median catchment distance of each provider, which may overstate distance for children farther away than the median, thus biasing the estimated effect of distance downward. In the sample of screened children, this procedure assigns 7.09 percent of tests to a minimum distance that is higher than the actual distance travelled to obtain the test.

<sup>17</sup>For organizations with multiple providers, I average the rankings.

<sup>18</sup>Only one percent of providers in the MPCF are pediatricians.

I also consider measures of quality that directly capture a provider’s lead screening behavior: I define providers as higher quality if they screen more children and/or screen them at the right times according to federal and state guidelines as follows. Because I do not observe a child’s provider if the child is not screened, I calculate a provider’s screening rate as the screening rate for children born within the median distance households travel to see that provider, and I weigh unscreened children by the inverse of their distance.<sup>19</sup> Because federal guidelines mandate that all children on Medicaid must be screened for lead poisoning at ages one and two, I compute the share of Medicaid children a provider screened at age one who have a second test by age two. I also compute the share of EBLLs detected by each provider with a required follow-up within 90 days.<sup>20</sup> I then aggregate these screening-based measures into a summary quality index. Finally, I consider a provider’s ability to perform capillary tests as an indicator of quality, because capillary testing may reduce the barrier to screening if households are averse to venous blood draws.<sup>21</sup>

Providers’ screening-based quality measures and providers’ medical schools capture different provider’s characteristics and are indeed imperfectly correlated (Appendix Figure A.10). My empirical analysis is robust to using different quality measures.

### 3.3 Childhood Lead Exposure Pathways

Although children can be exposed to lead through several channels, deteriorating lead paint is the most common source in Illinois: children in homes built prior to 1930 have the highest BLLs, after controlling for demographic characteristics and zip code fixed effects (Abbasi et al. 2020). Indeed, HUD estimates that 87 percent of houses built before 1940 in the US have lead paint, compared to 69 percent of houses built between 1940 and 1959 and 24 percent of houses built between 1960 and 1977 (HUD, 2011). Thus, I define children born in homes built before 1930 as high-risk, using parcel-level data on construction year in the Zillow Transaction and Assessment Dataset.<sup>22</sup>

<sup>19</sup>For most providers, the median child’s address is within 7 kilometers of their provider’s address.

<sup>20</sup>Appendix Figure A.8 shows that only around 50 percent of EBLLs have a follow-up test.

<sup>21</sup>Appendix Figure A.9 shows the location of providers of different quality in Illinois.

<sup>22</sup>More information on accessing the data can be found at <http://www.zillow.com/ztrax>. The results and opinions are those of the author and do not reflect the position of Zillow Group.

### 3.4 Lead Hazard Remediations

To measure lead hazard abatement following EBLL detection, I use data on addresses that receive remediation funding under HUD’s lead hazard control programs.<sup>23</sup> Because these funds are targeted to low-income property owners, these data do not cover the universe of lead hazard remediations. Yet, they provide a useful picture of case management following EBLL detection in the absence of more complete data.

## 4 Empirical Analysis: Travel Costs and Child Lead Screening

This section investigates how travel costs affect screening. First, I estimate the elasticity of screening with respect to travel costs. Second, I study how travel costs affect selection into screening. Third, I estimate the effect of travel costs on EBLL detection and hazard remediation. Fourth, I investigate how the quality of nearby providers affects screening.

To study the relationship between screening take-up and travel costs, I exploit changes in distance to providers over time due to providers opening and closing. As providers open and close, children born at the same location but in different years face different travel costs. This approach is internally valid if the timing of openings and closings is exogenous to trends in screening rates. This condition would be violated if providers open in areas targeted by campaigns to increase screening rates, or if providers open in richer, low-risk areas with decreasing screening rates. To investigate the plausibility of this assumption, I estimate the following regression:

$$ScreeningRate_{gy} = \sum_{\tau} \beta_{\tau} Entry_{g,y-\tau} + \sum_{\tau} \gamma_{\tau} Exit_{g,y-\tau} + \eta_g + \xi_y + \varepsilon_i, \quad (7)$$

where  $ScreeningRate_{gy}$  is the screening rate in neighborhood  $g$  and birth cohort  $y$ ;  $Entry_{g,y-\tau}$  and  $Exit_{g,y-\tau}$  are leads and lags of providers’ entries and exits, defined as changes in the distance

---

<sup>23</sup>The data were collected for a project with Stephen Billings, Michael Greenstone, and Kevin Schnepel, titled “National Evaluation of the Housing and Neighborhood Impact of the HUD Lead-Based Paint Hazard Control Program, 1993-2016” and funded by HUD.

between the neighborhood centroid and the closest provider;  $\eta_g$  is a set of neighborhood fixed effects and  $\xi_y$  is a set of birth cohort fixed effects. Figure 3 plots the  $\beta_\tau$  and  $\gamma_\tau$  coefficients from estimating equation (7) at the level of Census block and provider catchment area with a radius of 500 meters. My estimates suggest that providers' entries and exits are not correlated with pre-existing trends in screening rates. Moreover, Appendix Table A.8 shows no correlation between openings and closings and lagged neighborhood characteristics at the Census tract or block level.

Figure 3 suggests that providers' openings and closings provide exogenous variation in travel costs over time. I leverage this variation in a linear probability model that compares children born in the same location in different years, controlling for location and birth year fixed effects:

$$Y_{igy} = \beta d_i + \eta_g + \xi_y + \pi X_{igy} + \varepsilon_i, \quad (8)$$

where  $Y_{igy}$  is an outcome for child  $i$  born in neighborhood  $g$  in year  $y$ ,  $d_i$  measures a child's distance to the closest provider open in their birth year,  $\eta_g$  is a set of location fixed effects and  $\xi_y$  is a set of birth year fixed effects. My preferred specification defines location as Census block, but my results are robust to considering zip code, tract, block group, or address. My preferred specification omits individual-level controls  $X_{igy}$  as I aim to test for selection into screening using child characteristics as outcomes. I introduce neighborhood-level screening rates and siblings' screening outcomes as controls to test specific channels influencing screening, such as information. I cluster standard errors at the zip code level to allow for correlation in exposure sources and screening behavior, and my main results are robust to clustering at the county level (Appendix Table A.9).

The next sections examine the effect of distance on different outcomes  $Y_{igy}$ . Section 4.1 estimates the effect of travel costs on screening by looking at an indicator for whether a child is screened by age two. Section 4.2 studies selection by examining indicators for a screened child having certain characteristics, such as living in a home built prior to 1930, being black or hispanic, or having a single, teen, or low-education mother. Section 4.3 estimates the effect of travel costs on timely poisoning detection and remediation by looking at age at test and an indicator for a HUD-funded remediation at the address within three years.

## 4.1 Do Travel Costs Decrease Screening?

While the relationship between screening rates and distance to providers in Illinois is U-shaped in the raw data, it becomes closer to linear after controlling for neighborhood fixed effects suggesting a spurious correlation likely due to car ownership (Figure 4). In my main analysis, I drop the 31,178 children who are farther than 20 kilometers from a provider (2.6 percent of the sample), as they are very different from the rest of the sample.<sup>24</sup> Indeed, these outliers have a lower elasticity of screening with respect to travel costs (Appendix Table A.10).

Panel A of Table 1 estimates that being one kilometer farther away from a lead-screening provider, a 30 percent increase over the mean distance, decreases the likelihood that a child is screened by age two by 0.4 percentage points, or 0.9 percent relative to the mean. These estimates imply an elasticity of screening with respect to distance to the closest provider of -0.03. Einav et al. (2016) estimate an elasticity of -0.02 for take-up of cancer treatment, while Herbst & Tekin (2012) estimate an elasticity of -0.13 for take-up of childcare subsidy. Non-linear estimates in the right panel of Figure 4 imply that the screening rate in Illinois would have been 2 percentage points higher (4.35 percent) if every child in my sample had a provider within one kilometer.

Because most households do not visit their closest provider, I also estimate a two-sample two-stage-least-squares (2SLS) model. 2SLS scales the reduced form estimates by the share of compliers, that is the households who change provider when their closest provider opens or closes as compared to households in the same neighborhood who face different available providers at another time.<sup>25</sup> Moreover, providers' availability changes over time: providers take on new children only half the years they are open (Appendix Table A.7). Using the sample of screened children, Panel B of Table 1 estimates that being 1,000 meters farther away to the closest provider translates into an extra 75-280 meters travelled to get a child's first lead test. All but the most stringent specification with house fixed effects exhibit a strong first stage. Bootstrapping this first stage relationship to predict distance from provider of choice for all children in the sample, Panel C estimates

<sup>24</sup>On average, the closest provider is at 3.3 kilometers and the distribution is right skewed (Appendix Figure A.6).

<sup>25</sup>Appendix Table A.1 shows that households who themselves switch providers between their children's tests are more disadvantaged than the average household with a screened child.



that being one kilometer farther away from the provider of choice, a 13.4 percent increase over the mean, reduces screening by 1.3-7.2 percentage points, yielding an elasticity of -0.21 to -1.01.

Interpreting the magnitude of the effect of distance on screening take-up requires data on households' transportation mode, which I do not observe. Thus, I use car travel times, at 1–1.5 minutes per kilometer in Illinois (Agbodo & Nuss 2017).<sup>26,27</sup> The estimates in Table 1 imply that a \$6.25 increase in travel costs (a fifteen-minute two-way trip to the doctor at 7.5 kilometers each way and \$25 hourly wage), decreases screening take-up by 9 percent.<sup>28</sup>

This section argues that providers' openings and closings affect screening for lead exposure through changes in travel costs, but providers' openings and closings could also affect the salience of lead screening. Table 2 investigates the role of information in mediating the effect of travel costs. First, Columns 1-3 show that the estimated effect of travel costs on screening is robust to controlling for screening rates at the neighborhood-cohort level. This finding suggests that peer effects cannot fully explain the effect of providers' openings and closings. Second, Column 5 uncovers heterogeneous effects of travel costs on screening rates for siblings: travel costs have stronger effects for younger siblings whose older siblings have already been screened, unless the older sibling has an EBLL. Moreover, Column 6 shows that travel costs have a smaller, if not reversed, impact on screening when an immediate neighbor has an EBLL. Thus, it appears that households update their willingness to travel for screening based on updated information.

#### 4.1.1 Robustness Checks

The estimated effect of travel costs is robust to different specifications, sample selection criteria, travel costs measures, functional forms, and outcome definitions. Table 1 shows robustness to controlling for different sets of location fixed effects, suggesting that providers' openings and clos-

<sup>26</sup>Using the HERE API to compute travel times for a 12 percent random subsample of the data, Appendix Table A.11 estimates a reduction in screening likelihood per minute of travel time of 0.4 percentage points. This estimate is statistically indistinguishable from the estimate using distance in kilometers in Table 1. I am limited to using 12 percent of my sample for this exercise as the free version of the API only allows for up to 250,000 queries in a month.

<sup>27</sup>Appendix Table A.12 shows that households in Chicago are more sensitive to distance, suggesting that transit availability does not mitigate ordeals in this case. Indeed, Appendix Figure A.11 shows that households in tracts with low car ownership rates see larger effects of provider distance on screening rates.

<sup>28</sup>Source: Bureau of Labor Statistics.

ings are not correlated with children’s characteristics that also affect their likelihood of screening. Moreover, I estimate similar elasticities for children at different distances from providers, although absolute changes appear to have larger effects at smaller distances (Appendix Table A.13).

Appendix Table A.10 explores different specifications and distance measures. Column 3 controls for Census block group trends to assuage concerns that neighborhood changes, such as gentrification, drive the estimated relationship between screening rates and distance to providers. Column 4 uses average distance from the closest five providers as households do not always visit the closest provider. While attenuated with respect to my preferred estimate, the coefficient on this variable is negative and significant. Column 5 uses distance from the Census block centroid to remove variation in travel costs due to children living in different buildings within the same block, yielding estimates that are not statistically distinguishable from my preferred estimate. Related, Appendix Table A.14 includes both distance to the closest provider and distance to the five closest providers: because distance to closest provider has a higher explanatory power, provider density does not appear to mediate my findings. Finally, proximity to providers who accept new patients and patients on Medicaid matters most for screening take up (Appendix Table A.15).

Logistic and ordinary-least-square regressions that include regressors’ block-level means but omit block fixed effects yield similar results to my preferred linear probability model (Appendix Table A.16). This approach avoids the incidental parameters problem (Neyman & Scott 1948) and is equivalent to the linear fixed effects model if there is no correlation between the relevant regressors and the fixed effects (Mundlak 1978, Chamberlain 1984, Bafumi & Gelman 2016). This equivalence is important because Section 5.1 uses the logit framework to estimate households’ willingness-to-pay for screening. This table also shows robustness to measuring screening at different ages, as most screening happens by age two (Appendix Figure A.12).

## 4.2 Do Travel Costs Affect Selection into Screening?

Travel costs decrease screening take-up, but for whom is theoretically ambiguous. On one hand, families with low exposure risk will not be willing to pay the higher travel cost. On the other hand,

children facing high travel costs, who may also be at high risk, may forego screening.

To assess how the composition of screened children changes with travel costs, I estimate equation (8) on the sample of screened children, with children's characteristics as outcomes. I consider ex-ante observable and unobservable exposure risk, as measured by housing age and lead levels. Consider two children, one in an old house and one in an adjacent new house. There is a clinic 250 meters away, and both get screened. Years later, two new children move in; the clinic is closed and the closest provider is now a kilometer away. Only the child in the old house gets screened. Among the screened children in this example, the probability that a child lives in an old home increases with distance: it is 0.5 at 250 meters and 1 at one kilometer. Data from this example would suggest that hassles improve targeting based on observable risk, as illustrated in Figure 2.

In contrast to this example, Table 3 does not support the hypothesis that children screened at farther distances have higher observable or unobservable risk: they are less likely to live in a home built prior to 1930, and have slightly lower BLLs (only significant when controlling for Census tract fixed effects). Children screened at higher distances are also slightly less likely to be black or hispanic, with significant estimates only when controlling for tract fixed effects. These findings are largely robust to including time-varying neighborhood controls (Appendix Table A.17).

### 4.3 Does Proximity to Providers Improve Children's Outcomes?

Because travel costs decrease screening for high- and low-risk children alike, increased travel costs may hinder detection of lead-poisoning. If lower detection rates lead to remediating fewer homes, future residents may face increased poisoning risk. I investigate how travel costs affect prevention outcomes by estimating equation (8) with the likelihood and timeliness of detecting an EBLL, as well as the likelihood of remediations and future EBLLs at the same location, as outcome variables.

Table 4 shows that children who live one kilometer closer to a provider are 3.3 percent more likely to be diagnosed with an EBLL (Column 1). Moreover, these children are screened six days earlier, and are younger when their highest BLL is recorded (Columns 2-3). Early detection may improve long-term outcomes by reducing exposure (Billings & Schnepel 2018). Yet, Column 4

finds no evidence that proximity to providers is associated with higher HUD-funded remediation activity at a child's home within three years of birth. Consistent with the lack of impact of travel costs on remediations, Column 5 shows no evidence of lower future EBLR rates for homes closer to providers.<sup>29</sup> Temporary remediations could also explain the lack of impact of travel costs on future poisoning rates. These findings suggest that travel costs may affect poisoned children, but not future residents, questioning the prevention value of screening policies.

#### 4.4 Does Providers' Quality Affect Screening?

Providers' entries and exits affect children's access to health care generally, not just lead screening. Yet, the disparity between well-child visits and screening rates discussed in Section 2.1 suggests that providers and families may also exercise discretion in screening during a visit. As providers' practices differ greatly (Mullainathan & Obermeyer 2019, Kwok 2019, Fadlon & Van Parys 2020, Silver 2019, Currie et al. 2016, Van Parys 2016, Fletcher et al. 2014, Epstein & Nicholson 2009), I ask whether access to high-quality providers affects screening take-up.

I regress a child's screening indicator on indicators for the presence of any providers and of high-quality ones within given distances of a child's birth address. Screening quality measures include whether providers offer less-invasive capillary tests, adherence to screening guidelines, and screening rates. Moreover, I test for sorting by investigating whether proximity to a provider with a good screening record has additional explanatory power over proximity to a provider who attended a top 20 medical school, because parents may more easily observe a provider's alma mater and select on that, than screening record. I estimate

$$Y_{igy} = \sum_k \beta_k \text{ProviderInK}_i + \sum_k \gamma_k \text{HighScreenQualityInK}_i + \sum_k \delta_k \text{Top20SchoolInK}_i + \pi X_{igy} + \eta_g + \xi_y + \varepsilon_i, \quad (9)$$

where  $k \in \{< 1\text{km}, 2 - 5\text{km}, 5 - 10\text{km}, 10 - 20\text{km}\}$  and  $X_{igy}$  are individual and neighborhood level

<sup>29</sup>I observe over 2,000 remediations and repeated EBLRs in the same home. My findings are robust to limiting the sample to children with a higher incidence of these events, as well as to correcting for small sample bias (Appendix Table A.18). My sample size allows for detection of a 7 percent effect on remediations, meaning 37 percent of the EBLR cases detected due to reduced distance would have to take up remediations for this analysis to be powered.

characteristics that may correlate with parents' and providers' screening propensity.

Figure 5 shows that children closer to providers have higher screening rates, and the more so if they are closer to high-quality providers, as measured by all quality variables. The effect of demand-side proximity is 29-58 percent the effect of supply-side providers' quality, depending on the quality measure. Moreover, screening-based quality measures have additional predictive power beyond a provider's alma mater, suggesting that these results are not driven by households with a higher propensity to screen selecting to visit providers with better education.

## 5 Benefits of Counterfactual Prevention Policies

Travel costs decrease screening and poisoning detection without improving targeting. Could policies that increase screening improve outcomes for poisoned children and society? I exploit variation in travel costs to estimate households' willingness-to-pay (WTP) for screening and simulate the impact of five counterfactual policies aimed at increasing screening and/or remediations.

### 5.1 Exposure Risk and Willingness-to-Pay for Screening

I derive the WTP for screening by defining household  $i$ 's utility from screening as

$$u_i = \alpha_i - \beta_i(\theta_i d_i + p), \quad (10)$$

where  $d_i$  is distance from provider,  $\theta_i$  is opportunity cost of travel time,  $p$  is the price of a test, and  $\alpha_i$  and  $\beta_i$  are preference parameters (Einav et al. 2016). Assuming that  $\alpha_i = \delta^\alpha X_i + \varepsilon_i$ ,  $\beta_i = \delta^\beta X_i$  and that  $\varepsilon_i$  follows a Type I Extreme value distribution,  $i$ 's WTP for screening is  $\frac{\alpha_i}{\beta_i} - \theta_i d_i - p$ .

Table 5 estimates the marginal effect of distance on screening take-up of households with different observable characteristics using both the linear probability model in equation (8) and a logit model. To recover  $\alpha_i$  while avoiding the incidental parameters problem (Neyman & Scott 1948), I include block-level means of relevant regressors and omit block fixed effects as discussed in

Section 4.1.1. I then compute WTP for different households using the logit estimates. Column 1 reports estimates for the whole sample, while other columns report estimates for subsamples, obtained by interacting distance to the closest provider with indicators for household characteristics.

The average household has a negative WTP for screening, yet households in the riskiest homes, those built prior to 1930, are willing to pay \$6.14 for screening on average (Table 5). Similarly, households with low socioeconomic status have a higher WTP for screening than better off households, consistent with their heightened risk even after controlling for housing age. Because Panels A and B of Table 5 do not show large differences in the elasticity to travel costs,  $\beta_i$ , these different WTPs suggest households have different valuations of screening benefits,  $\alpha_i$ .

Because I do not observe the price each household pays for a test, I perform a bounding exercise. If all households face the same price, Table 5 implies that households in pre-1930 homes are willing to pay up to \$29.16 more than households in newer homes. If, instead, households in pre-1930 homes have no co-pay while low-risk households pay full price (\$43), the difference in WTP between high- and low-risk households becomes negative. Conversely, the difference widens to \$72.16 if riskier households pay full price due to lack of insurance. Because households often do not visit their closest provider, I can further divide the WTP estimates by the average ratio between minimum and actual distance, 75-281 meters per kilometer (Table 1), yielding a difference in WTP of \$103.77-388.80. Still, my definition of travel costs likely overestimates WTP as high-risk households are less likely to drive meaning they need more time to travel a given distance.<sup>30</sup>

To interpret the magnitude of these WTP estimates, I need a measure of screening benefits.<sup>31</sup> Section 2 discusses how under risk-neutrality and perfect information, perceived benefits are the converse of the expected costs of lead poisoning. By contrast, perceived benefits exceed expected poisoning costs under risk aversion and fall short of them if households underestimate treatment effectiveness or overestimate treatment costs. Households in pre-1930 homes have a 0.8 percentage point higher likelihood of an EBLI than households in new homes (Column 4, Appendix Table

<sup>30</sup>Appendix Figure A.13 shows a negative correlation between car ownership rates and the share of homes built prior to 1930 for Census tracts with fewer than 50 percent of homes built prior to 1930.

<sup>31</sup>Using data on chelation treatment for severe lead poisoning, Agee & Crocker (1996) estimate that parents are willing to pay \$16.11 to reduce their child's lead levels by one percent.

A.19), but estimates of the cost of an EBLL vary widely. On the one hand, [Gazze et al. \(2020\)](#) find that children with EBLs have test scores that are 0.031 standard deviations lower than their siblings, implying a net present value of lifetime earnings lost to lead poisoning of \$5,616 and an expected lifetime cost of living in a pre-1930 home relative to a new home of \$45.<sup>32</sup> On the other hand, the correlation between IQ losses and BLLs implies an expected lifetime cost of living in a pre-1930 home relative to a new home of \$910 ([Schwartz 1994](#)), but this estimate does not account for unobserved innate ability correlated with lead exposure. While most parameter values for benefits and WTP suggest parents may overvalue screening, these estimates omit the opportunity cost of the additional time parents spend caring for a poisoned child and additional damages not measured by test scores.

## 5.2 Policy Counterfactuals

This section simulates the societal benefits of four screening policies and one remediation policy in the 2014 cohort as modeled in equation (6). First, I look at incentives for households and providers. Then, I look at a policy opening screening locations in each zip code. Finally, I evaluate a 100 percent screening requirement for children in homes built prior to 1930. Moreover, I compare these policies to subsidizing full remediation for addresses with EBLs.

Table 6 reports the number of additional children screened and additional poisoning cases detected under each policy. I compute additional detection rates assuming that marginal children have the average poisoning rate in the 2014 cohort, based on my finding that hassles do not improve targeting (Section 4.2). When evaluating the screening mandate for old homes, I use the poisoning probability among children living in old homes. I compute the private benefits of each policy by summing the WTP for screening of the marginal households,  $b_i - c_i$ , estimated in Section 5.1.<sup>33</sup> I assume no prevention benefits from screening based on the lack of evidence that proximity

<sup>32</sup>I use estimates by [Chetty et al. \(2014\)](#) that a one-standard-deviation-decrease in test scores is associated with a 12 percent decrease in earnings at 28 and 2018 Current Population Survey data to compute a lifetime earnings profile, assuming a growth rate of real labor productivity of 1.9 percent and a discount rate of 3.38 (that is, the 30-year Treasury bond rate).

<sup>33</sup>The reported private benefits estimates are not rescaled by the relationship between actual and closest distance

to providers reduces future exposure (Section 4.3). Examining the opportunity cost of using public funds for these policies is outside the scope of this paper.

Because estimates of the externality of lead exposure  $e_i$  are not available, for each policy I compute the per-child difference between the policy's private benefits and its costs. This difference indicates the minimum value of the average externality that would make each policy cost-effective. All the screening policies I study appear to be cost-effective for externality values lower than \$5,617, a figure obtained from estimates in [Gazze et al. \(2020\)](#) that a lead-poisoned child decreases all of their peers' test scores by 0.01 standard deviations per grade.<sup>34</sup> Because this value omits the crime costs of lead poisoning, it likely underestimates the total externality of lead poisoning.<sup>35</sup>

First, I simulate the effect of giving households incentives for screening, following a large literature on immunization incentives ([Banerjee et al. 2010](#), [Bronchetti et al. 2015](#)). I assign variable incentives based on the zip code average realized travel distance, valued at 1.2 minutes per kilometer and \$25 per hour (\$10.5 on average). I identify the marginal children screened under this policy as those whose WTP turns from negative to positive under the counterfactual policy, weighting by the realized probability of screening for a given WTP. Column 1 of Table 6 shows that this policy's private benefits are positive but lower than the incentives disbursed as many inframarginal households receive subsidies.

Second, I consider a pay-for-performance (PFP) incentive for low-performing providers, a policy with mixed success ([Li et al. 2014](#), [Alexander & Schnell 2019](#)). For providers in high-risk zip codes with screening rates lower than 50 percent, I assume PFP leads them to screen an additional random 10 percent of children in their catchment area. Column 2 of Table 6 shows that PFP would lead to screening around four times more children than the household incentive, but achieve a similar private benefit, due to poorer targeting. Dividing the policy's private benefits among the 216

---

discussed in the previous section, which would imply smaller private benefits for each policy.

<sup>34</sup>I use estimates by [Chetty et al. \(2014\)](#) that a one-standard-deviation decrease in test scores is associated with a 12 percent decrease in earnings at 28 and 2018 Current Population Survey data to compute a lifetime earnings profile, assuming a growth rate of real labor productivity of 1.9 percent and a discount rate of 3.38 (that is, the 30-year Treasury bond rate).

<sup>35</sup>[Heckman et al. \(2010\)](#) estimate that 38–66 percent of the value of preschool programs is attributable to crime reductions.



providers affected yields an incentive of \$1,230, or \$5.24 per additional child.

Third, I simulate a provider opening at the centroid of each zip code without providers in 2014. In the past, lead screening was offered at the Special Supplemental Nutrition Program for Women, Infants, and Children (WIC), the single largest point of access to health-related services for low-income preschool children in the US ([General Accounting Office 1999](#)), and WIC status appears associated with increased screening ([Vaidyanathan et al. 2009](#)). Alternatively, pharmacies could acquire lead screening kits at a cost of \$382 for 48 tests. Column 3 of Table 6 shows that this policy would only screen 882 more children, consistent with households responding to even small distances. Yet, the benefits for these marginal children appear higher than the program's cost because capillary screening kits are cheap.

Fourth, I consider a mandate to screen all children in homes built prior to 1930, which leverages observable exposure risk to target screening. Column 4 of Table 6 shows that, compared to the screening incentive in Column 1, this policy yields fewer additional screenings and lower private benefits, but similar rates of poisoning detection. This result is consistent with the finding in Section 4.2 that households do not self-select into screening based on better information about unobservable risk. Thus, the social planner may be able to target screening based only on observable risk. However, it may be prohibitively costly to screen all children in old homes.

Fifth, I consider a policy that keeps screening constant but assumes perfect remediation after EBLL detection, preventing new lead poisoning cases at homes with previous cases. In the 2014 cohort, 638 homes had an EBLL. Because 10.3 percent of addresses with EBLs in the 2001–2003 cohorts have another child with EBLs within 10 years, I assume that remediating these 638 homes would prevent 66 new cases. The average remediation cost in the HUD data for the 2010–2016 period is \$10,646, suggesting lead poisoning externalities need to be on the order of \$100,000 for remediations to be cost-effective in terms of prevention benefits only. Importantly, I do not have estimates of averted case management costs that would factor in prevention benefits.

This section evaluates the social benefits of five screening and remediation policies. Overall, policies increasing screening rates have modest private benefits for marginal children, but may be

cost-effective after taking into account lead-poisoning externalities as small as \$3,500. Specifically, I consider a screening subsidy, which allows households with the highest WTP at the margin to select into screening, and find that even this policy has small private benefits. Then, I consider supply-side policies such as a PFP incentive and an increase in provider locations, and find that while both have worse targeting outcomes than the screening subsidy, PFP leads to higher screening rates and thus higher poisoning detection rates. To better study targeting, I next consider a screening mandate in old homes, and find that it leads to similar poisoning detection rates as the subsidy, suggesting that households do not have private information on unobservable risks. Finally, I examine perfect remediation and find it not to be cost-effective because of the uncertainty in turnover of residents at each address.

## 6 Conclusion

Lead paint in millions of US homes potentially endangers children's health. Families with small children need blood lead screening to identify these hazards. This paper examines screening take-up in Illinois and evaluates counterfactual prevention policies. I find that travel costs decrease screening rates but do not affect selection into screening based on either observable or unobservable exposure risk. Policies incentivizing screening have low private benefits, yet may be cost-effective when considering averted poisoning externalities.

My findings suggest that decreasing travel costs, for example through subsidies, could increase screening and lead poisoning detection without reducing targeting efficiency. However, increased provider access is not associated with higher remediation activity, suggesting case management may need improvement. Outside the scope of this paper, provider training may also increase screening, as provider quality affects screening.

## References

- AAP (American Academy of Pediatrics) (2012), ‘AAP data shed new light on pediatricians’ participation in medicaid, chip’, *AAP News* **33**(7), 21–21.
- Abbasi, A., Pals, B. & Gazze, L. (2020), ‘Policy changes and child blood lead levels by age 2 years for children born in illinois, 2001–2014’, *American Journal of Public Health* (0), e1–e7.
- Agbodo, N. & Nuss, J. (2017), Real travel time in Illinois, Report prepared for policy review, Health Facilities and Services Review Board.
- Agee, M. D. & Crocker, T. D. (1996), ‘Parental altruism and child lead exposure: Inferences from the demand for chelation therapy’, *Journal of Human Resources* pp. 677–691.
- Aizer, A. & Currie, J. (2019), ‘Lead and juvenile delinquency: new evidence from linked birth, school, and juvenile detention records’, *Review of Economics and Statistics* **101**(4), 575–587.
- Aizer, A., Currie, J., Simon, P. & Vivier, P. (2018), ‘Do low levels of blood lead reduce children’s future test scores?’, *American Economic Journal: Applied Economics* **10**(1), 307–41.
- Alatas, V., Purnamasari, R., Wai-Poi, M., Banerjee, A., Olken, B. A. & Hanna, R. (2016), ‘Self-targeting: Evidence from a field experiment in Indonesia’, *Journal of Political Economy* **124**(2), 371–427.
- Alexander, D. & Schnell, M. (2019), The impacts of physician payments on patient access, use, and health, Working Paper 26095, National Bureau of Economic Research.
- Allcott, H. & Greenstone, M. (2017), Measuring the welfare effects of residential energy efficiency programs, Working Paper 23386, National Bureau of Economic Research.
- Bafumi, J. & Gelman, A. (2016), Fitting multilevel models when predictors and group effects correlate, Working Paper 1010095, SSRN.

- Baicker, K., Mullainathan, S. & Schwartzstein, J. (2015), 'Behavioral hazard in health insurance', *The Quarterly Journal of Economics* **130**(4), 1623–1667.
- Banerjee, A. V., Duflo, E., Glennerster, R. & Kothari, D. (2010), 'Improving immunisation coverage in rural India: Clustered randomised controlled evaluation of immunisation campaigns with and without incentives', *Bmj* **340**, c2220.
- Billings, S. B. & Schnepel, K. T. (2018), 'Life after lead: Effects of early interventions for children exposed to lead', *American Economic Journal: Applied Economics* **10**(3), 315–44.
- Boomhower, J. & Davis, L. W. (2014), 'A credible approach for measuring inframarginal participation in energy efficiency programs', *Journal of Public Economics* **113**, 67 – 79.
- Brito, D. L., Sheshinski, E. & Intriligator, M. D. (1991), 'Externalities and compulsory vaccinations', *Journal of Public Economics* **45**(1), 69–90.
- Bronchetti, E. T., Huffman, D. B. & Magenheimer, E. (2015), 'Attention, intentions, and follow-through in preventive health behavior: Field experimental evidence on flu vaccination', *Journal of Economic Behavior & Organization* **116**, 270–291.
- Caldwell, J. & Berdahl, T. (2013), Trends in well-child visits: United States, 2002-2009, Statistical Brief (Medical Expenditure Panel Survey (US)), Agency for Healthcare Research and Quality (US), Rockville (MD).
- Carpenter, C. S. & Lawler, E. C. (2019), 'Direct and spillover effects of middle school vaccination requirements', *American Economic Journal: Economic Policy* **11**(1), 95–125.
- Chamberlain, G. (1984), Panel data, Vol. 2 of *Handbook of Econometrics*, Elsevier, pp. 1247 – 1318.
- Chandra, A., Handel, B. & Schwartzstein, J. (2019), Behavioral economics and health-care markets, in 'Handbook of Behavioral Economics: Applications and Foundations 1', Vol. 2, Elsevier, pp. 459–502.

- Chetty, R., Friedman, J. N. & Rockoff, J. E. (2014), 'Measuring the impacts of teachers ii: Teacher value-added and student outcomes in adulthood', *American Economic Review* **104**(9), 2633–79.
- Currie, J. (2006), The take-up of social benefits, in A. Auerbach, D. Card & J. Quigley, eds, 'Poverty, the Distribution of Income, and Public Policy', Russell Sage, New York, pp. 80–148.
- Currie, J., MacLeod, W. B. & Parys, J. V. (2016), 'Provider practice style and patient health outcomes: The case of heart attacks', *Journal of Health Economics* **47**, 64 – 80.
- Deshpande, M. & Li, Y. (2019), 'Who is screened out? application costs and the targeting of disability programs', *American Economic Journal: Economic Policy* **11**(4), 213–48.
- Einav, L., Finkelstein, A., Oostrom, T., Ostriker, A. J. & Williams, H. L. (forthcoming), 'Screening and selection: The case of mammograms', *American Economic Review* .
- Einav, L., Finkelstein, A. & Williams, H. (2016), 'Paying on the margin for medical care: Evidence from breast cancer treatments', *American Economic Journal: Economic Policy* **8**(1), 52–79.
- Epstein, A. J. & Nicholson, S. (2009), 'The formation and evolution of physician treatment styles: An application to cesarean sections', *Journal of Health Economics* **28**(6), 1126 – 1140.
- Fadlon, I. & Van Parys, J. (2020), 'Primary care physician practice styles and patient care: Evidence from physician exits in Medicare', *Journal of Health Economics* **71**, 102304.
- Feigenbaum, J. J. & Muller, C. (2016), 'Lead exposure and violent crime in the early twentieth century', *Explorations in economic history* **62**, 51–86.
- Ferrie, J., Rolf, K. & Troesken, W. (2015), 'Lead exposure and the perpetuation of low socioeconomic status', *cliometrics.org* .
- Finkelstein, A. & Notowidigdo, M. J. (2019), 'Take-up and targeting: Experimental evidence from SNAP', *The Quarterly Journal of Economics* **134**(3), 1505–1556.

- Fletcher, J. M., Horwitz, L. I. & Bradley, E. (2014), Estimating the value added of attending physicians on patient outcomes, Working Paper 20534, National Bureau of Economic Research.
- Galai, D. & Sade, O. (2006), 'The "ostrich effect" and the relationship between the liquidity and the yields of financial assets', *The Journal of Business* **79**(5), 2741–2759.
- Gazze, L., Persico, C. & Spirovska, S. (2020), 'The spillover effects of pollution: How exposure to lead affects everyone in the classroom', *IZA Discussion Paper 13133*.
- General Accounting Office (1999), 'Lead poisoning: Federal health care programs are not effectively reaching at-risk children'.
- Gerking, S. & Dickie, M. (2013), 'Valuing reductions in environmental risks to children's health', *Annu. Rev. Resour. Econ.* **5**(1), 245–260.
- Gerking, S., Dickie, M. & Veronesi, M. (2014), 'Valuation of human health: An integrated model of willingness to pay for mortality and morbidity risk reductions', *Journal of Environmental Economics and Management* **68**(1), 20–45.
- Grönqvist, H., Nilsson, J. P. & Robling, P.-O. (forthcoming), 'Understanding how low levels of early lead exposure affect children's life-trajectories', *Journal of Political Economy*.
- Hanna, R. & Olken, B. A. (2018), 'Universal basic incomes versus targeted transfers: Anti-poverty programs in developing countries', *Journal of Economic Perspectives* **32**(4), 201–26.
- Heckman, J. J., Moon, S. H., Pinto, R., Savelyev, P. A. & Yavitz, A. (2010), 'The rate of return to the highscope perry preschool program', *Journal of Public Economics* **94**(1-2), 114–128.
- Hendren, N. & Sprung-Keyser, B. (2020), 'A unified welfare analysis of government policies', *The Quarterly Journal of Economics* **135**(3), 1209–1318.
- Herbst, C. M. & Tekin, E. (2012), 'The geographic accessibility of child care subsidies and evidence on the impact of subsidy receipt on childhood obesity', *Journal of Urban Economics* **71**(1), 37–52.

- Hoffmann, B. (2018), ‘Do non-monetary prices target the poor? Evidence from a field experiment in India’, *Journal of Development Economics* **133**, 15 – 32.
- HUD (U.S. Department of Housing and Urban Development) (2011), ‘American healthy homes survey lead and arsenic findings’.
- IDHS (Illinois Department of Human Services) (2010), ‘The reduction of infant mortality in Illinois - Fiscal Year 2010’.
- Kaplowitz, S. A., Perlstadt, H., D’Onofrio, G., Melnick, E. R., Baum, C. R., Kirrane, B. M. & Post, L. A. (2012), ‘The predictive value of self-report questions in a clinical decision rule for pediatric lead poisoning screening’, *Public health reports* **127**(4), 375–382.
- Kleven, H. J. & Kopczuk, W. (2011), ‘Transfer program complexity and the take-up of social benefits’, *American Economic Journal: Economic Policy* **3**(1), 54–90.
- Kószegi, B. (2003), ‘Health anxiety and patient behavior’, *Journal of Health Economics* **22**(6), 1073–1084.
- Kuwayama, Y. & Olmstead, S. (2015), Water quality and economics: Willingness to pay, efficiency, cost-effectiveness, and new research frontiers, in R. Halvorsen & D. F. Layton, eds, ‘Handbook on the Economics of Natural Resources’, Edward Elgar (Northampton, MA:, pp. 474–501.
- Kwok, J. H. (2019), How do primary care physicians influence healthcare? Evidence on practice styles and switching costs from Medicare.
- URL: [https://static1.squarespace.com/static/5bd6632951f4d49caf7eecd5/t/5d431bd1176c5e000127e772/1564679228053/JenniferKwok\\_PrimaryCarePhysicians\\_JMP.pdf](https://static1.squarespace.com/static/5bd6632951f4d49caf7eecd5/t/5d431bd1176c5e000127e772/1564679228053/JenniferKwok_PrimaryCarePhysicians_JMP.pdf)
- Li, J., Hurley, J., DeCicca, P. & Buckley, G. (2014), ‘Physician response to pay-for-performance: Evidence from a natural experiment’, *Health Economics* **23**(8), 962–978.

- Lindo, J., Myers, C., Schlosser, A. & Cunningham, S. (2019), 'How far is too far? New evidence on abortion clinic closures, access, and abortions', *Journal of Human Resources* pp. 1217–9254R3.
- Lu, Y. & Slusky, D. J. (2016), 'The impact of women's health clinic closures on preventive care', *American Economic Journal: Applied Economics* **8**(3), 100–124.
- Lu, Y. & Slusky, D. J. (2017), 'The impact of women's health clinic closures on fertility', *American Journal of Health Economics* pp. 1–73.
- Mullainathan, S. & Obermeyer, Z. (2019), Who is tested for heart attack and who should be: Predicting patient risk and physician error, Working Paper 26168, National Bureau of Economic Research.
- Mundlak, Y. (1978), 'On the pooling of time series and cross section data', *Econometrica* **46**(1), 69–85.
- Myerson, R. M., Lakdawalla, D., Colantonio, L. D., Safford, M. & Meltzer, D. (2018), Effects of expanding health screening on treatment - What should we expect? What can we learn?, Working Paper 24347, National Bureau of Economic Research.
- Neyman, J. & Scott, E. L. (1948), 'Consistent estimates based on partially consistent observations', *Econometrica* **16**(1), 1–32.
- Nguyen, H.-L. Q. (2019), 'Are credit markets still local? Evidence from bank branch closings', *American Economic Journal: Applied Economics* **11**(1), 1–32.
- Nichols, A. L. & Zeckhauser, R. J. (1982), 'Targeting transfers through restrictions on recipients', *The American Economic Review* **72**(2), 372–377.
- Oster, E., Shoulson, I. & Dorsey, E. (2013), 'Optimal expectations and limited medical testing: Evidence from Huntington disease', *American Economic Review* **103**(2), 804–30.
- Raval, D. & Rosenbaum, T. (2018), Why is distance important for hospital choice? Separating home bias from transport costs.

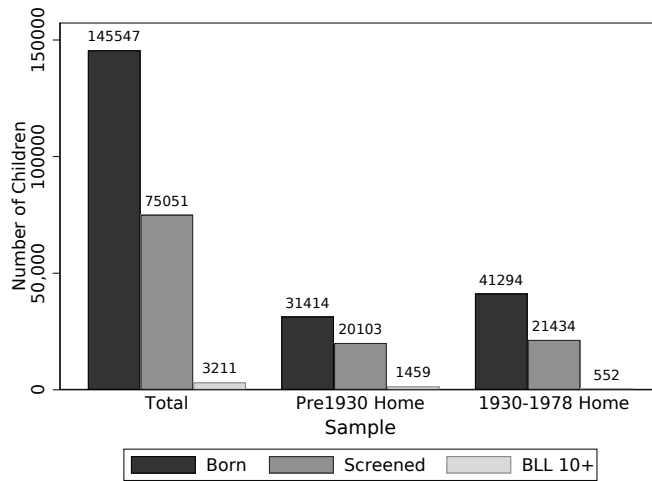


- Reyes, J. (2007), ‘Environmental policy as social policy? The impact of childhood lead exposure on crime’, *The BE Journal of Economic Analysis & Policy* .
- Reyes, J. (2015a), ‘Lead policy and academic performance: Insights from Massachusetts’, *Harvard Educational Review* **85**(1), 75–108.
- Reyes, J. W. (2015b), ‘Lead exposure and behavior: Effects on antisocial and risky behavior among children and adolescents’, *Economic Inquiry* **53**(3), 1580–1605.
- Rossin-Slater, M. (2013), ‘Wic in your neighborhood: New evidence on the impacts of geographic access to clinics’, *Journal of Public Economics* **102**, 51–69.
- Schnell, M. & Currie, J. (2018), ‘Addressing the opioid epidemic: Is there a role for physician education?’, *American journal of health economics* **4**(3), 383–410.
- Schwartz, J. (1994), ‘Societal benefits of reducing lead exposure’, *Environmental Research* **66**, 105–124.
- Silver, D. (2019), Haste or waste? Peer pressure and the distribution of marginal returns to health care.  
**URL:** <https://www.dropbox.com/s/ril7ugyvenb0ogk/jmp.pdf?dl=0>
- Vaidyanathan, A., Staley, F., Shire, J., Muthukumar, S., Kennedy, C., Meyer, P. A. & Brown, M. J. (2009), ‘Screening for lead poisoning: a geospatial approach to determine testing of children in at-risk neighborhoods’, *The Journal of pediatrics* **154**(3), 409–414.
- Van Parys, J. (2016), ‘Variation in physician practice styles within and across emergency departments’, *PloS one* **11**(8), e0159882.
- Venator, J. & Fletcher, J. (2019), Undue burden beyond Texas: An analysis of abortion clinic closures, births, and abortions in Wisconsin, Working Paper 26362, National Bureau of Economic Research.

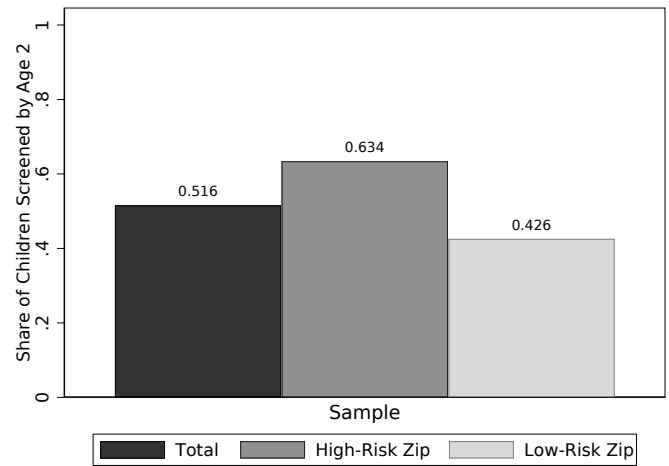
- Vivier, P. M., Hogan, J. W., Simon, P., Leddy, T., Dansereau, L. M. & Alario, A. J. (2001), ‘A statewide assessment of lead screening histories of preschool children enrolled in a medicaid managed care program’, *Pediatrics* **108**(2), e29–e29.
- Walker, R., Norbeck, T., Price, G., Libby, R. & Jones, P. (2016), Survey of America’s physicians: Practice patterns & perspectives, Technical report.
- Weber, S. & Peclat, M. (2016), ‘GEOROUTE: Stata module to calculate travel distance and travel time between two addresses or two geographical points’, Statistical Software Components, Boston College Department of Economics.
- Young, A., Chaudhry, H. J., Pei, X., Halbesleben, K., Polk, D. H. & Dugan, M. (2015), ‘A census of actively licensed physicians in the United States, 2014’, *Journal of Medical Regulation* **101**(2), 7–22.

Figure 1: Lead Screening and Exposure Rates in the Illinois 2014 Cohort

(a) Number of Children Born, Screened, and with BLLs 10+

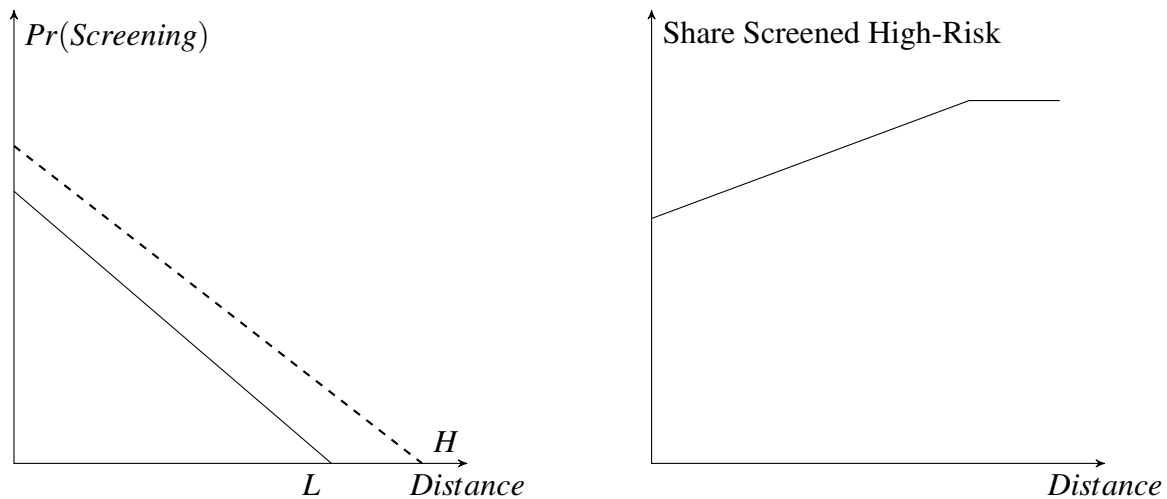


(b) Screening Rates by Zip Code Risk



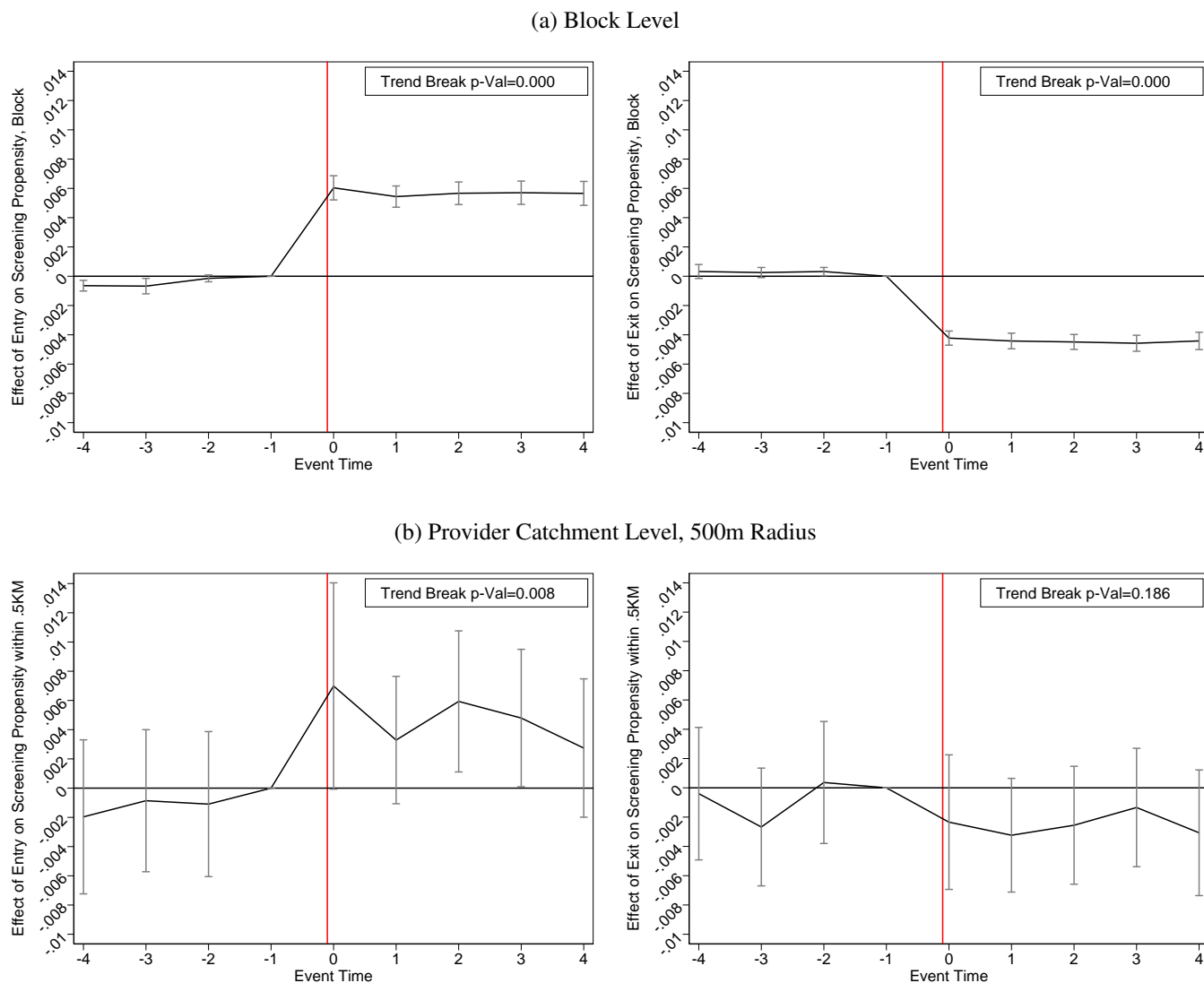
Notes: The figure plots screening and exposure rates for children born in Illinois in 2014. Panel A plots the number of children born, screened, and with blood lead levels (BLLs)  $\geq 10\mu\text{g}/\text{dL}$  in the whole sample and for the sample of children in pre-1930 and 1930-1978 homes. Panel B plots screening rates by age two by risk-level of the birth zip code.

Figure 2: Relationship between Distance and Screening Rates, by Risk



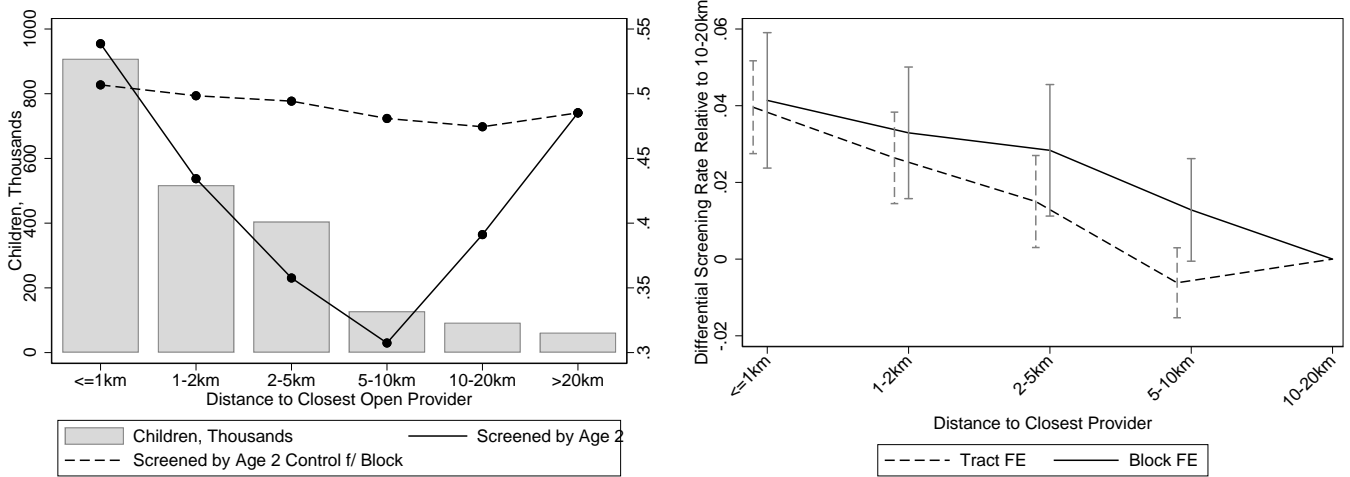
Notes: The figure illustrates the screening predictions from the ordeals model. The left panel plots hypothetical screening rates by distance for low risk (L) and high risk (H) households. The right panel plots the share of screened children who are high risk by distance as implied by the relationships plotted in the left panel.

Figure 3: Year-by-Year Effects of Openings and Closings



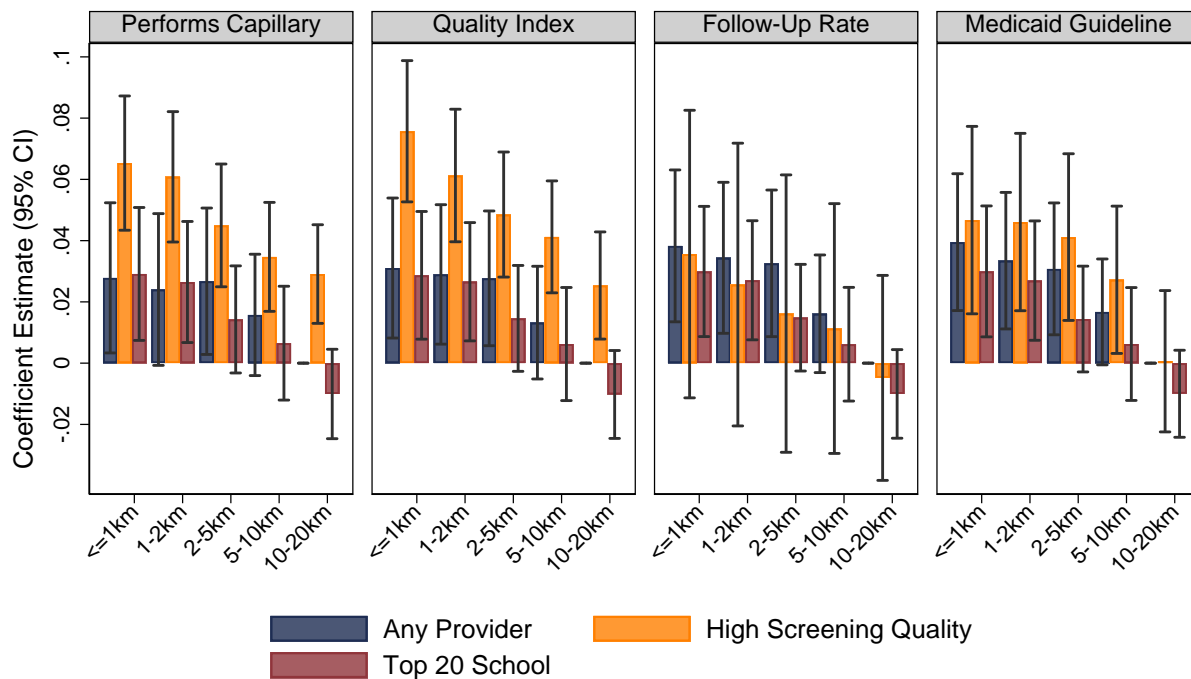
Notes: The figure plots DD coefficients on year-by-year entry and exit dummies, at the block (Panel A) level, and provider catchment with 500 meter radius level (Panel B). At the block level, entries and exits are defined as changes in distance from the area centroid to the closest provider. The outcome variable is the predicted screening probability of a child born in each level-year (levels are block and provider catchment), as predicted by a linear probability model including distance to closest open provider, level and birth cohort fixed effects. Coefficients on entry and exit in each panel are estimated in a single regression. The vertical line indicates the entry or exit period. For areas with entries or exits the sample is limited to a balanced panel in the  $[-4, 4]$  window around the entry or exit. Year and block or provider fixed effects are included. T-1 is the omitted category. The vertical bars are 95 percent confidence intervals. Standard errors are clustered at the zip code and provider level, respectively.

Figure 4: Determinants of Screening: Distance to Providers



Notes: The figure plots the average likelihood of a child being screened by age two by distance to closest open provider. The bars in the left panel show the number of children in each distance bin on the left y-axis, and the lines represents their screening rates on the right y-axis: the solid line plots raw means, while the dashed line plots residualized screening rates after controlling for Census block fixed effects. The right panel plots screening rates for each distance bin relative to children born 10 to 20 kilometers away from open providers controlling for tract fixed effects (dashed line) and block fixed effects (solid line), with vertical bars indicating 95% confidence intervals based on standard errors clustered at the zip code level.

Figure 5: Determinants of Screening: Provider Quality



Notes: The figure plots the effect of having any provider (blue bars), a high-quality provider based on the definition in each panel (orange bars) and a provider who attended a top 20 medical school (maroon bars) within each concentric buffer indicated on the x-axis on screening take-up. The quality index includes screening rates in a provider's catchment area, as well as a provider's rate of follow up within 90 days on cases of EBLs and a provider's rate of adherence to Medicaid guidelines, that is the rate at which children on Medicaid screened by that provider at age one have a second test at age two. Providers' catchment areas are computed based on the median distance of children to their screening providers in my sample. Within catchment areas, I compute provider-level screening rates by weighting unscreened children by the inverse of their distance to the provider. The sample includes all geocoded children born 2001-2014 whose closest provider is within 20 kilometers. Each regression includes child-level controls, as well as birth year and block fixed effects. Vertical bars display 95% confidence intervals based on standard errors clustered at the zip code level.

Table 1: Determinants of Screening: Distance to Provider

	(1)	(2)	(3)	(4)	(5)
<i>Panel A: Reduced Form. Dependent Variable: Screened by Age 2</i>					
Distance to Closest Open Provider	-0.008*** (0.001)	-0.004*** (0.000)	-0.004*** (0.001)	-0.004*** (0.001)	-0.003*** (0.001)
Mean Outcome Variable	0.46	0.46	0.46	0.46	0.47
N	2050536	2050553	2050533	2018383	1463352
<i>Panel B: First Stage. Dependent Variable: Distance Travelled</i>					
Distance to Closest Open Provider	0.366*** (0.035)	0.281*** (0.029)	0.166*** (0.028)	0.075** (0.032)	0.082* (0.048)
F-statistic	111.31	92.65	35.58	5.38	2.92
Mean Outcome Variable	7.74	7.74	7.74	7.54	7.25
N	585587	585638	585544	543689	312669
<i>Panel C: Second Stage. Dependent Variable: Screened by Age 2</i>					
Predicted Distance to Choice Provider	-0.023*** (0.001)	-0.013*** (0.001)	-0.022*** (0.002)	-0.073*** (0.019)	-0.147 (0.265)
Mean Outcome Variable	0.46	0.46	0.46	0.53	0.76
N	2047415	2048444	2040705	1527937	697536
Zip Code FE	X				
Tract FE		X			
Block Group FE			X		
Block FE				X	
Home FE					X

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . The table displays the impact of distance to providers on the likelihood of a child being screened by age two. The sample includes all geocoded children born 2001-2014 whose closest provider is within 20 kilometers. Panel A reports the effect of distance to the closest provider open during a child's birth year in kilometers. Panel B estimates the impact of distance to the closest provider open during the year of a test on the actual distance travelled to get the child's first test for the sample of screened children. Panel C estimates the effect of predicted distance from provider of choice on the likelihood a child is screened in the whole sample, where each column predicts distance from provider of choice using the first stage in Panel B in that column. Each column includes birth year fixed effects and a set of location fixed effects as indicated at the bottom of each column. Panels A and B include standard errors clustered at the zip code level in parentheses. Panel C includes standard errors from predictions bootstrapped in 500 iterations.

Table 2: Determinants of Screening: Information

Dependent Variable:	Screened by Age 2					
	(1)	(2)	(3)	(4)	(5)	(6)
Distance to Closest Open Provider	-0.003*** (0.000)	-0.003*** (0.001)	-0.005*** (0.001)	-0.003*** (0.000)	-0.002*** (0.001)	-0.004*** (0.001)
Share Tested in Neighborhood-Cohort	0.427*** (0.017)	0.171*** (0.011)	-0.109*** (0.005)			
Distance to Closest Open Provider X Born after Tested Sibling					-0.006*** (0.002)	
Distance to Closest Open Provider X Born after Sibling with BLL 10+ Born after Tested Sibling					0.007*** (0.003) -0.553*** (0.009)	
Born after Sibling with BLL 10+					-0.022* (0.012)	
Distance to Closest Open Provider X BLL 10+ within a Year of Birth within 15m BLL 10+ within a Year of Birth within 15m						0.012*** (0.001) 0.058*** (0.004)
Mean Outcome Variable	0.46	0.46	0.46	0.46	0.46	0.46
N	2050429	2049377	1404918	1052305	1052305	2018383
Children in Neighborhood-Cohort	69.02	26.00	4.95			
Tract FE	X					
Block Group FE		X				
Block FE			X			X
Mother FE				X	X	

Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table displays the impact of distance to the closest provider open during a child birth year on the likelihood a child is screened by age two. Columns 1-3 control for the share of screened children born in a child's neighborhood-cohort, with neighborhood defined at the bottom of each column. Columns 4-5 limit the sample to children with siblings and control for mother and birth order fixed effects. Column 6 controls for neighboring cases of blood lead level (BLL) above the intervention threshold of  $10\mu\text{g}/\text{dL}$ . Standard errors clustered at the zip code level in parentheses.



Table 3: Selection into Screening Conditional on Distance

Dependent Variable:	BLL 10+ By Age 2 (1)	BLL By Age 2 (2)	Home Pre1930 (3)	Black (4)	Hispanic (5)	Single Mother (6)	Mother 20 or Younger (7)	Mother High School or Less (8)
<i>Panel A: Tract and Year FE</i>								
Distance to Closest Open Provider	-0.0003** (0.000)	-0.0052* (0.003)	-0.0047*** (0.001)	-0.0023*** (0.000)	-0.0021*** (0.001)	-0.0019*** (0.001)	-0.0002 (0.000)	-0.0008 (0.001)
<i>Panel B: Block and Year FE</i>								
Distance to Closest Open Provider	-0.0004 (0.000)	-0.0084 (0.007)	-0.0012** (0.001)	-0.0004 (0.000)	-0.0002 (0.001)	0.0006 (0.001)	0.0002 (0.001)	-0.0004 (0.001)
Mean Outcome Variable	0.02	2.93	0.46	0.23	0.33	0.49	0.12	0.16
N	890091	890091	645177	890091	890091	890091	890091	890091

Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table displays the impact of distance to the closest provider open during a child's birth year on selection into screening by age 2. The sample includes all geocoded children born 2001-2014 whose closest provider is within 20 kilometers and who are screened. Outcome variables are indicated in each column. BLL is short for blood lead level. Panel A reports the effects controlling for the child's birth tract, Panel B controls for child's birth block. Each regression includes birth year fixed effects. Standard errors clustered at the zip code level in parentheses.

Table 4: Effect of Proximity to Providers on EBLL Detection, Detection Timing, and Prevention

Dependent Variable:	BLL 10+ Detected (1)	Age at First Test (2)	Age at Highest Test (3)	Remediation within 3 Years (4)	Future BLL 10+ Detected (5)
Distance to Closest Open Provider	-0.000269*** (0.000)	0.193423*** (0.051)	0.181126*** (0.049)	-0.000007 (0.000)	0.000024 (0.000)
Mean Outcome Variable	0.009	20.434	21.325	0.001	0.017
N	2018383	1194748	1194748	2018383	421920
Block FE	X	X	X	X	X

Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table displays the impact of distance to the closest provider open during a child birth year on the outcome indicated in each column. BLL is short for blood lead level. The sample includes all geocoded children born 2001-2014 whose closest provider is within 20 kilometers. Each column includes birth year and block fixed effects. Standard errors clustered at the zip code level in parentheses.

Table 5: Heterogeneity in Willingness to Pay for Screening

Sample:	All	Home Vintage			Black		Hispanic		Single Mother		Mother 20 or Younger	
	(1)	Pre1930 (2)	1930-1978 (3)	Post1978 (4)	No (5)	Yes (6)	No (7)	Yes (8)	No (9)	Yes (10)	No (11)	Yes (12)
<i>Panel A: Logit Marginal Effects</i>												
Distance to Closest Open Provider	-0.010*** (0.002)	-0.008*** (0.002)	-0.009*** (0.002)	-0.008*** (0.001)	-0.010*** (0.002)	-0.005** (0.002)	-0.010*** (0.002)	-0.004* (0.002)	-0.014*** (0.002)	0.001 (0.002)	-0.011*** (0.002)	0.002 (0.002)
<i>Panel B: OLS Coefficients</i>												
Distance to Closest Open Provider	-0.006*** (0.001)	-0.005*** (0.001)	-0.005*** (0.001)	-0.006*** (0.001)	-0.006*** (0.001)	-0.005*** (0.002)	-0.006*** (0.001)	-0.005*** (0.001)	-0.008*** (0.001)	0.000 (0.001)	-0.006*** (0.001)	0.001 (0.001)
<i>Panel C: Average Willingness to Pay</i>												
Average WTP (\$)	-5.258*** (0.001)	7.808*** (0.002)	-5.140*** (0.001)	-23.632*** (0.027)	-5.892*** (0.001)	5.017*** (0.002)	-6.468*** (0.001)	5.101*** (0.001)	-5.362*** (0.000)	1.218*** (0.000)	-3.436*** (0.000)	2.721*** (0.000)
P-value Equality w/ Ref		0.000	0.000		0.000		0.000		0.000		0.000	
Mean Outcome Variable	0.463	0.600	0.453	0.288	0.438	0.572	0.406	0.604	0.391	0.585	0.449	0.602
N	1451137	505167	578901	367069	1189347	261790	1036904	414233	916396	534741	1323733	127404

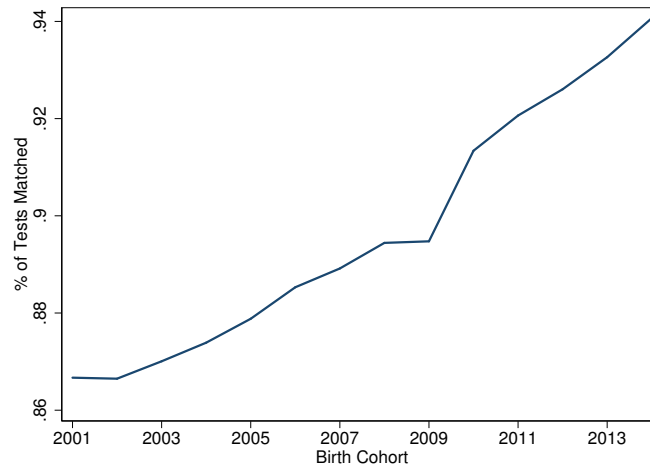
Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table displays the marginal effects of distance to providers on the likelihood of a child being screened by age two from logit (Panel A) and OLS (Panel B) models on different subsamples indicated in each column. Estimates for each set of columns, that is home vintages (Columns 2-4), race (Columns 5-6), ethnicity (Columns 7-8), mother's marriage status (Columns 9-10), and mother's age (Columns 11-12), are estimated in a single regression that interacts distance with the characteristic indicator in each column. Panel C reports average willingness-to-pay for screening for the average household in each subsample as estimated by the logit model in Panel A. The sample includes all geocoded children born 2001-2014 whose birth address matched a parcel record, and whose closest provider is within 20 kilometers and either opened or closed during their birth year. Each column includes birth year indicators, child-level demographic controls, and block-level averages of all included regressors. Standard errors clustered at the zip code level in parentheses.

Table 6: Policy Counterfactuals

Policy:	Household Incentive (1)	Provider Incentive (2)	Diffused Screening (3)	Pre1930 Screening Mandate (4)	Remediation Follow-Through (5)
Additional Children Screened, 1,000	15.91	50.70	0.88	11.31	
Additional BLLs 10+ Detected, 1,000	0.14	0.43	0.01	0.15	
Change in Private Welfare, \$1,000	370.09*** (3.55)	265.70*** (3.83)	9.92*** (0.34)	194.83*** (2.35)	
Prevention, \$1,000					393.72*** (4.70)
Cost, \$1,000	434.71	1774.47	7.02		6792.15
Private Benefit - Cost, Per Child with BLL 10+, \$1,000	-0.48	-3.50	0.39		

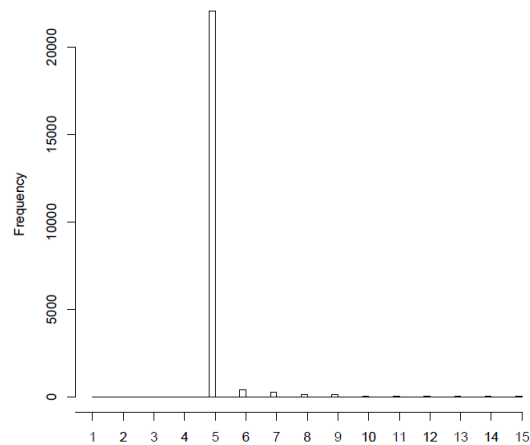
Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table displays the impact of the counterfactual policies in each column on the 2014 cohort. Additional cases detected are the product of additional children screened and the poisoning probability in the 2014 cohort (0.0085) except in Column 4 which uses the poisoning probability conditional on living in an old home (0.0131). The sum of the additional children's willingness-to-pay (WTP) yields the private benefits of each policy. WTP is estimated in a logit model that includes demographic and block-group level controls. Household incentives average \$10.5. Columns 1 and 3 count children whose WTP turns positive under the policy as additionally screened. Column 2 simulates increases in screening rates for low-screening providers in high-risk zip codes of 10 percentage points. Column 3 simulates providers opening at the zip code centroid for each zipcode-year cell without open providers, at \$7.96 per test. Column 4 assumes remediations in 638 homes with blood lead levels (BLLs)  $\geq 10\mu\text{g}/\text{dL}$  in 2014 prevent 66 new cases in the following ten years, at the baseline re-poisoning rate of 10.3 percent. Average remediation cost are \$10,646 per house.

Figure A.1: Match Rate between Blood Lead Levels and Birth Records



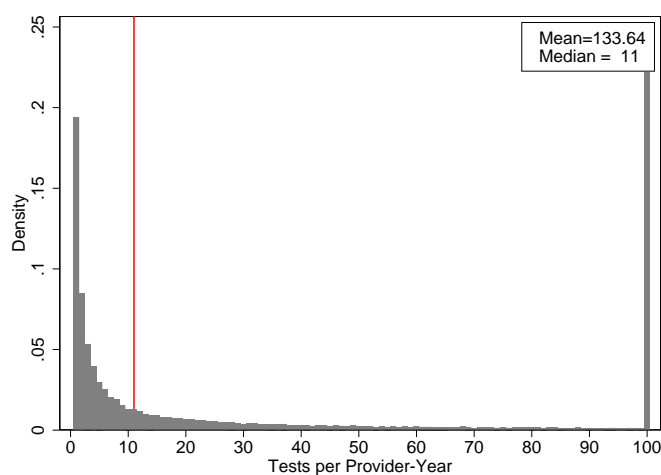
Notes: The figure plots the percent of tests successfully linked to birth records by birth cohort as recorded in the test data.

Figure A.2: Distribution of Test Results of Laboratory with Cutoff at  $5 \mu g/dL$



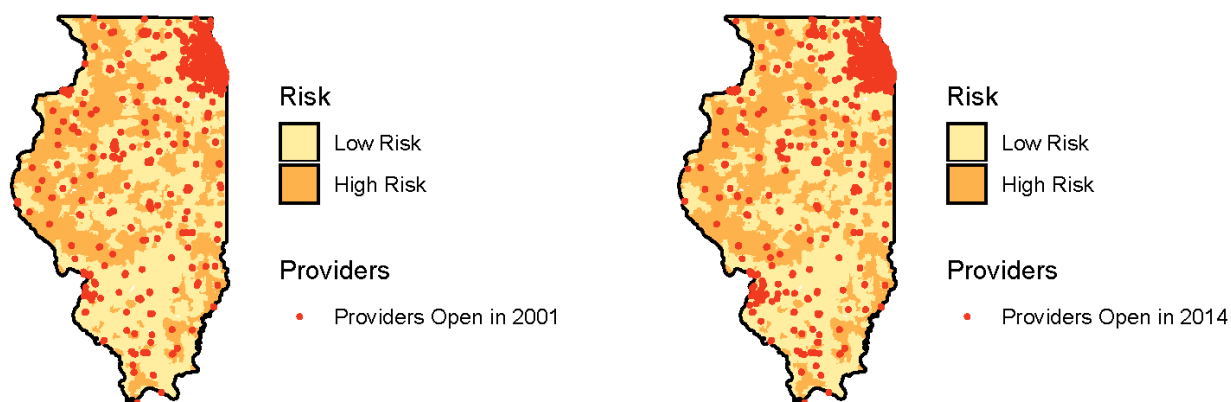
Notes: The figure plots the number of tests on the y-axis by blood lead level result on the x-axis for one laboratory in our sample.

Figure A.3: Distribution of Tests per Provider-Year



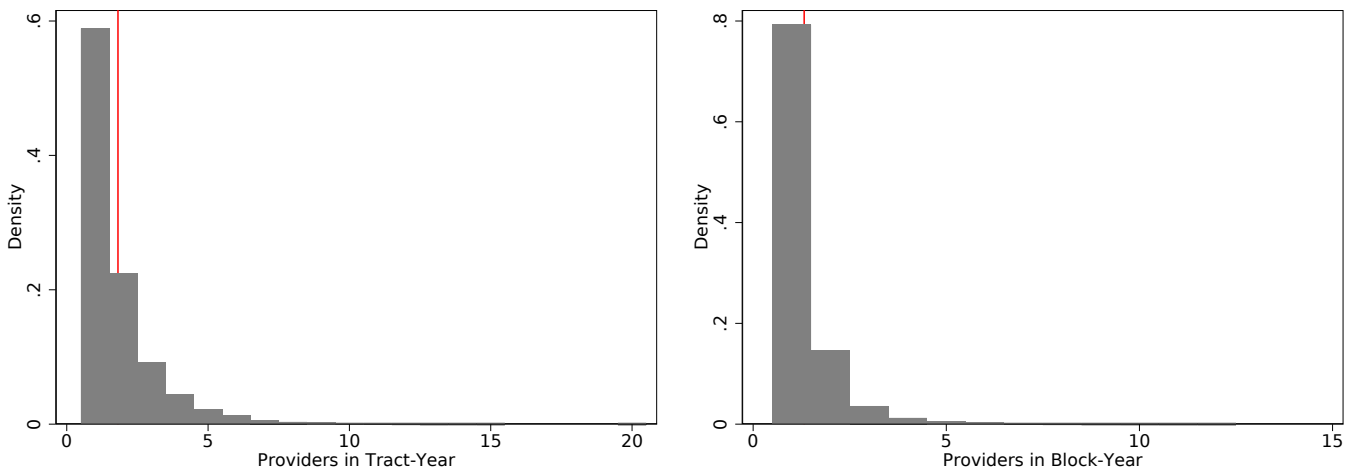
Notes: The figure plots the distribution of number of tests performed by providers in a year. The number of tests is censored at 100 for ease of visualization. The red vertical line indicates the median of the variable in the uncensored data.

Figure A.4: Location of Providers Operating in IL in 2001 and 2014



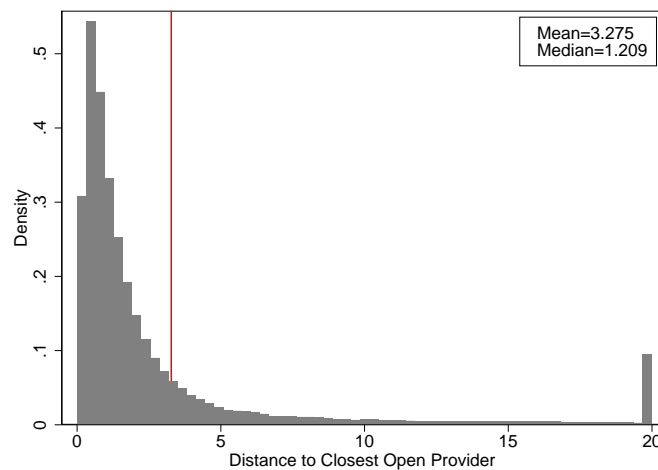
Notes: The figure plots the distribution of open providers in Illinois in high and low risk zip codes in the years 2001 (left) and 2014 (right).

Figure A.5: Distribution of Providers within Neighborhoods



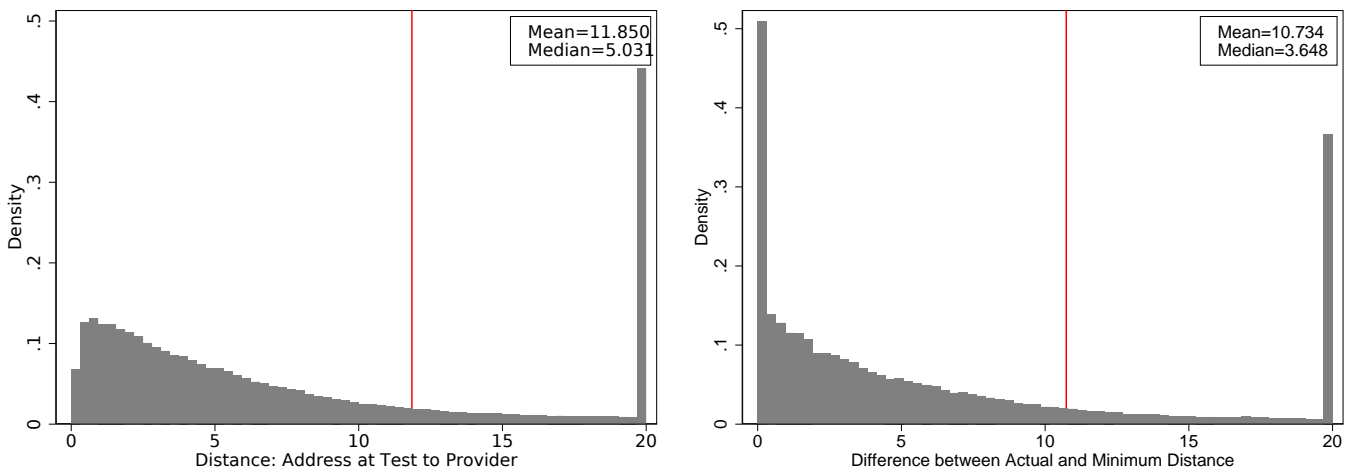
Notes: The figure plots the distribution of open providers at the tract-year (left panel) and block-year (right panel) level in Illinois, conditional on a neighborhood having a provider. The vertical red line indicates the mean number of provider in a neighborhood-year.

Figure A.6: Distance to Closest Providers



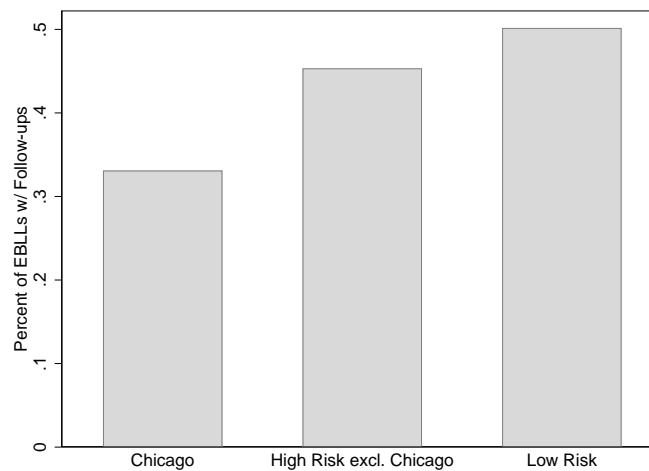
Notes: The figure plots the distribution of distance in kilometers from children's birth address to the closest provider open during the child's birth year. Distance is censored at 20km for ease of visualization. The red vertical line indicates the mean of the variable in the uncensored data.

Figure A.7: Distance to Providers



Notes: The left panel plots the distribution of distance in kilometers between children's address at test and the provider associated with the test. The right panel plots the distribution of the difference in kilometers between distance traveled at test and minimum distance between address at test and the closest active provider during the test's year. In both graphs, distance is censored at 20km for ease of visualization. The red vertical line indicates the mean of the variable in the uncensored data.

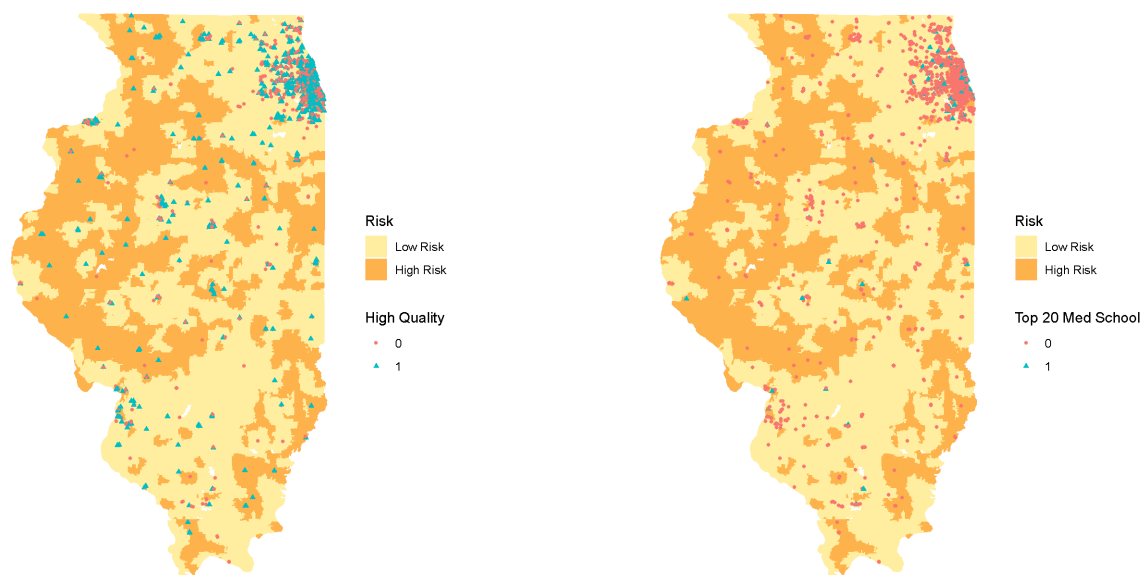
Figure A.8: Follow-up Rates



Notes: The figure plots follow-up rates in IL for tests that identify an elevated blood lead level (EBLL) by risk-level in birth zip code.

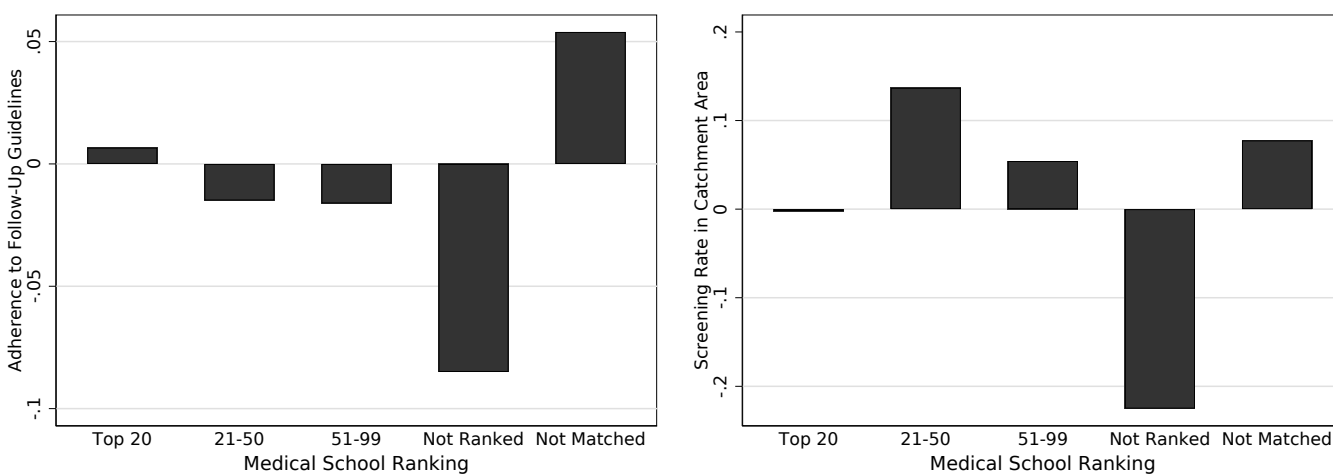


Figure A.9: Location of Providers, by Quality



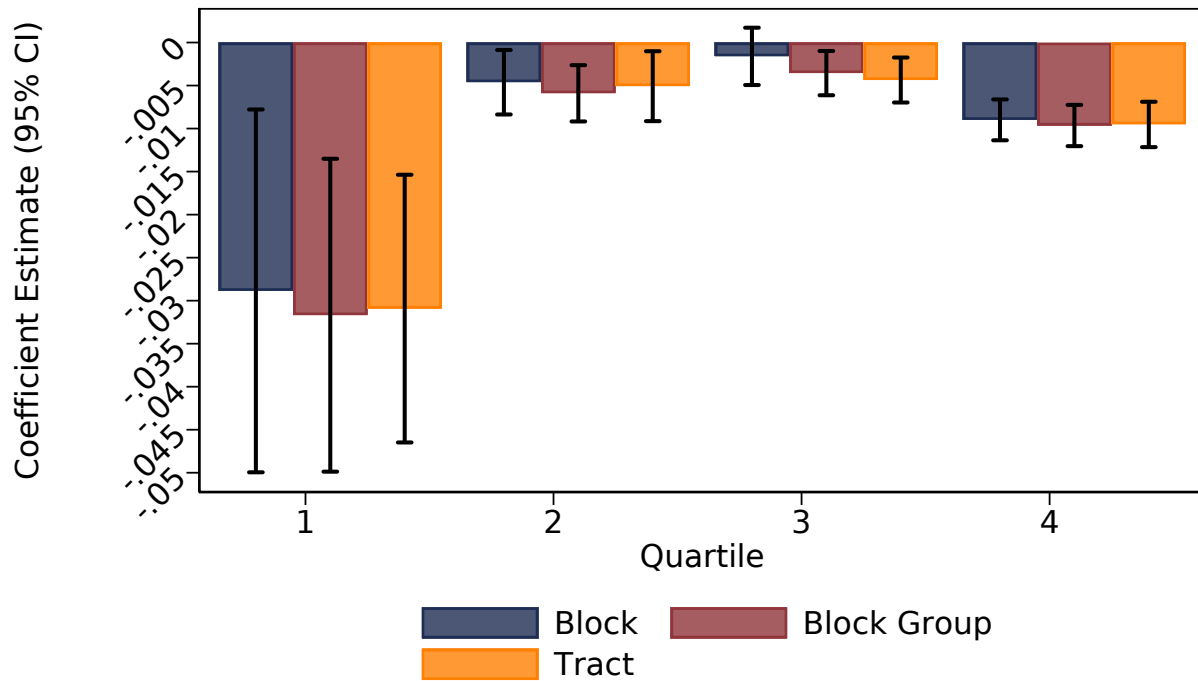
Notes: The figure plots the distribution of open providers by quality (left panel) and ranking of medical school of record (right panel) in high and low risk zip codes over the years 2001-2014. High-quality providers are defined as having a quality index above median.

Figure A.10: Providers: Correlation in Quality Measures



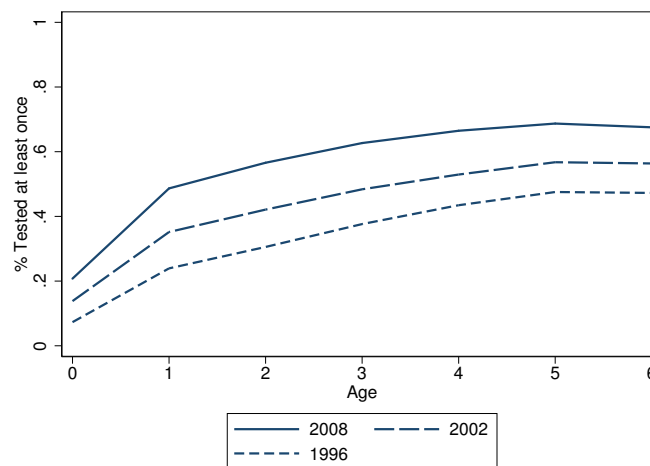
Notes: The figure plots on the y-axis the average z-scores of adherence to follow-up guideline (left panel) and screening rate (right panel) by ranking of the medical school each provider earned their degrees at on the x-axis.

Figure A.11: Determinants of Screening: Providers Distance, by Car Ownership



Notes: The figure plots the regression estimates of the effect of distance to closest provider on the likelihood that a child is screened by age two by quartile of car ownership rates in the child's Census tract, controlling for tract, block group, or block fixed effects. Tract-level car ownership in 2000 is measured in Census data. Vertical bars indicate 95% confidence intervals based on standard errors clustered at the zip code level.

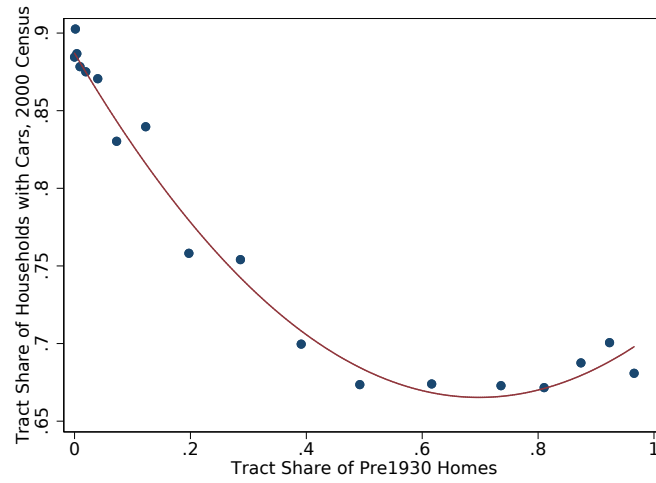
Figure A.12: Cumulative Distribution of Age at First Blood Lead Test



Notes: The figure plots the cumulative distribution of age of first test in Illinois over time.

---

Figure A.13: Correlation between Car Ownership and Housing Age



Notes: The figure plots the average car ownership rates by quantiles of share of pre-1930 homes in Census tract using 2000 Census data, and fits a quadratic line.

Table A.1: Summary Statistics: Children

Sample:	Whole Sample (1)	Screened Children (2)	Children Switching Providers (3)	Siblings Switching Providers (4)
Home Pre1930	0.348 (0.476)	0.451 (0.498)	0.539 (0.498)	0.508 (0.500)
Home 1930-1977	0.399 (0.490)	0.391 (0.488)	0.362 (0.481)	0.379 (0.485)
Low Income	0.278 (0.448)	0.366 (0.482)	0.464 (0.499)	0.436 (0.496)
Black	0.179 (0.383)	0.226 (0.418)	0.318 (0.466)	0.307 (0.461)
Hispanic	0.246 (0.431)	0.319 (0.466)	0.407 (0.491)	0.366 (0.482)
Single Mother	0.384 (0.486)	0.490 (0.500)	0.617 (0.486)	0.585 (0.493)
Mother 20 or Younger	0.091 (0.287)	0.119 (0.324)	0.170 (0.376)	0.135 (0.342)
Mother Less than High School	0.012 (0.109)	0.019 (0.137)	0.020 (0.139)	0.018 (0.131)
Mother High School, No Diploma	0.103 (0.304)	0.135 (0.342)	0.183 (0.386)	0.173 (0.378)
BLL 10+ within a Year of Birth within 15m	0.054 (0.226)	0.079 (0.269)	0.118 (0.322)	0.102 (0.302)
BLL 10+ within a Year of Birth 15-100m	0.104 (0.305)	0.139 (0.346)	0.193 (0.395)	0.174 (0.379)
Chicago Born	0.283 (0.450)	0.380 (0.485)	0.446 (0.497)	0.413 (0.492)
High Risk Zip excl. Chicago	0.169 (0.375)	0.204 (0.403)	0.213 (0.409)	0.216 (0.411)
Screened by Age 2	0.456 (0.498)	1.000 (0.000)	0.865 (0.341)	0.713 (0.452)
Highest BLL by Age 2	2.919 (2.596)	2.919 (2.596)	3.408 (3.279)	3.243 (3.029)
BLL 10+ by Age 2	0.020 (0.140)	0.020 (0.140)	0.035 (0.184)	0.030 (0.169)
Distance to Closest Open Provider	2.279 (3.195)	1.934 (3.004)	1.627 (2.576)	1.754 (2.757)
Has Provider w/ Capillary in 1Km	0.308 (0.462)	0.382 (0.486)	0.424 (0.494)	0.400 (0.490)
Has High Quality Provider in 1Km	0.295 (0.456)	0.374 (0.484)	0.424 (0.494)	0.398 (0.489)
Has Provider w/ Top 20 Degree in 1Km	0.033 (0.178)	0.039 (0.193)	0.040 (0.197)	0.040 (0.195)
N	2050536	934099	391985	705976

Notes: The table displays summary statistics for the covariates in the sample. BLL is short for blood lead level. Column 1 include all geocoded children whose birth address matched a parcel record for birth cohorts 2001-2014. Column 2 limits the sample to children whose birth address is within 2 kilometers of a provider opening or closing during their birth year. Column 3 limits the sample to children with multiple tests who switch providers across tests. Column 4 limits the sample to children in households with multiple siblings who switch providers across siblings.

Table A.2: Summary Statistics: Matched and Unmatched Blood Lead Tests

Sample:	Unmatched		Matched	
	Mean (1)	Standard Deviation (2)	Mean (3)	Standard Deviation (4)
Geocoded	0.729	0.445	0.762	0.426
Home Pre1930	0.540	0.498	0.545	0.498
Home 1930-1977	0.350	0.477	0.349	0.477
Low Income Block Group	0.458	0.498	0.457	0.498
Share Black in Tract	0.262	0.352	0.283	0.372
Share Hispanic in Tract	0.226	0.291	0.243	0.300
Fraction Less than High School	0.561	0.184	0.577	0.176
Chicago Residence	0.428	0.495	0.468	0.499
N	715273		4707326	

Notes: The table displays summary statistics for the unmatched (Columns 1-2) and matched (Columns 3-4) tests in the sample. Housing age and Census characteristics of block group and tracts are based on the child's address at time of test.

Table A.3: Sample Size and Linkages

	Tests Linked to Test Address		Test Linked to Birth Address		Children with Birth Records
	# Tests (1)	# Children (2)	# Tests (3)	# Children (4)	# Children (5)
Total	5,403,722	2,653,402	5,403,722	2,653,402	4,465,487
Matched to Birth Record	4,692,618	2,166,694	4,685,569	2,160,081	4,465,487
Geocoded	3,587,020	1,820,517	4,167,897	1,903,385	3,847,728
Born between 2001-2014	2,664,302	1,392,758	2,935,018	1,281,933	2,123,496
Linked to Parcel Data	1,926,388	1,007,129	2,144,859	890,637	1,466,015
Drop follow-up	1,851,106	1,004,026	2,064,753	890,637	1,466,015
Linkage with Census Block Data	1,850,783	1,003,859	1,722,482	780,980	1,465,336

Notes: The table displays the number of tests and unique children in my original sample (first row) and those remaining after each data cleaning and linkage step.

Table A.4: Screening Rates and Average Blood Lead Levels

	Illinois		Chicago	
	Geocoded	Non-Geocoded	Geocoded	Non-Geocoded
Screening Rate (%)	60%	58%	76%	74%
Avg. Blood Lead Level (ug/dL)	2.55	2.52	2.40	2.39

Notes: The table displays the screening rates and average blood lead levels in Illinois and Chicago, respectively, in the sample of geocoded (Columns 1 and 3) and non-geocoded (Columns 2 and 4) births (for screening rates) and tests (for average blood lead levels).

Table A.5: Sample Size and Extent of Lead Exposure

	Number of Tests, Excl. Follow-Up (1)	Number of Tests, Excl. Follow-Up, Linked to Covariates (2)	Number of Children (3)
<i>Panel A: Any Test Type</i>			
Total	2,557,184	1,594,313	953,749
Elevated (>10ug/dL)	77,919	37,310	27,175
Confirmed Elevated	70,171	32,319	22,579
<i>Panel B: Capillary Tests</i>			
Total	990,734	729,945	512,185
Elevated (>10ug/dL)	25,463	15,384	14,125
Confirmed Elevated	17,715	10,393	11,305
<i>Panel C: Venous Tests</i>			
Total	1,566,449	864,367	538,225
Elevated (>10ug/dL)	52,456	21,926	14,827

Notes: The table displays the number of tests (Column 1), number of tests excluding those that are within 90 days of a previous test (Column 2), and the number of children (Column 3) in my sample (Total) and those that display elevated levels, for any test (Panel A), capillary (Panel B), and venous (Panel C). I show separately the number of confirmed capillary tests, that is capillary tests that are followed up by another elevated level within 90 days, be it venous or capillary.



Table A.6: Estimates of Fraction of Movers in Sample

Sample:	Siblings (1)	Tested Child (2)
Moved	0.356	0.331
Moved to House with Different Risk	0.154	0.150
Moved to Zip Code with Different Risk	0.082	0.069
N	480865	883816

Notes: The table displays the share of households estimated to move within a two year period in my sample. Column 1 identifies movers among households with multiple children as those with a change in birth address between births. Column 2 identifies movers among households with a tested child as those whose residence address at time of test differs from the birth address. Houses are defined as having different risk if one is built before 1930 and one after.

Table A.7: Summary Statistics: Providers

	All Providers (1)	Opening Providers (2)	Closing Providers
Years Open	8.172 (6.051)	4.648 (3.697)	11.274 (6.391)
Individual Provider	0.242 (0.428)	0.220 (0.414)	0.180 (0.384)
Top 20 Degree	0.029 (0.168)	0.028 (0.165)	0.029 (0.169)
Top 21-50 Degree	0.174 (0.379)	0.151 (0.359)	0.190 (0.392)
Unranked Degree	0.685 (0.465)	0.709 (0.454)	0.649 (0.477)
Performs Capillary	0.636 (0.481)	0.541 (0.498)	0.780 (0.414)
Fraction Years Accepting New Patients	0.500 (0.404)	0.408 (0.406)	0.608 (0.369)
High Quality	0.703 (0.457)	0.758 (0.429)	0.667 (0.471)
N	4542	2060	2189

Notes: The table displays summary statistics for the providers in the whole sample (Column 1) and for providers who enter or exit between 2001 and 2014 (Columns 2-3, respectively).

Table A.8: Lagged Determinants of Providers' Entry and Exit, Neighborhood Level

Dependent Variable:	Entry	Exit	Distance To Closest Provider	Entry	Exit	Distance To Closest Provider
Neighborhood Level		Tract			Block	
	(1)	(2)	(3)	(4)	(5)	(6)
Number of Providers	-0.0445*** (0.009)	0.1794*** (0.012)	-0.1424*** (0.036)	-0.0509** (0.022)	0.2575*** (0.029)	-0.2172*** (0.080)
Number of Births	0.0001 (0.000)	0.0001 (0.000)	-0.0024 (0.002)	0.0000 (0.000)	0.0000 (0.000)	0.0014 (0.001)
Share Screened	0.0158 (0.012)	0.0126 (0.013)	-0.3001 (0.191)	0.0001 (0.000)	-0.0003 (0.000)	-0.0063 (0.012)
Average BLL	0.0005 (0.001)	0.0009 (0.002)	-0.0598** (0.027)	0.0000 (0.000)	0.0000 (0.000)	-0.0017 (0.001)
Share Homes Pre-1930	0.0152 (0.012)	-0.0052 (0.014)	-0.2432 (0.303)	-0.0003 (0.000)	0.0000 (0.000)	0.0135 (0.017)
Share Black	0.0122 (0.025)	0.0732*** (0.028)	0.1365 (0.243)	0.0003 (0.000)	0.0003 (0.000)	0.0024 (0.013)
Share Hispanic	0.0227 (0.021)	0.0087 (0.023)	-0.1596 (0.203)	-0.0001 (0.000)	0.0001 (0.000)	-0.0129 (0.009)
Share Single Mothers	0.0006 (0.015)	0.0150 (0.016)	-0.4121 (0.342)	0.0002* (0.000)	0.0000 (0.000)	-0.0232** (0.011)
Share Mothers 20 or Younger	-0.0486** (0.020)	-0.0159 (0.026)	0.3604 (0.392)	-0.0003** (0.000)	-0.0001 (0.000)	0.0139 (0.013)
Share Mothers High School or Less	0.0368** (0.019)	0.0368** (0.018)	0.0280 (0.247)	0.0000 (0.000)	0.0003 (0.000)	-0.0159 (0.011)
Mean Outcome Variable	0.0398	0.0535	2.8021	0.0005	0.0008	1.6101
N	32019	32019	32019	361900	361900	361830

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . The table displays the correlates of the likelihood that a provider opens (Columns 1,4) or closes (Columns 2,5) and average distance to providers (Columns 3,6) in a given year at different neighborhood levels. Observations in Columns 1-3 are at the tract-year level and in Columns 4-6 at the block-year level. Characteristics are lagged by one year, and all reflect births except for blood lead levels (BLLs) and number of providers. Each column includes year fixed effects and the neighborhood fixed effects indicated at the top of each column. Standard errors clustered at the neighborhood level in parentheses.

Table A.9: Determinants of Screening: Distance to Provider

Dependent Variable:	Screened by Age 2				
	(1)	(2)	(3)	(4)	(5)
Distance to Closest Open Provider	-0.008*** (0.001)	-0.004*** (0.001)	-0.004*** (0.001)	-0.004*** (0.001)	-0.003*** (0.001)
Mean Outcome Variable	0.46	0.46	0.46	0.46	0.47
N	2050535	2050553	2050533	2018383	1463352
Zip Code FE	X				
Tract FE		X			
Block Group FE			X		
Block FE				X	
Home FE					X

Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table displays the impact of the closest provider open during a child's birth year in kilometers on the likelihood of a child being screened by age two. The sample includes all geocoded children born 2001-2014 whose closest provider is within 20 kilometers. Each column includes birth year fixed effects and a set of location fixed effects as indicated at the bottom of each column. Standard errors clustered at the county level are in parentheses.

Table A.10: Determinants of Screening: Provider Access, Robustness Checks

Dependent Variable: Screened by Age 2	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Distance to Closest Open Provider	-0.0005** (0.000)	-0.0028*** (0.001)	-0.002*** (0.001)	-0.001*** (0.000)	-0.004*** (0.001)	-0.003*** (0.001)	-0.005*** (0.001)
Distance to Closest Open Provider X 20+km Away		0.0027*** (0.001)					
20+km Away		-0.0239 (0.015)					
Black						0.047*** (0.004)	0.051*** (0.005)
Hispanic						0.110*** (0.005)	0.110*** (0.005)
Single Mother						0.051*** (0.004)	0.042*** (0.004)
Mother 20 or Younger						0.016*** (0.002)	0.013*** (0.002)
Mother High School or Less						0.005 (0.003)	0.006* (0.003)
Home Pre1930							0.050*** (0.006)
Home 1930-1977							0.050*** (0.004)
BLL 10+ within a Year of Birth within 15m							0.067*** (0.005)
BLL 10+ within a Year of Birth 15-100m							0.014*** (0.003)
Mean Outcome Variable	0.46	0.46	0.46	0.46	0.46	0.46	0.46
N	2076225	2076225	2050533	2018383	2018351	2018383	1434900
Block FE	X	X		X	X	X	X
Block Group FE & Trend			X				
Distance Measure: Avg of 5 Closest Providers				X			
Distance Measure: From Block Centroid					X		

Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table displays the impact of distance to the closest provider open during a child birth year on the likelihood of a child being screened by age two. Columns 4 and 5 use different distance measures per the bottom of those columns. The sample includes all geocoded children born 2001-2014 whose birth address matched a parcel record. Columns 3-7 limit the sample to children within 20km of an open provider. BLL is short for blood lead level. Each column includes birth year fixed effects and location fixed effects per the bottom of each column. Standard errors clustered at the zip code level in parentheses.

Table A.11: Determinants of Screening: Car Travel Times

Dependent Variable:	Screened by Age 2			
	(1)	(2)	(3)	(4)
<i>Panel A: Travel Distance</i>				
Distance to Closest Open Provider	-0.008*** (0.001)	-0.004*** (0.001)	-0.004*** (0.001)	-0.005*** (0.002)
<i>Panel B: Travel Time</i>				
Travel Time to Closest Open Provider (Minutes)	-0.006*** (0.001)	-0.004*** (0.001)	-0.004*** (0.001)	-0.003** (0.001)
Mean Outcome Variable	0.46	0.46	0.46	0.46
N	244954	245018	244930	193050
Zip Code FE	X			
Tract FE		X		
Block Group FE			X	
Block FE				X

Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table examines the relationship between car travel times and screening likelihood in a 12 percent random sample stratified by block group and birth year. For this sample, I used the STATA command `georoute` (Weber & Peclat 2016), based on the HERE API which limits free requests to 250,000 observations, to estimate travel times by car to the closest open provider. Panel A estimates the impact of distance to the closest provider open on the likelihood a child is screened in this subsample. Panel B estimates the effect of travel time in minutes on the likelihood a child is screened. Each column includes year fixed effects and a set of location fixed effects for location indicated at the bottom of each column. Standard errors clustered at the zip code level in parentheses.

Table A.12: Determinants of Screening: Provider Access, by Zip Code Risk

Dependent Variable: Sample:	Chicago (1)	Screened by Age 2 High-Risk w/out Chicago (2)	Low-Risk (3)
<i>Panel A: Tract and Year FE</i>			
Distance to Closest Open Provider	-0.011** (0.005)	-0.002* (0.001)	-0.003*** (0.001)
<i>Panel B: Block and Year FE</i>			
Distance to Closest Open Provider	-0.008 (0.006)	-0.001 (0.002)	-0.003*** (0.001)
Mean Outcome Variable	0.61	0.55	0.34
N	576731	330241	1100179

Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table displays the impact of distance to the closest provider open during a child birth year on the likelihood of a child being screened by age two for different subsamples indicated in each column. The sample includes all geocoded children born 2001-2014 whose closest provider is within 20 kilometers. Each column includes birth year and block fixed effects. Standard errors clustered at the zip code level in parentheses.

Table A.13: Determinants of Screening: Provider Access, Different Samples

Dependent Variable:	Screened by Age 2				
Sample:	10KM	5KM	2KM	1KM	0.5KM
	(1)	(2)	(3)	(4)	(5)
Distance to Closest Open Provider	-0.006*** (0.001)	-0.009*** (0.002)	-0.018*** (0.004)	-0.027*** (0.008)	-0.030 (0.019)
Mean Outcome Variable	0.46	0.47	0.50	0.54	0.57
N	1933096	1809487	1407149	893976	412857
Block FE	X	X	X	X	X

Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table displays the impact of distance to the closest provider open during a child birth year on the likelihood of a child being screened by age two. The sample includes all geocoded children born 2001-2014 whose birth address matched a parcel record and who are born within the distance indicated at the top of each column from an open provider. Each column includes birth year and block fixed effects. Standard errors clustered at the zip code level in parentheses.



Table A.14: Determinants of Screening: Provider Access and Density

Dependent Variable:	Screened by Age 2		
	(1)	(2)	(3)
Distance to Closest Open Provider	-0.003*** (0.001)	-0.003*** (0.001)	-0.003*** (0.001)
Distance to 5 Closest Open Providers	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Mean Outcome Variable	0.46	0.46	0.46
N	2050535	2050515	2018367
Tract FE	X		
Block Group FE		X	
Block FE			X

Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table displays the impact of distance to the closest and the five closest providers open during a child birth year on the likelihood of a child being screened by age two. The sample includes all geocoded children born 2001-2014 whose birth address matched a parcel record within 20km of an open provider. Each column includes birth year fixed effects and location fixed effects per the bottom of each column. Standard errors clustered at the zip code level in parentheses.

Table A.15: Determinants of Screening: Provider Availability

Access Variable: Sample:	Accepts New Patients		Accepts Medicaid Patients		Accepts New & Medicaid Patients	
	All	Low Income	All	Low Income	All	Low Income
	(1)	(2)	(3)	(4)	(5)	(6)
Closest Open Provider within 1Km	0.018** (0.008)	0.019 (0.021)	0.025*** (0.009)	0.020 (0.021)	0.031*** (0.008)	0.031 (0.021)
Closest Open Provider 1-2Km	0.014* (0.008)	0.019 (0.021)	0.021** (0.008)	0.018 (0.021)	0.026*** (0.008)	0.028 (0.021)
Closest Open Provider 2-5Km	0.015* (0.008)	0.017 (0.023)	0.020** (0.008)	0.017 (0.023)	0.024*** (0.008)	0.025 (0.022)
Closest Open Provider 5-10Km	0.007 (0.007)	0.023 (0.021)	0.010 (0.007)	0.020 (0.021)	0.012* (0.007)	0.021 (0.020)
Closest Open Provider within 1Km, High Quality	0.054*** (0.009)	0.051*** (0.015)	0.060*** (0.008)	0.053*** (0.012)	0.051*** (0.007)	0.040*** (0.012)
Closest Open Provider 1-2Km, High Quality	0.046*** (0.008)	0.043*** (0.015)	0.052*** (0.007)	0.046*** (0.011)	0.044*** (0.006)	0.035*** (0.011)
Closest Open Provider 2-5Km, High Quality	0.035*** (0.008)	0.033** (0.014)	0.041*** (0.007)	0.038*** (0.011)	0.031*** (0.006)	0.025** (0.011)
Closest Open Provider 5-10Km, High Quality	0.019*** (0.007)	0.005 (0.015)	0.023*** (0.005)	0.011 (0.012)	0.018*** (0.005)	0.005 (0.012)
Closest Open Provider 10-20Km, High Quality	0.009 (0.006)	0.015 (0.015)	0.014** (0.005)	0.001 (0.013)	0.011** (0.005)	-0.002 (0.012)
Mean Outcome Variable	0.46	0.60	0.46	0.60	0.46	0.60
N	2018383	563938	2018383	563938	2018383	563938
Block FE	X	X	X	X	X	X

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . The table displays the impact of distance to the closest provider operating during a child birth year and distance to a provider possessing the characteristic indicated in each column on the likelihood of a child being screened by age two. The sample includes all geocoded children born 2001-2014 whose closest provider is within 20 kilometers (odd columns) or among those, only children living in low-income block groups (even columns). Each column includes birth year and block fixed effects. Standard errors clustered at the zip code level in parentheses.

Table A.16: Determinants of Screening: Provider Access, Logit Model

Dependent Variable: Specification:	Screened by Age 1		Screened by Age 2		Screened by Age 6	
	OLS (1)	Logit (2)	OLS (3)	Logit (4)	OLS (5)	Logit (6)
Distance to Closest Open Provider	-0.003*** (0.001)	-0.027*** (0.006)	-0.005*** (0.001)	-0.027*** (0.006)	-0.005*** (0.001)	-0.023*** (0.005)
Home Pre1930	0.037*** (0.004)	0.197*** (0.023)	0.050*** (0.006)	0.225*** (0.026)	0.063*** (0.006)	0.281*** (0.029)
Home 1930-1977	0.037*** (0.003)	0.203*** (0.019)	0.050*** (0.004)	0.226*** (0.021)	0.064*** (0.005)	0.280*** (0.022)
Black	0.024*** (0.004)	0.135*** (0.020)	0.051*** (0.005)	0.219*** (0.021)	0.094*** (0.005)	0.417*** (0.023)
Hispanic	0.089*** (0.005)	0.428*** (0.022)	0.109*** (0.005)	0.477*** (0.023)	0.127*** (0.005)	0.590*** (0.024)
Single Mother	0.029*** (0.003)	0.130*** (0.015)	0.042*** (0.004)	0.184*** (0.016)	0.050*** (0.004)	0.256*** (0.017)
Mother 20 or Younger	0.003 (0.002)	0.017* (0.010)	0.013*** (0.002)	0.060*** (0.009)	0.019*** (0.002)	0.132*** (0.012)
Mother Less High School or Less	0.002 (0.003)	-0.009 (0.014)	0.006* (0.003)	0.024* (0.014)	0.012*** (0.003)	0.091*** (0.017)
BLL 10+ within a Year of Birth within 15m	0.049*** (0.005)	0.252*** (0.020)	0.067*** (0.005)	0.311*** (0.022)	0.043*** (0.003)	0.265*** (0.020)
BLL 10+ within a Year of Birth 15-100m	0.009*** (0.003)	0.069*** (0.013)	0.014*** (0.003)	0.065*** (0.013)	0.013*** (0.002)	0.063*** (0.012)
Marginal Effect of Distance to Closest Open Provider		-0.006*** (0.001)		-0.007*** (0.001)		-0.005*** (0.001)
Mean Outcome Variable	0.32	0.32	0.46	0.46	0.61	0.61
N	1451137	1451137	1451137	1451137	1451137	1451137

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . The table displays OLS coefficients and coefficients and marginal effects from logit models of the impact of distance to the closest provider operating during a child birth year on the likelihood of a child being screened by age 1 (Column 1-2), age 2 (Column 3-4), and age 6 (Column 5-6). The sample includes all geocoded children born 2001-2014 whose birth address matched a parcel record, and whose closest provider is within 20 kilometers. BLL is short for blood lead level. Each column includes birth year indicators and block-level averages of all included regressors. Standard errors clustered at the zip code level in parentheses.

Table A.17: Selection into Screening Conditional on Distance: Robustness Checks

Dependent Variable:	BLL 10+ By Age 2 (1)	BLL By Age 2 (2)	Home Pre1930 (3)	Black (4)	Hispanic (5)	Single Mother (6)	Mother 20 or Younger (7)	Mother High School or Less (8)
<i>Panel A: Tract and Year FE</i>								
Distance to Closest Open Provider	-0.0003** (0.000)	-0.0044** (0.002)	-0.0043*** (0.001)	-0.0024*** (0.001)	-0.0026*** (0.001)	-0.0026*** (0.001)	0.0000 (0.000)	-0.0009** (0.000)
<i>Panel B: Block and Year FE</i>								
Distance to Closest Open Provider	-0.0001 (0.000)	-0.0003 (0.001)	0.0000 (0.000)	0.0001 (0.000)	0.0009** (0.000)	0.0010* (0.001)	0.0001 (0.000)	-0.0002 (0.000)
Mean Outcome Variable	0.02	2.99	0.46	0.24	0.38	0.48	0.12	0.16
N	697482	697482	645177	697482	697482	697482	697482	697482

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . The table displays the impact of distance to the closest provider open during a child birth year on selection into screening by age two. The sample includes all geocoded children born 2001-2014 whose closest provider is within 20 kilometers and who are screened. Outcome variables are indicated in each column. Panel A reports the effects controlling for the child's birth tract, Panel B controls for child's birth block. Each regression includes birth year fixed effects as well as tract or block level time-varying controls such as average blood lead levels (BLLs) by age 2, share of pre1930 homes, share black, share hispanic, share single mothers, share teen mothers, and share of mothers with high school education or less. Standard errors clustered at the zip code level in parentheses.

Table A.18: Effect of Proximity to Providers on Prevention, Robustness Checks for Rare Events

Specification:	Low Income Block (1)	Block with Remediation (2)	Logit (3)	Penalized Logit (4)
<i>Panel A: Remediation within 3 Years</i>				
Distance to Provider	0.000043 (0.000)	0.000079 (0.002)	-0.009173 (0.031)	-0.008720 (0.031)
Mean Outcome Variable	0.003	0.052	0.001	0.001
N	563938	54134	1636204	1636204
<i>Panel B: Future BLL 10+ Detected</i>				
Distance to Provider	-0.000717** (0.000)	-0.003795** (0.002)	0.008860 (0.011)	0.008919 (0.010)
Mean Outcome Variable	0.073	0.136	0.035	0.035
N	437433	43008	1199562	1199562

Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table displays the impact of distance to the closest provider open during a child birth year on the likelihood of remediation within three years (Panel A) and of future poisoning (Panel B). BLL is short for blood lead level. The sample includes all geocoded children born 2001-2014 whose closest provider is within 20 kilometers, with further constraints indicated in each column. Standard errors clustered at the zip code level in parentheses, except for Column 4 which reports standard errors under the assumption of homoscedasticity.

Table A.19: Determinants of Lead Exposure

Dependent Variable:	Highest BLL by Age 2		BLL 10+ by Age 2	
	(1)	(2)	(3)	(4)
Home Pre1930	0.415*** (0.025)	0.316*** (0.023)	0.010*** (0.001)	0.008*** (0.001)
Home 1930-1977	0.030* (0.016)	0.067*** (0.019)	-0.001* (0.001)	0.000 (0.001)
Low Income	0.003 (0.014)		-0.002*** (0.001)	
Black	0.255*** (0.050)	0.180*** (0.026)	0.003 (0.002)	0.002** (0.001)
Hispanic	-0.161*** (0.024)	-0.115*** (0.017)	-0.007*** (0.001)	-0.004*** (0.001)
Single Mother	0.025** (0.010)	0.026** (0.011)	0.001* (0.000)	0.001*** (0.001)
Mother 20 or Younger	0.038*** (0.014)	0.020 (0.015)	0.000 (0.001)	-0.001* (0.001)
Mother Less than High School	0.039 (0.028)	0.055* (0.031)	0.003*** (0.001)	0.005*** (0.001)
Mother High School, No Diploma	0.154*** (0.017)	0.149*** (0.017)	0.005*** (0.001)	0.005*** (0.001)
BLL 10+ within a Year of Birth within 15m	2.281*** (0.133)	2.078*** (0.135)	0.167*** (0.010)	0.157*** (0.010)
BLL 10+ within a Year of Birth 15-100m	0.231*** (0.024)	0.128*** (0.028)	0.008*** (0.001)	0.004** (0.002)
Mean Outcome Variable	2.97	2.99	0.02	0.02
N	671156	645177	671156	645177
Zip FE	X		X	
Block FE		X		X

Notes: \* $p < 0.10$ , \*\* $p < 0.05$ , \*\*\* $p < 0.01$ . The table displays estimates of the impact of various variables on a child's maximum blood lead level (Columns 1-2) and likelihood of having an elevated blood lead level (BLL) (Columns 3-4) by age two. The sample includes all geocoded children born 2001-2014 whose birth address matched a parcel record, and whose closest provider is within 20 kilometers. Each column includes birth year and block fixed effects. Standard errors clustered at the zip code level in parentheses.