Supporting Information for

Economic Consequences of Childhood Exposure to Urban Environmental Toxins

Dustin Frye, Gisella Kagy

This PDF file includes:

- Supporting text
- SI References

Supporting Information Text

1 Data Sources and Sample Construction

Figure 1 illustrates the data integration process for constructing our analysis sample. The figure reads left to right, from treatment assignment in childhood to outcome measurement in adulthood. On the left, hand-digitized historical sources on municipal water infrastructure (Baker, 1889 and 1897) and water chemistry (USGS) are linked by town name to the Census Place Project [Berkes et al., 2023], which provides geographic identifiers for individuals in the full-count census files. These town-level exposure measures are then matched via IPUMS HistIDs to individuals observed as children (ages 0–10) in the 1900, 1910, or 1920 Censuses (green boxes), assigning each child a town of residence and corresponding waterborne lead exposure classification. On the right, these childhood records are linked forward via HistIDs to the same individuals observed as adults (ages 20–50) in the 1940 Census (grey box), which provides the labor market outcomes we examine: earnings, employment, and occupation. The following subsections describe each data component in detail.

1.1 Individual-Level Census Data and Linking Methodology

To study the adult labor market consequences of childhood waterborne lead exposure, we use the full-count 1940 Census as the source of adult outcome data and then link these individuals backward to their childhood census records in 1900, 1910, or 1920. The 1940 Census is the first decennial census to report both wage and salary income and completed years of schooling, along with employment status, weeks worked, occupation, and other individual characteristics.

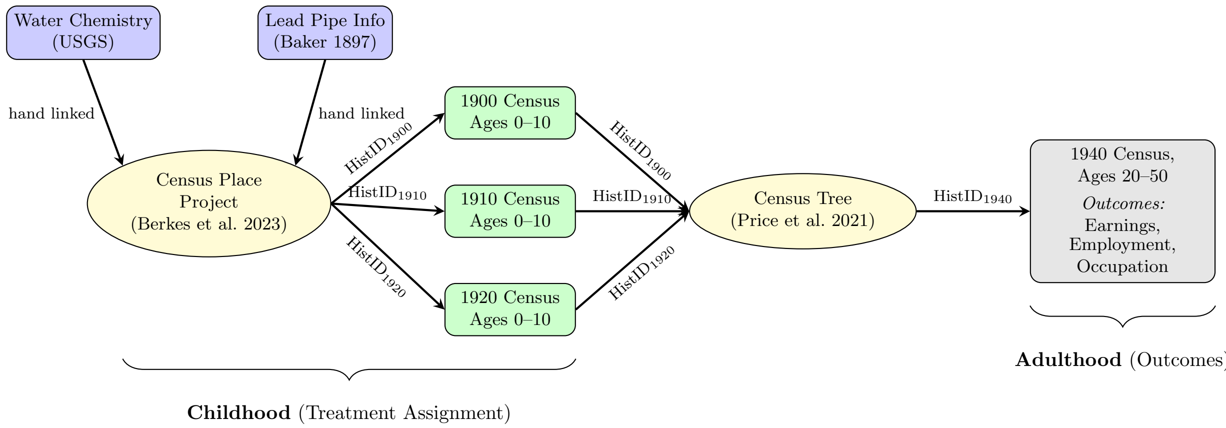


Figure 1: Data integration process showing town-level sources for waterborne lead exposure measures and linkage of 1940 Census cohort to childhood Census records.

Adult cohort. From the 1940 Census, we identify men ages 20–50, corresponding to cohorts still in or near prime working ages in 1940. We focus on men because wage and salary income captures a limited share of women's economic activity during this period. As shown in Figure 1, this 1940 cohort forms the core individual-level sample to which earlier childhood records are linked.

Linking to childhood census records. We then link individuals in the 1940 cohort backward to their earlier records in the 1900, 1910, or 1920 full-count Censuses, using crosswalks of HistIDs over time constructed from modern record-linkage methods [Price et al., 2021, Buckles et al., 2023a]. These earlier census records allow us to recover childhood town of residence and family background characteristics, including father's occupation. Because historical censuses do not contain stable person identifiers across decades, linked records are constructed using identifying characteristics that appear in multiple census waves, including names, ages, birthplaces, and other demographic information [Price et al., 2021, Buckles et al., 2023a].

All census files are drawn from the Integrated Public Use Microdata Series (IPUMS) full-count data [Ruggles et al., 2021]. The full-count files are especially important in our setting because many towns in our sample are small and would be sparsely represented, or absent altogether, in standard public-use census samples. Within each census cross-section, individuals are identified by IPUMS HistIDs, which provide the record-level identifiers used to reference linked observations across files.[1]

Assigning childhood exposure. Our measure of childhood exposure requires identifying the town in which each individual lived during early childhood. We therefore assign childhood location using the earliest linked census observation in which the individual is observed between ages 0 and 10. If an individual is linked in more than one childhood census within this age window, we use the earliest such observation. This procedure is designed to capture exposure during the period of greatest biological vulnerability to lead, when young children absorb it more efficiently

[1] HistIDs uniquely identify individuals within a given census file. Longitudinal connections across census years are established through the linked historical census files rather than by HistIDs alone.

than adults and their developing nervous systems are particularly susceptible to neurotoxicity [Hammond, 1982, Bellinger, 2004, Needleman, 2004], while also reducing measurement error from moves later in childhood or adolescence.[2]

Resulting retrospective panel. This procedure yields a retrospective panel linking adult labor market outcomes in 1940 to childhood location observed in one of three earlier censuses:

- Men born 1890–1900: ages 0–10 in the 1900 Census and ages 40–50 in 1940
- Men born 1900–1910: ages 0–10 in the 1910 Census and ages 30–40 in 1940
- Men born 1910–1920: ages 0–10 in the 1920 Census and ages 20–30 in 1940

Figure 2 reports linkage rates by age in the 1940 Census. Linkage rates range from 70–80% on average, which is in line with linkage rates reported in other literature [Buckles et al., 2025].

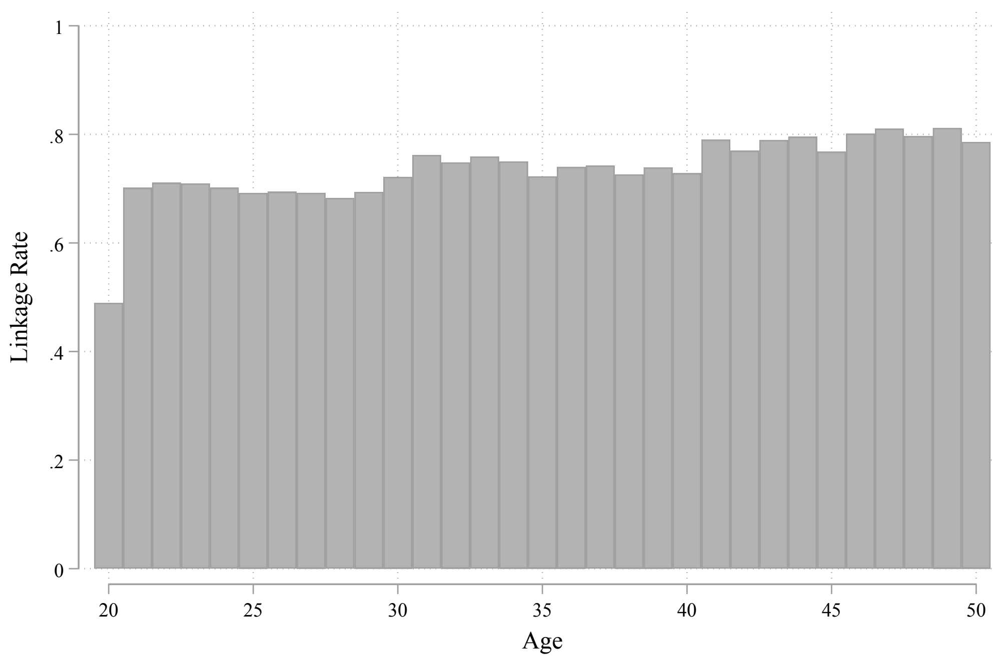


Figure 2: Linkage rates by age in the 1940 Census. Each point shows the share of men of a given age in the 1940 Census who are successfully linked to a childhood record in 1900, 1910, or 1920.

Geographic Identifiers and Town Matching. To link individuals' childhood locations to the town-level infrastructure and water chemistry data shown in Figure 1, we must identify each individual's town of residence in the historical census records. The standard IPUMS geographic identifiers include state, county, and, for larger incorporated places, city or town name. However, these standard identifiers do not capture all municipalities that appear in the original census records.

[2]We exclude individuals whose earliest linked census observation occurs after age 10. We also exclude cases with large inconsistencies in reported age across linked census records, defined as discrepancies greater than two years between age reported in the childhood census and age implied by the 1940 record.

In particular, many smaller towns are not represented as distinct places in the IPUMS geographic variables.

To improve coverage of small municipalities, we supplement the IPUMS geographic identifiers using data from the Census Place Project (CPP) [Berkes et al., 2023]. The CPP systematically extracts place names recorded in historical census manuscripts and assigns approximate geographic coordinates to these locations. Because CPP observations retain the underlying census record identifiers, they can be linked back to individuals in the full-count census data using IPUMS HistIDs. This linkage allows us to recover town-level identifiers for many additional municipalities that are not represented in the standard IPUMS place variables.

1.2 Town-Level Pipe Material Data

As illustrated in Figure 1, assigning childhood exposure requires combining individual census records with town-level information on water infrastructure. This section describes the historical sources used to measure the materials used in municipal service pipes, which form the first component of our treatment assignment.

The Manual of American Water-Works. Our town-level pipe material data come from *The Manual of American Water-Works*, compiled by M. N. Baker, a sanitary engineer who documented municipal water systems in the late nineteenth century. We use the 1889/1890 and 1897 editions.

The Manual reports information on municipal water systems across the United States, including water sources, treatment practices, distribution infrastructure, and pipe materials. The information was collected from local water superintendents and engineers and provides one of the most comprehensive historical sources on municipal water infrastructure available for this period. We systematically digitize these records to construct town-level measures of service pipe materials for nearly 4,000 municipalities.

Digitization Process. We digitize pipe material information primarily from the 1897 edition of the Baker Manual and supplement it with information from the 1889/1890 edition. For each town with a documented water system, we extract two pieces of information:

- Materials used for service pipes, that is, the connections from water mains to individual properties
- Year of system construction

Service pipe material is the most relevant determinant of residential lead exposure because service pipes connect the distribution mains directly to households, carrying water into homes where it is consumed. Our classification therefore focuses on service pipe materials rather than other components of the distribution system. Previous literature has also relied on service pipe material when discussing residential waterborne lead exposure [Troesken, 2008, Clay et al., 2014, Ferrie et al., 2012, Feigenbaum and Muller, 2016].

Pipe Material Classification. Using the information reported in the Manual, we classify towns into three mutually exclusive categories based on service pipe materials:

- **Pure Lead:** Towns reporting lead as the only pipe material in use.
- **Mixed Lead:** Towns reporting both lead and non-lead materials in service pipes.
- **Non-Lead:** Towns reporting no lead pipes.

Use of Multiple Editions. The 1897 edition is our primary source because it is the most comprehensive. We use the 1889/1890 edition as a supplement for two reasons. First, a small number of towns appear in the 1889/1890 edition but not in 1897. Second, and more importantly, some towns that appear in both editions are reported with different service pipe materials.

When a town appears in both editions, we compare the reported service pipe materials across the two sources. If the editions disagree in a way that indicates the presence of both lead and non-lead service pipes, we classify the town as Mixed Lead. We assign the Pure Lead and Non-Lead categories only when all of the available information indicates those materials were the only pipe materials reported. This approach treats disagreement across editions conservatively and avoids assigning a town to a pure category when the historical record suggests a mixed system or uncertainty in the underlying reports.

Supplementary Lead Service Line Inventories. To expand coverage of service pipe materials, we supplement the historical classifications from *The Manual of American Water-Works* with recently released municipal lead service line (LSL) inventories. These inventories were compiled by water utilities as part of the U.S. Environmental Protection Agency's Lead and Copper Rule Revisions, which require utilities to document the materials used in service line connections [Marcus, 2026].

Because modern inventories may not perfectly reflect historical pipe materials, we apply conservative classification rules and use the inventories only when they provide strong evidence about the presence or absence of lead service lines. Because lead service line replacement has historically occurred gradually in most municipalities, the presence or absence of lead lines in modern inventories can still provide informative evidence about historical system composition. We use these inventories to classify pipe materials for towns where water chemistry data are available but pipe material information is not recorded in the Baker Manual.

First, we classify a town as *Non-Lead* if its modern inventory reports zero lead service lines and at least 90% of service lines have been surveyed. Under these conditions, the absence of observed lead service lines strongly suggests that the system historically relied on non-lead materials.

Second, we classify a town as *Mixed Lead* when modern inventories indicate that at least 25% of surveyed service lines are lead. Because lead service line replacement has historically been incomplete in most U.S. municipalities, the presence of a substantial share of lead lines in modern inventories provides strong evidence that lead service connections were historically used.

Using these criteria, we assign pipe classifications to 74 additional towns in the final analytic sample. The remaining towns obtain pipe classifications directly from the Baker Manual data. In total, 897 towns in the analysis sample are classified using the historical Baker data and 74 towns are classified using modern LSL inventories. Our main estimates are not sensitive to including these towns; dropping the 74 modern-inventory towns leaves the Pure Lead × High Leaching Risk coefficient unchanged (Section 7.4).

Timing of Water System Construction. Figure 3 shows the number of water systems constructed each year from 1800 to 1900, distinguished by whether the system used Pure Lead pipes, Mixed Lead pipes, or Non-Lead pipes, along with the cumulative population served by these systems. The figure is based on the subset of towns in our final analytic sample for which we observe both pipe material data and water chemistry measurements. The figure is a stacked bar chart in which the height of the gray bar is the total number of systems completed that year.

Municipal water infrastructure expanded rapidly during the late nineteenth century, with all three types of systems (Pure Lead, Mixed Lead, Non-Lead) being constructed throughout this period. By 1900, the towns in our analytic sample collectively contained municipal water systems serving roughly 27 million residents, or about 35% of the US population, illustrating the widespread adoption of municipal water infrastructure during this period.

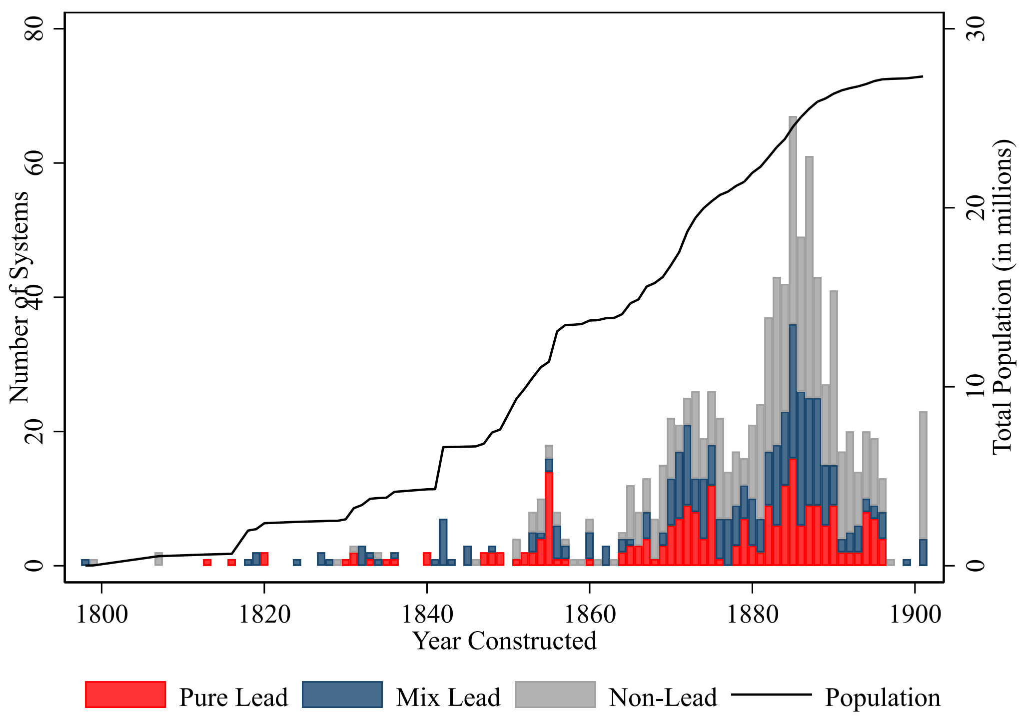


Figure 3: Number of municipal water systems constructed by year (1800–1900) and cumulative population served, distinguished by service pipe material type (Pure Lead, Mixed Lead, and Non-Lead). The figure uses towns in the final analytic sample for which both pipe material and water chemistry data are available. Data source: *The Manual of American Water-Works*.

Geographic Coverage and Representativeness. Figure 4 shows the geographic distribution of towns in the analytic sample by pipe material type. Coverage is concentrated in the Northeast and Midwest, reflecting earlier urbanization and earlier adoption of municipal water systems in these regions.

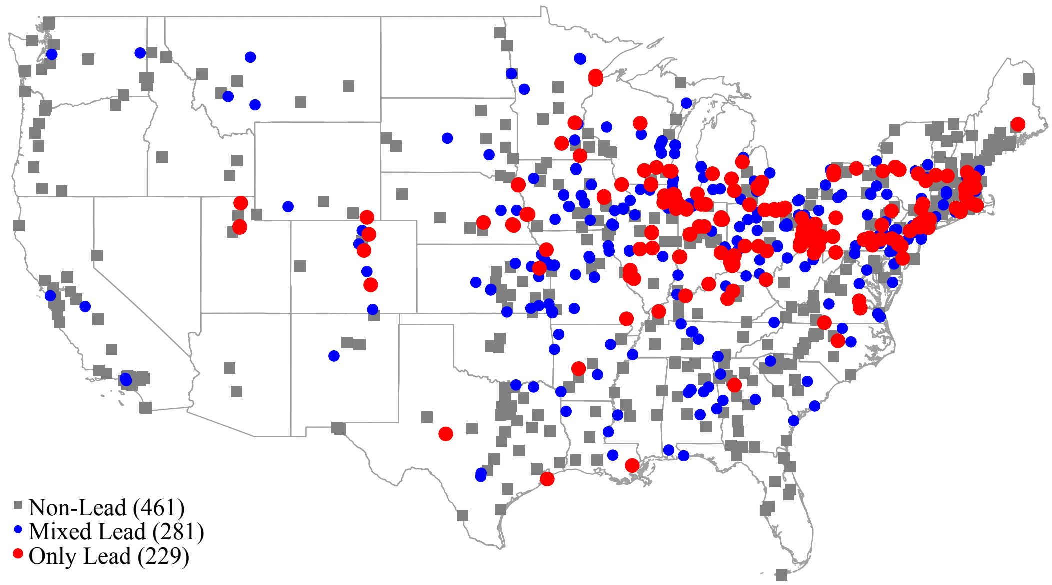


Figure 4: Geographic distribution of towns in the analytic sample by pipe material classification. Figure reproduced from the main manuscript.

1.3 Water Chemistry Data

As shown in Figure 1, the second component of treatment assignment combines pipe material data with historical measurements of municipal water chemistry. While pipe material information is available for nearly 4,000 towns, water chemistry measurements are only available for a subset of towns where the U.S. Geological Survey (USGS) conducted water quality sampling. Our final analytic sample therefore consists of towns that appear in both the Baker Manuals and the USGS chemistry records.

USGS Data Sources. Historical water chemistry data come from U.S. Geological Survey (USGS) publications documenting water quality measurements conducted between 1950 and 1952 [Lohr and Love, 1954]. These surveys were conducted by USGS in cooperation with state geological surveys and local water utilities to characterize water quality for industrial and municipal uses. Reported measurements include pH, hardness, and other chemical properties of municipal water supplies. For our analysis, we focus on pH and hardness because these variables interact to determine the corrosivity of water and therefore the extent to which lead leaches from pipe infrastructure into drinking water.

Variation Across Reports and Our Coding Procedure. The USGS reports characterize water chemistry in one of three forms: raw (untreated) source water, finished water without acidity treatment, or finished water with treatment for acidity. When finished water is reported, the documentation typically describes the treatment process, which most commonly involves chlorination but in some cases includes treatment for acidity, typically through lime addition. Some towns have measurements for more than one form, and within a form some have a single average while others report multiple sources separately. We record raw water values whenever they are available, averaging across sources when multiple raw measurements are reported. When raw water is not reported, we record finished water values and note whether the treatment process included any pH- or acidity-based treatments. Across the 971 towns in our analytic sample, this coding yields raw water values for 734 towns (76%), finished water values without acidity treatment for 205 (21%), and finished water values with treatment for acidity for 32 (3%). Our main specifications condition on a three-category water source indicator corresponding to these groups; a robustness check that drops the finished-water and acidity-treated towns is reported in Section 7.4.

Temporal Stability of Water Chemistry. Water chemistry was relatively stable during the late nineteenth and early twentieth centuries because most municipal water systems had not yet adopted chemical treatment processes that alter the natural properties of water. As a result, pH and hardness largely reflected underlying geological characteristics of local watersheds, including the mineral composition of surrounding soils and rock formations. Measurements taken during the 1950–1952 period therefore provide a reasonable approximation of water chemistry conditions during our cohorts' childhood years [Durfor and Becker, 1964, Ferrie et al., 2012].

Predicted Lead Solubility Model. As discussed in the main paper, pH and alkalinity interact in non-linear ways to determine the extent of lead leaching from pipe infrastructure. Because our historical chemistry data record hardness rather than alkalinity, we construct a predicted lead solubility model that uses hardness as a proxy for alkalinity, following established water chemistry relationships. The technical details of the model are described in Section 2 of this SI Appendix.

1.4 Final Sample

Our data construction combines linked individual census data with town-level information on pipe materials and water chemistry. We begin with men ages 20–50 observed in the 1940 Census who can be linked to a childhood census record when they were ages 0–10 in the 1900, 1910, or 1920 Census. We then merge these individuals to town-level pipe material data from *The Manual of American Water-Works* and water chemistry measurements from U.S. Geological Survey sources.

The primary constraint on the final sample is the availability of both pipe material and water chemistry data. While pipe material information is available for nearly 4,000 towns, water chemistry measurements can be matched to a smaller subset of municipalities. Restricting to towns with both data sources yields a final analysis sample of ∼6.2 million men whose childhood residence can be assigned to one of 971 towns with complete infrastructure and water chemistry data.

1.5 Inverse Probability Weights

The Census Tree links used in this study combine genealogy records from FamilySearch with algorithmic matching [Price et al., 2021, Buckles et al., 2023a]. Because the genealogy seed records reflect the research interests of FamilySearch users, the resulting linked sample over-represents white, native-born, married, and geographically stable men relative to the full census population [Price et al., 2021]. If these selection patterns correlate with adult labor market outcomes or with childhood pipe-type exposure, estimates from the unweighted linked sample may be biased.

To address this, we follow Bailey et al. [2020] and construct inverse probability weights (IPW) that reweight the linked sample to better represent the full population of men observed in the relevant childhood census. For each census wave (1900, 1910, and 1920) separately, we estimate a probit model for the probability of appearing in the linked sample as a function of age, race, state of residence, urban status, metro status, relationship to the household head, and birthplace [Wooldridge, 2007]. The probit is estimated on all male children and household dependents (ages $\leq$ 15) in each full-count census, restricted to counties where at least one linked observation is observed. This restriction effectively reweights the sample to look more representative of the initial set of locations. Weights are constructed as $\hat{w}_i = 1/\hat{p}_i$, where $\hat{p}_i$ is the estimated probability of being linked, upweighting individuals from groups with low baseline linking rates. Results are robust to alternative covariate sets in the linking probability model, including specifications that include an interaction between state of residence and urban status and exclude birthplace. The resulting weights have a median of 1.61, a mean of 1.87, and a 95th percentile of 3.20, indicating that most observations receive modest reweighting. The distribution has a right tail (99th percentile of 6.03, maximum of 118.3), reflecting individuals with demographic characteristics that make linking unlikely. We do not trim the weights in the primary specification. The unweighted results reported in Section 7 are qualitatively similar, indicating that the findings are not driven by extreme weights.

Importantly, childhood household economic characteristics, father's occupation score, family income rank, and homeownership, do not enter the linking probability model. The weights are identified from demographic and geographic predictors of linking success rather than from the parental traits that appear in our balance tests and outcome regressions. This ensures that the IPW-weighted balance assessment in Table 4 reflects genuine childhood-baseline comparability rather than mechanically improved balance on the variables being tested.

IPW and Mortality Selection. The IPW weights reweight the linked sample to resemble the population of male children observed in the relevant childhood census (1900, 1910, or 1920). This corrects for differential attrition between the childhood observation and the 1940 adult observation, including mortality that occurred between these two dates. Recent evidence suggests this concern is empirically relevant: Fletcher and Noghanibehambari [2023] find that childhood exposure to lead through water pipes reduces old-age longevity by approximately 2.7 months, with effects that are 3.5 times larger among nonwhite populations and twice as large among low-SES families. If lead-exposed children were more likely to die before 1940, the unweighted linked sample would over-represent healthier survivors from lead towns relative to non-lead towns. By reweighting to match the childhood population, the IPW partially addresses this form of selection: the weighted estimates can be interpreted as the average effect of lead exposure on earnings among the population of children alive and enumerated at ages 0–10.

The IPW cannot, however, address mortality that occurred before the initial childhood census. If lead exposure increased infant or early-childhood mortality before a child's first census enumeration, those children never enter the data and no reweighting scheme can recover their outcomes. The direction of any remaining mortality selection is ambiguous: if the most severely affected individuals died before observation, the surviving sample is positively selected on resilience and our estimates understate the true effect, but surviving individuals may also carry chronic health burdens caused by lead exposure that contribute directly to the labor market penalties we measure. These chronic effects are part of the causal channel our estimates capture rather than a selection artifact, and the two channels can operate simultaneously. Whether pre-census mortality selection differs by father's income quartile is unobservable in our data, but it bears on the interpretation of the SES heterogeneity results: if lead-induced infant mortality fell more heavily on Q1 families, the surviving Q1 population in lead towns may be unusually resilient, compressing the observed earnings gap between Q1 and higher quartiles for reasons unrelated to the buffering capacity of household resources.

2 Predicted Lead Solubility Model

2.1 Overview and Motivation

The amount of lead released from pipes into drinking water depends on water chemistry, particularly the interaction between pH and mineral content. This process, known as "plumbosolvency," is well established in the literature [Schock, 1990, Moore, 1973, Schock and Gardels, 1983, Cardew, 2003]. Simple classifications such as "acidic" versus "neutral" water fail to capture the dynamic nature of this interaction and therefore can miss important variation in potential water-borne lead exposure.

Our data exhibit substantial variation in water chemistry across towns. Figure 5 plots water hardness against pH for the 971 towns in our analysis sample, distinguishing between towns with Non-Lead, Mixed Lead, and Pure Lead service pipes. Two features of the data stand out. First, water chemistry varies widely across towns, with pH ranging from approximately 5 to 9 and hardness ranging from near zero to more than 700 mg/L. Second, there is substantial overlap in water chemistry across pipe material types. Towns with lead pipes do not systematically differ from towns with non-lead pipes in either pH or hardness. This pattern is consistent with our identifying assumption that pipe material adoption was not systematically related to water chemistry.

Because lead solubility depends on the interaction between these water chemistry parameters, we construct a continuous measure of predicted lead leaching potential. The theoretical literature models lead solubility as a function of pH and total alkalinity, which measures the buffering capacity of water. Our historical water chemistry data report hardness rather than alkalinity. Hardness and alkalinity are closely related measures of dissolved mineral content, particularly calcium and magnesium carbonates [Moore, 1973, Boyd, 2015]. We therefore use hardness as a proxy for alkalinity while following the theoretical framework developed in the water chemistry literature.[3] This predicted solubility measure provides a ranking of towns by theoretical lead leaching potential based on observed water chemistry.

[3]Because hardness typically exceeds alkalinity in natural waters, using hardness as a proxy tends to overstate buffering capacity and therefore produces conservative predictions of lead solubility, biasing predicted lead concentrations downward.

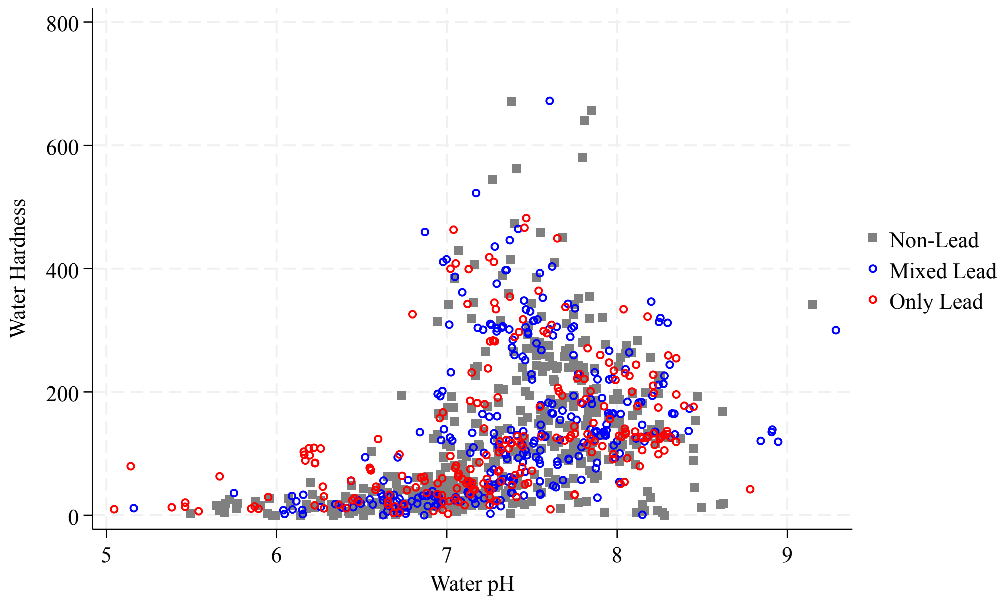


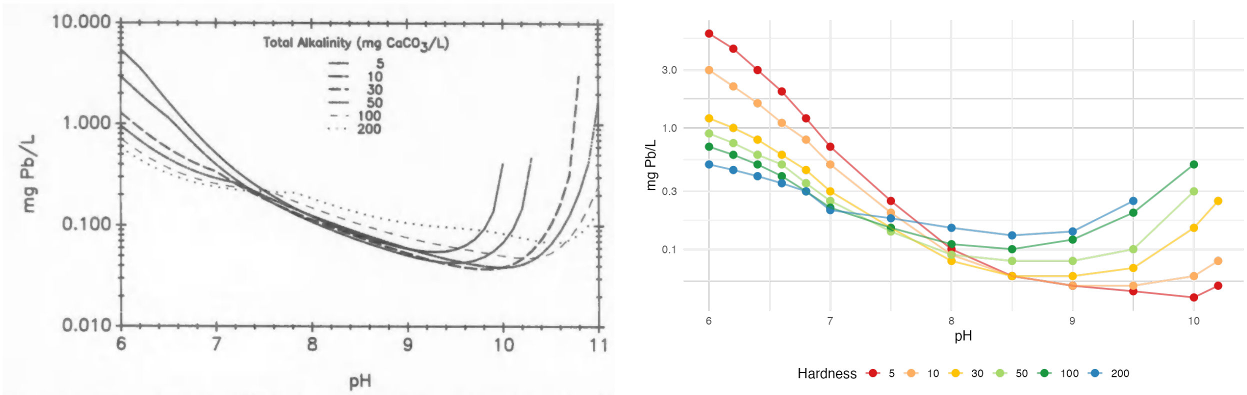
Figure 5: Water chemistry (pH and hardness) for all 971 towns in the analysis sample, distinguished by pipe material type. Gray squares indicate Non-Lead towns, blue circles indicate Mixed Lead towns, and red circles indicate Pure Lead towns. Water chemistry varies widely across towns and exhibits substantial overlap across pipe types.

2.2 Theoretical Background and Data Sources

Lead solubility in drinking water depends primarily on pH and alkalinity, which together determine the chemical equilibrium between dissolved lead and solid lead carbonate compounds that can form protective scales on pipe surfaces [Schock, 1990, Schock and Gardels, 1983, Moore, 1973]. At low pH, lead dissolves readily as free ions. At intermediate pH values with sufficient alkalinity, lead carbonate minerals precipitate and form protective scales that reduce lead release. At very high pH levels, lead can again become soluble through hydroxide complex formation. As a result, theoretical models predict a U-shaped relationship between pH and lead solubility, with the lowest lead concentrations occurring at intermediate pH levels when protective carbonate scales can form.

To quantify this relationship, we rely on the theoretical lead solubility model developed by Schock [1990]. Schock calculates equilibrium lead concentrations as a function of pH and total alkalinity using thermodynamic constants for the relevant chemical species, with the resulting predictions representing the upper bound on lead release for a given water chemistry environment. Figure 6a reproduces Schock's solubility curves: higher alkalinity substantially reduces lead solubility across the low-to-intermediate pH range, where sufficient buffering allows protective carbonate scales to form.

Because Schock presents the model as graphical curves rather than closed-form equations, we digitize the curves and extract predicted lead concentrations at multiple combinations of pH and alkalinity (Figure 6b). We record predicted concentrations for six alkalinity levels (5, 10, 30, 50, 100, and 200 mg/L as $CaCO_3$) across a grid of pH values, yielding approximately 74 observations of (pH, alkalinity, predicted lead concentration) used to estimate the statistical model described next.



(a) (b)

Figure 6: Theoretical lead solubility as a function of pH and alkalinity from Schock [1990] is shown in sub-figure (a). Digitized curves are shown in sub-figure (b).

2.3 Generalized Additive Model (GAM)

We approximate Schock's surface with a Generalized Additive Model (GAM) that predicts the base-10 logarithm of lead concentration (in mg/L) as a smooth bivariate function of pH and log-transformed alkalinity:

$$\log_{10}(\mathrm{Pb}) = f(\mathrm{pH}, \log_{10}(\mathrm{Alkalinity})) + \epsilon \tag{1}$$

where $f(\cdot)$ is estimated using tensor-product splines fit to the digitized Schock points. The basis dimension for pH is set to $k = 10$ to capture the nonlinear pH–solubility relationship, while the basis dimension for alkalinity is limited to $k = 6$ to avoid overfitting given the small number of alkalinity levels in the training data.

Model Fit. The fitted GAM closely reproduces Schock's theoretical solubility curves, explaining 99.9% of the variation in the digitized data ($R^2 = 0.999$). The root mean squared error is 0.015 in $\log_{10}$ concentration units, corresponding to a typical multiplicative prediction error of approximately 3.6% (0.023 mg/L in original units). Figure 6b compares the fitted values to Schock's original figure across the full range of the training data.

2.4 Extrapolation Strategy

A subset of towns have water chemistry values outside the range of Schock's original figure (pH 6.0–10.2 and alkalinity up to 200 mg/L; recall that we use observed hardness as our proxy for alkalinity). Among the 971 towns in our analysis sample, 236 towns (24%) fall outside this range: 28 towns have pH below 6.0, and the remaining 208 towns have hardness exceeding 200 mg/L.

Because spline-based models can exhibit unstable behavior when extrapolating beyond the training domain, we apply simple extrapolation rules designed to preserve the qualitative relationships implied by the chemistry of lead dissolution. Importantly, most extrapolation occurs along the high-hardness dimension, where predicted lead concentrations change relatively slowly with additional buffering. By contrast, only a small number of towns require extrapolation into the low-pH region where lead dissolution increases rapidly with acidity. In practice, these extrapolated regions correspond to extreme predicted lead values: very acidic water implies very high predicted lead concentrations, while very high hardness implies strong buffering and therefore very low predicted lead concentrations. As a result, towns requiring extrapolation lie far from the portion of the distribution used to classify high- and low-leaching environments.

2.4.1 Low pH (pH < 6.0): Monotonic Slope Extension

For towns with pH values below the range of Schock's figure, we extrapolate predicted lead concentrations by extending the local slope of the fitted GAM at the lower pH boundary. This approach preserves the monotonic increase in predicted lead concentrations as water becomes more acidic, consistent with the theoretical chemistry of lead dissolution.

Specifically we:

1. Calculate the slope of the log predicted lead concentration between pH 6.0 and pH 6.1 for the given alkalinity level.
2. Extend this slope linearly to the target pH value.

Formally:

$$\log_{10}(\text{Pb}_{\text{target}}) = \log_{10}(\text{Pb}_{6.0}) + \text{Slope}_{6.0} \times (6.0 - \text{pH}_{\text{target}}) \tag{2}$$

where

$$\text{Slope}_{6.0} = \frac{\log_{10}(\text{Pb}_{6.1}) - \log_{10}(\text{Pb}_{6.0})}{6.1 - 6.0}. \tag{3}$$

This rule ensures that towns with highly acidic water receive appropriately high predicted lead concentrations while preserving the monotonic relationship implied by the underlying chemistry.

2.4.2 High Hardness (Hardness > 200 mg/L): Log-Linear Trend

For towns whose hardness exceeds 200 mg/L (the upper alkalinity bound of Schock's figure), we extrapolate predicted lead concentrations using the log-linear trend implied by the interior of the model. At high alkalinity, carbonate buffering increasingly suppresses lead release but with diminishing returns, implying a gradual decline in predicted lead concentrations.

Specifically we:

1. Calculate predicted lead concentrations at alkalinity levels of 100 and 200 mg/L for the observed pH.

2. Estimate the slope of the decline in predicted lead concentrations in log-log space:

$$\beta_{\text{slope}} = \frac{\log_{10}(\text{Pb}_{200}) - \log_{10}(\text{Pb}_{100})}{\log_{10}(200) - \log_{10}(100)}. \tag{4}$$

3. Apply a conservative lower bound to prevent unrealistically rapid declines in predicted concentrations:

$$\beta_{\text{slope}} = \max(\beta_{\text{slope}}, -1.5).$$

4. Project predicted lead concentration to the target alkalinity level H:

$$\log_{10}(\text{Pb}_H) = \log_{10}(\text{Pb}_{200}) + \beta_{\text{slope}}\big(\log_{10}(H) - \log_{10}(200)\big). \tag{5}$$

The choice of the 100–200 mg/L baseline range anchors the extrapolation in the stable interior of the model and avoids boundary artifacts from the spline fit. Sensitivity analyses using alternative baseline ranges (150–200 and 120–200 mg/L) yield qualitatively similar results.

2.5 Classification of Towns into High and Low Leaching Environments

Using the GAM model and extrapolation rules described above, we compute predicted lead concentrations (in parts per billion) for each of the 971 towns in the analysis sample based on their observed pH and hardness values.

Figure 7 displays the distribution of predicted lead concentrations across towns, plotting predicted lead (on a logarithmic scale) against water pH with points colored by hardness range. Predicted lead concentrations vary widely across municipalities, ranging from less than 10 ppb to more than 1,000 ppb. Importantly, towns with similar pH values can have very different predicted lead concentrations depending on water hardness, underscoring the importance of accounting for both chemistry parameters.

Primary Classification. The predicted concentrations reflect theoretical equilibrium solubility under laboratory conditions [Schock, 1990], not direct measurements of lead in household tap water. They are best interpreted as a ranking of towns by leaching potential rather than as precise exposure levels. Actual lead dissolution in early twentieth-century water systems depended on factors the model cannot capture, including pipe age, flow rates, temperature, and the condition of interior mineral scale. We therefore use the predicted values to classify towns into binary leaching-risk categories rather than treating them as a continuous measure of dose. Our primary specification defines high-leaching towns as those in the top quartile:

- **High Leaching:** predicted lead $\geq$ 75th percentile
- **Low Leaching:** predicted lead $<$ 75th percentile

The 75th percentile threshold aligns closely with the pH ranges used in prior work. As illustrated in Figure 7, all towns with strongly acidic water (pH < 6.5), which prior work would classify as high-leaching environments based on pH alone, fall above the 75th percentile of predicted lead concentrations. Conversely, towns with pH ≥ 7.2, which are typically treated as low-leaching environments, all fall below this threshold. The predicted solubility model primarily differentiates leaching potential within the intermediate pH range (6.5–7.2), where water hardness and buffering capacity substantially influence lead release. The quartile-based cutoff preserves the conventional pH-based classification while incorporating additional variation from water hardness to separate towns in this ambiguous middle range. Towns requiring extrapolation in the previous section lie well above or below this threshold and do not influence the classification boundary.

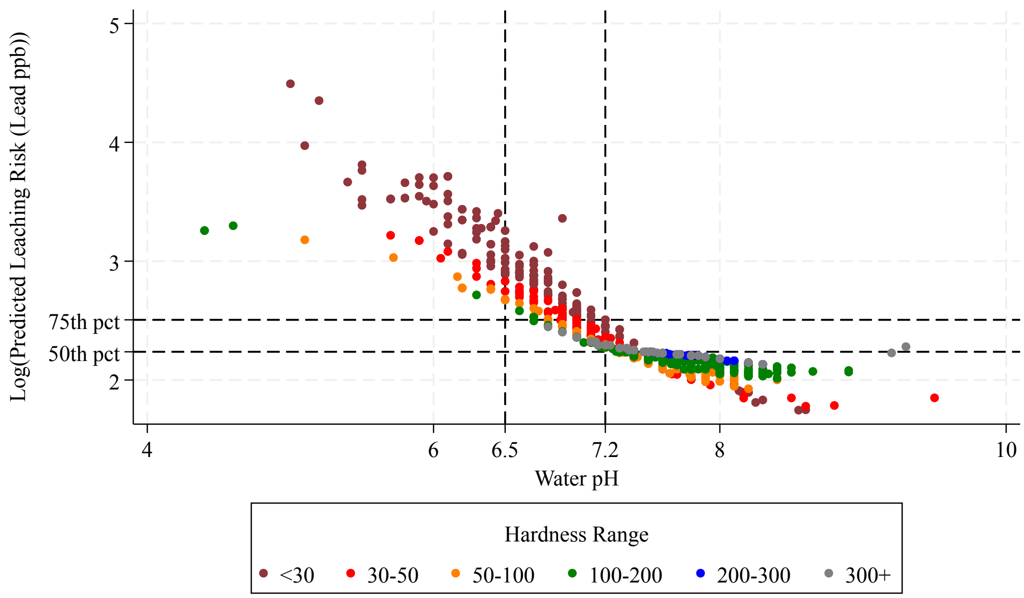


Figure 7: Distribution of predicted lead concentrations across 971 towns as a function of water pH, with points colored by hardness range. Predicted lead concentration is shown on a logarithmic scale. Horizontal lines indicate the 75th percentile (primary threshold) and 50th percentile (robustness threshold) of predicted lead concentrations. Substantial variation in predicted lead exists even among towns with similar pH, reflecting the importance of water hardness in determining lead leaching potential.

3 Variable Construction and Measurement

3.1 Lead Exposure Variables

Our primary independent variables measure childhood exposure to waterborne lead arising from the interaction between municipal water infrastructure and local water chemistry. Lead exposure occurs when water flows through lead service pipes and chemically leaches lead into drinking water.

We therefore construct exposure measures based on both pipe material and the chemical conditions governing lead leaching.

Pipe Material Types. Using historical engineering reports and municipal records described in Section 1.2, towns are classified as: (i) *Pure Lead*, where lead was reported as the pipe material in use; (ii) *Mixed Lead*, where both lead and non-lead materials were reported; and (iii) *Non-Lead*, where no lead pipes were reported.

Leaching Risk Levels. We classify towns according to their predicted lead leaching potential based on water chemistry variables described in Section 2.5. Towns are classified as *High Leaching Risk* if predicted lead concentration lies at or above the 75th percentile of the distribution, and *Low Leaching Risk* otherwise.

Exposure Categories. Our empirical specifications interact pipe material type with predicted leaching risk to construct six exposure categories, shown in Table 1. These categories capture variation in both the presence of lead infrastructure and the chemical conditions governing lead leaching. Historical evidence suggests that pipe material choices were made largely without reference to the underlying water chemistry, making the interaction of these two dimensions a natural way to characterize predicted lead exposure across towns [Troesken, 2006].

Table 1: Exposure Categories by Pipe Material and Predicted Leaching Risk (Number of Towns)

		Predicted Leaching Risk	
		Low	**High**
Pipe Material	**Non-Lead**	Non-Lead × Low Risk (336)	Non-Lead × High Risk (125)
	Mixed Lead	Mixed Lead × Low Risk (218)	Mixed Lead × High Risk (63)
	Pure Lead	Pure Lead × Low Risk (174)	Pure Lead × High Risk (55)

3.2 Outcome Variables

Outcome variables are measured in the 1940 Census. Panel A of Table 2 reports summary statistics for the 1940 outcome measures used in the analysis. These outcomes capture different dimensions of adult labor market attainment, including earnings, annual work intensity, labor force status at the time of enumeration, and occupation.

Wage and Salary Income. The 1940 Census asked respondents to report the total amount of money wages or salary received during 1939. This measure includes earnings from wage and salary employment, including commissions, but excludes self-employment income, business profits, investment income, and other non-wage sources. Approximately 78.6% of men aged 20–50 in our sample report positive wage and salary income.[4]

[4] Income responses were top-coded at $5,000 in the original census schedule. For values above this threshold, enumerators recorded "$5,000+" rather than the exact amount. Approximately 0.16% of observations are top-coded. We exclude these observations from the wage and salary income analysis; including them at the top-coded value of $5,000 leaves our estimates unchanged.

Weeks Worked, Unemployment, and Labor Force Participation. The 1940 Census records labor market attachment using both annual information for 1939 and labor force status during a specific reference week. Weeks worked refers to the number of equivalent full-time weeks worked during 1939. This measure is not a simple count of calendar weeks with any work; respondents who worked part-time during the year were instructed to convert their work into full-time-week equivalents and report the resulting total in whole weeks. We therefore interpret weeks worked as a measure of annual work activity during 1939, expressed in a common full-time metric.

By contrast, unemployment and labor force participation are measured for the reference week of March 24–30, 1940. We define an individual as unemployed if he was in the labor force during that week but classified as seeking work rather than at work. We define an individual as in the labor force if he was either working or seeking work during the reference week, and as out of the labor force otherwise. These measures therefore capture short-run labor market status at the time of enumeration rather than annual work activity during 1939.

Occupation. The 1940 Census reports an individual's current occupation if he was working at the time of enumeration and his most recent occupation if he was unemployed. We use the harmonized 1950 occupation codes provided by IPUMS, which classify respondents into 223 detailed occupation categories in our analytical sample. For our occupational sorting analysis, we collapse these detailed occupations into two broad tiers: lower-tier occupations (those with below median occupation-level earnings) and upper-tier occupations (those with above median occupation-level earnings).

3.3 Control Variables

Our empirical specifications include two layers of controls aimed at distinct margins of selection. First, lead-pipe towns differed from non-lead-pipe towns in size, industrial composition, and economic structure; our town-level covariates ensure that comparisons are made within similar town environments. Second, these same economic differences across towns could produce compositional differences in the households that sorted into or remained in each town type, even though families could not knowingly avoid lead exposure (the health risks were not yet understood). Our individual and household-level covariates address this second, compositional margin.

Town-Level Controls

Region Fixed Effects. We include fixed effects for seven Census regions based on IPUMS regional classifications. The standard IPUMS classification includes nine divisions; to ensure sufficient variation across treatment groups within each region, we combine the West South Central and East South Central divisions into a single South Central region and the Pacific and Mountain divisions into a single West region.

Town Population Fixed Effects. We control for town size using seven fixed effects based on 1900 population:

Table 2: Individual Summary Statistics

	(1) Mean	(2) SD	(3) Min	(4) Max	(5) Count
Panel A: 1940 Census					
Wage and Salary Income	1139.479	1016.441	0	5,000	5,896,429
Weeks Worked in Prior Year	40.102	18.021	0	52	6,006,388
Not in Labor Force	0.050	0.218	0	1	6,006,388
Unemployed	0.095	0.293	0	1	5,706,425
Below Median Wage Occup.	0.391	0.488	0	1	5,675,407
Completed Years of Education	11.397	3.192	0	21	6,113,616
Panel B: Census Characteristics at Initial Observation					
Age at Exposure	4.470	2.957	0	10	6,229,412
In School (Age>5)	0.851	0.356	0	1	2,408,203
Non-White	0.037	0.188	0	1	6,229,412
Number of Siblings	2.427	1.965	0	9	6,229,412
Male Head of HH	0.954	0.209	0	1	6,229,412
Father Immigrant	0.409	0.492	0	1	6,229,412
Mother Immigrant	0.378	0.485	0	1	6,229,412
Family Rents	0.675	0.469	0	1	6,229,412
Family Owns with Mortgage	0.191	0.393	0	1	6,229,412
In Group Quarters	0.004	0.063	0	1	6,229,412
Parents Married	0.949	0.219	0	1	6,229,412
Not in Metro	0.221	0.415	0	1	6,229,412
In Central City	0.654	0.476	0	1	6,229,412
HH Head Occup. Income Score	25.064	12.470	0	80	6,229,412
Fathers Income Rank	61.517	20.504	1	100	5,507,555

Notes: Panel A reports variables measured in the 1940 complete-count census. Panel B reports characteristics measured at the individual's initial childhood observation in the 1900, 1910, or 1920 complete-count census.

- Bin 1: Population $<$ 2,500
- Bin 2: 2,500 to 10,000
- Bin 3: 10,000 to 25,000
- Bin 4: 25,000 to 50,000
- Bin 5: 50,000 to 100,000
- Bin 6: 100,000 to 250,000
- Bin 7: 250,000+ (includes all New York City boroughs)

Proximity to Metropolitan Areas. We include two controls for proximity to large cities, since proximity affects both treatment probability and long-run economic outcomes. First, we include an indicator for whether the town was a secondary location within a larger water system. Second, we calculate the distance from each town to the nearest metropolitan city and assign towns to five distance-bin fixed effects: distance = 0 (the town is itself metropolitan), 0–25th percentile, 25th–50th percentile, 50th–75th percentile, and above the 75th percentile. Percentiles are calculated using the distribution of non-zero distances.

Water System Controls. Pipe classifications for 74 of the 971 towns in the analysis sample are based on modern lead service line inventories rather than the Baker Manual (see Section 1). We include an indicator variable for towns with this modern-inventory assignment. We also control for water source characteristics from USGS reports using a three-category fixed effect for water treatment status (raw water reported, finished water without acidity treatment, and finished water with acidity treatment); the fixed effects absorb systematic differences in water treatment practices across towns.

Town Industry and Occupation Composition Controls

To account for differences in local economic structure, we construct town-level measures from complete-count Census microdata during the childhood period.

Industry Employment Shares. For manufacturing, services, and farming, we create fixed effects based on the share of the town's workforce employed in each industry, using quartile bins: 0–25%, 25–50%, 50–75%, and 75–100%. These controls ensure we compare towns with similar industrial structures. Mining employment has a more skewed distribution, so we instead create indicators for towns in the top 10% and top 5% of mining employment share.

Occupational Composition. We control for the share of workers in white-collar occupations using quartile bin fixed effects: 0–25%, 25–50%, 50–75%, and 75–100%. This captures variation in the occupational composition of local labor markets that industry shares alone do not capture, since towns with similar industry mixes can have very different occupational distributions.

Home Ownership Rates. We control for home ownership rates using quartile bin fixed effects based on the share of households owning their homes: 0–25%, 25–50%, 50–75%, and 75–100%. Home ownership rates serve as a town-level proxy for local economic stability and household wealth in this period and are a key control in Feigenbaum and Muller [2016].

Individual and Household Controls

We include detailed controls for individual, household, and parental characteristics measured when individuals are children (ages 0–10).

Individual Characteristics. We control for race (white, black, other) and a full set of age fixed effects, with one indicator per year of age.

Household Location. We include indicators for whether the household resides in the metropolitan or urban portion of the town, as well as an indicator for group quarters residence (institutions such as boarding schools, orphanages, or reformatories).

Household Head and Parent Characteristics. We control for sex of the household head, marital status, parental immigration status (whether one or both parents were born outside the

United States), home ownership, and fixed effects for household head's occupation using single-digit occupation codes. These variables capture household socioeconomic status during childhood.

4 Validation of Research Design

4.1 Balance Tests

Our empirical strategy compares individuals exposed to lead through municipal water systems with individuals whose local water systems did not generate meaningful lead leaching. A central concern for this comparison is selection into the types of cities that adopted lead service pipes. If towns that installed lead plumbing differed systematically in economic conditions, demographics, or other determinants of children's long-run outcomes, estimated effects could reflect these confounding differences rather than the causal impact of lead exposure.

Historical evidence suggests that health considerations played little role in pipe material selection during the period we study. Instead, the adoption of lead service pipes was primarily driven by engineering convenience and the broader development of municipal water infrastructure. Engineers widely favored lead pipes for service connections because they were flexible, durable, and relatively easy to install and maintain in dense urban environments [Troesken, 2006, 2008]. At the same time, the prevailing engineering view, the "Doctrine of Protective Power," held that mineral coatings would rapidly form inside lead pipes, preventing sustained leaching into drinking water. Although the doctrine acknowledged that corrosive water could increase lead dissolution, engineers generally assumed these protective coatings would form quickly and eliminate any risk [Troesken, 2006]. As a result, municipal planners typically selected pipe materials based on engineering and infrastructure considerations rather than detailed assessments of local water chemistry.

Prior research confirms that the primary predictors of lead pipe adoption were city size and wealth rather than health concerns or environmental conditions. Both Clay et al. [2014] and Feigenbaum and Muller [2016] show that larger and more economically developed municipalities were substantially more likely to adopt lead plumbing as they built out municipal water systems. This pattern is also evident in our data. Table 3 presents baseline town characteristics measured in the 1900 Census for towns with Pure Lead, Mixed Lead, and Non-Lead pipes. Columns (1)–(3) report raw means and standard deviations for each pipe type, columns (4)–(5) present unconditional differences relative to Non-Lead towns, and columns (6)–(7) report conditional differences after controlling for the set of city-level traits we discuss in Section 3.

The unconditional comparisons reveal large differences across pipe types. Pure Lead towns are substantially larger (mean population 45,600 compared to 13,100 for Non-Lead towns), more urbanized (76% versus 65%), located closer to major metropolitan areas, and exhibit greater industrial employment. These towns also have higher median occupational income scores and lower home ownership rates. These patterns closely mirror historical accounts of municipal water system development: larger and more industrialized cities tended to build water infrastructure earlier and were more likely to install lead service pipes [Troesken, 2008, Clay et al., 2014].

Columns (6)–(7) present comparisons across towns after conditioning on the factors that prior research identifies as central to lead pipe adoption. In particular, we compare towns within similar structural environments by controlling for region, town population bins, proximity to metropolitan areas, and the industrial composition of the local economy (see Section 3). Once towns are compared within these environments, the observable differences across pipe types shown in the balance table become small and statistically insignificant.

Importantly, these historical adoption patterns also imply that any remaining endogeneity in pipe material choice is likely to bias estimated lead effects toward zero. Because larger and more economically developed cities were more likely to adopt lead pipes, children growing up in these cities would plausibly have experienced better labor market outcomes even in the absence of lead exposure. Simple comparisons across pipe types would therefore tend to attenuate the estimated negative effects of lead exposure.

Individual-Level Balance. Table 4 examines balance in household characteristics measured during childhood (ages 0–10), comparing individuals who grew up in Pure Lead, Mixed Lead, and Non-Lead towns. Columns (1)–(3) report mean characteristics by pipe type, while columns (4)–(5) report IPW-weighted differences conditional on the same set of structural town characteristics used in the town-level balance tests. The IPW weights are identical to those used in the primary analysis and correct for differential selection into the linked sample, ensuring that this balance assessment reflects the same effective sample as the outcome regressions.

The economically important measures of childhood socioeconomic status are balanced across pipe types. Father's income rank, family homeownership, immigration background, and literacy show no significant differences between Pure Lead and Non-Lead towns. Occupational income score is not significantly different in the IPW-weighted specification.

A few smaller differences emerge. Pure Lead towns have slightly more children per household (coefficient of 0.092, significant at 1%) and a slightly lower share of white-collar fathers (coefficient of -0.011, significant at 5%), though both differences are small in magnitude relative to their underlying distributions. Importantly, these variables, along with all other household characteristics in the table, are included as covariates in the main regression specifications. Any residual imbalance correlated with pipe type but unrelated to leaching risk would affect outcomes in lead-pipe towns regardless of local water chemistry. That the estimated effects are concentrated where water chemistry promotes lead dissolution, and that water chemistry variation has no effect on outcomes in Non-Lead towns, indicates that confounds correlated with pipe adoption are not driving the results.

Taken together, the balance tests show that the critical SES indicators are not systematically related to pipe type once structural town characteristics are controlled, and any remaining small differences are absorbed by covariate adjustment.

Table 3: Differences in Town Demographic Characteristics by Pipe Material

				Unconditional Differences		Conditional Differences	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Pure Lead	Mixed Lead	Non-Lead	PL - NL	ML - NL	PL - NL	ML - NL
pH	7.30	7.39	7.31	-0.01	0.08	-0.07	-0.05
	(0.72)	(0.58)	(0.64)	[0.07]	[0.05]	[0.06]	[0.05]
Hardness	129.01	142.32	117.96	11.05	24.36**	-7.61	-10.00
	(105.43)	(120.52)	(116.68)	[10.21]	[10.35]	[8.65]	[8.02]
Predicted Lead ppb	445.32	441.19	510.79	-65.47	-69.60	-58.02	-5.89
	(972.42)	(1463.38)	(1612.23)	[101.66]	[119.83]	[197.96]	[194.29]
High Leaching Risk	0.24	0.22	0.27	-0.03	-0.05	0.03	0.05
	(0.43)	(0.42)	(0.45)	[0.04]	[0.04]	[0.04]	[0.03]
Population (in 000s)	45.59	38.68	13.08	32.52***	25.60**	-5.22	3.21
	(134.62)	(155.70)	(28.32)	[9.06]	[10.15]	[5.72]	[4.41]
Pct. Urban	76.09	76.97	65.42	10.67***	11.55***	0.06	0.24
	(38.96)	(37.18)	(42.02)	[3.27]	[2.98]	[2.72]	[2.18]
Dist. to Nearest Metro	56.52	106.93	162.65	-106.13***	-55.72***	-0.56	-1.73
	(71.12)	(109.74)	(151.60)	[10.13]	[11.23]	[4.63]	[5.57]
Median Occup. Income Score of Town	22.49	21.85	21.34	1.15***	0.51**	0.27	0.18
	(2.91)	(3.10)	(3.39)	[0.25]	[0.24]	[0.18]	[0.18]
Pct. in White Collar Occup.	22.32	21.55	22.16	0.16	-0.60	0.06	-0.31
	(9.07)	(7.93)	(8.07)	[0.73]	[0.62]	[0.36]	[0.34]
Pct. Emp in Manuf	26.35	21.02	19.69	6.66***	1.33	-0.19	-0.16
	(15.94)	(14.26)	(15.68)	[1.41]	[1.23]	[0.52]	[0.44]
Pct. Emp in Retail	13.23	12.97	12.75	0.48	0.22	0.36	0.06
	(4.59)	(4.40)	(4.64)	[0.41]	[0.36]	[0.27]	[0.23]
Pct. Emp in Service	23.73	26.26	26.54	-2.81***	-0.28	-0.45	0.02
	(9.77)	(10.02)	(9.84)	[0.84]	[0.78]	[0.45]	[0.40]
Pct. Emp in Mining	2.84	3.06	1.75	1.09	1.30*	0.44	0.40
	(9.56)	(9.41)	(6.11)	[0.76]	[0.78]	[0.41]	[0.34]
Pct. Emp in Other Ind.	24.68	24.53	23.47	1.20	1.06	0.68	0.72
	(9.45)	(8.87)	(8.63)	[0.79]	[0.67]	[0.55]	[0.51]
Pct. Own Home	40.22	41.69	42.97	-2.75**	-1.29	0.03	0.13
	(12.90)	(13.93)	(11.98)	[1.10]	[1.08]	[0.57]	[0.49]
Observations	229	281	461	971	971	971	971

Notes: This table reports on baseline differences in town characteristics. Columns (1)–(3) report raw means and standard deviations for each pipe type, columns (4)–(5) present unconditional differences relative to Non-Lead towns, and columns (6)–(7) report conditional differences after controlling for the set of city-level traits discussed in Section 3. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Differences in Household Head Characteristics by Pipe Type

				Town Conditional Differences	
	(1)	(2)	(3)	(4)	(5)
	Pure Lead	Mixed Lead	Non–Lead	PL-NL	ML-NL
Male HH Head	0.958	0.954	0.949	-0.004^{**}	-0.005^{**}
	(0.202)	(0.209)	(0.219)	[0.002]	[0.002]
White	0.976	0.962	0.943	-0.005	-0.019^{*}
	(0.152)	(0.192)	(0.232)	[0.009]	[0.011]
Father Immig.	0.444	0.439	0.296	0.011	-0.005
	(0.497)	(0.496)	(0.457)	[0.020]	[0.017]
Mother Immig.	0.408	0.408	0.273	0.008	-0.001
	(0.491)	(0.491)	(0.446)	[0.019]	[0.016]
Number of Children	2.465	2.392	2.425	0.093^{***}	-0.013
	(1.975)	(1.940)	(1.990)	[0.033]	[0.030]
White Collar Job	0.240	0.255	0.246	-0.011^{**}	-0.005
	(0.427)	(0.436)	(0.431)	[0.005]	[0.005]
Literate	0.953	0.946	0.948	-0.003	-0.001
	(0.212)	(0.227)	(0.222)	[0.004]	[0.004]
Occ. Income Score	27.871	27.971	27.369	-0.098	0.128
	(9.510)	(9.763)	(10.419)	[0.117]	[0.132]
Fathers Income Rank	62.059	62.368	59.147	-0.275	-0.274
	(19.929)	(20.016)	(22.041)	[0.474]	[0.525]
Family Owns Home	0.321	0.285	0.386	-0.007	-0.004
	(0.467)	(0.452)	(0.487)	[0.009]	[0.007]
In Group Quarters	0.004	0.004	0.005	-0.000	0.000
	(0.062)	(0.062)	(0.068)	[0.001]	[0.001]
Observations	2,387,381	2,453,813	1,388,218	6,229,412	6,229,412

Notes: Columns (1)–(3) report unconditional means and standard deviations for individuals who grew up in Pure Lead, Mixed Lead, and Non-Lead towns, respectively. All variables are measured at the time of the childhood census record (1900, 1910, or 1920). Columns (4)–(5) report IPW-weighted conditional differences (Pure Lead minus Non-Lead; Mixed Lead minus Non-Lead), estimated by regressing each characteristic on pipe-type indicators while absorbing the town-level categorical fixed effects described in Section 3. IPW weights are identical to those used in the primary analysis and correct for differential selection into the linked sample. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Water Chemistry Balance and Independence from Pipe Type. A central feature of our research design is that the chemical conditions governing lead leaching are independent of the infrastructure choices that determine pipe materials. While larger and more industrialized cities were more likely to install lead service pipes, historical evidence indicates that municipal engineers did not systematically consider water chemistry when selecting pipe materials. As a result, the chemical conditions that determine lead leaching vary largely independently of pipe adoption.

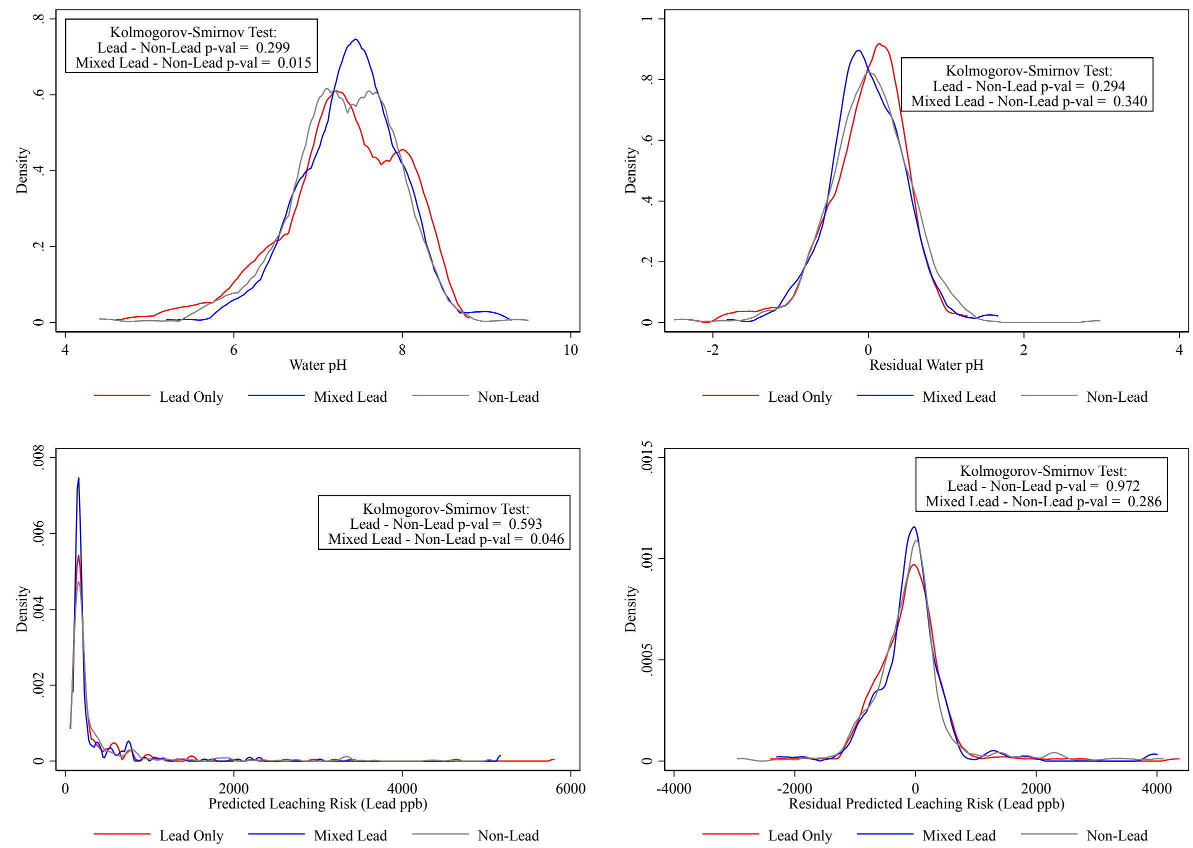
Figure 8 provides confirmation of this independence. The figure presents kernel density plots for pH and predicted lead concentration, with separate distributions shown for Pure Lead (red), Mixed Lead (blue), and Non-Lead (gray) towns. The left column displays raw distributions, while the right column shows residualized distributions after conditioning on town-level characteristics described in Section 3.

The pH distributions overlap closely across pipe types. Kolmogorov–Smirnov tests fail to reject equality of the pH distributions for Pure Lead versus Non-Lead towns ($p = 0.299$), while Mixed Lead versus Non-Lead towns show a marginal difference ($p = 0.015$). After conditioning on the town characteristics used in our empirical specifications, the distributions become even more similar: the p-values rise to 0.264 for Pure Lead versus Non-Lead and 0.375 for Mixed Lead versus Non-Lead. Visually, the kernel density plots display substantial overlap across all pipe types both before and after conditioning.

A similar pattern emerges for predicted lead concentration, the measure summarizing predicted lead leaching potential. The raw distributions show no difference for Pure Lead versus Non-Lead towns ($p = 0.593$) and a marginal difference for Mixed Lead versus Non-Lead towns ($p = 0.0496$). After conditioning on town characteristics, however, the distributions become statistically indistinguishable: the Kolmogorov–Smirnov p-values rise to 0.946 and 0.263 respectively. These results indicate that any small differences in water chemistry across pipe types are explained by observable town characteristics and disappear once towns are compared within similar structural environments.

This independence plays a critical role in our identification strategy. Lead exposure in drinking water arises only when lead service pipes interact with corrosive water chemistry. Because pipe adoption and water chemistry vary independently across towns, the interaction between pipe infrastructure and local chemical conditions generates variation in lead exposure that is plausibly exogenous to other determinants of long-run outcomes.

Figure 8: Water Chemistry Density Plot



Notes. Distribution of water chemistry characteristics by pipe type. Left panels show raw distributions; right panels show distributions of residuals after conditioning on town-level characteristics discussed in Section 3. Kernel density plots are shown for Pure Lead (red), Mixed Lead (blue), and Non-Lead (gray) towns. Kolmogorov–Smirnov test p-values for equality of distributions are reported in each panel. Water chemistry distributions overlap substantially across pipe types, particularly after conditioning on town size, confirming independence of pipe material choice and water chemistry.

5 Empirical Specifications

5.1 Main Specification: Labor Market Outcomes

Our main empirical specification estimates the effect of childhood exposure to waterborne lead, determined by the interaction between pipe material type and water chemistry, on adult labor market outcomes. In the main paper, Figure 2 reports coefficient plots for three outcomes: log wage and salary income, weeks worked, and unemployment. In the appendix, we present the underlying regression results, including labor force participation as an additional outcome, along with supplementary estimates that compare within pipe type across leaching risk environments, which more directly reflects the identifying variation in the design.

We estimate the following specification:

$$
\begin{aligned}
Y_{ic} = &\beta_1 \text{LowLR}_c + \beta_2 \text{PureLead}_c \times \text{LowLR}_c + \beta_3 \text{PureLead}_c \times \text{HighLR}_c \\
&+ \beta_4 \text{MixLead}_c \times \text{LowLR}_c + \beta_5 \text{MixLead}_c \times \text{HighLR}_c \\
&+ X_i' \gamma + Z_c' \theta + \alpha_r + \varepsilon_{ic},
\end{aligned} \tag{6}
$$

where Y_{ic} is the adult labor market outcome for individual i who grew up in town c. $LowLR_c$ is a binary variable indicating if the water chemistry environment is low leaching risk, while $HighLR_c$ is a binary variable indicating if the water chemistry is a high leaching risk environment. $PureLead_c$ indicates if pure lead pipes are used and $MixLead_c$ indicates if lead was one of the types of materials used in the pipes. X_i denotes individual and household controls measured during childhood, including age fixed effects with one indicator per year of age in 1940, race, household location (urban or metropolitan), group quarters residence, household head characteristics (sex, marital status, parental immigration status, and home ownership), and household head occupation. The vector Z_c includes town-level controls measured during childhood, including population bins, proximity to metropolitan areas, water system controls, local industrial composition, occupational composition, and home ownership rates. Region fixed effects, α_r, absorb broad differences across childhood Census regions. These controls are discussed at length in Section 3. Standard errors are clustered at the pipe district level (824 clusters), which groups towns sharing a common water system, to account for correlation in outcomes among individuals exposed to the same local water infrastructure and chemistry.

The omitted category in equation 6 is *Non-Lead* towns in *High Leaching Risk* environments. The coefficient β_1 captures the difference between Non-Lead towns in low- versus high-leaching environments. Because Non-Lead towns contain no lead infrastructure, β_1 near zero indicates that water chemistry does not independently affect labor market outcomes in the absence of lead pipes, providing a placebo test for the chemistry channel. Because β_1 enters as a main effect for all Low Leaching Risk towns, the coefficients β_2 and β_4 measure the effect of lead pipes within Low LR environments relative to *Non-Lead* $\times$ *Low LR*, while β_3 and β_5 measure the effect of lead pipes within High LR environments relative to the omitted *Non-Lead* $\times$ *High LR* group. Each lead-pipe coefficient therefore reflects a within-leaching-environment comparison between lead and non-lead towns.

The coefficients from equation 6 can be interpreted in two complementary ways. First, they describe differences across pipe types within a given chemical environment, which is the specification used in the main paper coefficient plots. These comparisons are straightforward to visualize because all coefficients are reported relative to a common omitted category. Second, the same specification yields differences across leaching environments within a given pipe type through linear combinations of the estimated coefficients. These within-pipe chemistry contrasts more directly reflect the identification strategy, since water chemistry provides plausibly exogenous variation in lead exposure conditional on pipe type.

Table 5 presents the regression results in these two forms. Panel A reports the coefficients directly from equation 6, corresponding to the specification used in the main-paper coefficient plots. Panel B reports the within-pipe-type chemistry contrasts, computed from equation 6 as $\beta_3 - \beta_2$ for Pure Lead towns and $\beta_5 - \beta_4$ for Mixed Lead towns. These contrasts indicate that the differences observed across towns are driven by elevated leaching risk within pure lead towns, providing direct empirical support for the role of water chemistry in generating the labor market penalties.

5.2 Occupational Sorting: Two-Tier Analysis

Figure 3, Panel A in the main paper examines whether childhood lead exposure affects occupational sorting by changing the probability that individuals work in lower-tier occupations in adulthood. Lower-tier occupations are defined as those with median earnings below the overall occupation-level median in 1940.

We estimate the same specification as in equation 6, but replace the dependent variable with an indicator for employment in a lower-tier occupation. Formally, we estimate

$$\begin{aligned}\mathbf{1}[\text{LowerTier}]_{ic} = {} & \beta_1 \text{LowLR}_c + \beta_2 \left(\text{PureLead}_c \times \text{LowLR}_c\right) + \beta_3 \left(\text{PureLead}_c \times \text{HighLR}_c\right) \\ & + \beta_4 \left(\text{MixLead}_c \times \text{LowLR}_c\right) + \beta_5 \left(\text{MixLead}_c \times \text{HighLR}_c\right) + X_i'\gamma + Z_c'\theta + \alpha_r + \varepsilon_{ic},\end{aligned} \tag{7}$$

The coefficients are interpreted in the same way as in the main labor market specification. For example, β_3 captures the difference in the probability of working in a lower-tier occupation for individuals from Pure Lead $\times$ High Leaching Risk towns relative to the omitted category of Non-Lead $\times$ High Leaching Risk towns. All control variables and fixed effects are identical to those used in equation 6. All estimates are IPW weighted.

Table 5: Lead and Labor Market Outcomes

	(1) Ln(Income)	(2) Weeks Worked	(3) Not in Labor Force	(4) Unemployed | LFP=1
Panel A: Weighted by Baseline IPW				
Low Leaching Risk	0.003	0.177	-0.004	0.002
	(0.015)	(0.265)	(0.005)	(0.003)
Pure Lead × Low LR	-0.012	0.206	-0.001	-0.001
	(0.012)	(0.163)	(0.001)	(0.003)
Pure Lead × High LR	-0.047***	-0.671**	-0.002	0.011***
	(0.016)	(0.265)	(0.004)	(0.004)
Mix Lead × Low LR	0.004	0.145	0.001	-0.001
	(0.013)	(0.153)	(0.001)	(0.003)
Mix Lead × High LR	-0.000	0.293	-0.004	0.002
	(0.016)	(0.278)	(0.004)	(0.003)
Observations	4,896,348	6,006,388	6,006,388	5,706,425
Adj. R-squared	0.222	0.047	0.006	0.026
Non-Lead Mean	974.503	40.460	0.052	0.082
Panel B: Difference Within Pipe Type and Across Chemistry Categories				
	Ln(Income)	Weeks Worked	Not in Labor Force	Unemployed | LFP=1
Pure Lead × High LR	-0.036**	-0.877***	-0.000	0.011***
	(0.018)	(0.306)	(0.005)	(0.004)
Mix Lead × High LR	-0.005	0.148	-0.005	0.003
	(0.020)	(0.322)	(0.005)	(0.004)
N	4896348	6006388	6006388	5706425

Notes: Panel A reports IPW-weighted estimates from equation 6. Panel B reports within-pipe-type chemistry contrasts (High LR minus Low LR) computed as linear combinations of Panel A coefficients. All estimates IPW weighted. Standard errors clustered at the pipe district level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

5.3 Within-Occupation Wage Effects

Figure 3 Panel B in the main paper examines whether the earnings penalty associated with childhood lead exposure persists after conditioning on occupational assignment, that is whether lead-exposed workers earn less than unexposed workers employed in the same occupation.

To isolate the wage component not explained by occupational sorting, we augment the main wage specification with detailed occupation fixed effects. Formally, we estimate

$$\begin{aligned} Y_{ic} =& \beta_1 \text{LowLR}_c + \beta_2 \text{PureLead}_c \times \text{LowLR}_c + \beta_3 \text{PureLead}_c \times \text{HighLR}_c \\ &+ \beta_4 \text{MixLead}_c \times \text{LowLR}_c + \beta_5 \text{MixLead}_c \times \text{HighLR}_c \\ &+ X_i'\gamma + Z_c'\theta + \alpha_r + \delta_{k(i)} + \varepsilon_{ic}, \end{aligned} \tag{8}$$

where Y_{ic} is log wage and salary income for individual i who grew up in town c. The addition over equation 6 is $\delta_{k(i)}$ which denotes fixed effects for the detailed 1950-basis occupation code of individual i. The occupation fixed effects absorb mean wage differences across job categories, so the lead exposure coefficients are identified from within-occupation variation in wages across individuals from towns with different pipe materials and water chemistry. The coefficient β_3 therefore captures the average wage penalty for men from Pure Lead × High Leaching Risk towns relative to those from Non-Lead × High Leaching Risk towns, holding occupation fixed. The vectors X_i, Z_c, and region fixed effects α_r are identical to those in equation 6. Standard errors are clustered at the pipe district level. All estimates are IPW weighted.

Because occupational placement is itself affected by lead exposure, conditioning on occupation introduces sample selection: among workers in a given occupation, the lead-exposed group excludes men displaced into lower-tier occupations by lead and retains those who achieved that occupation despite exposure. The within-occupation coefficient should therefore be interpreted descriptively, as evidence that the earnings penalty is not fully accounted for by occupational sorting, rather than as a clean causal estimate of a distinct channel operating on valued worker attributes in isolation.

Panel B of Figure 3 in the main paper shows that the wage penalty does not disappear after conditioning on detailed occupation. Among men in the same occupation, those who grew up in Pure Lead × High Leaching Risk towns earn 2.6% less than those from Non-Lead × High Leaching Risk towns ($p < 0.05$). The within-occupation penalty is larger in High Leaching Risk environments than in Low Leaching Risk environments for Pure Lead towns, consistent with the chemistry channel operating on valued worker attributes as well as occupational sorting. The Non-Lead chemistry contrast (β_1) remains near zero, confirming that water chemistry does not affect within-occupation wages in the absence of lead pipes. These results indicate that the earnings consequences of childhood lead exposure extend beyond occupational sorting: even conditional on working in the same occupation, men raised in environments with greater lead leaching earn measurably less.

5.4 Childhood SES and Adult Earnings Across Exposure Environments

The main labor market results document average earnings penalties from childhood lead exposure. A natural question is whether these penalties vary by childhood socioeconomic status, and in particular whether families with greater economic resources could partially shield their children from the consequences of exposure.

Measuring Childhood Socioeconomic Status. We proxy for childhood socioeconomic status using the father's predicted income rank when the individual was ages 0–10. Because the Census did not collect individual income data before 1940, we follow the approach developed in the recent intergenerational mobility literature [Abramitzky et al., 2021, Collins and Wanamaker, 2022, Ward, 2023, Buckles et al., 2023b] and predict father's income from observable characteristics recorded in the childhood census.

Following the methodology in Abramitzky et al. [2021] and Collins and Wanamaker [2022], we use the 1940 full-count Census, the first to report wage and salary income, to regress log earnings on father's detailed 1950-basis occupation code, a quadratic in age, state × race interactions, metro

Table 6: Occupation Sorting and Within-Occupation Wages

	(1) Below Median Occup.	(2) Ln(Income)
Low Leaching Risk	0.004	0.013
	(0.006)	(0.011)
Pure Lead	0.005	-0.015
	(0.005)	(0.010)
Mix Lead	0.001	-0.002
	(0.004)	(0.011)
Pure Lead × High LR	0.015*	-0.011
	(0.008)	(0.015)
Mix Lead × High LR	-0.007	0.001
	(0.008)	(0.016)
Observations	5,675,407	4,896,348
Adj. R-squared	0.115	0.360
Non-Lead Mean	0.414	974.503

Notes: Table reports IPW-weighted estimates from equation 6 and within-pipe-type chemistry contrasts (High LR minus Low LR) computed as linear combinations of coefficients. Standard errors clustered at the pipe district level. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

× race interactions, and occupation × region interactions. We then apply this model to predict income for fathers observed in the 1900, 1910, and 1920 Censuses based on their occupation, age, state, race, and metropolitan status. This approach captures geographic and demographic variation in the returns to a given occupation, rather than assigning a single national income value to all workers in the same job. We additionally adjust predicted values for self-employment and farming income using occupation × class-of-worker scalars constructed from the 1960 5% sample, following Collins and Wanamaker [2022]. Our primary measure uses the version that adjusts for both self-employment and local labor market conditions.

Fathers are then ranked within the full-count Census distribution in their respective census wave (1900, 1910, or 1920) on a 0–100 percentile scale and divided into quartiles, where Q1 represents the lowest 25% and Q4 the highest 25%. This rank-based approach is standard in the historical mobility literature and is less sensitive to changes in the shape of the income distribution over time than level-based measures [Collins and Wanamaker, 2022, Ward, 2023]. We use a single observation of father's occupation from the earliest linked census rather than averaging across multiple observations. Ward [2023] shows that averaging reduces classical measurement error and improves estimates of permanent income. Our interest, however, is in the household's economic position during the period of childhood lead exposure, not in a permanent income measure. The single-observation rank better reflects the resources available to the family at the time of exposure, which is the relevant margin for assessing whether household circumstances could buffer against lead's developmental effects. Because the predicted income values are based on occupation × geography × demographic cells rather than individual earnings, fathers in the same occupation, state, and demographic group receive the same predicted value. The quartile classification therefore captures broad differences in household economic standing rather than fine-grained individual-level variation.

Empirical Specification. To examine how the earnings penalty varies across the SES distribution, we extend the main specification by interacting the full set of pipe type × leaching-risk categories with indicators for father's income quartile. The quartile index runs from $q = 1$ to $q = 4$, with a single omitted category of Non-Lead × High Leaching Risk × Q1. Every coefficient therefore measures earnings relative to that common baseline, allowing direct comparisons of both levels and gradients across all six town types:

$$\begin{aligned} Y_{ic} = & \sum_{q=1}^{4} \phi_{1q} \, \text{NonLead}_c \times \text{LowLR}_c \times \mathbf{1}[Q_i = q] + \sum_{q=2}^{4} \phi_{2q} \, \text{NonLead}_c \times \text{HighLR}_c \times \mathbf{1}[Q_i = q] \\ & + \sum_{q=1}^{4} \phi_{3q} \, \text{PureLead}_c \times \text{LowLR}_c \times \mathbf{1}[Q_i = q] + \sum_{q=1}^{4} \phi_{4q} \, \text{PureLead}_c \times \text{HighLR}_c \times \mathbf{1}[Q_i = q] \\ & + \sum_{q=1}^{4} \phi_{5q} \, \text{MixLead}_c \times \text{LowLR}_c \times \mathbf{1}[Q_i = q] + \sum_{q=1}^{4} \phi_{6q} \, \text{MixLead}_c \times \text{HighLR}_c \times \mathbf{1}[Q_i = q] \\ & + X_i'\gamma + Z_c'\theta + \alpha_r + \varepsilon_{ic}, \end{aligned} \tag{9}$$

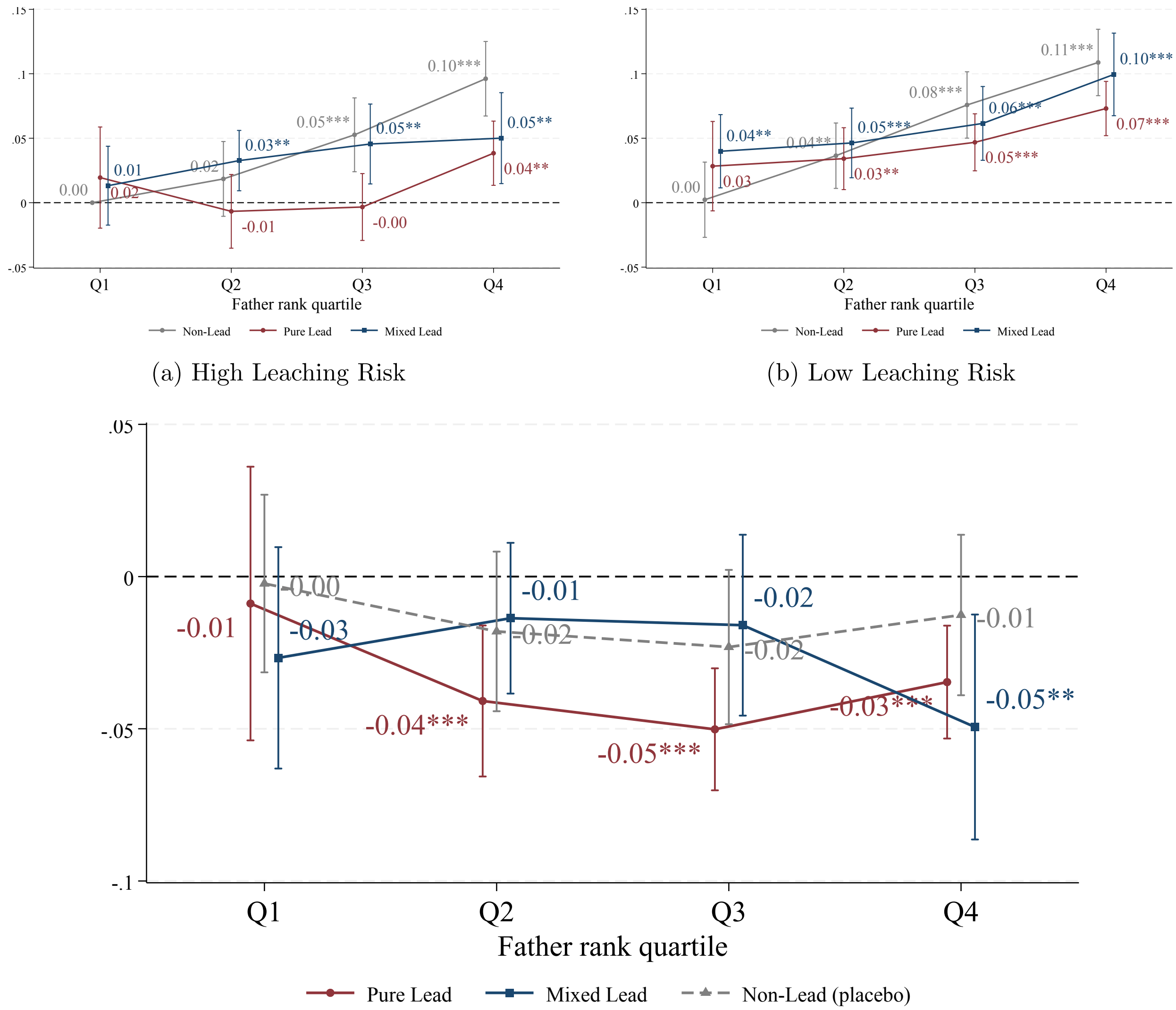
where $\mathbf{1}[Q_i = q]$ indicates father's income quartile, and the summation for the Non-Lead × High Leaching Risk group begins at $q = 2$ because the $q = 1$ cell is the omitted category. The vectors X_i, Z_c, and α_r are identical to those in equation 6. Each group's Q1 indicator enters as an explicit regressor (except the omitted cell), so the level of Q1 earnings varies freely across town types. Because all 23 included cells are measured against a common baseline, the specification reveals both how earnings levels differ across town types at each quartile and how the SES gradient (the slope from Q1 to Q4) varies with pipe material and water chemistry.

Figure 9 displays the results from equation (9). Panels (A) and (B) show earnings levels at each father-income quartile for all three pipe types, separately for High and Low Leaching Risk environments. Panel (C) plots the within-pipe-type chemistry contrast: the difference in earnings between High and Low Leaching Risk environments at each quartile, holding pipe type fixed.

In High Leaching Risk environments (Panel A), the three pipe types start at comparable Q1 earnings levels, confirming that the baseline is similar after conditioning on town characteristics. The Non-Lead gradient then rises sharply: Q4 sons earn 10% more than Q1 sons ($p < 0.01$). The Pure Lead gradient is compressed: Q2 and Q3 sons show no earnings advantage over Q1, and Q4 sons earn only 4% more ($p < 0.05$). Mixed Lead towns fall between the two, consistent with partial pipe coverage.

In Low Leaching Risk environments (Panel B), all three pipe types show upward-sloping gradients of broadly similar magnitude: Non-Lead Q4 = 11%, Pure Lead Q4 = 7%, Mixed Lead Q4 = 10% (all $p < 0.01$). The roughly parallel structure across pipe types in less corrosive water environments indicates that the substantial gradient compression observed in Panel (A) is driven by

Figure 9: Earnings by Father's Income Quartile Across Exposure Environments



(a) High Leaching Risk

(b) Low Leaching Risk

(c) Within-Pipe-Type Chemistry Contrast (High LR − Low LR)

Notes. Panels (A) and (B): each point shows estimated log wage and salary income relative to Non-Lead × High LR × Q1 from equation (9). Panel (C): difference in earnings between High and Low Leaching Risk environments at each quartile, within each pipe type. Non-Lead (dashed gray) serves as a placebo. All estimates IPW weighted. 90% confidence intervals. Standard errors clustered at the pipe district level.

chemistry-induced leaching rather than by lead infrastructure alone. The compression is sharpest in environments where water chemistry produces greater lead dissolution.

An important caveat applies to the across-pipe-type comparisons in Panels (A) and (B). The quartile composition of father's income rank differs across treatment cells. Lead pipe towns, which tend to be larger and more industrial, have fewer Q1 families: Q1 represents approximately 5% of the Pure Lead × High Leaching Risk cell compared to 17% of the Non-Lead × High Leaching Risk cell. "Bottom quartile" therefore describes a different population in lead vs. non-lead towns, and cross-pipe-type comparisons at Q1 should be interpreted accordingly. This compositional concern does not apply to Panel (C), which compares within the same pipe type across chemistry environments: Q1 families in Pure Lead × High Leaching Risk towns and Q1 families in Pure Lead × Low Leaching Risk towns are drawn from the same types of cities and face the same compositional selection. The within-pipe-type chemistry contrast is therefore the more reliable basis for assessing how lead exposure interacts with childhood SES.

Panel (C) isolates the chemistry effect within each pipe type. For Pure Lead towns, high-leaching chemistry reduces earnings at Q2 (-4%, $p < 0.01$), Q3 (-5%, $p < 0.01$), and Q4 (-3%, $p < 0.01$), with a small and insignificant penalty at Q1. The Non-Lead placebo is near zero at every quartile, confirming that chemistry matters only in the presence of lead pipes. Mixed Lead towns show a significant penalty at Q4 (-5%, $p < 0.05$) but no clear pattern at lower quartiles. The concentration of Pure Lead penalties at Q2 and Q3 indicates that corrosive water chemistry in lead-pipe cities harms sons from the broad middle of the SES distribution most severely, while Q1 sons, who face the weakest intergenerational earnings transmission even absent lead exposure, experience a smaller penalty. Beyond this weak intergenerational transmission, unobserved within-town differences in actual lead exposure and differential pre-1940 mortality may both contribute to the Q1 null, though neither channel is directly testable in our data.

The within-pipe-type chemistry contrast in Panel (C) applies the same identification logic as the main labor market results: holding infrastructure fixed, variation in water chemistry generates the earnings penalties. That this pattern extends across the SES distribution, compressing the earnings gradient in lead-pipe cities with corrosive water but not in lead-pipe cities with protective water chemistry, provides evidence that actual lead exposure, not city characteristics correlated with pipe adoption, drives the observed compression of intergenerational earnings transmission.

6 Additional Results

6.1 Educational Attainment

The clinical literature documents that childhood lead exposure impairs cognitive development, reduces IQ, and lowers academic performance [Canfield et al., 2003, Bellinger et al., 1992, Needleman, 2004]. A natural question is whether these effects translate into reduced educational attainment in our sample. We examine two margins: school attendance during childhood (ages 6–10, observed in the 1900, 1910, or 1920 Census) and completed years of education (observed in the 1940 Census).

Why Additional Education Controls. Lead-pipe towns were the largest and wealthiest cities in the sample, and unconditional attendance rates and mean years of completed schooling are both higher in lead-pipe towns than in non-lead towns at every age. Without additional controls for education supply, the regression would absorb the lead chemistry penalty into this educational advantage, potentially masking a true negative effect.

We therefore augment the baseline specification (equation 6) with three additional town-level controls that capture local education supply and norms. First, a five-category fixed effect for teachers per 100 children ages 6–20 (<1.5, 1.5–4, 4–6, 6–10, and 10+), measured from the childhood census. Second, an eight-category fixed effect for the number of college faculty in the town (0, 1–4, 5–10, 11–25, 26–50, 51–100, 101–200, and 200+), also from the childhood census. Third, for the completed years of education specification, a five-category fixed effect for the average years of schooling among rural residents in a 10–50 km ring around each town (<7, 7–8, 8–9, 9–10, and 10+ years), measured from the 1940 Census. This donut-ring measure captures the education norms of the surrounding rural population, net of urban composition effects. The third control is omitted from the childhood attendance specification because it is measured in 1940, after the attendance outcome is observed.

Childhood School Attendance. Panel (A) of Figure 10 displays estimates of childhood school attendance (an indicator for whether the child was reported as attending school in the childhood census, restricted to ages 6–10). The primary treatment cell, Pure Lead × High Leaching Risk, shows a negative and marginally significant effect: children in high-leaching lead-pipe towns were 2.3 percentage points less likely to be attending school ($p = 0.065$), a 2.8% decrease relative to the baseline attendance rate of 83%. The Non-Lead chemistry contrast (β_1) is near zero, confirming that water chemistry does not affect attendance in the absence of lead pipes. Mixed Lead × High Leaching Risk shows a smaller and insignificant effect (−1.0 pp), consistent with lower doses from partial pipe coverage.[5]

Completed Years of Education. Panel (B) of Figure 10 displays estimates using completed years of education in 1940 as the outcome. The primary treatment cell, Pure Lead × High Leaching Risk, is near zero and statistically indistinguishable from the omitted group. The remaining cells show no coherent pattern. Lead exposure through the water supply does not appear to have reduced the quantity of schooling that affected cohorts ultimately completed.

Interpreting the Contrast. The pattern across panels of Figure 10 suggests that lead exposure reduced school attendance during childhood but did not reduce completed years of education by adulthood. Several features of the institutional environment during this period are consistent with this pattern. Compulsory attendance laws, enacted and strengthened by states throughout this period, required attendance until age 14–16 and were actively enforced in the urban, lead-

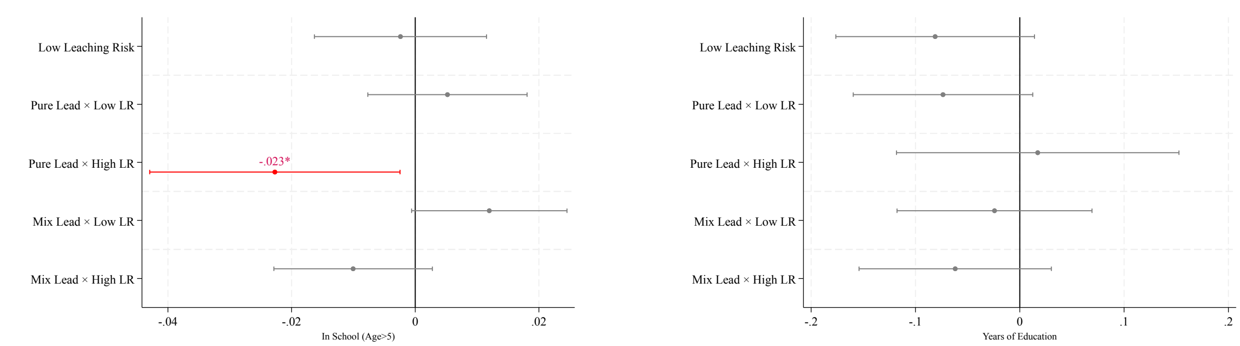
[5] The attendance variable is drawn from the childhood census (1900, 1910, or 1920), and the definition of school attendance varies across waves. In 1900, the reference period was the past 12 months; in 1910, approximately 7.5 months (since September 1); and in 1920, approximately 4 months. The shorter reference period in 1920 would tend to understate attendance gaps, since children who attended school earlier in the fall but withdrew by January would still be counted as attending. Birth cohort fixed effects absorb level differences across census waves.

pipe cities where our treatment effects are concentrated [Clay et al., 2012, 2021, Goldin and Katz, 2008a]. The rapid expansion of secondary education during the High School Movement further compressed educational attainment across the population, driven by local institutional investment and community wealth rather than individual cognitive capacity [Goldin, 1998, Goldin and Katz, 2008b]. These institutional forces may have pushed lead-affected children through the schooling system despite lower attendance rates during the early grades.

An alternative possibility is that the attendance gap reflects measurement rather than behavior. The childhood attendance variable captures a point-in-time snapshot from the census enumeration, while completed years of education reflects cumulative attainment over two decades. A child who was temporarily absent at the time of enumeration, whether due to illness, seasonal labor, or other disruption, would appear as not attending without experiencing any reduction in lifetime schooling. We cannot distinguish chronic absence from temporary disruption in these data.

Regardless of the explanation, the contrast between the two panels bears on the interpretation of the main labor market results. The null result on completed years of education does not rule out differences in the quality of learning during the same reported years of schooling, but it indicates that the earnings penalty is not explained by differences in reported educational attainment. The labor market effects of childhood lead exposure in this setting are consistent with evidence linking lead to reduced cognitive functioning [Needleman, 2004, Bellinger, 2004] and behavioral problems that persist into adulthood [Billings and Schnepel, 2018, Reyes, 2015, Gazze et al., 2024].

Figure 10: Lead Exposure and Education



(a) School Attendance, Ages 6–10 (b) Completed Years of Education

Notes. Coefficient estimates and 90% confidence intervals from equation (6) augmented with education supply controls. Panel (A) uses an indicator for school attendance during childhood (ages 6–10, observed in the 1900, 1910, or 1920 Census). Panel (B) uses completed years of education measured in the 1940 Census. Standard errors clustered at the pipe district level. Omitted category: Non-Lead × High Leaching Risk. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.2 Differences by Race

In the linked sample, nonwhite men are sparsely represented in lead-pipe cities. In the primary group of interest (Pure Lead × High Leaching Risk, 55 towns), nonwhite men comprise only 2.6% of the cell in the raw linked sample. Most towns in the cell contain at least one linked nonwhite resident, but the absolute number per town is very small. IPW weights correct for differential Census linkage rates at the individual level based on observable characteristics, but their capacity to remedy sparseness is limited when so few nonwhite individuals were successfully linked. The thinness of nonwhite representation in the linked treatment cell reflects the limits of the linked sample rather than a structural zero in the underlying population. Race heterogeneity analysis is not feasible in this setting: there is insufficient linked data in the relevant cells to estimate differential treatment effects by race, regardless of the weighting scheme applied.

6.3 Back-of-Envelope Cost Calculation

To illustrate the economic magnitude of the estimated effects, we combine the wage and unemployment results into a per-person measure of annual earnings losses. Because the log wage specification conditions on positive earnings, the 4.7% wage penalty and the 1.1 percentage point unemployment effect capture separate margins: the intensive margin (lower wages when working) and the extensive margin (higher probability of earning nothing in a given year). Expected annual earnings combine these two channels:

$$\text{E[earnings]} = \Pr(\text{employed}) \times \text{E[wage} \mid \text{employed]} \tag{10}$$

We present calculations using two baseline populations and two treatment effect estimates, yielding a range of per-person costs. For the baseline, we use either the Non-Lead town means ($\overline{W}$ = \$974.50; unemployment rate = 8.2%), which represent the counterfactual for the omitted category, or the full sample means ($\overline{W}$ = \$1,139.48; unemployment rate = 9.5%). For the treatment effect, we use either the across-pipe-type estimate (−4.7% wage penalty) or the more tightly identified within-pipe-type chemistry contrast (−3.6%). Table 7 reports the resulting annual earnings losses in both 1940 and 2026 dollars (CPI multiplier = 22.34).

Table 7: Annual Earnings Losses from Childhood Lead Exposure

	Non-Lead Baseline		Full Sample Baseline	
	1940 \$	2026 \$	1940 \$	2026 \$
Across-pipe-type (−4.7%, +1.1pp)	\$52	\$1,168	\$60	\$1,350
Within-pipe-type (−3.6%, +1.1pp)	\$43	\$950	\$49	\$1,099

Notes. Each cell reports the annual expected earnings loss per exposed individual: $\text{E[earnings]}_{\text{baseline}} - \text{E[earnings]}_{\text{exposed}}$, where expected earnings = Pr(employed) × E[wage | employed]. The across-pipe-type row applies the Pure Lead × High Leaching Risk coefficients from the main specification (−4.7% wage penalty, +1.1 percentage point unemployment). The within-pipe-type row uses the chemistry contrast within Pure Lead towns (−3.6%, +1.1pp). CPI adjustment uses Bureau of Labor Statistics CPI-U, 1940 to 2026 (×22.34).

The annual per-person earnings loss ranges from approximately \$950 to \$1,350 in 2026 dollars, depending on the baseline and treatment effect estimate used. To scale these annual losses over a working career, the relevant horizon depends on the persistence of the effect across the lifecycle. Our sample observes men at ages 20–50; the estimates therefore reflect average effects across this age range. If the penalty persists at similar magnitudes beyond age 50, cumulative losses over a full working life would be correspondingly larger. Applying the annual losses over the 30-year window directly observed in our data yields cumulative per-person losses ranging from approximately $28,500 to $40,500 in 2026 dollars.

Several features of this calculation make it conservative. The sample includes only men; women are excluded because fewer than 20% report positive wage income in 1940. The estimates capture only waterborne lead from municipal pipes, excluding exposure from lead paint, occupational sources, or other environmental pathways. The calculation reflects earnings losses only, omitting health costs, reduced life expectancy, and behavioral consequences documented in other studies. The ITT design assigns treatment at the town level; individual-level exposure almost certainly varied within towns, so the per-person cost for those with higher actual consumption of municipal water would exceed these estimates.

7 Robustness

7.1 Unweighted Results

To confirm that the IPW weights are not driving the main findings, Table 8 reproduces the main specification (equation 6) without inverse probability weights. The pattern of results is qualitatively similar to the IPW-weighted estimates, though some coefficients are slightly attenuated, consistent with the unweighted sample over-representing groups with higher baseline linking rates.

Table 8: Lead and Labor Market Outcomes: Unweighted

	(1) Ln(Income)	(2) Weeks Worked	(3) Not in Labor Force	(4) Unemployed | LFP=1
Low Leaching Risk	0.016	0.310	-0.002	-0.001
	(0.016)	(0.213)	(0.003)	(0.003)
Pure Lead × Low LR	-0.012	0.260	-0.002	-0.001
	(0.011)	(0.168)	(0.001)	(0.003)
Pure Lead × High LR	-0.040**	-0.587**	-0.000	0.009**
	(0.017)	(0.230)	(0.002)	(0.004)
Mix Lead × Low LR	-0.002	0.178	0.001	-0.001
	(0.013)	(0.156)	(0.001)	(0.003)
Mix Lead × High LR	0.010	0.362	-0.002	0.001
	(0.018)	(0.237)	(0.003)	(0.004)
Observations	4,896,348	6,006,388	6,006,388	5,706,425
Adj. R-squared	0.204	0.052	0.005	0.029
Non-Lead Mean	974.503	40.460	0.052	0.082

Notes: Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7.2 Alternative Chemistry Cutoffs

Robustness: Median Threshold. As a robustness check, we also estimate specifications using the median (50th percentile) predicted lead concentration as the classification threshold. This alternative rule divides towns into equal-sized High and Low leaching risk groups. Panel A of Table 9 reports results using this median cutoff. Estimated effects are generally smaller and less precisely estimated than in the main specification, which defines high-leaching towns as those in the top quartile of predicted lead concentrations. This attenuation is expected because the median cutoff groups together towns with moderate and high predicted lead leaching potential, reducing the contrast between treatment and comparison environments. Nevertheless, the signs of the estimated coefficients remain broadly consistent with the main specification.

Alternative Classification: pH Only. Prior literature examining lead exposure from water infrastructure has often classified towns based solely on water pH, without accounting for alkalinity or hardness [Clay et al., 2014]. To facilitate comparison with this earlier work, we also estimate specifications using a classification based only on water pH:

- **High Leaching (pH only):** pH < 6.5
- **Medium Leaching (pH only):** $6.5 \leq$ pH < 7.2
- **Low Leaching (pH only):** pH ≥ 7.2

Panel B of Table 9 presents results using this pH-only classification. The largest effects appear in towns with highly acidic water, where prior chemistry research predicts the greatest potential for lead dissolution. However, the pH-only classification collapses substantial variation in buffering capacity across towns with similar pH values. The predicted solubility model that incorporates both pH and hardness therefore provides a more nuanced measure of lead leaching potential, capturing variation in exposure within the intermediate pH range where water chemistry differences meaningfully influence lead release.

7.3 State Fixed Effects

The main specification controls for seven Census region fixed effects, an aggregation we prefer because it preserves potentially valuable comparisons across cities in different states that share similar industrial and economic structures. A more restrictive alternative is to absorb time-invariant state-level differences directly: cities within a state may behave similarly through shared influences such as state health offices, sanitary regulators, or engineering norms that shape municipal water-system decisions. Much of what could differ across cities within a region is already captured by our town-level controls for population, industrial composition, and metropolitan proximity. We nevertheless re-estimate the main wage specification replacing region fixed effects with state fixed effects as a stricter robustness check.

With state fixed effects, the identifying variation comes from towns within the same state that differ in pipe type or chemistry. Figure 11 reports the results. The Pure Lead × High Leaching

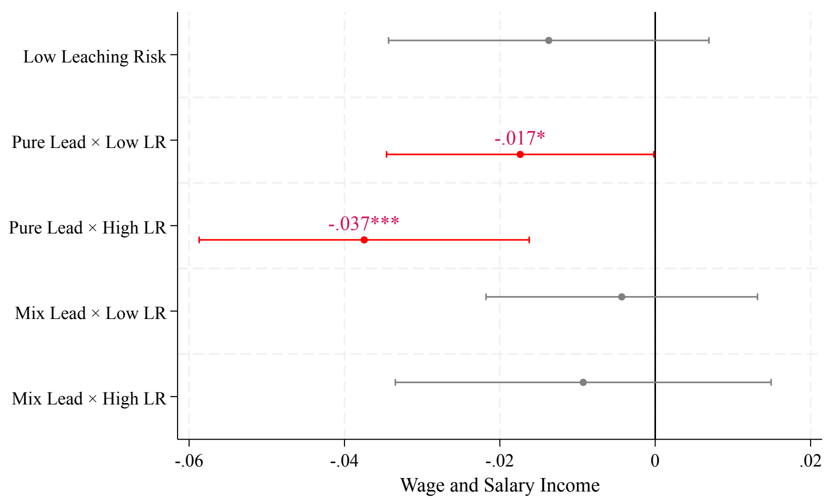
Table 9: Lead and Labor Market Outcomes Across Alternative Chemical Property Cutoffs

	(1) Ln(Income)	(2) Weeks Worked	(3) Not in Labor Force	(4) Unemployed | LFP=1
Panel A: Median Leaching Risk Cutoff				
Low Leaching Risk	0.004	0.022	0.001	0.002
	(0.014)	(0.194)	(0.003)	(0.002)
Pure Lead × Low LR	-0.018	0.055	-0.001	0.001
	(0.015)	(0.197)	(0.002)	(0.003)
Pure Lead × High LR	-0.032***	-0.260	-0.002	0.005*
	(0.012)	(0.209)	(0.002)	(0.003)
Mix Lead × Low LR	0.012	0.131	0.001	-0.002
	(0.015)	(0.175)	(0.001)	(0.003)
Mix Lead × High LR	-0.010	0.110	-0.000	0.003
	(0.013)	(0.191)	(0.002)	(0.003)
Observations	4,896,348	6,006,388	6,006,388	5,706,425
Adj. R-squared	0.222	0.047	0.006	0.026
Non-Lead Mean	974.503	40.460	0.052	0.082
Panel B: pH Acidity				
	Ln(Income)	Weeks Worked	Not in Labor Force	Unemployed | LFP=1
Med. Acidity	0.002	-0.256	0.005	0.003
	(0.019)	(0.357)	(0.005)	(0.004)
Low Acidity	0.004	-0.049	0.002	0.003
	(0.019)	(0.299)	(0.002)	(0.004)
Pure Lead × High Acidity	-0.050**	-0.904**	0.003	0.015***
	(0.021)	(0.386)	(0.003)	(0.005)
Pure Lead × Med. Acidity	-0.031**	-0.165	-0.004	0.004
	(0.015)	(0.300)	(0.004)	(0.004)
Pure Lead × Low Acidity	-0.016	0.077	-0.001	0.000
	(0.013)	(0.181)	(0.002)	(0.003)
Mix Lead × High Acidity	0.007	-0.095	0.002	0.010
	(0.024)	(0.507)	(0.003)	(0.007)
Mix Lead × Med. Acidity	-0.015	0.027	-0.001	0.005
	(0.015)	(0.250)	(0.004)	(0.003)
Mix Lead × Low Acidity	0.010	0.235	0.001	-0.003
	(0.015)	(0.160)	(0.001)	(0.002)
Observations	4,896,348	6,006,388	6,006,388	5,706,425
Adj. R-squared	0.222	0.047	0.006	0.026
Non-Lead Mean	974.503	40.460	0.052	0.082

Notes: Table reproduces main specification, equation 6, using the median leaching risk cutoff in Panel A, and for pH acidity cutoffs in Panel B on the main labor market outcomes of interest. Significance levels are denoted by * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Risk coefficient is -0.037 ($p < 0.01$), compared to -0.047 in the main specification, an attenuation of approximately 21%. This reduction is consistent with some state-level factors being correlated with both lead adoption patterns and labor market outcomes within regions. The estimate remains economically meaningful and statistically significant. Pure Lead $\times$ Low Leaching Risk is -0.017 ($p < 0.10$), indicating that lead pipes impose some wage penalty even in low-leaching environments once state-level heterogeneity is absorbed. The Non-Lead chemistry contrast (β_1) remains near zero and insignificant, confirming that water chemistry variation does not affect wages in the absence of lead pipes even within these narrower within-state comparisons. Both Mixed Lead coefficients are near zero and insignificant.

Figure 11: Replacing Region with State Fixed Effects



Notes. Main wage specification with state fixed effects replacing region fixed effects. 90% confidence intervals. Standard errors clustered at the pipe district level.

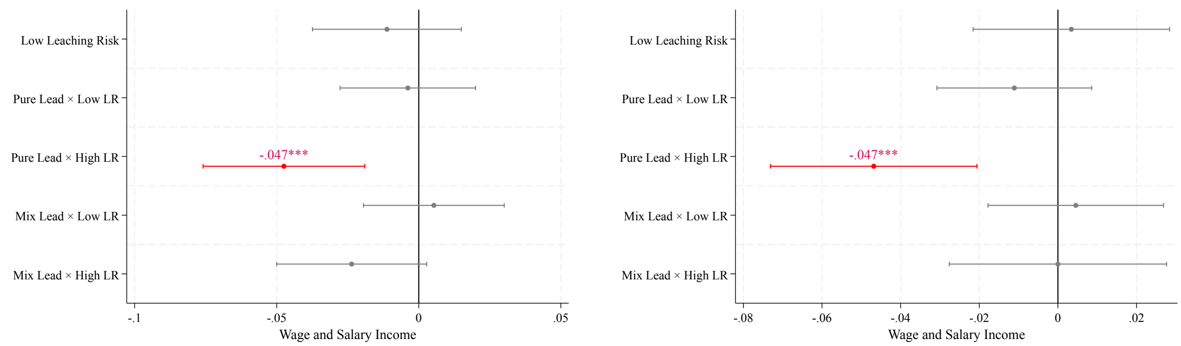
7.4 Excluding Finished Water and Modern Pipe Assignment

Two features of the data construction warrant direct sensitivity checks. First, water chemistry measurements for 32 towns (3% of the sample) reflect finished (treated) water rather than raw source water. Because water treatment, particularly lime addition, raises pH, finished water measurements may understate the corrosiveness of the raw water that historically entered lead service pipes. Misclassifying these towns into lower leaching risk categories would attenuate our chemistry contrast and bias estimates toward zero. Second, pipe material classifications for 74 towns come from modern lead service line inventories rather than the Baker Manual (see Section 1.2). Although the classification rules are conservative, modern inventories may not perfectly reflect historical pipe materials if lead lines were replaced at different rates across towns.

Figure 12 addresses both concerns. The left panel re-estimates the main specification after dropping all towns with finished water chemistry measurements. The right panel drops towns classified using modern LSL inventories. In both cases, the Pure Lead $\times$ High Leaching Risk coefficient is

-0.047 ($p < 0.01$), identical to the full-sample estimate. The remaining coefficients are also stable. These results confirm that neither the inclusion of treated-water chemistry measurements nor the use of modern pipe inventories drives the main findings.

Figure 12: Excluding Finished Water Towns (left) and Modern Pipe Assignments (right)



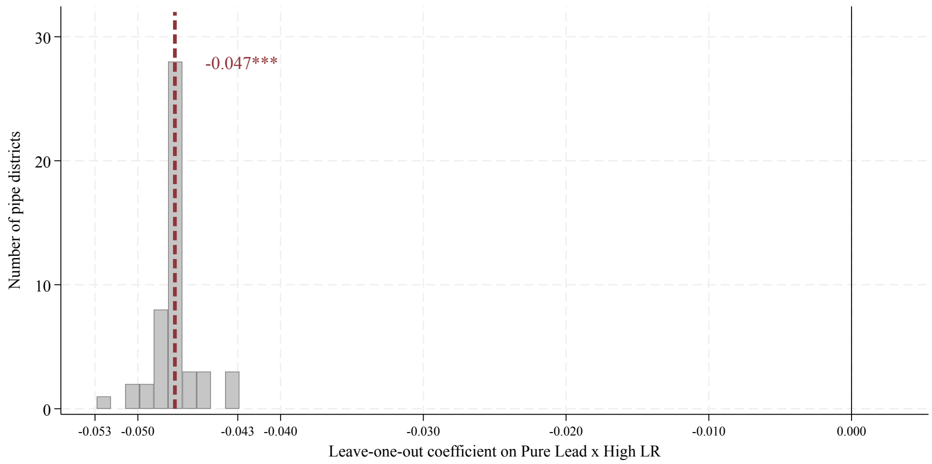
Notes. Main wage specification excluding towns with finished (treated) water chemistry measurements (left panel) and towns with pipe classifications from modern lead service line inventories (right panel). 90% confidence intervals. Standard errors clustered at the pipe district level.

7.5 Leave-One-Out for Highest Exposure Locations

The Pure Lead × High Leaching Risk cell contains 55 towns organized into 50 pipe districts. If the main result were driven by a single large or unusual pipe district, the finding would be less credible as evidence of a general relationship between lead exposure and labor market outcomes. To assess this, we sequentially drop each of the 50 pipe districts and re-estimate the main specification on the remaining sample.

Figure 13 reports the distribution of the 50 leave-one-out coefficients. The estimates range from -0.053 to -0.043, tightly clustered around the full-sample estimate of -0.047. No single pipe district, when removed, shifts the coefficient by more than 0.006 log points in either direction. The result is not driven by any individual location.

Figure 13: Leave-One-Out: Pure Lead × High Leaching Risk Pipe Districts



Notes. Distribution of Pure Lead × High Leaching Risk coefficients from 50 leave-one-out regressions, each dropping one of the 50 pipe districts in the highest exposure cell. Red dashed line indicates the full-sample estimate (−0.047). Main specification with IPW weights. Standard errors clustered at the pipe district level.

References

Ran Abramitzky, Leah Boustan, Elisa Jácome, and Santiago Pérez. Intergenerational mobility of immigrants in the united states over two centuries. *American Economic Review*, 111(2):580–608, 2021.

Martha Bailey, Connor Cole, and Catherine Massey. Simple strategies for improving inference with linked data: A case study of the 1850–1930 IPUMS linked representative historical samples. *Historical Methods: A Journal of Quantitative and Interdisciplinary History*, 53(2):80–93, 2020.

David C Bellinger. Lead. *Pediatrics*, 113(3):1016–1022, 2004.

David C Bellinger, Karen M Stiles, and Herbert L Needleman. Low-level lead exposure, intelligence and academic achievement: a long-term follow-up study. *Pediatrics*, 90(6):855–861, 1992.

Enrico Berkes, Ezra Karger, and Peter Nencka. The census place project: A method for geolocating unstructured place names. *Explorations in Economic History*, 87:101477, 2023.

Stephen B Billings and Kevin T Schnepel. Life after lead: Effects of early interventions for children exposed to lead. *American Economic Journal: Applied Economics*, 10(3):315–344, 2018.

Claude E Boyd. Total hardness. In *Water Quality: An Introduction*, pages 179–187. Springer, 2015.

Kasey Buckles, Joseph Price, Zachary Ward, and Haley E.B. Wilbert. Family trees and falling apples: Historical intergenerational mobility estimates for women and men. Working Paper 31918, National Bureau of Economic Research, 2023a.

Kasey Buckles, Joseph Price, Zachary Ward, and Haley EB Wilbert. Family trees and falling apples: Historical intergenerational mobility estimates for women and men. *NBER Working Paper*, 31918, 2023b.

Kasey Buckles, Adrian Haws, Joseph Price, and Haley EB Wilbert. Breakthroughs in historical record linking using genealogy data: The census tree project. *Explorations in Economic History*, page 101717, 2025.

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

PT Cardew. A method for assessing the effect of water quality changes on plumbosolvency using random daytime sampling. *Water Research*, 37(12):2821–2832, 2003.

Karen Clay, Jeff Lingwall, and Jr. Stephens, Melvin. Do schooling laws matter? evidence from the introduction of compulsory attendance laws in the united states. Working Paper 18477, National Bureau of Economic Research, October 2012. URL `http://www.nber.org/papers/w18477`.

Karen Clay, Werner Troesken, and Michael Haines. Lead and mortality. *Review of Economics and Statistics*, 96(3):458–470, 2014.

Karen Clay, Jeff Lingwall, and Melvin Stephens Jr. Laws, educational outcomes, and returns to schooling evidence from the first wave of us state compulsory attendance laws. *Labour Economics*, 68:101935, 2021.

William J Collins and Marianne H Wanamaker. African american intergenerational economic mobility since 1880. *American Economic Journal: Applied Economics*, 14(3):84–117, 2022.

Charles Norman Durfor and Edith Becker. *Public water supplies of the 100 largest cities in the United States, 1962*. Number 1812. US Government Printing Office, 1964.

James J Feigenbaum and Christopher Muller. Lead exposure and violent crime in the early twentieth century. *Explorations in Economic History*, 62:51–86, 2016.

Joseph P Ferrie, Karen Rolf, and Werner Troesken. Cognitive disparities, lead plumbing, and water chemistry: Prior exposure to water-borne lead and intelligence test scores among world war two us army enlistees. *Economics & Human Biology*, 10(1):98–111, 2012.

Jason Fletcher and Hamid Noghanibehambari. Toxified to the bone: Early-life and childhood exposure to lead and men's old-age mortality. Working Paper 31957, National Bureau of Economic Research, 2023.

Ludovica Gazze, Claudia Persico, and Sandra Spirovska. The long-run spillover effects of pollution: How exposure to lead affects everyone in the classroom. *Journal of Labor Economics*, 42(2): 357–393, 2024.

Claudia Goldin. America's graduation from high school: The evolution and spread of secondary schooling in the twentieth century. *The Journal of Economic History*, 58(2):345–374, 1998.

Claudia Goldin and Lawrence F Katz. Mass secondary schooling and the state: the role of state compulsion in the high school movement. In *Understanding long-run economic growth: Geography, institutions, and the knowledge economy*, pages 275–310. University of Chicago Press, 2008a.

Claudia Dale Goldin and Lawrence F Katz. *The race between education and technology*. harvard university press, 2008b.

Paul B Hammond. Metabolism of lead. *Lead Absorption in Children: Management, Clinical, and Environmental Aspects, Urban and Schwarzenberg, Baltimore, Munich*, 1982.

EW Lohr and SK Love. Industrial utility of public water supplies in the united states 1952: Parts 1 and 2. *Geological Survey Water-Supply Papers*, 1299, 1954.

Michelle Marcus. Can we get the lead out? updates to the lead and copper rule for public drinking water. *Review of Environmental Economics and Policy*, 20(1), 2026.

MR Moore. Plumbosolvency of waters. *Nature*, 243(5404):222–223, 1973.

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

Joseph Price, Kasey Buckles, Jacob Van Leeuwen, and Isaac Riley. Combining family history and machine learning to link historical records: The Census Tree data set. *Explorations in Economic History*, 80:101391, 2021. doi: 10.1016/j.eeh.2021.101391.

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

Steven Ruggles, Catherine A Fitch, Ronald Goeken, J David Hacker, Matt A Nelson, Evan Roberts, Megan Schouweiler, and Matthew Sobek. Ipums ancestry full count data: Version 3.0 [dataset]. *Minneapolis, MN: IPUMS*, 2021.

Michael R Schock. Causes of temporal variability of lead in domestic plumbing systems. *Environmental monitoring and assessment*, 15(1):59–82, 1990.

Michael R Schock and Marvin C Gardels. Plumbosolvency reduction by high ph and low carbonate—solubility relationships. *Journal-American Water Works Association*, 75(2):87–91, 1983.

Werner Troesken. *The great lead water pipe disaster*. MIT Press, 2006.

Werner Troesken. Lead water pipes and infant mortality at the turn of the twentieth century. *Journal of Human Resources*, 43(3):553–575, 2008.

Zachary Ward. Intergenerational mobility in american history: Accounting for race and measurement error. *American Economic Review*, 113(12):3213–3248, 2023.

Jeffrey M Wooldridge. Inverse probability weighted estimation for general missing data problems. *Journal of Econometrics*, 141(2):1281–1301, 2007.