The Impact of Lead Water Pollution on Birth Outcomes: A Natural Experiment in Scotland[1]

Anthony Higney*[2], Nick Hanley[3] and Mirko Moro[4]

Abstract

We explore whether maternal lead exposure affects birthweights and child mortality in a setting where average blood lead levels were higher than in any country today. We analyse two drinking water interventions in Scotland that reduced lead levels in Glasgow and Edinburgh from 1978 onwards. Using both a staggered difference-in-differences design and a regression discontinuity design, we examine administrative data of over 350,000 births between 1975 and 2000. Our findings indicate this lead abatement did not significantly increase birthweights or reduce under-5 mortality.

JEL codes: I18; Q53

Keywords: Under-5 mortality; pollution; lead; Difference-in-Differences

1 Introduction

An estimated 2.4 million children die within their first year of life globally (UN IGME 2021). A further estimated 2 million are stillborn (UNICEF 2022a). High levels of lead in the environment have, historically, been linked with poor health outcomes for children. With an estimated 1-in-3 children having elevated levels of lead in their systems (GBD, 2019), and the global burden of lead estimated to be responsible for as many as 900,000

[1]We thank Jeffrey Wooldridge, Nils Braakmann, Markus Gehrsitz, Danny Campbell and Hector Gutierrez Rufrancos for their comments on this paper. We also thank participants at the European Association of Environmental and Resource Economists (EAERE) conference and the Scotland and Northern England Conference in Applied Microeconomics. Any remaining errors are our own. We are also grateful for the financial support from the UK Economic and Social Research Council.
* Corresponding author. Anthony.Higney@glasgow.ac.uk
[2] School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow.
[3] School of Biodiversity, One Health and Veterinary Medicine, University of Glasgow.
[4]Division of Economics, University of Stirling, Stirling.

deaths a year (UNICEF, 2020), reducing lead pollution may be one route to prevent infant death and morbidity.

In this paper, we follow Troesken (2006) in examining the impact of lead water pollution on infant health outcomes. Lead can contaminate drinking water through chemical reactions to plumbing. Metal from lead pipes and fixtures can dissolve or erode into the water supply. This reaction is particularly severe when the water has low mineral content or high acidity, and it is said to be highly *plumbosolvent.*

We focus on the effect of lead pollution on *birthweight* and *under-5 mortality* (birthweight being a proxy for a wide range of future life outcomes, see section 2.1). Previous evidence shows mixed findings on birthweights and on under-5 mortality (see section 2.1), with most studies relying on correlational estimates of the relationship between lead pollution and birth outcomes, or low sample sizes. The very few quasi-experimental papers, reviewed in Clay et al. (2024), tend to find an effect, but the quasi-experimental paper with the largest sample size (Grönqvist et al., 2020, appendix E), does not. In their review of lead and infant health papers, Clay et al. (2024) cite the paucity of evidence and call for more research on this topic. We directly answer that call.

Our contribution is to examine the effects of lead pollution on infant health outcomes in a setting where blood lead levels were higher than the average levels found in *any* country today (Ericson et al., 2021). We use rich administrative data containing all births in Scotland's two largest cities – Glasgow and Edinburgh – and the surrounding areas over the period 1975-2000. We link this data with home address at time of birth, mother's characteristics, and infant health outcomes up to 5 years later. We combine this administrative data with plausibly exogenous variation in lead exposure from two separate interventions by public drinking water supply agencies in Glasgow and Edinburgh. This combination of data allows us to credibly identify the effect of lead on birthweight and under-5 mortality. The interventions we examine sharply reduced water lead levels and blood lead levels in both cities. Linking the data to home address allows us to capture if a mother and child lived in an area subject to the lead reduction treatment at the time of birth.

Our main research design, based on a robust, staggered difference-in-differences approach, improves upon most of the previous literature which is based on selection on

observables as an identification strategy. Selection on observables may result in biased estimates, as lead pollution is correlated with socio-economic factors. Additionally, many of the studies have small sample sizes, while our sample includes over 650,000 children.

The case studies of Glasgow and Edinburgh are noteworthy as these are areas with historically high levels of lead before the interventions. Measured average blood lead levels in Glasgow were 18μg/dl for everyone (Quinn, 1985) and 15μg/dl for mothers (Watt et al., 1996a), whereas the highest average blood lead levels today in any country is 11.4μg/dl in Pakistan (Ericson et al., 2021). Edinburgh and Glasgow were characterised by acidic soft water which made them especially plumbosolvent. In 1975, 33% of households in Scotland had water lead levels above 50μg/l, compared to only 10% in England. Glasgow was particularly affected, with 50% of households surveyed having water lead levels above 100μg/l, mainly due to the nature of the soil chemistry from which drinking water was collected (Quinn, 1985; Potter, 1997; Richards and Moore, 1984). We rely extensively on the long-running Glasgow (Watt et al. 1996a) and Edinburgh (Macintyre et al. 1998) lead studies, which meticulously detailed and researched the reductions in the water and blood lead levels over the 1980s and 1990s.

We do not find evidence for an effect of lead in drinking water on birthweights or under-5 mortality. We perform a large variety of robustness checks, including an alternative regression discontinuity design, and find similar results in all cases.

We believe that, although high lead levels clearly have an effect on many outcomes, our study found no discernible effect on birth and infant outcomes at the lead levels present in Scotland during the period under investigation. Our findings suggest that when lead significant strides have already been made towards improving infant health, as was the case in Scotland from the mid-1970s onwards, there are limited short-run effects of lead remediation on acute health outcomes. However, it is plausible that exposure to even low levels of lead at birth or during early childhood could have long-term impacts on critical outcomes such as educational attainment, propensity to commit crime, and productivity (see section 2.1). We conclude that while lead remediation is still worthwhile, our study does have implications for the importance allocated to lead remediation compared to other infant outcome interventions in specific settings. If such "low-hanging fruit" policies as improved nutrition or better neonatal healthcare are still

to be implemented, and resources are constrained, then it may be better to prioritise these interventions over lead remediation.

2 Background

2.1 Lead Pollution and Birth Outcomes

A child is first exposed to lead pollution through the placenta (Dorea and Donangelo, 2006). A mother's exposure to lead can in turn expose a foetus to lead. Furthermore, due to increased bone remodelling, previous maternal lead pollution can affect the foetus, as both lead and calcium (chemically similar) are released from the bones at an increased rate during pregnancy (Yurdakök, 2012). Maternal and infant lead levels are of similar magnitudes and highly correlated (Al-Saleh et al., 1995), but the relationship between exposure and absorption of lead is complex. For example, it is mitigated by maternal calcium intake (Dorea and Donangelo, 2006). Therefore, there are mediators between lead exposure and the damage it may cause.

A large literature has found diverse impacts of lead pollution. Biological harms include damaged nerve system and brain development when young (Cecil et al., 2008, Brubaker et al., 2009), and at higher levels abdominal pain, headaches, and seizures (WHO, 2010). Behavioural harms include aggressiveness (Marcus et al., 2010), worse memory, and lower attention span (Vlasak et al., 2019). The wider socio-economic impacts resulting from these include increased propensity to commit crime (Higney et al., 2022), lower educational attainment (Hollingsworth et al, 2022, Zheng, 2021), and possibly lower productivity due to health damage (He and Ji, 2021).

Exposure to lead pollution can have significant negative impacts on the development of children, both before and after birth. In severe cases, it can even result in stillbirth or death. Numerous studies have been conducted to determine the extent of these damages. In this paper, we focus on the effect of lead pollution on *birthweight* and *under-5 mortality*. We use birthweight because it is a generally accepted proxy for future health outcomes. It is associated with a wide range of health outcomes such as higher cardiovascular and cancer deaths, diabetes, and obesity, as well as more immediate health outcomes such as under-5 mortality and morbidity (Law, 2002, Wilcox, 2001;

Chatterji et al., 2014; Behrman and Rosenzweig, 2004;Royer 2006). However, it is generally not thought to be low birthweight itself that *causes* these harms, rather it is a proxy for underlying biological conditions, such as low nutrient ingestion in the womb or premature birth (Wilcox, 2001). We use under-5 mortality as it is the most damaging of possible harms from lead ingestion.

A number of studies estimate the relationship between lead exposure and birthweight. Xie et al. (2013) find a negative correlation between maternal or cord lead levels and birthweight in 252 infants. Similarly, Bornschein, R.L. et al. (1989) find a negative link between maternal blood lead levels and birthweight in 202 inner city infants. Taylor et al. (2014) find that 12% of infants whose mothers have elevated levels of lead (>5μg/dl) have low birthweight compared to 10% when lead levels are lower. In contrast, Azayo et al. (2009) find no association between maternal blood lead levels and birthweight in 150 women in Tanzania, but the average lead levels were below 5μg/dl, which is the threshold used by the WHO guidelines (WHO, 2021). Golmohammadi et al. (2007) use a sample with much higher average lead levels but also find no association in their sample of 89 infants in Iran. McMichael et al. (1986) found no association with birthweight for 749 mothers in Australia, although they do find an association with other outcomes such as spontaneous abortion.

In summary, the correlational findings on birthweight are mixed, in terms of both disagreeing on the presence of an effect and on the level of lead at which an effect is found. However, one problem with the previous papers is that they rely on selection on observables as an identification strategy. This likely biases estimates, as lead exposure is often confounded with poverty, race, and education. Many of the studies also have low sample sizes, and so may be inadequately powered. Recently, studies with improved identification strategies have examined the relationship between lead and birthweights. Most of these quasi-experimental studies are reviewed in Clay et al. (2024) and tend to find an effect. For example, Dave and Yang (2022) look at a setting where the pH on one side of the water supply in Newark fell sharply, and therefore began leeching lead from pipes again, while it remained steady on the other side of the city. They found an effect on birthweight, but it became smaller and not significant when more post treatment years are added. They rationalise this as showing the effects of mitigation strategies by mothers, such as moving to bottled water, once the increased lead levels were widely

known. Missing from the main results of the review was Grönqvist et al. (2020) who use an instrument of local lead moss levels for blood lead levels with a sample of 800,000 children in Sweden and find no effect on birthweights or premature births.

For spontaneous abortion (before 28 weeks) and stillbirths (after 28 weeks) high levels of lead have long been known to have an effect. So much so, that lead oxide was described as being used as an abortifacient by women in the 1800s (Hall and Ransom, 1906). In some cases, the amounts of lead ingested were strong enough to cause lead poisoning in the mother (Ransom, 1900). There are many papers which have examined the effects on spontaneous abortions, stillbirths and their correlates. Falcon et al (2003) find that premature births and pregnancy anomalies tended to have higher levels of lead in the placentas of 83 births (although they find no association with birthweight). Wibberly et al (1977) found that lead levels were higher in placentas where a neonatal death occurred in Birmingham. In contrast, McMichael et al. (1986) do not find any difference in pre-pregnancy maternal blood lead levels for neonatal deaths and other births. Angell and Lavery (1982) collected cord blood lead levels in 635 cases and found no relationship with lead levels and pregnancy complications that might lead to death such as preterm delivery or premature membrane rupture, although they did not look at spontaneous abortions/stillbirths directly. Vinceti et al. (2001) examine historical birth anomalies in a heavily lead polluted area of northern Italy. They find increased oral clefts and other disabilities but no increase in neural tube defects.

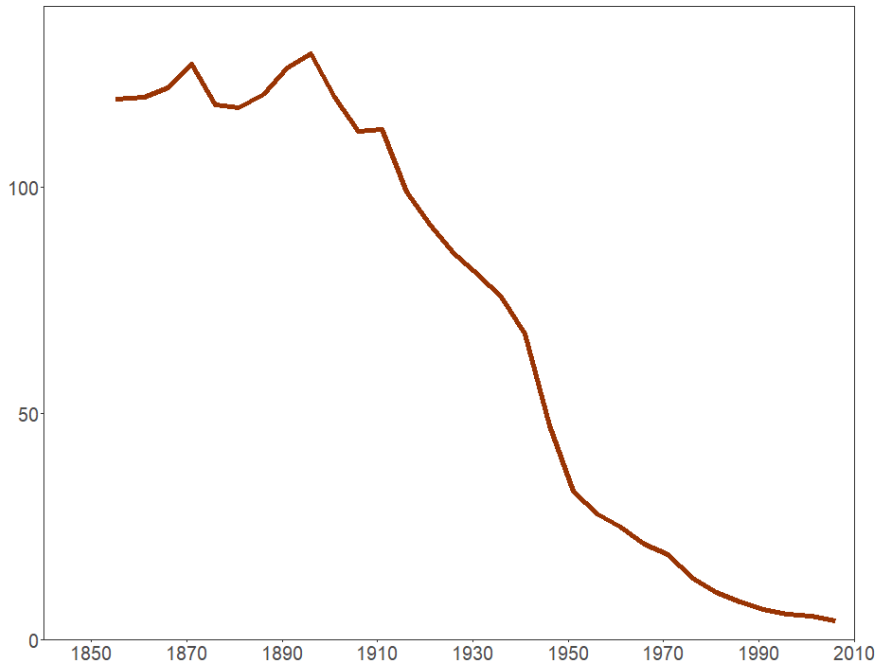
Looking specifically at studies which use quasi-experiments, these are again reviewed in Clay et al. (2024). They find 3 studies that looked at mortality and all find an effect. As an example, Clay et al. (2014) use the differences in city water pH levels as an instrument for lead exposure and find lead increased deaths. Although not included in the review, Edwards (2014) also finds that a short-term spike in lead water pollution in Washington DC (due to a change in the chemical treatment) resulted in an increase of the foetal death rate. Grönqvist et al. (2020) do not look specifically at deaths but do not find any effect on premature births.

2.2 Lead Plumbing and Water Treatment in Glasgow and Edinburgh

Scotland has seen greatly reduced infant deaths from all causes since 1900 (figure 1). Starting in the 1970s, interventions to reduce the amount of lead in drinking water

supplies began in Edinburgh and Glasgow, and were improved upon in the 80s and 90s. This was after infant deaths and stillbirths had already sharply reduced, thanks to improved nutrition, hygiene, and health practices. Therefore, this is a setting where the relatively easy gains had already been exhausted, and lead might be thought to account for a larger share of the remaining deaths and pregnancy complications.

Figure 1 – Deaths Within First Year of Life, per 1000 Births in Scotland



Source: National Records of Scotland (2022)

Lead piping was widely used in Scotland before being banned for new work in 1968 (Richards et al., 1980). Lead is malleable, relatively cheap, and has an extremely long life as infrastructure (Feigenbaum and Muller, 2014, and Krebs, 2019). Lead piping began to be phased out from the 70s in Scotland, but still, in the 90s, as many as 589,000 homes in Scotland were estimated to contain lead pipes (Potter, 1997), around 30% of the total. There were also as many as 60,000 water storage tanks made of lead, mostly in Glasgow and Edinburgh (Krebs, 2019). These were used because water service was still intermittent in the first half of the 20th century. The tanks allowed households to store and use water during any non-flowing periods.

The reason lead water pipes have not been entirely replaced are twofold: 1) It is expensive, and 2) Homeowners do not know they have lead pipes and that they are responsible for their replacement. Communication pipes are owned by the water

supplier and those made of lead have now all been replaced (Akoumianaki, 2017). Internal lead piping still exists in many households but has also been gradually replaced and is estimated to only account for 20-30% of the remaining lead pollution in home water supplies in the UK (Akoumianaki, 2017). The main pollution burden is thought to come from the lead supply pipes (also called service pipes). These are the responsibility of the property owner to replace, but are underground, and therefore difficult to see. An added complication is that property owners may not know they have the responsibility to obtain replacement, even though grants are available. Today there are estimated to be 273,000 homes out of 2.6 million in Scotland with lead piping (Robertson et al., 2020). Watt et al. (2000) estimated as many as 160,000 households out of 300,000 in Glasgow alone had a lead service pipe in 2000.

The dangerous combination of naturally acidic water chemistry and lead pipes began to be taken seriously in the 1970s. The UK's Department for the Environment carried out a series of surveys of blood lead levels in the 70s and 80s. The findings were that "The highest blood lead concentrations were related to plumbosolvent water" and not distance to roads (Quinn, 1985). The acidic soft water in Scotland's two largest cities made piping there extremely plumbosolvent. In 1975 surveys found 33% of households in Scotland had water lead levels above 50µg/l, compared to 10% in England (Potter 1997). Glasgow was especially viewed as a problem, with 50% of household surveyed having water lead levels above 100µg/l (Richards and Moore, 1984). Average blood lead levels were higher in both cities than in any country today (figure 5 and Ericson et al., 2021).

This paper exploits two interventions to reduce lead in the water supply. We treat these as natural experiments, given the somewhat arbitrary assignment of water treatment to Glasgow and Edinburgh (treatment groups), while leaving many adjacent areas untouched until years later (control groups). These control areas were not only treated later, but also had higher pH levels for their untreated water than the treatment areas. In particular, we rely on these plausibly exogenous variation of lead to identify treatment effects of lead on birth outcomes:

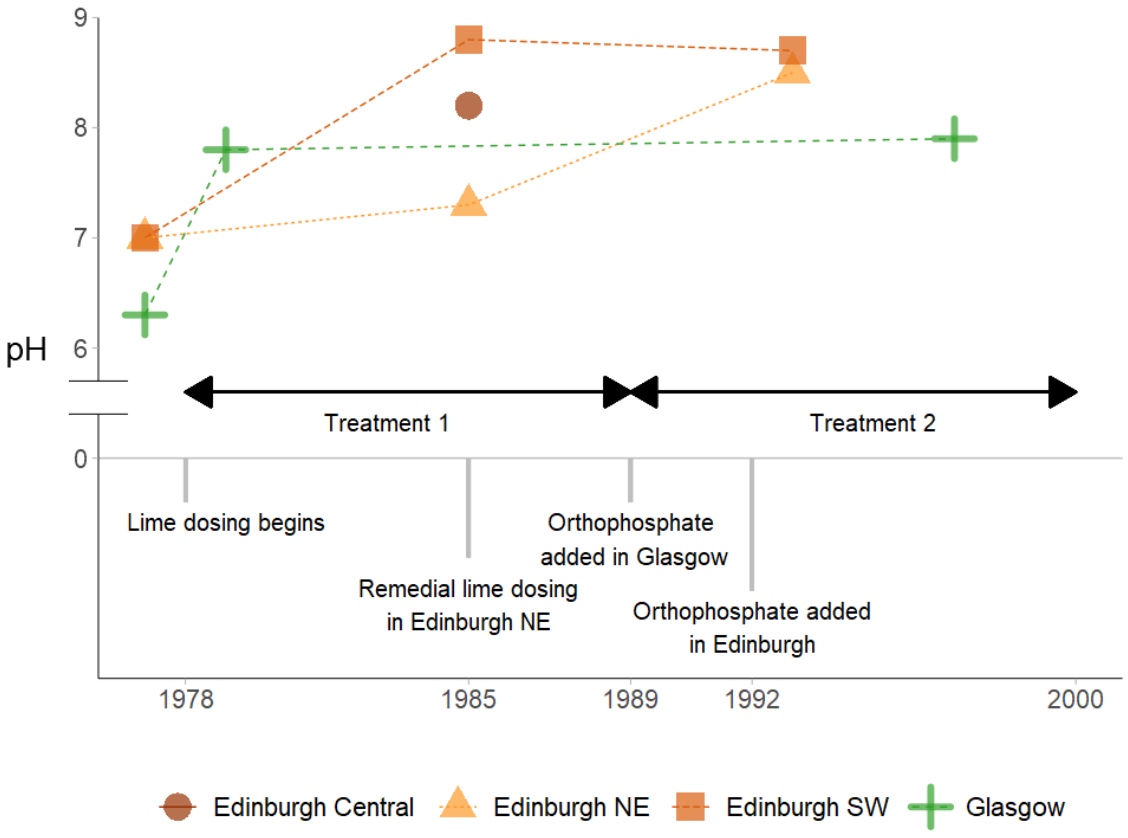
- Treatment 1: the staggered increase in pH levels through lime dosing in Glasgow and Edinburgh. This occurred in 1978 for Glasgow, Edinburgh Southwest and (partially) Edinburgh Central, and in 1985 for Edinburgh Northeast. Importantly,

these interventions were not carried out in the many surrounding areas of these cities until much later.

- Treatment 2: the staggered dosing with orthophosphate in both cities in 1989 in Glasgow and in 1991 in every area of Edinburgh. Also in this case, this treatment was not implemented in the neighbouring areas until years later.

A full timeline of the treatments is given in figure 2. The control group in all cases consists of adjacent areas that did not receive the treatment until later (depicted by the grey areas in figure 6). This timeline, along with different treatment groups, provides the scope for the implementation of staggered difference-in-differences methods, described in Section 4.

Figure 2 - pH Levels in Each Water Supply Area

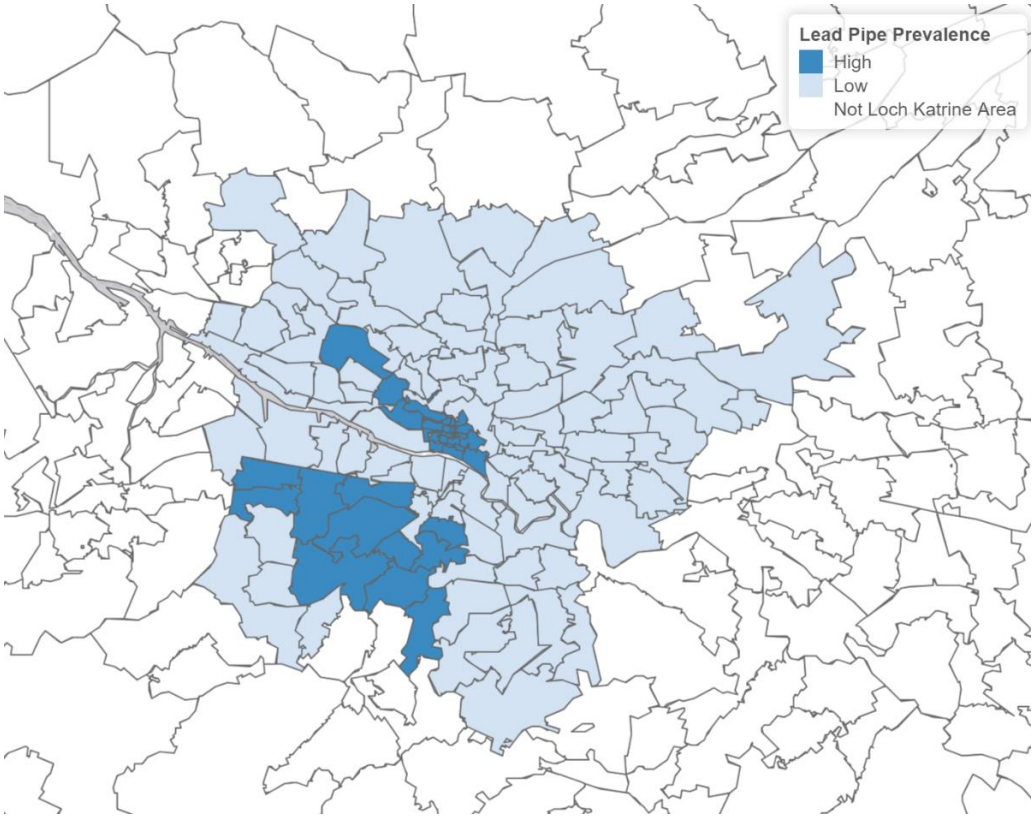


Notes: This chart show pH levels for the various water supply areas in different years. It also shows when the lime dosing (treatment 1) and orthophosphate dosing (treatment 2) began in each area. Water engineers targeted pH levels to be well above 7 to reduce plumbosolvency. Sources: Macintyre et al., (1998), Richards et al., (1980) and Watt et al., 1996.

Treatment 1, lime dosing was undertaken because experiments with the Glasgow water supply in 1973 showed it would raise the pH effectively and thus lower plumbosolvency (Richards et al., 1980). Following this, in both Glasgow and Edinburgh, an investment in an automatic lime-dosing system was considered worthwhile. These began operation in 1978. In Glasgow, which is supplied by Loch Katrine water, the pH was raised from 6.3 to 7.8 after this dosing (figure 2). The Glasgow water supply area at this time is mapped in figure 3. This map also highlights postcodes with a higher prevalence of lead piping in homes. The distinction between areas with high and low prevalence of lead pipes is based on a survey conducted in Glasgow in the early 1990s which findings are summarised in Watt et al. (1996a). We exploit the difference in high and low lead prevalence areas in Glasgow as a robustness check in the tables A.8 and A.9 in the appendix.

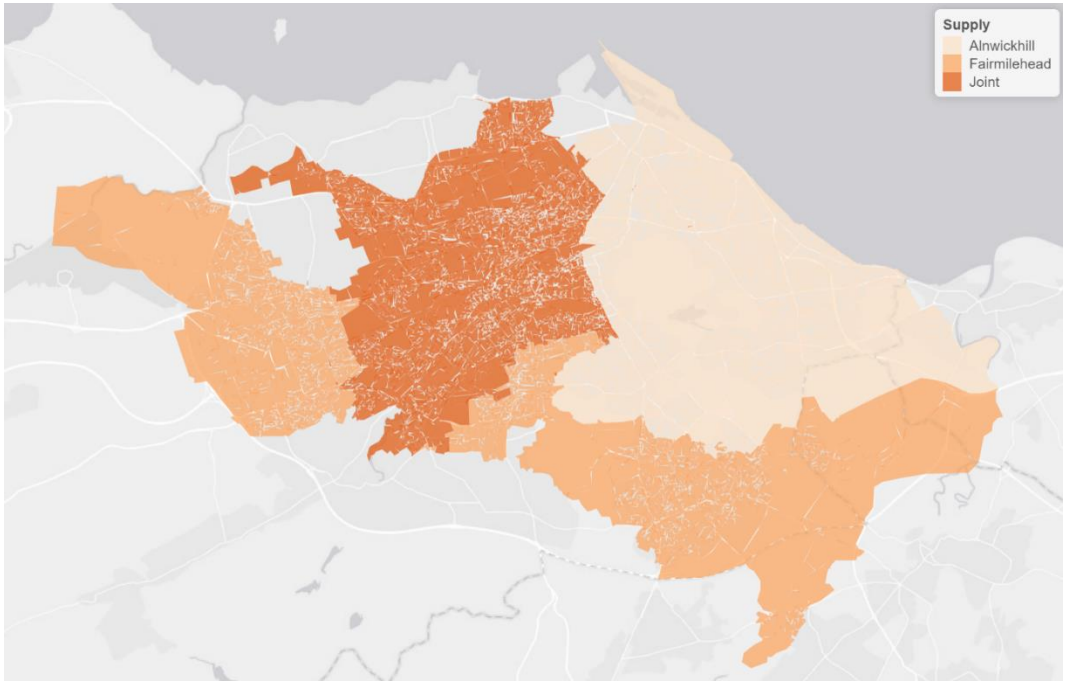
In Edinburgh, the city was supplied with water from two main sources: Alnwickhill, which served the north-eastern part of the city (referred to as Edinburgh NE for simplicity), and Fairmilehead, which catered to the south-western region (Edinburgh SW). The central area of the city (Edinburgh Central) received a combined supply from both these sources. See figure 4 for the map of these water supply areas. The water from both these sources was relatively soft, with a pH level of around 7 before dosing. In 1978, a successful lime dosing at Edinburgh SW raised the pH level to above 8. However, similar efforts at Edinburgh NE faced technical difficulties, delaying effective dosing until 1985. As a result, the pH level in Edinburgh NE remained below 8 until after 1985, when it eventually rose to around 8.5. Edinburgh Central was supplied jointly by Edinburgh NE and Edinburgh SW and therefore received a partial treatment. However, when measured in 1985 it was found that its pH level exceeded 8, aligning more closely with the pH level observed in Edinburgh SW, the area treated successfully in 1978. The pH levels before and after the lime dosing for each water supply area are depicted in figure 2. This data provides a clear comparison of the effectiveness of the treatments in Glasgow and Edinburgh SW in 1978 and in Edinburgh NE in 1985.

Figure 3 - Historical Loch Katrine Water Supply Area with High and Low Lead Piping Prevalence



Note: The areas in dark and light blue represent treated areas that underwent water lead reduction interventions in 1978 and 1989. The heterogeneous effect of these interventions on areas with higher and lower prevalence of lead piping is investigated in the Appendix. Source: Watt et al. (1996a).

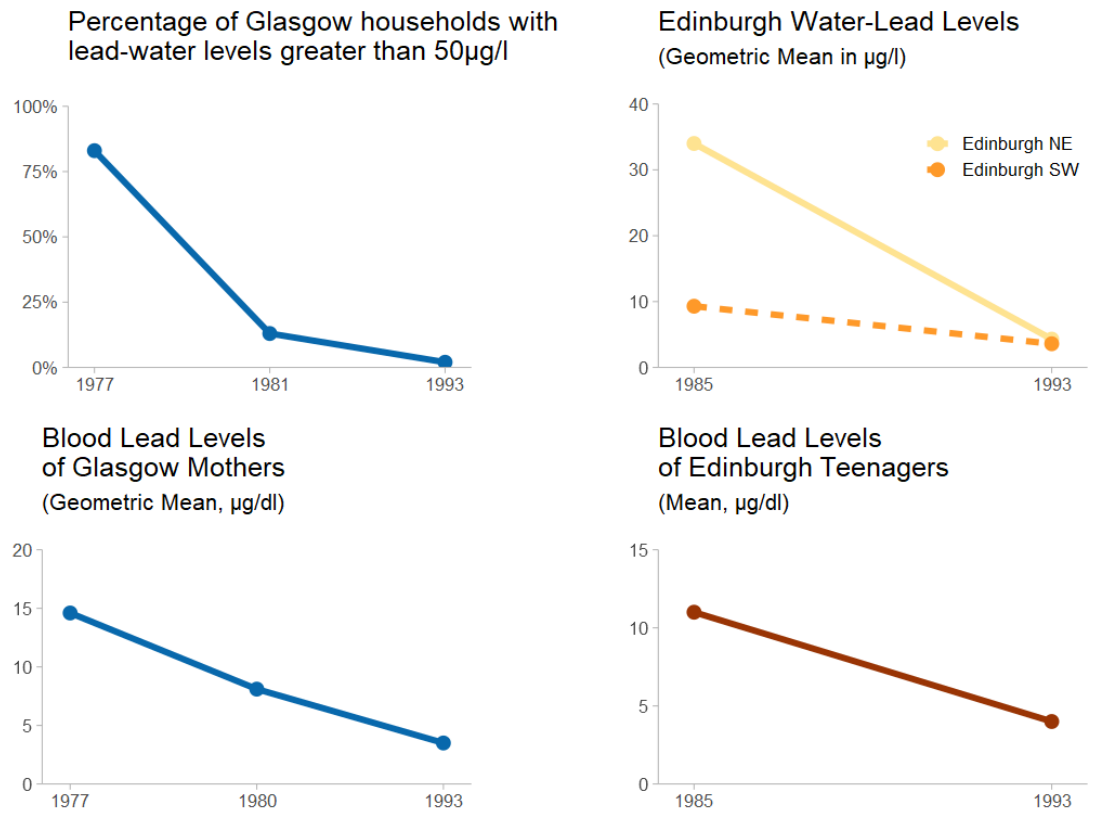
Figure 4 - Historical Water Supply Areas in Edinburgh



Note: The North-East area of Edinburgh (Edinburgh NE) was served by the Alnwickhill water supply. The South-West area of Edinburgh (Edinburgh SW) was served by the Farmilehead water supply. The central area (Edinburgh Central) is jointly served by both water supplies.

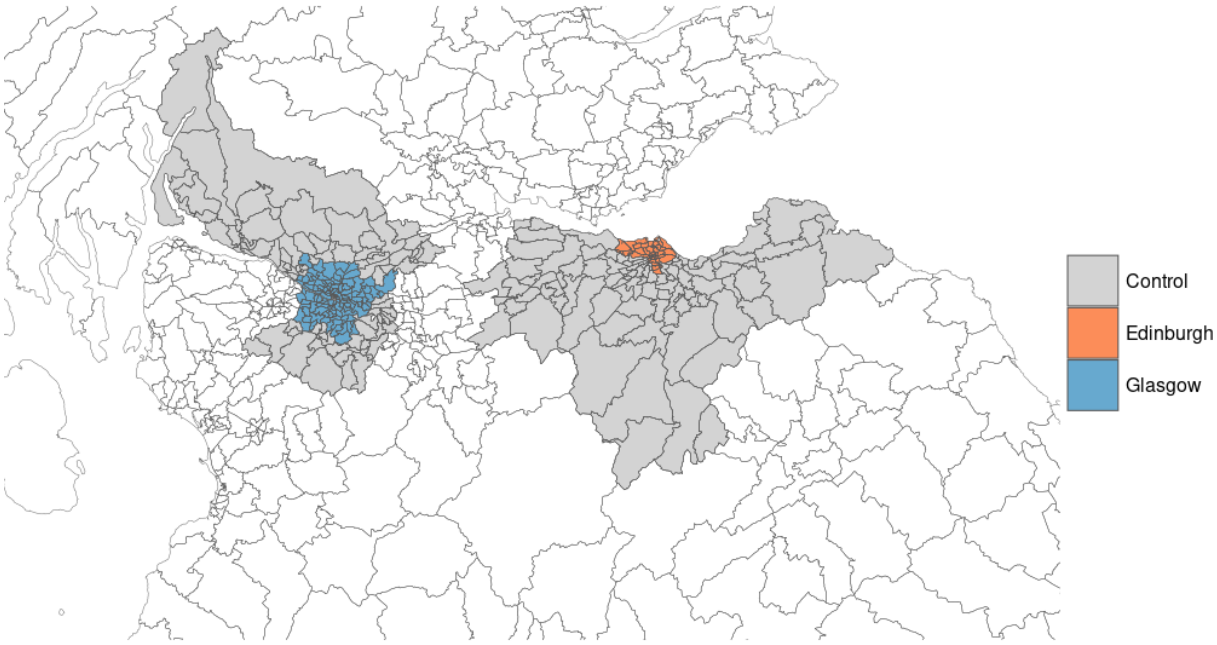
These treatments lowered both water and blood lead levels (figure 5), but not enough. Treatment 2, orthophosphate dosing, was carried out to further decrease lead in the water supply. Orthophosphate dosing reduces the solubility, and therefore bioavailability, of lead in the water supply (Comber et al, 2011). This is a different mechanism for reducing lead pollution compared to lime treatment, which raises the pH. This was done in Glasgow in 1989 and in all of Edinburgh in 1991. This further recued blood lead levels (figure 5). As the UK government brought in stricter lead-water maximum levels, other areas followed. Eventually, in the 2000s, 95% of the UK's water would be treated with orthophosphate (Hayes and Hydes, 2012).

Figure 5 – Water and Blood Lead Levels in Glasgow and Edinburgh



Sources: Watt et al., (1996a) and Moore (1998), Macintyre et al., (1998)

Figure 6 - Map of Treatment and Control Areas.



3 Data

We use health data from Public Health Scotland (PHS). The data covers all the pregnancy outcomes in Glasgow, Edinburgh, and the surrounding postcode areas of each city for the period 1975 to 2000. This data is from the Scottish Morbidity Records (SMR) and the Death, Birth and Stillbirths Registrations (NRS) Furthermore, from the NRS records, we link live births records with death registrations to identify if a child died before age five[5].

These data are matched to the historical Scottish Water Supply area maps for the relevant areas by using maternal postcodes for the relevant period. That is, the address of the house at the time of pregnancy is assigned to the postcode and coded as the relevant treatment/control group. The map of water supply areas, and the various treatment groups are included in figures 3, 4, and 6. Figure 3 shows the Loch Katrine water supply area during the period the data cover. There is a further split in the Loch Katrine supply area between postcode sectors with relatively high levels of lead piping compared to those with relatively low levels of lead piping, as given in Watt et al. (1996a). In high lead areas, 19% reported lead piping in a survey, while in low lead areas it was 9%. Figure 4 shows the water supply areas in Edinburgh during the period the data cover. The Fairmilehead source supplied mostly the west of Edinburgh (which we label Edinburgh SW), while Alnwickhill served Edinburgh NE. The "Joint" area, which we label for simplicity Edinburgh Central, is supplied by both water sources during this period. As explained in the methods section, the first treatment of liming was effective in 1978 in Edinburgh SW, but not effective till 1985 in Edinburgh NE. We therefore consider Edinburgh Central which is served by both sources as being treated at the same time as Edinburgh SW in 1978, but exclude it as a robustness check in the appendix.

Our two main outcomes are birthweight and under-5 mortality. We use only single births. Twins, and other multiple births are excluded as their outcomes tend to be very different, with lower birthweights in comparison to single births, as well as different probabilities of complications. However, multiple births are only around 1.5% of all births. Our mortality indicator includes all deaths and recorded non-viable pregnancies,

[5] We would like to thank the Electronic Data Research and Innovation Service (eDRIS) of Scotland for their help with data handling and access.

stillbirths and spontaneous abortions. We also link the data with Scottish Morbidity Records so that it includes any deaths up to age 5. Under-5 mortality is the commonly used definition of child mortality and is the indicator used in Sustainable Development Goal (SDG) target 3.2: "Newborn and child mortality: By 2030, end preventable deaths of newborns and children under 5 years of age, with all countries aiming to reduce neonatal mortality and under-5 mortality"[6]. Additional data used as controls at the individual level include the biological sex of the baby, and a series of mother's characteristics such as age, height, and previous obstetric history, such as the number of previous spontaneous abortions, and number of previous pregnancies. At the postcode level, the data is linked to Carstairs scores[7], which are material deprivation indices that rank different areas by using information from the 1981 census about car ownership, male unemployment, overcrowding and low social class. If a postcode is in the bottom two deciles, we code that postcode with an indicator variable as being in a deprived area. This link to deprivation allows us to control for higher income households undertaking more pollution avoidance behaviour. Table 1 includes descriptive statistics of the variables used.

[6] See the SDG targets and indicators here: https://sdgs.un.org/.
[7] These are rankings of areas by material deprivation. The variable takes into account material good ownership, such as car ownership, self-reported class, and unemployment amongst other variables to make an index. The Cairstairs scores were originally developed by Carstairs and Morris (1991) and are regularly generated and published by the MRC/CSO Social and Public Health Sciences Unit, the University of Glasgow (https://www.gla.ac.uk/schools/healthwellbeing/research/mrccsosocialandpublichealthsciencesunit/programmes/inequalities/healthinequalities/determinantsofhealthandhealthinequalitiesinscotland/carstairsscores/).

Table 1 - Descriptive Statistics

Variable	Group	Mean	Median	Std Dev	Obs
Birthweight (grams)					
	Control	**3318**	**3360**	**594**	**353,643**
	Edinburgh	**3320**	**3360**	**595**	**76,498**
	Edinburgh NE	3310	3350	603	26,172
	Edinburgh SW	3372	3410	567	8,315
	Joint	3317	3360	595	42,011
	Glasgow	**3257**	**3300**	**591**	**216,771**
Death Indicator Variable					
	Control	**0.01**	**0**	**0.09**	**353,643**
	Edinburgh	**0.01**	**0**	**0.09**	**76,498**
	Edinburgh NE	0.01	0	0.10	26,172
	Edinburgh SW	0.01	0	0.09	8,315
	Joint	0.01	0	0.09	42,011
	Glasgow	**0.01**	**0**	**0.09**	**216,771**
Proportion Living in Deprived Areas (Carstairs Index)					
	Control	**0.11**	**0**	**0.32**	**353,643**
	Edinburgh	**0.13**	**0**	**0.34**	**76,498**
	Edinburgh NE	0.06	0	0.23	26,172
	Edinburgh SW	0.04	0	0.19	8,315
	Joint	0.20	0	0.40	42,011
	Glasgow	**0.59**	**1**	**0.49**	**216,771**
Total Previous Pregnancies					
	Control	**1.18**	**1**	**1.28**	**353,643**
	Edinburgh	**1.12**	**1**	**1.29**	**76,498**
	Edinburgh NE	1.07	1	1.24	26,172
	Edinburgh SW	1.15	1	1.27	8,315
	Joint	1.15	1	1.31	42,011
	Glasgow	**1.24**	**1**	**1.41**	**216,771**
Mother's Age					
	Control	**27.72**	**28**	**5.29**	**353,643**
	Edinburgh	**28.43**	**29**	**5.46**	**76,498**
	Edinburgh NE	28.25	28	5.32	26,172
	Edinburgh SW	29.93	30	5.24	8,315
	Joint	28.25	28	5.54	42,011
	Glasgow	**26.90**	**27**	**5.58**	**216,771**
Number of Previous Spontaneous Abortions					
	Control	**0.22**	**0**	**0.57**	**353,643**
	Edinburgh	**0.23**	**0**	**0.59**	**76,498**
	Edinburgh NE	0.23	0	0.57	26,172
	Edinburgh SW	0.24	0	0.61	8,315
	Joint	0.23	0	0.59	42,011
	Glasgow	**0.23**	**0**	**0.59**	**216,771**
Male Infant Proportion					
	Control	**0.51**		**0.50**	**353,643**
	Edinburgh	**0.51**		**0.50**	**76,498**
	Edinburgh NE	0.51		0.50	26,172
	Edinburgh SW	0.52		0.50	8,315
	Joint	0.51		0.50	42,011
	Glasgow	**0.51**		**0.50**	**216,771**

4 Methods

We use the plausibly exogenous change in water treatments, at different points in time, to identify the effect of lead-water pollution on birth and early life health outcomes. Our main specifications use a staggered difference-in-differences design. We further discuss the estimands, assumptions necessary, and specifications below. The following section is largely based on the expositions in Athey and Imbens (2022), Wooldridge (2021), and Wooldridge (2010).

4.1 Estimands

Our main results focus on three estimands. First, the average effect of water treatment (and therefore lead reduction) at time t on the group which began treatment at time r.

We write this τ_{rt} and define it formally below.

(1) $$\tau_{rt} = E[\, y_{it}(r) - y_{it}(0) \mid d_{ir} = 1\,],\ r = q, \dots, T;\ t = r, \dots, T.$$

Where $y_{it}(r)$ is the outcome for child i at time t given their water supply began treatment at time r, and $r \leq t$, and $y_{it}(0)$ is the unobserved counterfactual outcome for child i at time t where they have not yet received treatment, q is the first period where any cohort is treated, and d_{ir} is a cohort indicator which equals 1 if individual i is in treatment group r. Simply, τ_{rt} is the average treatment effect on the treated (ATT) for that treatment cohort in that year.

Our second estimand is the average treatment for a specific treatment group, for all the years of treatment in our data.

(2) $$\bar{\tau}_r = E[\tau_{rt}]\,, r = q, \dots, T; t = r, \dots, T.$$

Which we estimate as:

(3) $$\hat{\bar{\tau}}_r = \frac{\sum_{t=r}^{T} \hat{\tau}_{rt}}{(T-r+1)}$$

Our third estimand, is the overall average for all groups and all years, which we estimate with:

(4) $$\hat{\tau} = \sum_{t=r}^{T} \hat{\tau}_r \ \times w_r$$

Where w_r is simply a weight that equals the proportion of treated units that are in group r.

4.2 Difference-in-Differences Design Models

Our main results are from models relying on difference-in-differences designs. In the baseline, reduced form model, lead levels, given by the variable $Lead_{it}$, are assumed to affect the birth outcome as shown in (5). The effect of lead is given by T .

$$(5) \qquad y_{it} = \ c_j + g_t + \ \boldsymbol{x_i}\boldsymbol{\beta} + (\mathrm{T} \times Lead_{it}) + \ u_{it}$$

where y_{ijt} is the outcome for individual i, at time t. There is a time-invariant postcode-level effect, c_j, a time trend in outcome, g_t, and a vector of other variables that affect the outcome, $\boldsymbol{x_i}$, which vary by individual. The final term u_{it} is the error term.

This model cannot be estimated for a number of reasons, not least because lead exposure of each individual at each time is unknown. Even if known, other variables may co-vary with lead and the outcome, leading to biased estimates due to endogeneity. We could estimate a two-way fixed effects model using a $post \ \times treatment$ indicator but, given the staggered timing of the intervention between Glasgow, Edinburgh SW, and Edinburgh NE, this could lead to the effect not being identified, due to the "forbidden comparisons" problem (Goodman-Bacon, 2021, and Calloway and Sant'Anna, 2021).

However, given the plausibly exogenous change in lead exposure outlined in section 2, we can identify the effect of the lead reduction if we are willing to accept certain assumptions. Following Wooldridge (2021), the main parallel trends assumption we rely on is a conditional common trends assumption.

Conditional Common Trends, Staggered Treatment (CCTS)

Following Athey and Imbens (2022) and Wooldridge (2021), we define the outcome for the never-treated group as $y_{it}(\infty)$. Given this, we formally state the CCTS assumption as:

$$(6) \qquad E\big[\, y_{it}(\infty) - y_{i1}(\infty) \mid d_{ir}, \boldsymbol{x_{ij}} \big] = \ E\big[\, y_{it}(\infty) - y_{i1}(\infty) \mid \ \boldsymbol{x_{ij}} \big],$$
$$r = q, \dots, T; t = 2, \dots, T.$$

The assumption states that for every cohort the trend in outcome if never treated is unrelated to being in any treatment cohort, after conditioning on the covariates. This can be tested to a degree, see section 5.

Under-5 mortality is a binary outcome and it is likely the CCTS assumption is violated. Therefore we replace this assumption when estimating the effect on under-5 mortality.

Conditional Parallel Relative Trends, Staggered Treatment (CPRTS)

$$(7) \qquad \frac{E\left[y_{it}(\infty)|d_{ir},\ \boldsymbol{x_i}\right]}{E\left[y_{i1}(\infty)|d_{ir},\ \boldsymbol{x_i}\right]} = \frac{E\left[y_{it}(\infty)|\boldsymbol{x_i}\right]}{E\left[y_{i1}(\infty)|\ \boldsymbol{x_i}\right]}, t = 2, \dots, T, r = q, \dots, T$$

The ratio of average outcome if never-treated at time *t* compared to the first period average outcome only depends on the covariates. There is no selection into or out of treatment.

4.3 Estimation

There are now several estimation methods to deal with difference-in-differences when there is staggered adoption (see Roth et al., 2023 for a recent review). Following Harmon (2023), these can be divided into two groups: "subgroup" types and "imputation" types. "Subgroup" types use a particular subset of the data to be a "clean" control group for a particular treatment group. For example, all not yet treated units before the treatment for a particular cohort starts, or just all never-treated units. "Imputation" types instead will impute counterfactuals with either inverse probability weighting or regression adjustment. For our main estimates, we use the Wooldridge (2021) two-way Mundlak estimation method, which is an "imputation" type. We primarily estimate using the Wooldridge (2021) two-way Mundlak because it can simultaneously handle a repeated cross-section and the use of a quasi-maximum likelihood logistic model that is needed due to the CPRTS assumption. However, as a robustness check we also include alternative "subgroup" type estimators in the appendix. Note that these alternatives use the CCTS assumption for mortality rather than the CPRTS assumption used for the two-way Mundlak and both cannot be true at the same time.

For birthweights we use a linear model and the CCTS assumption and estimate with a two-way Mundlak regression with robust errors clustered at postcode sector level following Wooldridge (2021).

$$(8)\quad y_{it} = \eta + , \boldsymbol{x_i}\boldsymbol{\kappa} + \sum_{r=q}^{T} \lambda_r d_{ir} + \sum_{r=q}^{T} \zeta_r (d_{ir} \times , \boldsymbol{x_i}) + \sum_{s=2}^{T} \theta_s f s_t + \sum_{s=2}^{T} (f s_t \times \boldsymbol{x_i})\boldsymbol{\pi_t} + \sum_{r=q}^{T} \sum_{s=r}^{T} \tau_{rt} (d_{ir} \times p_{irt} \times f s_t) + \sum_{r=q}^{T} \sum_{s=r}^{T} (d_{ir} \times p_{irt} \times f s_t \times \dot{\boldsymbol{x}}_{\boldsymbol{ir}})\boldsymbol{\rho_{rt}} + u_{it}$$

Where η is the intercept and $f s_t$ are indicators for every time period that equal 1 when $s = t$. The d_{ir} variable is the cohort indicator as in the CCTS assumption. Next, p_{irt} is a post-treatment indicator. It equals 1 for every period after that group first received treatment. Formally, $p_{irt} = 1 \; \forall \; t > q_r - 1$, where q_r is the period which the group first received treatment. For example, given $t = 1,2,3$ and group 1 was first treated in period 2, then $q_1 = 2$. If group 2 first received treatment in period 3 then $q_2 = 3$. Finally, $\dot{\boldsymbol{x}}_{\boldsymbol{ir}}$ is the deviation from the cohort average for individual *i*.

If we wish to allow for more heterogeneity in time trends, we can model the time trend as simply $d_{ir} \times t$. This allows for heterogeneity in time trends and also allows us to test the common trends assumption with a Wald test jointly on all the coefficients of $d_{ir} \times t$.

For under-5 mortality, we use the CPRTS assumption and estimate with a quasi-maximum likelihood logistic two-way Mundlak regression.

$$(9)\quad E[\, y_{it} \mid d_{ir}, \boldsymbol{x_i}\,] = \boldsymbol{\Lambda}\,[\, \eta + , \boldsymbol{x_i}\boldsymbol{\kappa} + \sum_{r=q}^{T} \lambda_r d_{ir} + \sum_{r=q}^{T} \zeta_r (d_{ir} \times , \boldsymbol{x_i}) + \sum_{s=2}^{T} \theta_s f s_t + \sum_{s=2}^{T} (f s_t \times \boldsymbol{x_i})\boldsymbol{\pi_t} + \sum_{r=q}^{T} \sum_{s=r}^{T} \tau_{rt} (d_{ir} \times p_{irt} \times f s_t) + \sum_{r=q}^{T} \sum_{s=r}^{T} (d_{ir} \times p_{irt} \times f s_t \times \dot{\boldsymbol{x}}_{\boldsymbol{ir}})\boldsymbol{\rho_{rt}}\,]$$

Where $\boldsymbol{\Lambda}$ represents the logistic function. The treatment effect estimated is an average partial effect (APE) of being treated. That is, we estimate the model and then take the coefficients applicable for a particular year and cohort. We take the expected value of the values with the treatment variable minus the values without including the treatment variable. We obtain standard errors for the APE with bootstrapping.

5 Results

We first plot the mean birthweight for Glasgow, Edinburgh and the control group of surrounding areas in figure 7. In all groups there is a clear upward trend. The Edinburgh average tracks closely with the control group, while the Glasgow average remains below

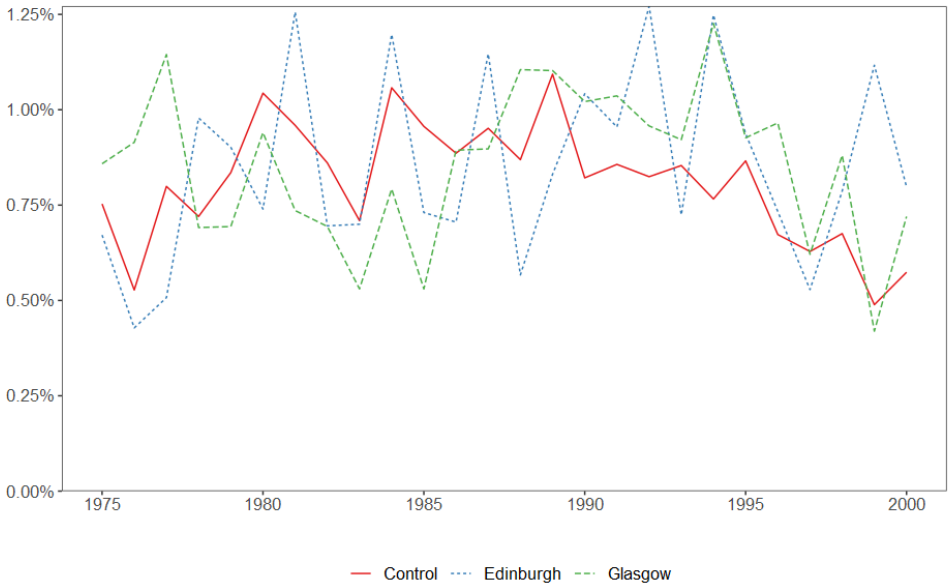
both at all times, and there is no convergence even after the 1978 and 1989 treatment. Both the Edinburgh and Glasgow averages are more volatile than the control group. The trends appear similar for all groups. There is no clear treatment effect to be seen in the raw averages, but this may be due to differences in group characteristics that affect the treatment effect. We explore this in section 5.1.

Similarly, in figure 8 we plot under-5 mortality rates for each group. The trends are again similar, but the percentage for Glasgow and Edinburgh is more volatile. No clear treatment effect is visible in the raw data, but once again this may be due to heterogeneity in group characteristics and therefore selection bias. We move on now to the difference-in-differences estimation.

Figure 7 - Average Birthweight by Treatment Cohort, Grams

Notes: Chart shows the mean birthweight in grams of each birth in Edinburgh, Glasgow and the control group.

Figure 8 - Under-5 Mortality rate by Treatment Cohort



Notes: Chart shows the total Under-5 mortality rates, the probability of all deaths and recorded non-viable pregnancies, including stillbirths and spontaneous abortions, for all births in each birth in Edinburgh, Glasgow and the control group.

5.1 Two-Way Mundlak Regressions

Table 2 presents estimates for the effect of the treatment 1 on both birthweights (columns 1 and 2), and deaths (columns 3 and 4). Standard errors are clustered by postcode sector for the ATTs in columns 1 and 2. For the APEs in Columns 3 and 4 we use a bootstrap for the standard errors. In column 1 we see the overall average effect on birthweights, without covariates, which has a small negative point estimate. This is the opposite sign from what we would expect, but the estimate is not significant at the 5% confidence level and is a precise null. The group level estimates are significant, but Glasgow has a small negative point estimate, while both the Edinburgh water areas have positive estimates. Column 2 includes covariates in the model, as described in section 3. The overall estimate is again negative, the opposite sign from expected, while it is statistically significant. This is mainly driven by the negative estimate from Glasgow. This time the Edinburgh SW estimate is also negative, but not significant, while the Edinburgh NE estimate is again positive. Overall, we find no evidence that treatment 1 increased birthweights.

In column 3 we show the estimates for the effects on under-5 mortality without including covariates. The overall average partial effect is negative, as we would expect,

and significant at the 5% confidence level. However, when we add covariates in column 4 it becomes smaller and not significant. We see a similar pattern for the Glasgow group, as this is the main cause of the sign for the overall estimate. The Edinburgh SW estimate is positive in column 3, implying the lead reduction increased mortality, but becomes negative and not significant in column 4 when we add covariates. The Edinburgh NE estimate is negative in both columns but only significant with covariates. Overall, we find no strong evidence for a decrease in under-5 mortality due to treatment 1.

We test the assumptions of common trends and no anticipation by carrying out similar regressions but also including a cohort indicator interacted with a continuous time variable in the regression, as suggested in Wooldridge (2021). We carry out a Wald test of joint null effects on each cohort and time interaction. We reject the hypothesis of no anticipation or no common trends if the Wald test fails and finds the coefficients to be jointly statistically significant. In the case of the birthweights regression without covariates (column 1) the p-value for the test is 0.01, so we reject the hypotheses of no anticipation or common trends. However, we do not reject the conditional no anticipation and conditional common trends assumptions used in column 2 when we include covariates (p-value 0.17). For columns 3 and 4, the regressions with mortality as the dependent variable, we do not reject the hypotheses of no anticipation or common trends (p-values 0.72 and 0.26 respectively).

In table 3, we only examine years for which all of the treatment areas already had treatment 1, but not yet treatment 2. Glasgow was treated in 1989, and all Edinburgh water areas in 1991. In columns 1 and 2 we see the estimates of the effect on birthweight of treatment 2. All point estimates are small and not significant, save the Edinburgh estimate in column 2 where we include covariates, but this is negative. Overall, we again conclude that the reduction in lead due to the treatment had no effect on birthweights. In column 3, we estimate the effect on mortality without covariates. The overall estimate is positive, but not significant. The Glasgow estimate is negative and not significant, while the Edinburgh effect has the opposite from the expected sign and is significant. In column 4, when we include covariates, the overall APE becomes negative but remains not significant. The Glasgow estimate remains negative but is significant, while the Edinburgh point estimate remains positive but is no longer

significant. Overall, we conclude there is no strong evidence for an effect of treatment 2 on under-5 mortality.

Table 2 - Effect of treatment 1 by Cohort (Two-Way Mundlak Method)

Dependent Variable	Birthweights (1)		Birthweights (2)		Under-5 Mortality (3)		Under-5 Mortality (4)	
	ATT	SE	*ATT*	SE	*APE*	SE	*APE*	SE
Overall	-3.5	(1.9)	-14.5	(2.5)	-0.0017	(0.0008)	-0.0017	(0.0012)
Glasgow	-9.2	(2.2)	-19.2	(3.1)	-0.0030	(0.0010)	-0.0018	(0.0015)
Edinburgh SW	13.5	(4.8)	-4.7	(4.4)	0.0030	(0.0010)	-0.0002	(0.0009)
Edinburgh NE	13.7	(5.7)	10.3	(5.9)	-0.0001	(0.0013)	-0.0040	(0.0014)
Observations	618,108		612,483		646,893		641,004	
Clusters	398		398		398		398	
Covariates	*No*		*Yes*		*No*		*Yes*	

Notes: Table shows cohort specific treatment effects from lime dosing using two-way Mundlak regressions. We use robust standard errors, clustered by postcode sector (columns 1 and 2) or bootstrapped (columns 3 and 4). ATT = Average Treatment on the Treated estimate. APE = Average Partial Effect estimate. Birthweights is the birthweight of the child in grams. Under-5 Mortality is the probability of all deaths and recorded non-viable pregnancies, including stillbirths and spontaneous abortion. Birthweight regression estimates are rounded to 1 decimal place, mortality estimates are rounded to 3 decimal places.

Table 3 - Effect of Treatment 2 by Cohort (Two-Way Mundlak Method)

Dependent Variable	Birthweights (1)		Birthweights (2)		Under-5 Mortality (3)		Under-5 Mortality (4)	
	ATT	SE	*ATT*	SE	*APE*	SE	*APE*	SE
Overall	1.7	(3.4)	-3.3	(3.1)	0.0006	(0.0004)	-0.0006	(0.0006)
Glasgow	3.7	(3.1)	2.7	(3.0)	-0.0004	(0.0005)	-0.0013	(0.0006)
Edinburgh	-1.4	(7.2)	-12.3	(6.2)	0.0020	(0.0007)	0.0005	(0.0011)
Observations - Overall	362,598		287,326		362,687		359,071	
Clusters - Overall	391		391		391		391	
Covariates	*No*		*Yes*		*No*		*Yes*	

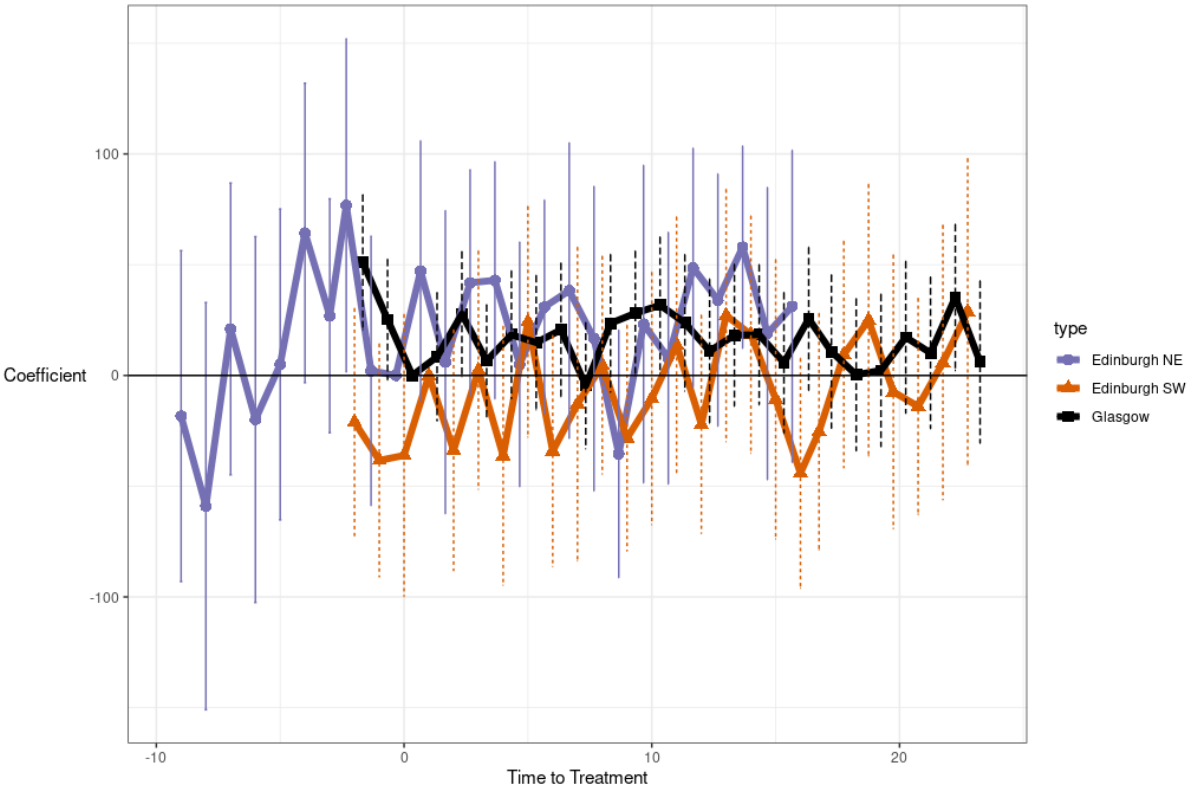
Notes: Table shows cohort specific treatment effects from orthophosphate dosing using two-way Mundlak regressions. We use robust standard errors, clustered by postcode sector. ATT = Average Treatment on the Treated estimate. APE = Average Partial Effect estimate. Birthweights is the birthweight of the child in grams. Under-5 Mortality is the probability of all deaths and recorded non-viable pregnancies, including stillbirths and spontaneous abortion. Birthweight regression estimates are rounded to 1 decimal place, mortality estimates are rounded to 3 decimal places.

In figure 9, we show event study estimates of treatment 1 using two-way Mundlak regressions for the effect on birthweight. The cohort-year interactions are shown as time-to-treatment, so Edinburgh NE, which was treated later, has more pre-treatment periods than Glasgow and Edinburgh SW. There does not appear to be any clear pre-trends before treatment for any cohort, but there is also no clear treatment effect after treatment for any cohort. Edinburgh SW estimates are mostly negative while Glasgow and Edinburgh NE mostly positive, but all have wide 95% confidence intervals covering zero. In figure 10 we show the same event study estimates of the effect on birthweight for treatment 2. Both Edinburgh and Glasgow Estimates are typically negative, but again very wide intervals covering zero. We see no strong evidence for either pre-trends or a treatment effect in these graphs.

In figure 11 we show event study APE's for each cohort year for the effect on mortality of treatment 1. Again no clear evidence of pre-trends. Glasgow has mostly negative point estimates after treatment while Edinburgh SW and Edinburgh NE mostly positive. The two Edinburgh water treatment areas have almost all intervals covering zero while

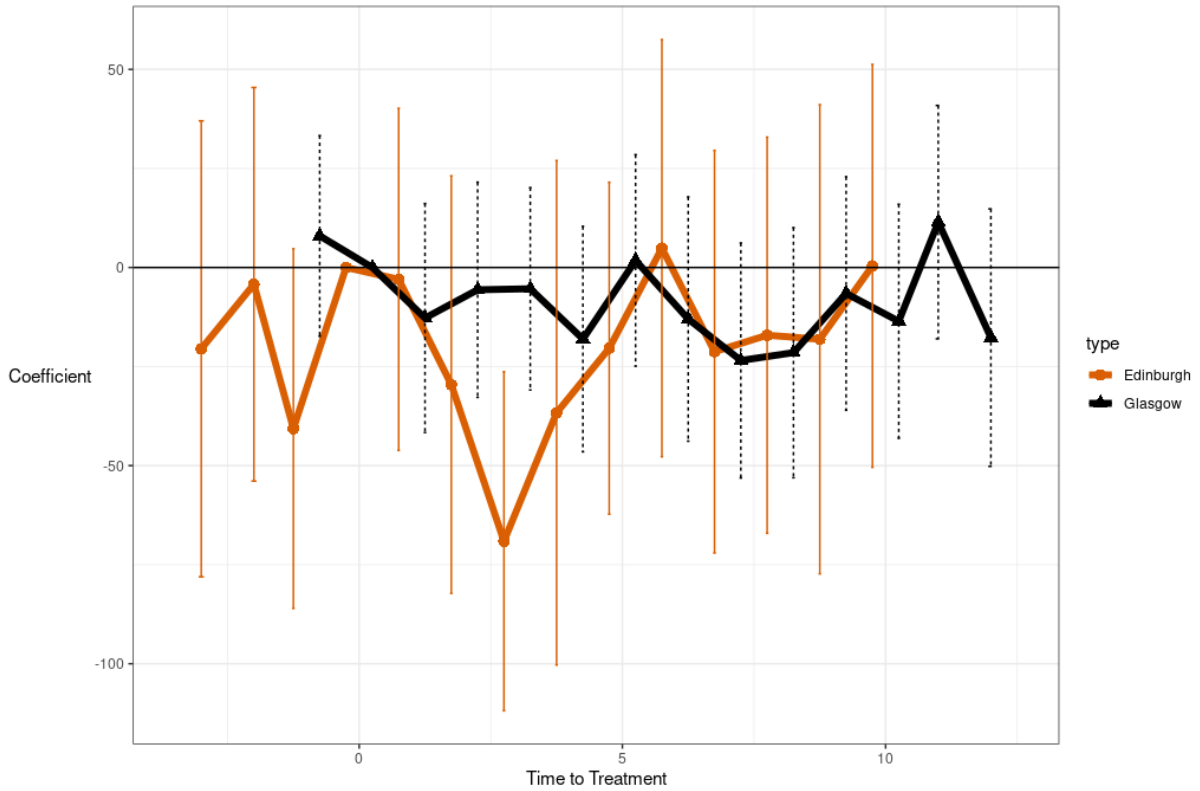
Glasgow does have some intervals that are statistically significant. When we look at the effect on mortality of treatment 2 in figure 12, we see perhaps some evidence for a negative pre-trend in Edinburgh. Point estimates in both Glasgow and Edinburgh tend to be negative but all except one post-treatment interval cover zero. Again, we conclude there is no strong evidence for an effect on under-5 mortality in these event studies.

Figure 9 - Effect of Treatment 1 on Birthweights (Event Study with Two-way Mundlak Estimator)



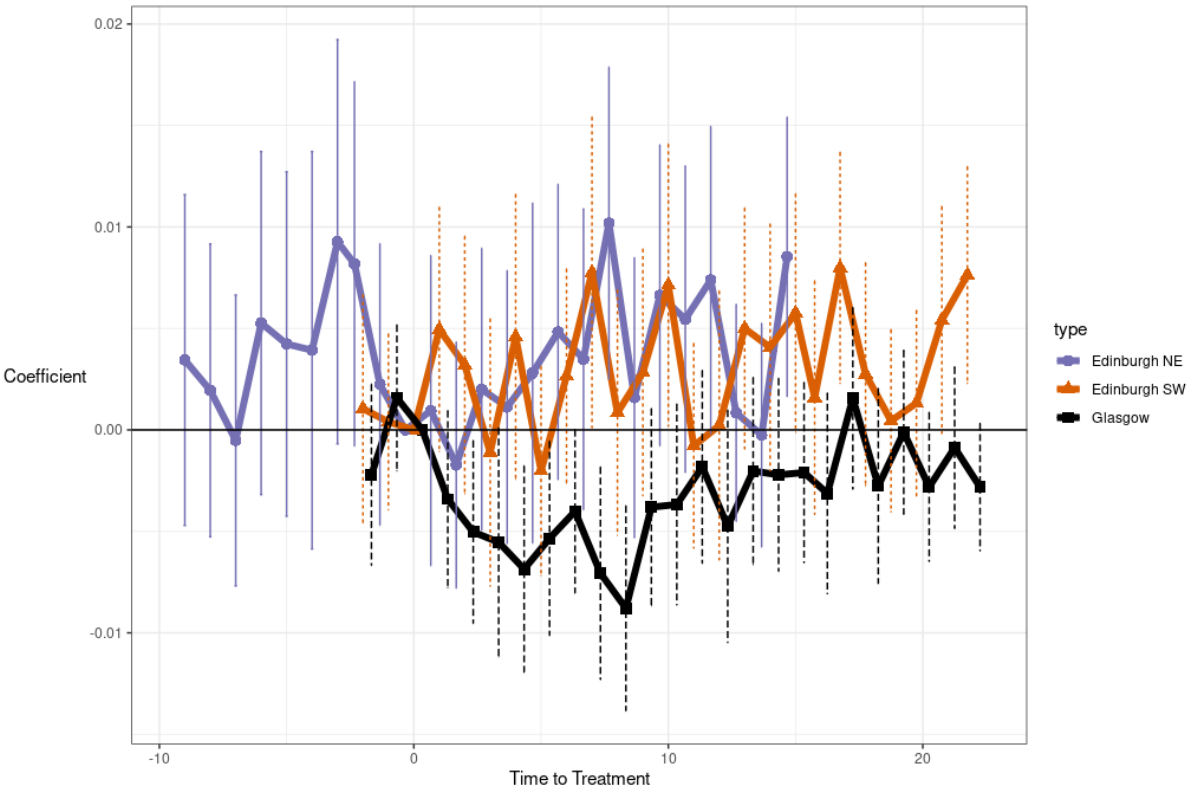
Notes: Table shows estimated treatment effects on different groups of lead reduction due to lime-dosing in Glasgow and Edinburgh compared to a never-treated control group. We use standard errors clustered by postcode sector. In all cases the estimates are the average treatment on the treated for that year on that group Birthweights is the birthweight in grams for a child.

Figure 10 - Effect of Treatment 2 on Birthweights (Event Study with Two-way Mundlak Estimator)



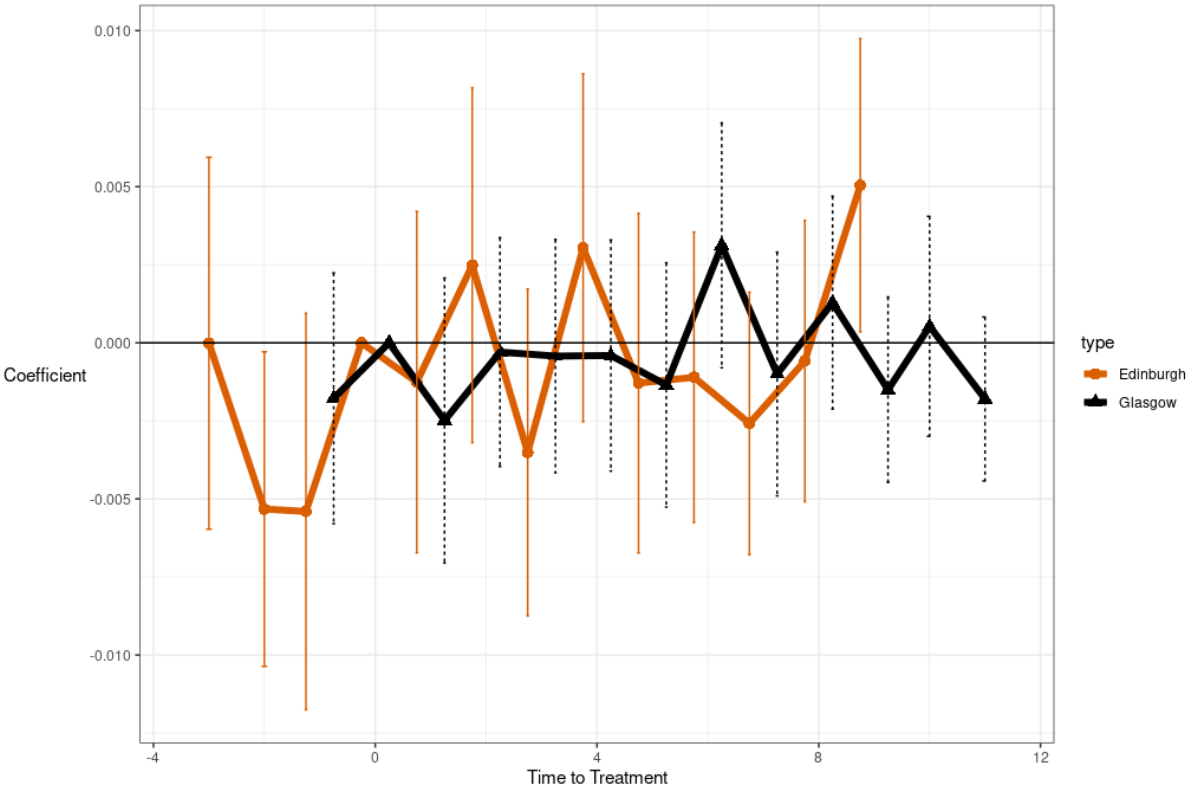
Notes: Table shows estimated treatment effects on different groups of lead reduction due to orthophosphate in Glasgow and Edinburgh compared to a never-treated control group. We use standard errors clustered by postcode sector. In all cases the estimates are the average treatment on the treated for that year on that group Birthweights is the birthweight in grams for a child.

Figure 11 - Effect of Treatment 1 on Under-5 Mortality (Event Study with Two-way Mundlak Estimator)



Notes: Table shows estimated treatment effects on different groups of lead reduction due to lime-dosing in Glasgow and Edinburgh compared to a never-treated control group. We use bootstrapped standard errors. In all cases the estimates are the average partial effect for that year on that group. Mortality is the probability of all deaths and recorded non-viable pregnancies, including stillbirths and spontaneous abortion for under-5s.

Figure 12 - Effect of Treatment 2 on Under-5 Mortality (Event Study with Two-way Mundlak Estimator)



Notes: Table shows estimated treatment effects on different groups of lead reduction due to orthophosphate in Glasgow and Edinburgh compared to a never-treated control group. We use bootstrapped standard errors. In all cases the estimates are the average partial effect for that year on that group. Mortality is the probability of all deaths and recorded non-viable pregnancies, including stillbirths and spontaneous abortion for under 5s.

5.2 Regression Discontinuity Design

As an alternative identification strategy, we use a sharp regression discontinuity design (RDD). Here we separate the cohorts and regress each individual's outcome on the date of birth with the cut-off being the treatment date. The running variable is the date. We use local linear regressions with a triangular kernel. We also use the optimal non-parametric bandwidth selection method with the robust bias corrected intervals of Calonico et al. (2020). This means that, within each cohort, we use a direct before and after treatment comparison. We no longer need to rely on the assumptions in section 4. Instead, we assume the expected value of the outcome is continuous in the neighbourhood of the treatment cut-off for both treated and untreated units. That is, mothers cannot perfectly manipulate birth dates so as to be one side of the treatment cut-off. This would be violated if mothers knew about the upcoming water treatment

and decided to delay birth until after treatment. Given there is always some randomness in birth dates (as many mothers will attest), we believe this is a reasonable assumption. See Cattaneo and Titiunik (2022) for a recent review of regression discontinuity design and its assumptions.

Given this assumption holding, we estimate the effect of treatment for individual pregnancies near the treatment cut-off. We do not use this as our main estimation strategy for two reasons, the actual difference in lead exposure near the cut-off may be small. Therefore, the estimate may be too noisy to find an effect. Secondly if the assumptions in section 4 hold, the RDD is less efficient because we are discarding so much of the variation.

Table 5 - Local Average Treatment Effect on Birthweights, Regression Discontinuity Design Results

Group	**Coefficient**	**Std Error**	**Observations**	**Bandwidth (days)**
Glasgow, 1	-3.5	(14.5)	216,556	1701
Edinburgh SW, 1	-10.2	(33.2)	50,291	1488
Edinburgh NE, 1	-108.0	(46.1)	26,151	1310
Glasgow, 2	8.8	(14.1)	216,556	1416
Edinburgh, 2	7.2	(28.4)	76,442	1461

Notes: This table reports the local average treatment effect from separate sharp regression discontinuity designs on birthweights in grams. Robust, bias corrected standard errors are reported in brackets. Top 3 rows show the effect of treatment 1, lime dosing. Bottom 2 show effect of treatment 2, orthophosphate dosing.

Table 6 -Local Average Treatment Effect on Under-5 Mortality, Regression Discontinuity Design

Group	Coefficient	Std Error	Observations	Bandwidth (days)
Glasgow, 1	-0.0002	(0.0025)	216,771	1230
Edinburgh SW, 1	0.0027	(0.0058)	50,326	1303
Edinburgh NE, 1	-0.0030	(0.0049)	26,172	1781
Glasgow, 2	-0.0019	(0.0024)	216,771	1722
Edinburgh, 2	-0.0019	(0.0039)	76,498	1757

Notes: This table reports the local average treatment effect from separate sharp regression discontinuity designs on deaths. Robust, bias corrected standard errors are reported in brackets. Top 3 rows show the effect of treatment 1, lime dosing. Bottom 2 show effect of treatment 2, orthophosphate dosing. Mortality is the probability of all for all deaths and recorded non-viable pregnancies, including stillbirths and spontaneous abortion.

Table 5 and table 6 shows the results for the various RDD estimations. We use each cohort and we show both treatments 1 and 2.W e also examine both outcomes: birthweights and deaths. For Glasgow, we see that none of the results are significant at the 5% level. For birthweights, both treatment 1 and 2 estimates are small in magnitude, and the first is positive, while the latter is negative. For mortality, both treatment estimates imply lowered deaths due to treatment, but neither is statistically significant at the 5% level.

For Edinburgh SW and treatment 1, both results are the opposite sign from expected, implying treatment lowered birthweights and raised deaths, but neither is statistically significant. In Edinburgh NE for treatment 2, the point estimates imply it lowered birthweights and lowered deaths. The birthweights estimate is large and statistically significant at the 5% level. The mortality estimate is not significant. When we look at Edinburgh as a whole for treatment 2, both estimates are of the expected sign. The point estimate implies birthweights increased and deaths decreased, but neither is statistically significant.

Overall, the RDD results are consistent with our main results. There is no clear effect on birthweights or deaths in either Glasgow or Edinburgh.

5.3 Further Robustness Checks

In the online appendix, we carry out a wide variety of further robustness checks. These confirm our main results. These include alternative staggered difference-in-differences estimators, comparing high and low lead areas in Glasgow, and alternative two-way Mundlak specifications. The results are qualitatively similar in all cases.

6 Discussion and conclusion

We examined the effect of reducing lead pollution in drinking water on health outcomes, (birthweights and under-5 mortality) in Glasgow and Edinburgh in the 20th century. We use a setting with plausibly exogenous staggered treatment and therefore utilise a difference-in-differences design that accounts for the staggered nature of the treatment, as well as a regression discontinuity design comparing births just before and just after treatment. Across a variety of specifications, and with multiple robustness checks, we find no clear evidence of a beneficial effect of lead water pollution reduction on infant health outcomes. These findings contribute to the existing literature on the impact of lead and birth outcomes, using a much larger sample, and plausibly exogenous variation in lead, and high levels of lead.

It is worth thinking about these results in light of the plausible mechanisms laid out in section 2. It may simply be that there is no effect on these outcomes for the level of lead in our setting. However, there are also alternative explanations. The first is that an effect does exist, but even with our sample size, given the low number of child deaths annually, – often less than 1 percent – it may not be sufficient to detect it with precision. Related to this is the possibility that there is an effect, but not at the levels of lead in our sample. It may be the level of exposure in our sample was simply too low to have a detectable average effect. Lead pollution at high levels causes very obvious and extreme health problems, but at lower levels it is much harder to see acute lead poisoning symptoms. However, the water and blood lead levels in our sample, especially for Glasgow, are much larger than in other studies which do find an effect.

A second possibility is that the literature on lead and child mortality may be affected by publication bias. The lead-crime literature does suffer from publication bias (Higney et

al., 2022). It can be the case that even if every paper estimated an unbiased causal effect, in the presence of publication bias we would still be left with a bias in the literature. It is beyond the scope of this paper to estimate if there is publication bias for the lead pollution and birth outcomes literature, but it is a problem across empirical science and cannot be ruled out here.

Thirdly, it is possible that lead exposure would typically have an effect on child mortality, but a mediating factor can reduce its impact. In the recent review of Clay et al. (2024) shows most quasi-experimental studies find an effect. However, all the quasi-experimental studies which find an effect are from North America, while our paper and Grönqvist et al. (2020) do not find an effect and are from Northern Europe. One potential difference is nutrition, specifically high milk intake. The UK and Northern Europe has some of the highest milk consumption rates in the world (FAO, 2022), more than two and a half times the global average (FAO, 2022). In the 1930s, the National Milk Scheme in Scotland promoted milk consumption and provided targeted subsidies for mothers of children under five years old. This, along with the provision of free milk in schools for much of the 20th century (Krebs, 2019), led to a significant increase in milk consumption. Studies have shown that high milk intake is associated with lower blood lead levels (Chuang et al., 2004). Thus, it is possible that better nutrition, particularly high milk consumption, may have played a role in reducing the impact of lead on under-5 mortality in Northern Europe compared to elsewhere.

Our study has several limitations. Firstly, the treatment variation is at a postcode sector level rather than an individual level and we observe most of our sample only once. This means that our confidence intervals are less precise than if the treatment was exogenously applied at an individual level. However, this is common in the literature that examines the effect of lead with quasi-experiments, because lead cannot be ethically given as part of a randomised control trial. We therefore must rely on coarser treatment variation. Secondly, our estimates of the average treatment on the treated imply that the effects we observe may only apply to larger urban areas. Perhaps there is something systematically different about rural areas that would lead to a stronger effect. For example, Parker and Wilby (2013) find that domestic water use per capita is much higher in rural areas compared to urban areas.

Failure to find a robust effect on infant outcomes in this data set does not mean lead remediation is pointless. Lead has been shown to affect a large variety of outcomes. Our findings do, however, have implications for the policy importance attached to lead remediation compared to other actions directed at infant health, such as improved nutrition or better neonatal healthcare. If such "low-hanging fruit" policies are still to be implemented, and resources are constrained, then it may be better to focus on those before lead remediation is undertaken.

References

Akoumianaki I, 2017. Lead in drinking water: public health, mitigation, and economic perspectives. CD2016_03. Available online at: crew.ac.uk/publications.

Al-Saleh, I., Khalil, M.A., Taylor, A., 1995. Lead, erythrocyte protoporphyrin, and hematological parameters in normal maternal and umbilical cord blood from subjects of the Riyadh region, Saudi Arabia. Arch. Environ. Health 50, 66–73. https://doi.org/10.1080/00039896.1995.9955014

Angell, N.F., Lavery, J.P., 1982. The relationship of blood lead levels to obstetric outcome. Am. J. Obstet. Gynecol. 142, 40–46. https://doi.org/10.1016/S0002-9378(16)32282-7

Athey, S., Imbens, G.W., 2022. Design-based analysis in Difference-In-Differences settings with staggered adoption. J. Econom. 226, 62–79. https://doi.org/10.1016/J.JECONOM.2020.10.012

Azayo, M.M., Manji, K., Kalokola, F., 2009. Lead Levels in Women at Delivery at the Muhimbili National Hospital: A Public Health Problem. J. Trop. Pediatr. 55, 138–139. https://doi.org/10.1093/TROPEJ/FMN085

Behrman, J.R., Rosenzweig, M.R., 2004. Returns to Birthweight. Rev. Econ. Stat. 86, 586–601. https://doi.org/10.1162/003465304323031139

Bellinger, D.C., 2005. Teratogen update: Lead and pregnancy. Birth Defects Res. Part A Clin. Mol. Teratol. 73, 409–420. https://doi.org/10.1002/BDRA.20127

Bornschein, R.L., Grote, J., Mitchell, T., Succop, P.A., Dietrich, K.N., Krafft, K.M., Hammond, P.B., 1989. Effects of Prenatal Lead Exposure on Infant Size at Birth. Lead Expo. Child Dev. 307–319. https://doi.org/10.1007/978-94-009-0847-5_18

Brubaker, C.J., Schmithorst, V.J., Haynes, E.N., Dietrich, K.N., Egelhoff, J.C., Lindquist, D.M., Lanphear, B.P., Cecil, K.M., 2009. Altered myelination and axonal integrity in adults with childhood lead exposure: A diffusion tensor imaging study. Neurotoxicology. https://doi.org/10.1016/j.neuro.2009.07.007

Callaway, B., Goodman-Bacon, A., Sant'Anna, P.H.C., 2021. Difference-in-Differences with a Continuous Treatment. arXiv. https://arxiv.org/abs/2107.02637

Callaway, B., Sant'Anna, P.H.C., 2021. Difference-in-Differences with multiple time periods. J. Econom. 225, 200–230. https://doi.org/10.1016/J.JECONOM.2020.12.001

Calonico S, Cattaneo MD, Farrell MH. 2020. Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. The Econometrics Journal. 23(2):192–210. doi:10.1093/ectj/utz022.

Carstairs V, Morris R, 1991. Deprivation and health in Scotland. Aberdeen: Aberdeen University Press.

Cattaneo, M. D. and R. Titiunik, 2022. Regression Discontinuity Designs. Annual Review of Economics 14(1): 821-851.

Cecil, K.M., Brubaker, C.J., Adler, C.M., Dietrich, K.N., Altaye, M., Egelhoff, J.C., Wessel, S., Elangovan, I., Hornung, R., Jarvis, K., Lanphear, B.P., 2008. Decreased brain volume in adults with childhood lead exposure. PLoS Med. https://doi.org/10.1371/journal.pmed.0050112

Chatterji, P., Kim, D., Lahiri, K., 2014. Birth weight and academic achievement in childhood. Health

Econ. 23, 1013–1035. https://doi.org/10.1002/HEC.3074

Chuang, H.Y., Tsai, S.Y., Chao, K.Y., Lian, C.Y., Yang, C.Y., Ho, C.K., Wu, T.N., 2004. The influence of milk intake on the lead toxicity to the sensory nervous system in lead workers. Neurotoxicology 25, 941–949. https://doi.org/10.1016/J.NEURO.2004.04.003

Clay K, Hollingsworth A, Severnini E. 2024. The Impact of Lead Exposure on Fertility, Infant Mortality, and Infant Birth Outcomes. Review of Environmental Economics and Policy. 18(2):301–320. doi:10.1086/730854.

Clay, K., Troesken, W., Haines, M., 2014. Lead and Mortality. Rev. Econ. Stat. 96, 458–470. https://doi.org/10.1162/REST_A_00396

Comber, S., Cassé, F., Brown, B., Martin, J., Hillis, P., Gardner, M., 2011. Phosphate treatment to reduce plumbosolvency of drinking water also reduces discharges of copper into environmental surface waters. Water Environ. J. 25, 266–270. https://doi.org/10.1111/J.1747-6593.2010.00219.X

Dave, D.M., Yang, M., 2022. Lead in drinking water and birth outcomes: A tale of two water treatment plants. J. Health Econ. 84, 102644. https://doi.org/10.1016/J.JHEALECO.2022.102644

Dorea, J.G., Donangelo, C.M., 2006. Early (in uterus and infant) exposure to mercury and lead. Clin. Nutr. 25, 369–376. https://doi.org/10.1016/J.CLNU.2005.10.007

Doucouliagos, H., Stanley, T.D., 2009. Publication Selection Bias in Minimum-Wage Research? A Meta-Regression Analysis. Br. J. Ind. Relations 47, 406–428. https://doi.org/10.1111/J.1467-8543.2009.00723.X

Edwards, M., 2014. Fetal death and reduced birth rates associated with exposure to lead-contaminated drinking water. Environ. Sci. Technol. 48, 739–746. https://doi.org/10.1021/ES4034952/SUPPL_FILE/ES4034952_SI_001.PDF

Ericson B, Hu H, Nash E, Ferraro G, Sinitsky J, Taylor MP, 2021. Blood lead levels in low-income and middle-income countries: a systematic review. The Lancet Planetary Health, Volume 5, Issue 3, e145 - e153. doi: 10.1016/S2542-5196(21)00281-3. PMID: 33713615.

Falcón, M., Viñas, P., Luna, A., 2003. Placental lead and outcome of pregnancy. Toxicology 185, 59–66. https://doi.org/10.1016/S0300-483X(02)00589-9

FAO [WWW Document], 2022 https://www.fao.org/faostat/en/#data/FBS (accessed 12.31.22).

Feigenbaum, J.J., Muller, C., 2016. Lead exposure and violent crime in the early twentieth century. Explor. Econ. Hist. 62, 51–86. https://doi.org/10.1016/j.eeh.2016.03.002

Gazze, L., Heissel, L., 2021. Infrastructure upgrades and lead exposure: Do cities face trade-offs when replacing water mains? Journal of Environmental Economics and Management, Volume 108,102472,ISSN 0095-0696, https://doi.org/10.1016/j.jeem.2021.102472.

Gazze, L., Persico, C., Spirovska, S., 2021. The Long-Run Spillover Effects of Pollution: How Exposure to Lead Affects Everyone in the Classroom. SSRN Electron. J. https://doi.org/10.2139/SSRN.3842757

GBD, 2020. Global Burden of Disease (GBD 2019) | Institute for Health Metrics and Evaluation [WWW Document]. Glob. Burd. Dis. - IMHE. URL

https://www.healthdata.org/gbd/2019 (accessed 1.26.22).

Gechter, M., Meager, 2022. Combining Experimental and Observational Studies in Meta-Analysis: A Debiasing Approach. *Working paper*.

Goldenberg RL, Culhane JF. 2007. Low birth weight in the United States. The American Journal of Clinical Nutrition. 85(2):584S-590S. doi:10.1093/ajcn/85.2.584S.

Golmohammadi, Tahere, Mohammad Reza Ansari, Abdorrahim Nikzamir, Reza S. Abhari and Shahriar Bin Elahi, 2007. The effect of maternal and fetal lead concentration on birth weight: polluted versus non-polluted areas of Iran. Medicine 74-78.

Goodman-Bacon, A., 2021. Difference-in-differences with variation in treatment timing. J. Econom. 225, 254–277. https://doi.org/10.1016/J.JECONOM.2021.03.014

Grönqvist H, Nilsson JP, Robling P-O. 2020. Understanding How Low Levels of Early Lead Exposure Affect Children’s Life Trajectories. Journal of Political Economy. 128(9):3376-3433. doi:10.1086/708725.

Grossman, D.S., Slusky, D.J.G., 2019. The Impact of the Flint Water Crisis on Fertility. Demography 56, 2005–2031. https://doi.org/10.1007/S13524-019-00831-0

Hall, A., Ransom, W.B., 1906. PLUMBISM FROM THE INGESTION OF DIACHYLON AS AN ABORTIFACIENT. Br. Med. J. 1, 428. https://doi.org/10.1136/BMJ.1.2356.428

Hansard HC Deb 23 May 1990 Lead in Water (Scotland) [Hansard], vol 173 cc389-96. https://api.parliament.uk/historic-hansard/commons/1990/may/23/lead-in-water-scotland

Hayes, C.R., Hydes, O.D., 2012. UK experience in the monitoring and control of lead in drinking water. J. Water Health 10, 337–348. https://doi.org/10.2166/WH.2012.210

He, Q., Ji, X., 2021. The Labor Productivity Consequences of Exposure to Particulate Matters: Evidence from a Chinese National Panel Survey. Int. J. Environ. Res. Public Health 18, 12859. https://doi.org/10.3390/IJERPH182312859

Hollingsworth, A., Rudik, I., 2021. The Effect of Leaded Gasoline on Elderly Mortality: Evidence from Regulatory Exemptions. Am. Econ. J. Econ. Policy 13, 345–73. https://doi.org/10.1257/POL.20190654

Krebs, C., 2019. Lead Poisoning and Intelligence: A Search for Cause and Effect in the Scottish Mental Surveys. J. Environ. Public Health 2019. https://doi.org/10.1155/2019/8980604

Law, C.M., 2002. Significance of birth weight for the future. Arch. Dis. Child. - Fetal Neonatal Ed. 86, F7–F8. https://doi.org/10.1136/FN.86.1.F7

Macintyre, C., Fulton, M., Hepburn, W., Yang, S., Raab, G., Davis, S., Heap, M., Halls, D., Fell, G., 1998. Changes in blood lead and water lead in Edinburgh. An eight year follow-up to the Edinburgh lead study. Environ. Geochem. Health 20, 157–167. https://doi.org/10.1023/A:1006593225113/METRICS

Marcus, D.K., Fulton, J.J., Clarke, E.J., 2010. Lead and conduct problems: A meta-analysis. J. Clin. Child Adolesc. Psychol. https://doi.org/10.1080/15374411003591455

McMichael, A.J., Vimpani, G. V., Robertson, E.F., Baghurst, P.A., Clark, P.D., 1986. The Port Pirie

cohort study: maternal blood lead and pregnancy outcome. J. Epidemiol. Community Heal. 40, 18–25. https://doi.org/10.1136/JECH.40.1.18

Moore, M.R., Goldberg, A., Fyfe, W.M., Richards, W.N., 1981. Maternal Lead Levels After Alterations To Water Supply. Lancet 318, 203–204. https://doi.org/10.1016/S0140-6736(81)90384-6

Moore, M.R., Robertson, S.J., Gilmour, W.H., Murray, G.D., Britton, A., Low, R.A., Watt, G.C.M., 1998. Decline of maternal blood lead concentrations in Glasgow. J. Epidemiol. Community Heal. 52, 672–673. https://doi.org/10.1136/JECH.52.10.672

National Records of Scotland [WWW Document], 2022. https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/vital-events/general-publications/vital-events-reference-tables/2020/list-of-data-tables#section4 (accessed 12.31.22).

Parker JM, Wilby RL. 2013. Quantifying Household Water Demand: A Review of Theory and Practice in the UK. Water Resour Manage. 27(4):981–1011. doi:10.1007/s11269-012-0190-2.

Potter, S., 1997. *Lead in Drinking Water*. Science and Environment Section House of Commons Library, 1997.

Quinn, M.J., 1985. Factors Affecting Blood Lead Concentrations in the UK: Results of the EEC Blood Lead Surveys, 1979–1981. Int. J. Epidemiol. 14, 420–431. https://doi.org/10.1093/IJE/14.3.420

Ransom, W.B., 1900. Diachylon as an Abortifacient. Br. Med. J. 2, 201.

Taylor, C.M., Golding, J., Emond, A.M., 2015. Adverse effects of maternal lead levels on birth outcomes in the ALSPAC study: a prospective birth cohort study. BJOG An Int. J. Obstet. Gynaecol. 122, 322–328. https://doi.org/10.1111/1471-0528.12756

Richards WN, Britton A, Cochrane A., 1980. Reducing plumbosolvency-the effect of added lime on the Loch Katrine supply to Glasgow. Journal of the Institute of Water Engineers and Scientists 1979;34:315-333

Richards, W. N., & Moore, M. R., 1984. Lead Hazard Controlled in Scottish Water Systems. Journal (American Water Works Association), 76(8), 60–67. http://www.jstor.org/stable/41272146

Robertson, G., Wilson, A., Papastathopoulos, Y., Graham, M., Eades, L., Chatzisymeon, E., 2020. Estimating the number of homes in Scotland with internal lead piping. [WWW Document]. https://www.dwqr.scot/media/1f1n3kv4/ueh-007-19-project-management-doc-048-final-project-report-20-november-2020.pdf (accessed 12.31.22).

Royer, H., 2009. Separated at Girth: US Twin Estimates of the Effects of Birth Weight. Am. Econ. J. Appl. Econ. 1, 49–85. https://doi.org/10.1257/APP.1.1.49

Scottish Water, 2021. Your guide to water pipework. [WWW Document] https://www.scottishwater.co.uk/-/media/ScottishWater/Document-Hub/Your-Home/Pipework-in-your-home/090821SWYourGuideToWaterPipework2021v1pageslr.pdf (accessed 12.31.22).

Sun L, Abraham S. 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. Journal of Econometrics. 225(2):175‑199. doi:10.1016/j.jeconom.2020.09

Troesken Werner. 2006. *The Great Lead Water Pipe Disaster*. Cambridge Mass: MIT Press.

UNICEF, 2020. The Toxic Truth: Children's exposure to lead pollution undermines a generation of potential | UNICEF [WWW Document]. URL https://www.unicef.org/reports/toxic-truth-childrens-exposure-to-lead-pollution-2020 (accessed 12.31.22).

UNICEF, 2022a. A Neglected Tragedy: The global burden of stillbirths - UNICEF DATA [WWW Document]. URL https://data.unicef.org/resources/a-neglected-tragedy-stillbirth-estimates-report/ (accessed 12.31.22).

UNICEF, 2022b. Child mortality data - UNICEF DATA [WWW Document]. https://data.unicef.org/resources/dataset/child-mortality/ (accessed 12.31.22).

UN IGME (United Nations Inter-agency Group for Child Mortality Estimation), 2021. [WWW Document]. https://www.who.int/publications/m/item/levels-and-trends-in-child-mortality-report-2021 (accessed 12.31.22).

Vinceti, M., Rovesti, S., Bergomi, M., Calzolari, E., Candela, S., Campagna, A., Milan, M., Vivoli, G., 2001. Risk of birth defects in a population exposed to environmental lead pollution. Sci. Total Environ. 278, 23–30. https://doi.org/10.1016/S0048-9697(00)00885-8

Vlasak, T., Jordakieva, G., Gnambs, T., Augner, C., Crevenna, R., Winker, R., Barth, A., 2019. Blood lead levels and cognitive functioning: A meta-analysis. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2019.03.052

Watt, G.C.M., Britton, A., Gilmour, H.G., Moore, M.R., Murray, G.D., Robertson, S.J., 2000. Public health implications of new guidelines for lead in drinking water: a case study in an area with historically high water lead levels. Food Chem. Toxicol. 38, S73–S79. https://doi.org/10.1016/S0278-6915(99)00137-4

Watt, G.C.M., Gilmour, W.H., Moore, M.R., Murray, G.D., Britton, A., Robertson, S.J., Womersley, J., 1996a. The Glasgow 93 Lead Study. Great Glasgow Health Board, Internal Report.

Watt, G.C.M., Gilmour, W.H., Moore, M.R., Murray, G.D., Britton, A., Robertson, S.J., Womersley, J., 1996b. Is lead in tap water still a public health problem? An observational study in Glasgow. BMJ 313, 979–981. https://doi.org/10.1136/bmj.313.7063.979

Wibberley, D.G., Khera, A.K., Edwards, J.H., Rushton, D.I., 1977. Lead levels in human placentae from normal and malformed births. J. Med. Genet. 14, 339. https://doi.org/10.1136/JMG.14.5.339

Wilcox, A.J., 2001. On the importance—and the unimportance— of birthweight. Int. J. Epidemiol. 30, 1233–1241. https://doi.org/10.1093/IJE/30.6.1233

Wooldridge, J.M., 2021. Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators. SSRN Electron. J. https://doi.org/10.2139/SSRN.3906345

Wooldridge, J. M., 2010. Econometric analysis of cross section and panel data (2nd ed.). MIT Press.

WHO, 1970. European standards for drinking-water, 2nd ed. World Health Organization. https://apps.who.int/iris/handle/10665/40025

WHO, 2010. Preventing Disease through Healthy Environments: a Global Assessment of the Burden of Disease from Environmental Risks. [WWW Document]. https://www.who.int/publications/i/item/9789241565196 accessed 12.31.22).

WHO, 2016. The public health impact of chemicals: knowns and unknowns. [WWW Document]. https://www.who.int/publications/i/item/WHO-FWC-PHE-EPE-16-01 (accessed 12.31.22).

WHO, 2021. Exposure to lead: a major public health concern, 2nd edition [WWW Document]. https://www.who.int/publications/i/item/9789240037656 (accessed 12.31.22).

Xie, X., Ding, G., Cui, C., Chen, L., Gao, Y., Zhou, Y., Shi, R., Tian, Y., 2013. The effects of low-level prenatal lead exposure on birth outcomes. Environ. Pollut. 175, 30–34. https://doi.org/10.1016/J.ENVPOL.2012.12.013

Yurdakök, K., 2012. Environmental pollution and the fetus. J. Pediatr. Neonatal Individ. Med. 1, 33–42. https://doi.org/10.7363/010116

Zheng, J., 2021. Lead exposure, human capital formation, and inequality: the impacts of drinking water lead on short-run and long-run educational outcomes. Working paper. https://www.jiamengzheng.com/assets/files/lead_JMP_zheng.pdf