CENTER FOR GLOBAL DEVELOPMENT

How Much Would Reducing Lead Exposure Improve Children's Learning in the Developing World?

Lee Crawfurd, Rory Todd, Susannah Hares, Justin Sandefur, and Rachel Silverman Bonnifield

ABSTRACT

Around half of children in low-income countries have elevated blood lead levels. What role does lead play in explaining low educational outcomes in these settings? We conduct a new systematic review and meta-analysis of observational studies on the relationship between lead exposure and learning outcomes. Adjusting for observable confounds and publication bias yields a benchmark estimate of a 0.12 standard deviation reduction in learning per natural log unit of blood lead. As all estimates are non-experimental, we present evidence on the likely magnitude of unobserved confounding, and summarize results from a smaller set of natural experiments. Our benchmark estimate accounts for over a fifth of the gap in learning outcomes between rich and poor countries, and implies moderate learning gains from targeted interventions for highly exposed groups (≈ 0.1 standard deviations) and modest learning gains (< 0.05 standard deviations) from broader public health campaigns.

KEYWORDS
Lead Poisoning, Child Education, Developing Countries

JEL CODES
P46, I18, K32

WORKING PAPER 650 · JULY 2023

How Much Would Reducing Lead Exposure Improve Children's Learning in the Developing World?

Lee Crawfurd, Rory Todd, Susannah Hares, Justin Sandefur, and Rachel Silverman Bonnifield
Center for Global Development

We're grateful to Bruce Lanphear and David Evans for helpful comments, and to Aisha Ali, Zack Gehan, and Maimouna Konate for research assistance. This work was supported by Givewell and the Effective Altruism Global Health and Development Fund.

Lee Crawfurd, Rory Todd, Susannah Hares, Justin Sandefur, and Rachel Silverman Bonnifield. 2023. "How Much Would Reducing Lead Exposure Improve Children's Learning in the Developing World?" CGD Working Paper 650. Washington, DC: Center for Global Development. https://www.cgdev.org/publication/how-much-would-reducing-lead-exposure-improve-childrens-learning-developing-world

CENTER FOR GLOBAL DEVELOPMENT

2055 L Street, NW Fifth Floor

Washington, DC 20036

202.416.4000

1 Abbey Gardens

Great College Street

London

SW1P 3SE

www.cgdev.org

Contents

List of Figures

List of Tables

1 Introduction

Over 600 million children in low- and middle-income countries have elevated blood lead levels (Rees and Fuller, 2020). This includes just 3 percent of children in high-income countries, and more than half of children in low-income countries (Figure A1).[1]

Children's scores on standardized tests of reading and mathematics in these same countries typically fall between one and three standard deviations below the level of performance of children in high-income settings (Angrist et al., 2021). The World Bank and UNESCO's modeled estimates suggest 91 percent of children in low-income countries cannot read and comprehend a simple text by age 10, compared to just 8 percent in high-income countries (Azevedo et al., 2022).

Are these two broad patterns connected? In this paper we explore the evidence of a causal link between lead exposure and children's learning outcomes. Since the relationship between lead and cognition varies with level of exposure, we place particular emphasis on the likely impacts at blood-lead levels observed among children in the developing world. We then ask what potential gains there are to learning from feasible interventions to reduce lead exposure in low- and lower-middle income countries. What proportion, if any, of the large learning gaps between countries might be explained by lead exposure? Further, is lead abatement a viable and cost-effective strategy to improving education performance?

We approach these questions by revisiting and extending existing meta-analyses of studies on the relationship between measures of blood lead and child learning outcomes. Much of the literature on the effects of lead on cognitive development has focused on impacts on IQ. A key contribution of our study is in additionally gathering effects on standardized test scores for reading and mathematics. We show that the magnitude of the association between blood lead levels is fairly consistent across these alternative outcomes. On average, we estimate that a one (natural) log unit increase in blood lead is associated with a $-.23\sigma$ change in reading and mathematics test scores. Ninety-five percent of estimates are between $-.28$ and $-.18\sigma$.

We then review how inclusion of controls for parental characteristics and socioeconomic status changes the association between lead and cognitive outcomes in

[1] These figures are derived from estimates produced by IHME/UNICEF (Rees and Fuller, 2020), who define elevated blood lead levels as exceeding the US CDC reference value of 5 micrograms per deciliter (the CDC lowered its reference value to 3.5 micrograms in 2021). We combine these exposure rates with population estimates from the UN (United Nations, Department of Economic and Social Affairs, Population Division, 2019).

associational studies. Standard controls typically, but not uniformly, lead to smaller associations with magnitudes that remain relevant for public health policy. We also assess the importance of publication bias in the literature. Standard tests based on funnel plots suggest publication bias may exaggerate the magnitude of the relationship between lead and cognitive outcomes. Lower-powered studies report larger effects, and an anomalous share of p-values fall just below conventional significance levels. However, standard approaches to correcting for this bias only moderately reduce our estimated average effect.

Accounting for study characteristics and correcting for publication bias, we find that a (natural) log unit increase in blood lead is associated with a $-.12\sigma$ decrease in reading and mathematics scores in the developing world. Magnitudes are similar for effects on IQ, reading, and mathematics scores.

While all of these estimates are based on observational studies, we also review the smaller number of quasi-experimental studies, all of which show larger effects for IV estimates than OLS estimates. By contrast, bounding procedures based on coefficient stability imply that unobserved heterogeneity could explain roughly one-third of the OLS association.

In order to quantify the overall potential learning gains from eliminating lead exposure, we combine our estimates of the association between lead and learning, with estimates of the prevalence of lead poisoning, and estimates of the effect of interventions targeting lead exposure. If given a causal interpretation, the overall results of this meta-analysis imply that reducing children's mean blood lead levels from average existing levels in LMICs (5.3 µg/dL) to high-income country levels (0.5µg/dL in the US according to the EPA) would close 21 percent of the learning gap between developing and developed countries for which we have data. Projects designed to reduce acute lead exposure at specific polluted sites have reduced mean blood lead levels by 1.35 - 2.3 log units, implying a 0.16 to to 0.26σ increase in learning for affected children. Projects targeting chronic low-level lead exposure in wider populations have reduced blood lead levels by 0.34 - 0.42 log units, imply learning gains of 0.04 to 0.05σ. Even these smaller effect sizes based on conservative assumptions about the effect of broad public health initiatives offer a potentially cost-effective means to improve learning outcomes (ignoring health benefits), if these programs can be implemented effectively and cheaply at scale in low- and middle-income countries.

2 Meta-analytic methods

We review the literature on the relationship between lead poisoning, measured through blood testing, and learning outcomes. We focus on studies that measure some kind of cognitive test outcome, and also have blood lead level measures from the same individuals, whether contemporaneously or from different points in time.

2.1 Search strategy

Our search strategy is summarised in the flowchart in Figure A2. We seek to identify all studies measuring the correlation between blood lead levels and IQ or test scores. We start with three recent systematic reviews. The first finds 27 papers published between 2010 and 2020 that measure child (aged 0–19 years) blood lead levels and 'full-scale' IQ scores (Galiciolli et al., 2022)[2], of which 9 had at least one eligible result according to the criteria described in the next section. The second reviews 8 papers published between 2000 and 2020, focused on effects on IQ, with a study population under the age of 12 years (Heidari et al., 2022), of which 7 had at least one eligible result. The third reviews 34 studies in low-and middle-income countries with measures of lead exposure and a standardized measurement of neurodevelopment, from which we extract an additional 7 studies (Heng et al., 2022). In addition, we conducted our own systematic search using Google Scholar. We search for articles with the following terms in their title: (lead) AND (exposure OR blood OR level) AND (intelligence OR intellectual OR cognitive OR cognition OR education OR achievement OR IQ OR score OR math OR reading OR school). From this we found 951 potential results. We also found several studies through citation searching from an initial relevant paper. These two methods identified an additional 24 papers that had at least one result fulfilling the eligibility criteria described below. This led to a combined total of 47 unique studies.

2.2 Extracting estimates from studies

Two researchers independently extracted estimates from studies following the same protocol, with disagreements resolved by a third researcher. We extract any coefficient relating maternal or child blood lead levels to any of three cognitive outcomes

[2]Full-scale IQ scores are the average of scores on five distinct abilities: Verbal Comprehension, Visual Spatial, Fluid Reasoning, Working Memory, and Processing Speed.

- IQ, math skills, or reading skills- on the full sample. We include estimates from sub-samples where subgroups are defined according to blood lead level. We focus on full-scale IQ rather than its separate components (performance and verbal IQ) where those are reportedly additionally. We include studies that use tests that are designed as general intelligence or IQ tests, such as the General Cognitive Index (Schnaas et al., 2006; Cooney et al., 1989), the British Ability Scales (Fulton et al., 1987), and the Kohs Block Design Test (Vega-Dienstmaier et al., 2006). We focus on general reading and math composite outcomes rather than any individual subcomponents. We exclude estimates which are entirely a combination of other extracted estimates. We exclude results which include blood measures for multiple ages in a model separately, as this has a different estimand: the effect of exposure at a particular age, relative to exposure at another age. We exclude estimates where the outcome is a binary re-coding of a continuous outcome, and a result using this continuous outcome has been extracted. We exclude estimates where the outcome was measured before the age of three. We exclude results of models which interact lead exposure with other covariates. For results to be eligible, studies need to report - or provide the information to calculate - an effect size, a standard error or confidence interval, a measure of blood lead exposure, and the standard deviation of the outcome. We also discarded results with less than 20 observations. After discarding studies that did not meet these criteria, we are left with 286 estimates from 47 unique studies.

We also code a number of other characteristics for each study: the country where the study was carried out; the specific outcome type (IQ, reading, or mathematics); the standard deviation of this outcome in the sample; the functional form for the estimation of the effect of blood lead on the cognitive outcome; the mean blood lead level of the sample[3]; whether exposure for a result was based on more than one blood measurement; the average age at blood sampling; the average age at outcome testing; and which variables were adjusted for in the analysis.

As we have multiple estimates from most studies, we account for the resulting dependence between results using robust variance estimation (Hedges et al., 2010), imputing a conservative value of 1 for the correlation between results from the same study, meaning that studies which contributed more results were not overweighted (Fisher and Tipton, 2015).

[3]Studies report averages as arithmetic means, geometric mean, and medians. We convert medians to arithmetic means using the method in Wan et al. (2014). Studies which report both arithmetic means and geometric means show the former are generally slightly larger as a result of the skewed distribution of blood lead levels. However, differences are small and so we treat both types as equivalent in analysis.

2.3 Harmonizing variables across studies

We follow the guidelines provided by the Cochrane Handbook for Systematic Reviews of Interventions (Higgins and Green, 2008) to calculate effect sizes and standard errors where these are not given in the required form. Where studies give effects as correlation coefficients, these are converted to raw regression coefficients using the formula:

$$\text{Effect Size} = \text{Correlation Coefficient} \times \frac{\sigma_y}{\sigma_x} \tag{1}$$

where σ_y and σ_x denote the standard deviation of the outcome (IQ or test scores, in our case) and the independent variable (blood lead levels), respectively.

Where studies only provide p-values for effect sizes, we convert these to standard errors using the formulas:

$$\text{Standard Error} = \text{Effect Size}/z \tag{2}$$

where z is calculated using the inverse CDF of the p-value divided by 2.

Where studies report only that a p-value is less than a particular value, we impute half the threshold value (e.g if a paper were to report $p < 0.01$, we would impute $p = 0.005$). For studies which do not report sample standard deviations, we impute population standard deviations where available. One study (Shadbegian et al., 2019) has as its outcome the percentile in which the child placed in math and reading tests; we convert this to a z-score, assuming test scores to be normally distributed.

2.4 Re-expressing coefficients from log and linear models

Existing meta-analyses have documented a roughly log-linear relationship between blood lead levels and IQ, whereby the natural logarithm of blood lead levels is proportional to IQ scores (see for example Figure 1).[4] Letting $\mathbf{X}$ denote a vector of

[4]In Figure 1 we collect published dose-response curves from 13 different studies. These include Lanphear et al. (2005) who combine data from seven different longitudinal studies that track a total of 1,333 children followed from birth or infancy until 5–10 years of age. Rothenberg and Rothenberg (2005) directly compare a linear and log-linear specification, finding a substantially better fit with the log-linear specification.

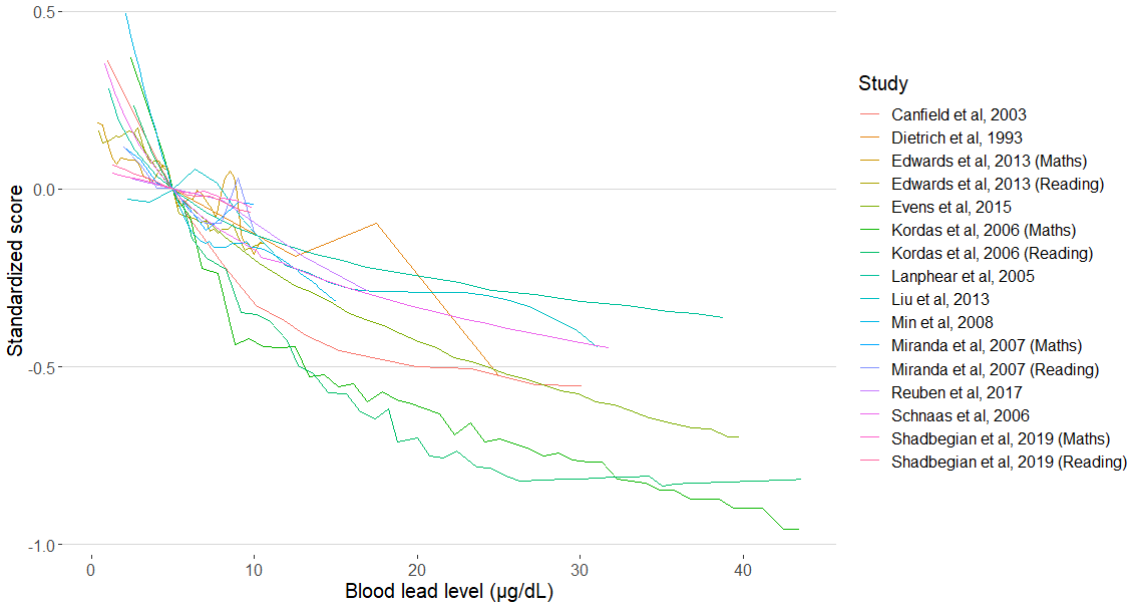
individual and family control variables, and subscript i denote individuals, this log-linear specification is our reference benchmark in what follows:

$$\frac{Y_i - \mu_y}{\sigma_y} = \alpha \ln(BLL_i) + \mathbf{X_i}\gamma + \varepsilon_i \tag{3}$$

where Y_i represents an individual's IQ or test score on mathematics or reading tests.

However, individual studies use a range of functional forms to model the relationship between blood lead exposure and the cognitive outcome. Some report a regression coefficient for a linear increase in blood lead, some report the coefficient for a log-unit increase—with bases of 2, e, or 10—and some report the difference between two groups defined by a value or range of values for blood lead levels. In order to compare the effects of lead exposure across studies, we first need to harmonize these results. Based on Lanphear et al. (2005) and other studies shown in Figure 1, we assume that the true relationship between blood lead and cognitive outcomes is log-linear, although we allow more flexibility in our meta-regression by including a linear term for a study's mean blood lead level. In order to compare across studies, we convert all effect sizes into natural log units. We do this using a 're-expression algorithm'. All such algorithms are necessarily imperfect. Linakis et al. (2021) show that re-expressions are likely to be biased where the distribution of the underlying variable has a skewed distribution, which is true in our case. In Table A3 we use microdata from the US NHANES study and three other studies in our review, to estimate both logarithmic and linear specifications using the same data, and compare the performance of three algorithms: Linakis et al. (2021), Rodríguez-Barranco et al. (2017), and Dzierlenga et al. (2020). The Linakis et al. (2021) method has the lowest root mean squared error, and so we proceed with this algorithm.

Figure 1: Log-linear dose-response curves



Note: This figure reproduces dose-response curves published from the individual studies named. They have been standardized and centred to converge at 5 µg/dL, to allow gradients to be compared. Curves from Edwards et al, 2013, refer to those for non-Hispanic Black children.

We first convert studies which compare discrete low and high exposure groups to a linear equivalent. Studies define "low" and "high" exposure groups using different thresholds. We therefore convert these effects to linear estimates by dividing the estimated effect size by the difference in average BLL between the high and low exposure groups.[5]

[5]Where results are from models with a wider range of covariates which excludes observations without selected variables, the mean blood lead level and outcome standard deviation may be slightly different to the values for the sample; however, these differences are likely to be negligible. For studies which express effects as the difference between two groups defined by a range of blood lead levels, we require mean blood lead levels for the two groups. Where this is not given, we impute these values where possible. Where the range of a group is only 1 µg/dL, we take the midpoint; for example, for a group with blood lead levels 4 to 5 µg/dL, we impute a mean of 4.5 µg/dL (Shadbegian et al., 2019; Edwards et al., 2013). For 7 studies, we simulate log-normal distributions using the overall mean and standard deviation given for the study sample, and then approximate group means given this simulated distribution (Liu et al., 2013; Surkan et al., 2007). For another we instead simulate a normal distribution as this better fits the reported moments of the data (Rasoul et al., 2012).

Following Linakis et al. (2021), we calculate a conversion factor I as a function of the desired logarithmic base α and median blood lead level (BLL).[6]

$$I = \alpha^{log_\alpha(meanBLL)+0.5} - \alpha^{log_\alpha(meanBLL)-0.5} \tag{4}$$

We then multiply coefficients estimated from linear models by this conversion factor to obtain an estimate of what the coefficient would have been under a logarithmic model.

Overall, we see larger effects for estimates from linear models than from logarithmic models, suggesting that some error may still result from this approximation. In order to mitigate against this, we adjust for whether a result was reported in a non-logarithmic form in our metaregression analysis, described below. We also include an interaction with the average blood lead level of the sample, as the magnitude of error will be a function of this variable.

[6]While this algorithm performs best using the median of the exposure distribution, we find it still performs well using arithmetic means, the most commonly reported average in our studies.

2.5 Adjustments for publication bias

A common concern in meta-analyses, particularly those involving observational studies, is publication or reporting bias. If only statistically significant results are reported, we will produce a biased estimate of the true average effect. In our case, funnel plots do indicate that publication bias may be an issue in this literature, with relatively few statistically insignificant results (Figure A4). Specifically, results are asymmetrically distributed, with "missing" results in the region of null or even positive associations, and the distribution of z-statistics shows a spike just below the significance threshold. The former pattern is confirmed by an Egger et al. (1997) asymmetry test, and appears not to be caused by errors arising from the re-expression algorithms (Figure A5).

There are several approaches to adjusting for publication bias, none of which are perfect. For our main estimate we include a control for the variance of each estimate, following Stanley and Doucouliagos (2014). We also show in Appendix Table A4 that adjusting our uncontrolled result using three other approaches, which generally only moderately weaken the effect. The simplest approach is the Egger regression intercept, which includes the standard error as a control variable. The Egger approach is shown by Stanley and Doucouliagos (2014) in simulations to be overly biased towards zero, and that a better approach ('PET-PEESE') is to use a nonlinear quadratic approximation to the true unknown nonlinear relationship between effect sizes and effect size variance, as we do in our main result. The trim-and-fill approach estimates the number of results 'missing' due to publication bias and impute these results. Finally, the p-uniform* method makes use of the principle that the p-values of estimates should be uniformly distributed at the true effect size (Aert and Assen, 2018). We can also show the sensitivity of average effects to different assumptions about the degree of selection on statistical significance (Copas and Shi, 2001). Results from this method, shown in table A6, indicate that treatment effects remain statistically significant with a range of reasonable assumptions about the potential degree of selection of estimates on statistical significance (Figure A6).

2.6 Meta-regression

Our focus in this review is estimating the effect of lead exposure on learning in developing countries. Our primary estimate of the average effect is therefore the constant term β_0 from a meta-regression in which an observed effect size $\hat{\theta}_{ij}$ for result i in study j is related to m result covariates, 1 through M. Estimates are

weighted inversely to their variance, and we use the Hedges et al. (2010) estimator which is robust to unknown correlation between multiple estimates from the same study.

$$\hat{\theta}_{ij} = \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + ... + \beta_m X_{Mij} + \epsilon_{ij} \tag{5}$$

The independent variables correspond to characteristics of the underlying studies. These are (i) whether the original result modelled the effect of blood lead on the cognitive outcomes using a logarithmic specification or not; (ii) the average blood lead level for the result sample, centred by the average of mean childhood BLLs for the 34 countries estimated by Ericson et al. (2021); (iii) an interaction between these two effects; (iv) an indicator for whether BLL was measured contemporaneously to the outcome measurement, as the highest BLL measurement among several taken for an individual, or prenatally, rather than as an average over several measurements over a lifetime or as a 'lagged' measurement during an earlier potentially critical developmental period; (v) an indicator variable for studies from high income countries; (vi) an indicator for results using IQ rather than reading or mathematics test scores as the outcome variable; (vii) an indicator for results which fail to control for parental education; (viii) an indicator for results which fail to control for family income or socioeconomic status;[7] (ix) the variance of the estimate, which may vary systematically with effect size in the presence of publication bias. All of these variables are coded such that a value of zero corresponds to our preferred specification: i.e., effects on reading or mathematics in a low- or middle-income country, in a context with average BLL, using a log-linear specification, where exposure is calculated as a lifetime average or from an earlier period than the time of outcome measurement, with controls for household income and parental education. Note that no single result fulfils all of these criteria.

[7]We define this kind of control broadly: for example, we code results which adjust for the Home Observation for Measurement of the Environment, which measures the physical and social environment of children, as controlling for this.

3 Meta-analysis results

3.1 Study characteristics

From the total of 47 studies included in our meta-analysis, 17 are from the United States, 12 are from other high-income countries (Australia, Canada, Italy, New Zealand, South Korea, Taiwan, and UK), and 18 are from low or middle-income countries (Brazil, China, Colombia, Ecuador, Egypt, India, Malaysia, Mexico, Nigeria, Peru, Pakistan, and the Philippines). All except one study are observational (the exception Aizer et al. (2018) employs an instrumental variable design), however many are longitudinal.

The median sample size in our estimates is 389. The median age at blood lead testing is four years old, and at cognitive testing is eight years old. The average blood lead level is 7.06µg/dL. The majority of results have controls for parent IQ or education and family background or income (Table 1).

Table 1: Result characteristics

	Mean	Median	SD
Age at Blood Test	4.70	4.00	3.00
Age at Outcome	7.74	8.00	2.70
Mean blood lead	7.06	6.70	4.31
Sample Size	13,869.04	389.00	40,211.86
Binary Controls			
Used logarithmic specification	0.24	0.00	0.43
Controlled for parent IQ or ed	0.69	1.00	0.46
Controlled for family income or wealth	0.61	1.00	0.49
Used average or lagged BLL measure	0.62	1.00	0.49
Outcome was Maths or English	0.40	0.00	0.49
From Low/Middle Income Country	0.37	0.00	0.48
Observations	286		

Note: This table shows descriptive statistics for the 286 estimates from 47 studies in our meta-analysis.

3.2 Average effects

Overall, we find that a one log unit reduction in BLL is associated with a $-.12$ standard deviation improvement in test scores. As discussed in Section 2.6, our preferred

meta-regression specification adjusts for the choice of original model specification, and control variables in each study, and is shown in Table 2.

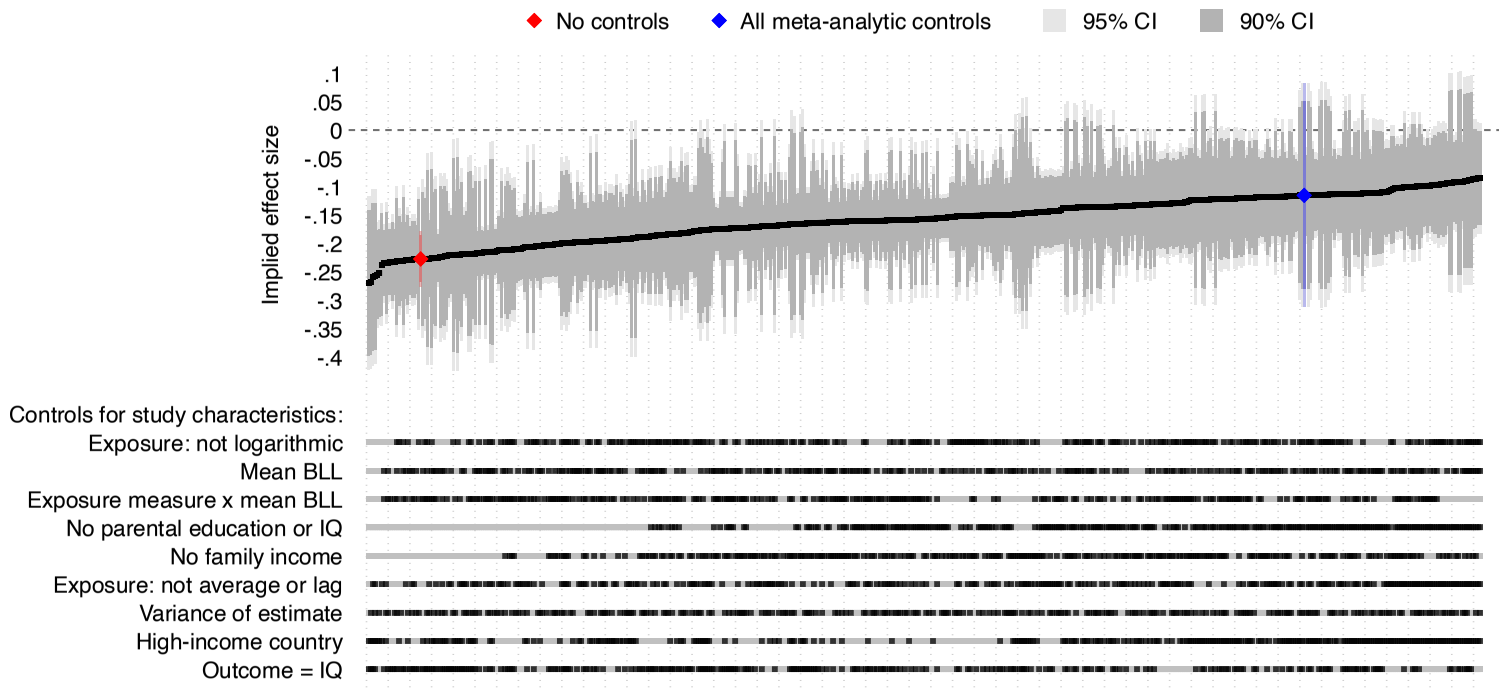
To explore the influence of specific sets of controls systematically, we report a specification curve a la Simonsohn et al. (2014). Figure 2 displays the average effect size from a meta-analytic regression controlling for each of all possible combinations of study characteristics listed above. All specifications yield a negative relationship between (log) BLL and learning outcomes, though some are statistically insignificant.

Table 2: Meta-regression of effect size on study characteristics

	(1)	(2)	(3)	(4)
Constant	-0.227***	-0.225***	-0.154***	-0.115
	(0.025)	(0.026)	(0.021)	(0.100)
Potential confounds and sources of bias:		ref.	ref.	ref.
Effect variance (for pub.bias)		-0.139	-0.143	-0.137
		(0.488)	(0.484)	(0.462)
No control for parent ed/IQ			-0.173***	-0.167**
			(0.058)	(0.072)
No control for family income			0.002	0.009
			(0.061)	(0.071)
Harmonization of specification and context:				ref.
Not logarithmic				0.000
				(0.057)
Mean BLL				0.000
				(0.007)
Not log spec X Mean BLL				-0.008
				(0.011)
Exposure: not average or lag				-0.036
				(0.047)
Outcome (IQ)				0.015
				(0.061)
High income country				-0.046
				(0.063)
N (Estimates)	286	286	286	286
N (Studies)	47	47	47	47

Note: * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses. We use the Hedges et al. (2010) estimator to account for dependence between multiple estimates from the same study. The dependent variable in each case is the standardized effect size of one log unit increase in lead exposure. The constant represents the average effect. Column 1 presents the unadjusted average. Column 2 includes the PET-PEESE (Stanley and Doucouliagos, 2014) adjustment for publication bias. Column 3 includes controls for potentially confounding study-level characteristics - so the constant represents the average for studies that estimate a log-linear relationship, in a sample with average blood lead levels, with controls for parental IQ or education, and with a lagged measure of blood lead. In column 4, we additionally include controls to adjust the constant to studies in low or middle-income country settings, and focused on reading or mathematics test scores rather than IQ.

Figure 2: Specification curve for meta-analytic regression of estimates of the relationship between (log) blood lead levels and learning outcomes



Note: Each coefficient corresponds to the mean relationship between log BLL and standardized learning outcomes from a meta-analytic regression controlling for the study characteristics listed in the bottom panel.

We also show our results graphically in the forest plots in Figures 3 and 4. These figures show the weighted average of effect sizes across studies for our two main outcomes, IQ and math/reading test scores, respectively. The mean of the pooled effect is -0.22 for results with IQ as the outcome, -0.20 for studies with math scores, and -0.24 for studies with reading scores.[8] Unlike our meta-regression estimate, however, these averages are unadjusted for study characteristics and potential publication bias.

Results are robust to dropping each individual study sequentially (Figure A3), and also to dropping studies in which we had to convert or impute different statistics (Table A6). For studies in which we use the population rather than sample standard deviation in outcomes, effects are 0.04 standard deviations larger, but not statistically significantly different.

3.3 Heterogeneity

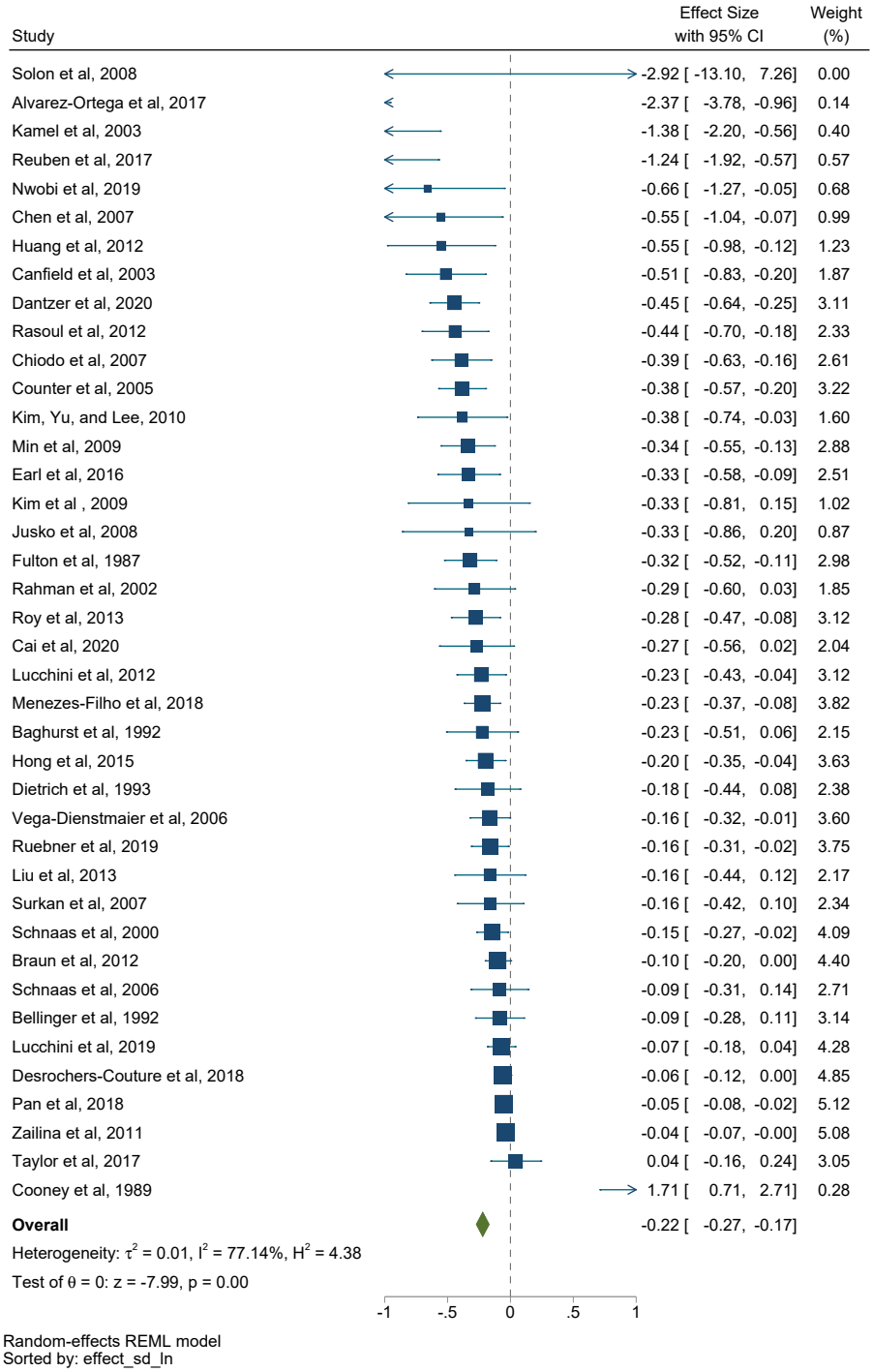
Our main meta-analysis shows substantial heterogeneity between studies. The I^2 statistics (the share of variability in effect sizes not caused by sampling error) from Figures 3 and 4 are all well above the 75% rule-of-thumb indicating "significant heterogeneity" (Higgins et al., 2003).

As we saw in Table 2 however, just one covariate was statistically significantly correlated with study effect size - whether that study controlled for parental education or IQ. The control for publication bias (the effect variance) also has a large coefficient, but is imprecisely estimated, and does not create a large shift in the overall mean effect size. We see no statistically significant differences for effects on math or reading compared with IQ, for high-income countries compared with middle-income countries, or between studies according to the timing of blood lead measurement, i.e., concurrent versus lagged or averaged over measurements at multiple time points.

We also show each of these controls individually in bivariate regressions in Table A5. While controlling for most of these study characteristics individually only alters the overall effect of log BLL on learning outcomes modestly, the combined effect of controlling for all of them is quite dramatic. The overall estimate of the BLL-learning link falls from $-.23\sigma$ in column 1 with no controls to -0.13σ with various controls intended to adjust for potential confounding, publication bias, and differences in measurement (Table 2). Finally, because our ultimate interest is in reading

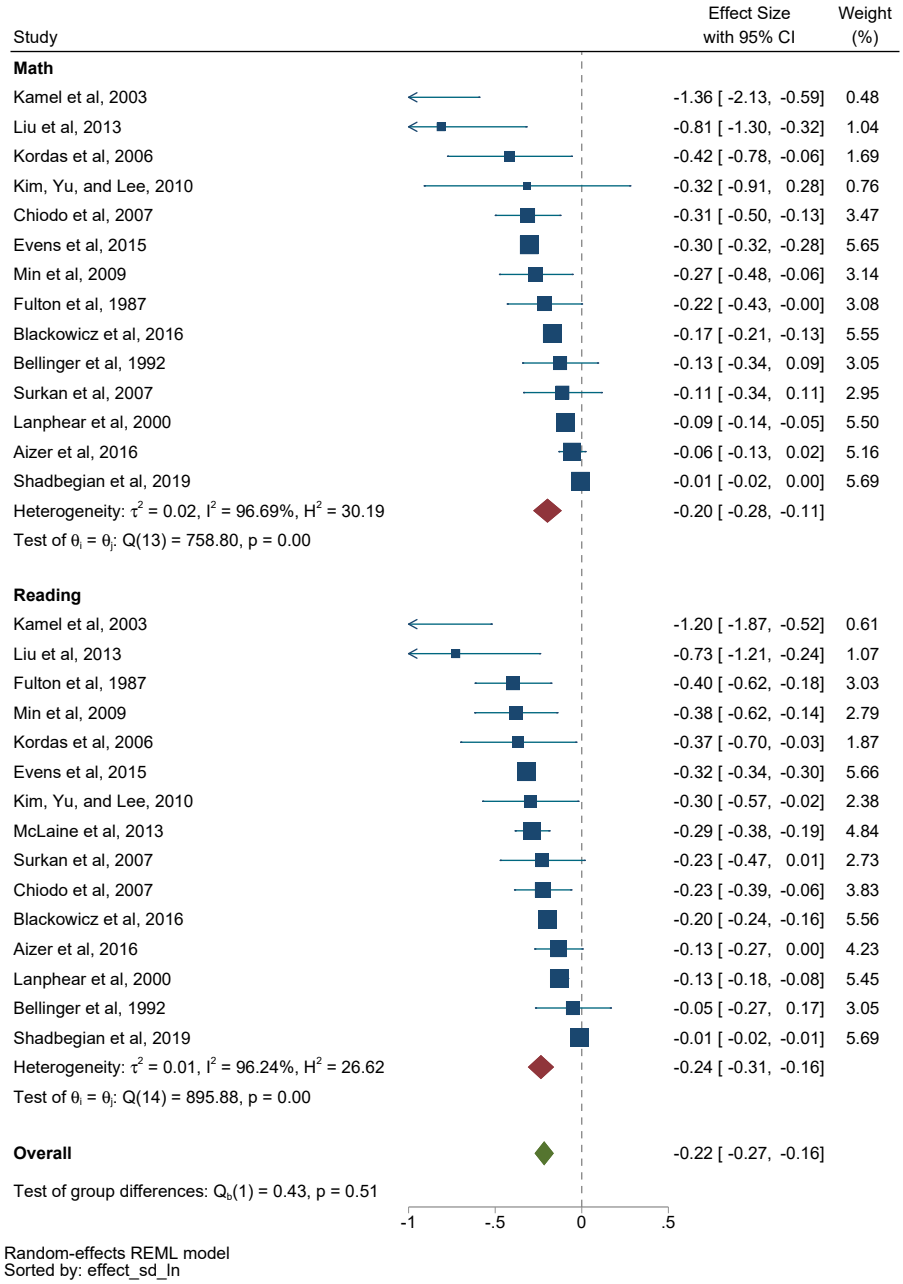
[8]Note that here we show results collapsed to their mean for each study where there are multiple estimates from a study. These results are similar to those obtained using robust variance estimation accounting for unknown correlation between estimates within studies (Hedges et al., 2010).

Figure 3: Effects of lead on IQ



Note: This figure shows average effects for each study. In cases where more than one effect is reported per study, we show here the mean value. The estimated overall effect size (of -0.22 with no moderators) based on these mean effect sizes is only marginally different to the effect size estimated from all individual estimates, with a robust variance estimator to account for unknown dependence within studies.

Figure 4: Effects of lead on reading and mathematics assessments



Note: This figure shows average effects for each study. In cases where more than one effect is reported per study, we show here the mean value. The estimated overall effect size (of -0.20 for math and -0.24 for reading with no moderators) based on these mean effect sizes is only slightly different to the effect size estimated from all individual estimates, with a robust variance estimator to account for dependence within studies.

and mathematics scores in developing countries, in column 4 we also control for an indicator of studies in high-income countries and results using IQ as the outcome variable, yielding an estimate of $-.12\sigma$.

3.4 Comparing our results to other reviews

How do our estimates compare to prior meta-analyses? Lanphear et al. (2005) estimate that an increase from 2.4 µ/dL to 30 µ/dl was associated with a 3.9 decrease in IQ, equivalent to a 0.18 standard deviation change per log unit increase in exposure. An earlier review found that a doubling of lead (from 10 to 20 µg/dl blood lead or 5 to 10 ug/g tooth lead) was associated with a 1 to 2 point reduction in IQ. This is equivalent to around a 0.14σ reduction per log unit of lead (Pocock et al., 1994). Heidari et al. (2022) find that the standardised difference in mean IQ scores between high (>10µg/dL) and low exposure groups (<10µg/dL) is 0.541σ. Similarly, Galicioli et al. (2022) find a difference between exposed and unexposed groups of 7.37 IQ points (equivalent to approximately 0.49σ). Heng et al. (2022) provide a narrative review but not a meta-analytic average effect size. Our results for effects on IQ are therefore well within the range of other reviews, whilst we also extend our review to reading and mathematics test scores.

4 Assessing the role of unobserved confounders

Since observational studies may suffer from omitted variable bias, what reason is there to believe the associations we report above reflect a causal relationship? One piece of evidence in support of a causal interpretation is that experimental studies conducted with animals show significant impacts of lead exposure on cognitive function (see for example Gilbert and Rice (1987) and Tena et al. (2019)). Furthermore, though the mechanisms by which lead exposure causes cognitive impairment are still not fully understood, lead has been shown to interfere with several processes involved in neurological development and functioning, adding credibility to a causal interpretation (Ramírez Ortega et al., 2021).

While animal studies and the identification of neurological mechanisms lend credence to the existence of some causal effect, they don't tell us much about the magnitude of this effect in humans, or whether observational associations are biased by unmeasured confounders. In this section we present two sources of evidence on potential bias which can help us pin down the magnitude of the true causal effect. First, we discuss studies focused on natural experiments in lead exposure. Second, we assess the sensitivity of observational estimates to selection on unobserved variables using the coefficient stability method (Oster, 2016).

First, a handful of natural experiments with a reasonable claim to causality have been published on the lead-cognition relationship, and these studies allow us to directly compare causal with observational estimates. In Table 3 below we collect observational ordinary least squares (OLS) and quasi-experimental instrumental variable (IV) estimates from five studies.[9] In all cases, the IV estimate is larger than the OLS estimate, sometimes substantially so. Aizer et al. (2018) find results two to three times larger in the IV specification than the OLS one. They argue that the IV results are larger due to measurement error attenuating the OLS estimates. An alternative explanation for seeing larger coefficients for IV estimates than OLS estimates is that they estimate treatment effects for slightly different populations. OLS estimates the average treatment effect across the entire population, whereas the IV estimates the *local* average treatment effect only amongst the subgroup of the population for whom the IV shifts their behaviour. We don't, however, have a compelling reason to believe that treatment effects would necessarily be different in the different

[9]We don't include all of these studies in our meta-analysis as they don't all have direct blood lead measures. We do include the one study that does have blood lead measures, which uses a lead remediation program in Rhode Island as a source of exogenous variation, showing that instrumental variable estimates are larger than observational estimates (Aizer et al., 2018).

sub-populations affected and unaffected by the instruments in these contexts. Aizer et al. (2018) report four estimates from linear specifications (effects per unit change in blood lead), for two outcomes (reading and math) and two different instrucmental variable strategies. Applying the re-expression algorithm we describe in section 2.4, these results are equivalent to an average effect size of -0.13 standard deviations per natural log unit of blood lead, very close to our central meta-analytic estimate of $-.12$ standard deviation.

Clay et al. (2019) use an instrumental variable strategy based on the layout of the 1944 Interstate Highway System Plan to estimate the relationship between county soil lead levels and cognitive difficulties amongst students. They find a small and statistically insignificant relationship in the observational OLS specification, but statistically significant effects with the IV specification. Counties with a highway recommended in the 1944 plan were 17 percent more likely to have above median lead concentrations in topsoil. In the IV specification, counties with above median soil lead concentrations had 4 percentage points higher levels of children with cognitive difficulties (10 times larger than in the OLS specification).

Table 3: IV Estimates Produce Larger Effects than OLS Estimates

Author	**Outcome**	**OLS**	**IV**	**Lead measure**	**Instrument**
Aizer et al 2018	Reading	0.026 (0.002)	0.073 (0.037)	Blood	Nearby home lead remediation
Aizer et al 2018	Math	0.017 (0.001)	0.030 (0.034)	Blood	Nearby home lead remediation
Clay et al 2019	Cognitive Difficulty	0.0029 (0.004)	0.0414 (0.019)	Soil	1944 Interstate Highway Plan
Grönqvist et al., 2020	GPA	0.164 (.030)	0.256 (.0769)	Moss	Leaded petrol ban
Feigenbaum et al., 2016	Homicides	0.219 (0.064)	1.022 (0.257)	Water pipes	Distance to lead refinery

Note: Results for Aizer et al (2018) take from Table 8, for Clay et al (2019) taken from Table 3, for Grönqvist et al. (2020) from Table 6., and for Feigenbaum et al (2016) from Table 2.

Grönqvist et al. (2020) study the phase-out of leaded petrol in Sweden in 1980-81. They show that reductions in lead measured in nationwide moss samples is associated

with improved test scores. They find IV estimates that are around 50 percent larger than OLS estimates.

A similar pattern of larger effect sizes in quasi-experimental estimates than in observational estimates has also been shown in studies looking at other outcomes besides test scores, including school suspensions and juvenile detention (Aizer and Currie, 2019) and homicide rates (Feigenbaum and Muller, 2016).

Other quasi-experimental studies also show causal effects of lead exposure, though without providing a direct comparison of observational and causal estimates. Hollingsworth et al. (2022) focus on the switch to unleaded petrol in US Nascar motor racing in 2007. They estimate that removing the exposure of a school 1 mile from a race track would be equivalent to increasing school spending per pupil by $750. Sorensen et al. (2019) evaluate lead hazard control programs, finding that each percentage point reduction in lead poisoning in early childhood led to 0.04 standard deviations higher math scores and 0.08 standard deviations reading scores. Rau et al. (2015) study the opening of a toxic waste dump in Chile in a difference-in-difference framework, showing that attending schools closer to the site after it opened lowered test scores. Higney et al. (2022) shows that water treatment in Scotland lead to improvements in test scores only in areas with high prevalence of lead pipes.

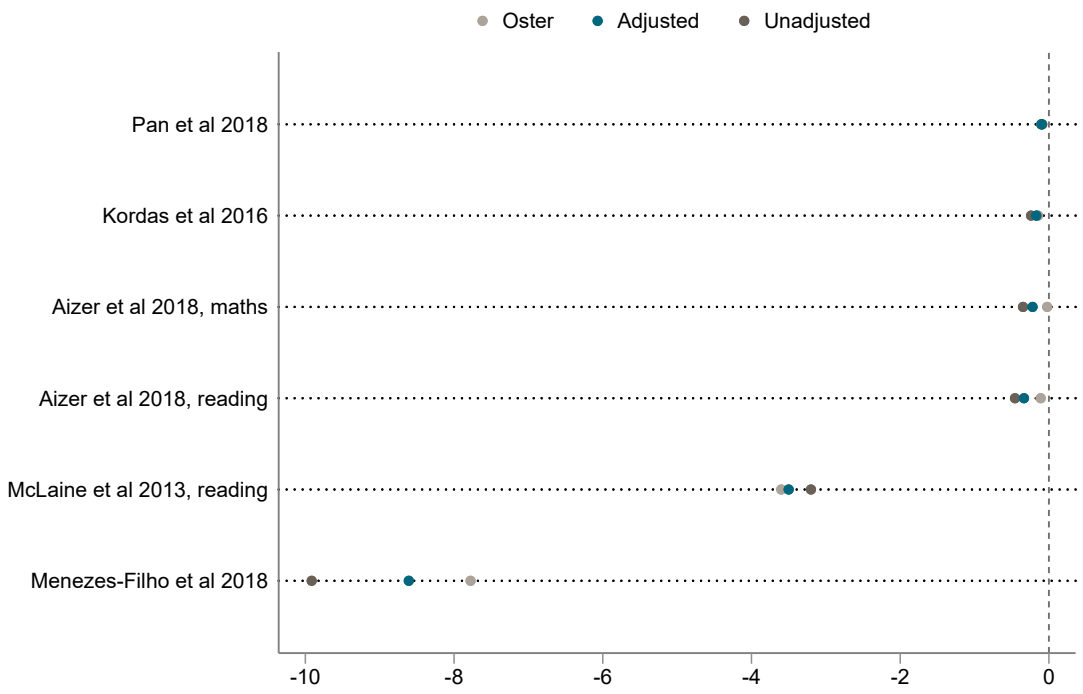
Second, we assess the robustness of observational estimates to selection on both observed covariates and proportional selection on unobserved covariates, based on the stability of estimated coefficients and model fit (R-squared). Specifically, we follow Oster (2016) in estimating the coefficient $\beta^\star$ that would result were we able to control for unobserved confounders. This equation (6) requires that we make assumptions about the proportional degree of selection δ on unobservables (typically set at equal to selection on observables or 1), and about the maximum plausible value of the model fit R_{max} (typically set at 1.3 times the model fit in the fully adjusted model $\tilde{R}$). Given these assumptions and our knowledge of the coefficients in the adjusted $\tilde{\beta}$ and unadjusted models $\mathring{\beta}$, and model fits of the adjusted $\tilde{R}$ and unadjusted models $\mathring{R}$, we can calculate the coefficient $\beta^\star$.

$$\beta^\star = \tilde{\beta} - \delta \frac{(\mathring{\beta} - \tilde{\beta})(R_{max} - \tilde{R})}{(\tilde{R} - \mathring{R})} \tag{6}$$

We identify 18 estimates (from 16 studies) in which both an unadjusted and adjusted estimate is provided. Only six of these estimates (from five studies) also report an R-squared which allows us to calculate the Oster-adjusted estimate. These

estimates are reported in Figure 5. Amongst the six estimates for which we can estimate adjustments for both observable and unobservable confounders, the average reduction in the coefficient adjusted for observables is 14 percent, whereas the average reduction for observables and unobservables is 33 percent.

Figure 5: Selection on observed and unobserved variables



Note: This figure shows the change in the coefficient in a regression of learning on blood lead before ("unadjusted") and after adjusting for observed covariates ("adjusted"), and proportional selection on unobserved covariates ("Oster"). We show here estimates from all studies in our meta-analysis that report both an adjusted and unadjusted estimate. We are only able to report the Oster-adjusted estimates for the five studies that report the R-squared of their regressions.

5 Policy simulations

In order to interpret these effects, we consider two benchmarks; first, what the effect on learning would be of reducing current lead exposure in low and middle income countries to the level in high-income countries. Second, the effect of different interventions to reduce blood lead, and what this implies for learning.

5.1 The effect of removing lead altogether

What would be the effect of reducing current lead exposure in low and middle income countries to the level in high-income countries? A recent systematic review shows that mean BLL levels (μ_{bll}) in 34 low and middle-income countries range from 1.7 to 9.3, with an unweighted average across countries of 5.3µg/dL (Ericson et al., 2021). Levels for under 18 year olds in the US are around 0.5µg/dL (US EPA, 2015).

To simulate the effect of reducing BLL levels in each country to the U.S. level, we follow Ericson et al. (2021) and assume a log-normal distribution of BLL within each country. Thus for each country we calculate the mean of the log of BLL as

$$\mu = \ln\left(\frac{\mu_{bll}^2}{\sqrt{\mu_{bll}^2 + \sigma_{bll}^2}}\right)$$

Then, because our preferred specification for the relationship between BLL and cognitive scores is log-linear, we can calculate the effect on learning as:

$$\text{Improvement in learning} = \beta \times (\mu - 0.5)$$

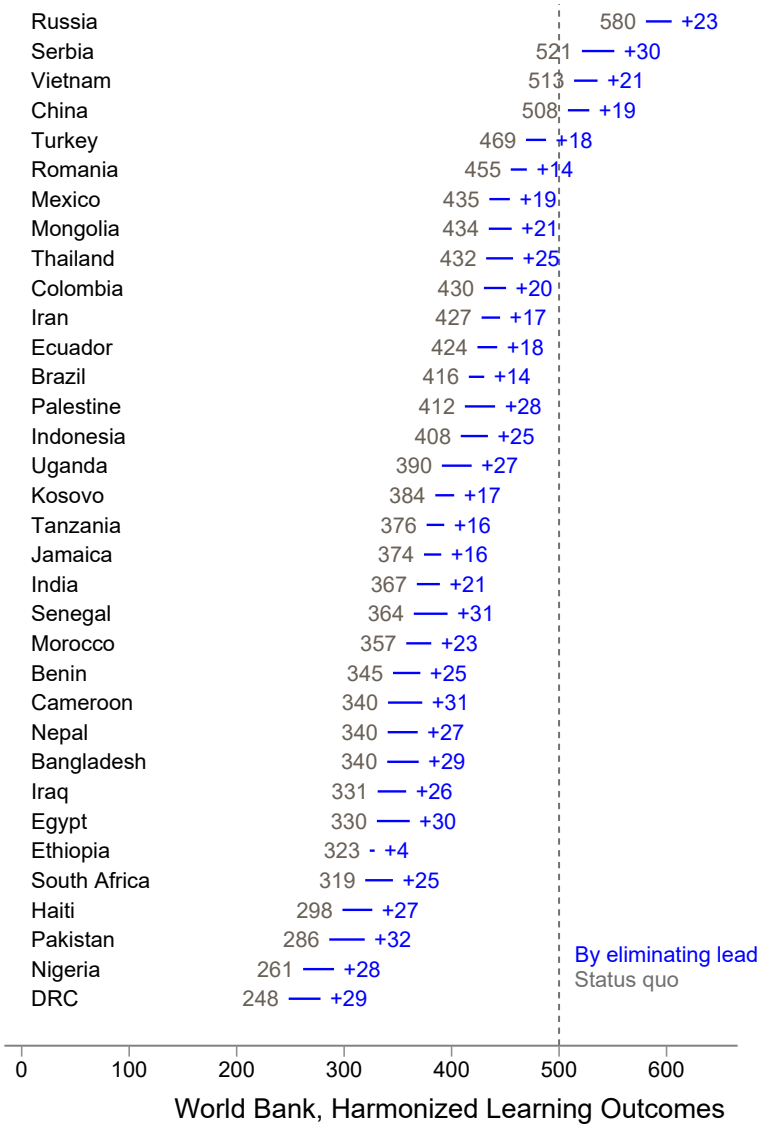
where β is the average effect size from the meta-analysis above, e.g., $-.12\sigma$ after covariate adjustments.

To measure gaps in learning outcomes around the world, we rely on the harmonized learning outcomes reported by the World Bank (Angrist et al., 2021), which are normed to have a mean of 500 and standard deviation of 100 points.

Figure 6 shows the implied impact on the World Bank harmonized learning outcomes for each of the 34 countries in the Ericson et al. (2021) sample. Using the adjusted estimates of lead's effect on learning from the meta-analysis ($-.12$) the increments are meaningful, with a magnitude ranging from 4 to 31 points depending on the country. The results suggest a major role for lead exposure in explaining

learning gaps between rich and poor countries. As an example, the lowest scoring country, the Democratic Republic of Congo, lies 250 points below the global mean of 500, and reducing BLL to 0.5µg/dL would improve scores by 29 points. On average for these 34 countries, reducing BLL to US levels improves learning by 23 points, equivalent to 21 percent of the 110 point learning gap to the global benchmark of 500 points on the World Bank scale.

Figure 6: Simulated effect of eliminating blood lead on national learning outcomes



Note: For each country we simulate the effect of reducing blood lead levels to 0.5 μg/dL, based on the distributions of BLL reported for 34 countries by Ericson et al. (2021), average learning outcomes reported by the World Bank, and the coefficient from our meta-regression of a $-.12\sigma$ in learning per natural logarithm increase in BLL.

5.2 The effect of interventions to reduce blood lead

Reducing lead exposure altogether down to US levels may be unrealistic, at least in the short-term. What kind of reductions are feasible? We review studies on four classes of intervention: regulatory, educational, medical, and targeted environmental interventions.

5.2.1 Regulation

The most important regulatory intervention has been the banning of leaded gasoline or petrol, which has shown that large reductions in blood lead are feasible at low cost in both developed and developing countries Angrand et al. (2022). Before-after studies around leaded petrol bans have shown large reductions in blood lead in both high and low-income countries, including in India (Singh and Singh, 2006), Kenya (United Nations Environment Programme (UNEP), 2014), Pakistan (Manser et al., 1990; Rahbar et al., 2002), and South Africa (Mathee et al., 2006). Whilst leaded gasoline has already now been banned in every country in the world, there may still be considerable room for increased regulatory action in other areas. Other regulatory actions taken by the majority of OECD countries include regulations on lead in water, air, food, batteries, food containers, and paint, with several countries having additional regulations on lead in dust, soil, sewage, waste, and pesticides (Silbergeld, 1997). The majority of low- and middle-income countries do not have regulations on the sale of lead in paint (UNEP, 2020). The evidence base though remains thin - a systematic review into exposure via consumer products found zero studies on either regulatory, educational, or environmental interventions, and no studies on regulatory interventions targeting exposure via drinking water (Pfadenhauer et al., 2016).

5.2.2 Targeted remediation

Second, some environmental interventions targeted at high risk populations have shown large reductions, though few efforts in developing countries have been well documented. One search for remediation efforts in developing countries (whether evaluated or not) found just 13 projects in total (O'Brien et al., 2021). Another systematic review of studies on remediation of lead-contaminated soil found just five studies, all in North America, and with mixed results (Dobrescu et al., 2022). In Table 4 we summarise results from three studies Nigeria, Dominican Republic, and Bangladesh. All were areas of acute exposure and remediation was expensive, but

did lead to large falls in blood lead. For example a huge reduction was documented from efforts in Zamfara, Nigeria, from 149 to 15 (Tirima et al., 2016), implying improvement in learning by 0.28 standard deviations. A soil remediation effort at a former lead smelter in Haina, the Dominican Republic, reduced lead levels from 20.6 µg/dL to 5.34 µg/dL, or 0.16 standard deviations in test scores (Ericson et al., 2018). Both of these large gains came from addressing acutely polluted sites. In Bangladesh, soil capping, household cleaning, and awareness-raising reduced lead levels from 22.6 µg/dL to 14.8 µg/dL Chowdhury et al. (2021). All of these studies rely on before-after designs and so may be biased estimates.

Several studies from the US show results of targeted paint remediation, demonstrating modest positive effects on blood lead levels (Billings and Schnepel, 2018; Leighton et al., 2003; Staes et al., 1994). Removing lead pipes has also shown promise (Pfadenhauer et al., 2016).

5.2.3 Education

Parental education efforts in Georgia and China where exposure is more diffuse and chronic, have found modest reductions in blood lead. Following striking results from a nationally representative blood lead testing in Georgia, a government program of action was implemented to educate parents on the issue. This involved initial BLL testing, followed by letters sent to all families of children with elevated BLLs, with advice on how to reduce lead exposure and on dietary habits that can help reduce BLLs (increased calcium, iron and vitamin C). Parents were advised to visit a pediatrician to assess physical and mental development and iron deficiency, and pediatricians were provided training in early detection and management of lead exposure. This reduced BLLs from 9.6 to 6.8 µg/dL (Ruadze et al., 2021), equivalent to a 0.04 standard deviation improvement in test scores. A parental education intervention in China reduced BLL from 10 to 8 (Shen et al., 2004). A recent study from Bangladesh has shown promising impacts on behaviour change, but doesn't yet report results on blood lead (Jahir et al., 2021).

5.2.4 Medical intervention

Finally, evidence on the provision of calcium supplements has shown mixed results, from no effect in the US (Markowitz et al., 2004) and Mexico (Ettinger et al., 2009), to small effects in Nigeria (Keating et al., 2011) and large effects in Indonesia (Syofyan et al., 2020; Haryanto et al., 2015). Calcium supplementation may therefore hold promise in contexts in which calcium deficiency is of greater concern.

Table 4: The implied effect of feasible blood lead reductions on learning

Change in Blood lead	**High BLL**	**Low BLL**	**Effect on Learning (SD)**	**Study Design**
One Log Unit (from meta-analysis)			0.12	
Reduction from LMIC to HIC levels	5.3	0.5	0.28	
Observed reduction in US (1976 to 1990)	14.6	2.8	0.20	Before-after
1. Regulatory intervention				
- Leaded Petrol Ban				
- India	18.1	12.1	0.05	Before-after
- Kenya	8	5.6	0.04	Before-after
- Pakistan	38	15.6	0.11	Before-after
- South Africa	16	6.4	0.11	Before-after
2. Targeted Remediation				
- Soil remediation				
- Zamfara, Nigeria	149	15	0.28	Before-after
- Dominican Republic	20.6	5.34	0.16	Before-after
- Bangladesh	22.6	14.8	0.05	Before-after
- Paint remediation				
- US (NY)	24.3	12.3	0.08	Diff-in-diff
- US (NC)	17.85	9	0.08	Diff-in-diff
- US (Missouri)	34	29.58	0.02	Diff-in-diff
3. Education (parental)				
- Georgia	9.6	6.8	0.04	Before-after
- China	10.1	7.9	0.03	RCT
4. Medical (calcium supplements)				
- Indonesia (Medan)	2.1	0.01	0.64	RCT
- Indonesia (Bandung)	13.7	4.95	0.12	RCT
- Nigeria	9.9	8.8	0.01	Case-control
- Mexico	4.1	3.649	0.01	RCT
- US	21.4	21.7	0.00	RCT

Note: The High and Low Blood Lead Levels (BLLs) shown are the average arithmetic mean for low- and middle-income countries or LMICs (Ericson et al., 2021), high-income countries or HICs (US EPA, 2015), and the study in China (Shen et al., 2004), geometric means for the studies in Nigeria (Tirima et al., 2016), Dominican Republic (Ericson et al., 2018), and Bangladesh (Chowdhury et al., 2021), and medians for Georgia (Ruadze et al., 2021). Geometric means are generally smaller than arithmetic means. The total effect on learning is calculated as the product of the log unit gap, and the estimated coefficient of the effect of one log unit on learning outcomes, taken from our meta-analysis.

6 Conclusion

Can widespread lead exposure explain low average learning outcomes in the developing world? While data coverage is limited in both cases, (a) mean lead exposure is over 5 µg/dL among children in low- and middle-income countries with reasonably representative samples, or roughly ten-times higher than the United States, and (b) learning levels among primary-school aged pupils lag more than one full standard deviations behind OECD levels in the same set of countries.

We extend existing meta-analyses of studies linking lead exposure to cognitive outcomes, expanding the traditional focus on IQ to include measures of reading and mathematics performance among primary-school students. Taken at face value, the association between lead and learning outcomes in individual-level data across 47 studies suggests that one log unit higher blood lead levels reduces learning levels by roughly $-.23$ standard deviations. In a simple model of global learning gaps, this effect size is sufficient to suggest that observed lead levels explain over half of the gap in learning outcomes between developing and developed countries.

Raw correlations likely overstate the true causal impact of lead on learning levels, for at least two reasons. First, we find evidence of publication bias in estimates of the lead-learning link. Funnel plots reveal striking asymmetry in the distribution of findings, and clustering of p-values just below conventional significance levels. Second, it is impossible to rule out a large role for unobserved confounding in the lead-learning link. We present various pieces of evidence suggesting the true causal link is much smaller than the published literature.

Meta-analytic regressions controlling for these and other potential sources of bias suggest that 1 log unit in blood lead levels reduces learning levels by roughly $-.12$ standard deviations, approximately one half of the naive estimate. Nevertheless, a true causal effect of this magnitude may remain a viable lever for policy action based only on education outcomes – in addition to myriad other health benefits beyond the scope of this paper.

Notably, even with our most conservative estimates, the magnitude of learning gains achievable through lead eradication is comparable to many popular policy initiatives to improve education quality in the developing world (eg, Evans and Yuan, 2022 show that the average effect size on learning from RCTs in global education is 0.1 standard deviations). In cost-benefit terms, a focus on lead exposure is likely justified solely on education grounds if countries are able to achieve significant reductions in lead exposure through low-cost, large scale policy reforms such as improved regulation of the lead paint and lead battery industries.

References

Aert, R. C. M. v. and Assen, M. A. L. M. v. (2018). Correcting for Publication Bias in a Meta-Analysis with the P-uniform* Method.

Aizer, A. and Currie, J. (2019). Lead and Juvenile Delinquency: New Evidence from Linked Birth, School, and Juvenile Detention Records. The Review of Economics and Statistics, 101(4):575–587.

Aizer, A., Currie, J., Simon, P., and Vivier, P. (2018). Do Low Levels of Blood Lead Reduce Children's Future Test Scores? American Economic Journal: Applied Economics, 10(1):307–341.

Alvarez-Ortega, N., Caballero-Gallardo, K., and Olivero-Verbel, J. (2017). Low blood lead levels impair intellectual and hematological function in children from Cartagena, Caribbean coast of Colombia. Journal of trace elements in medicine and biology: organ of the Society for Minerals and Trace Elements (GMS), 44:233–240.

Angrand, R. C., Collins, G., Landrigan, P. J., and Thomas, V. M. (2022). Relation of blood lead levels and lead in gasoline: an updated systematic review. Environmental Health, 21(1):138.

Angrist, N., Djankov, S., Goldberg, P. K., and Patrinos, H. A. (2021). Measuring human capital using global learning data. Nature, 592(7854):403–408. Number: 7854 Publisher: Nature Publishing Group.

Azevedo, J. P., Rogers, H., Ahlgren, E., Akmal, M., Cloutier, M.-H., Ding, E., Raza, A., Wong, Y. N., Montoya, S., Chakroun, B., Chang, G.-C., Guerriero, S., Dewan, P., Mizunoya, S., Reuge, N., Russell, K., Yao, H., Bronwin, R., Cohen-Mitchell, J., Dintilhac, C., and Boggild-Jones, I. (2022). The State of Global Learning Poverty: 2022 Update. Technical report, World Bank, UNESCO, UNICEF, FCDO, USAID, Bill & Melinda Gates Foundation.

Baghurst, P. A., McMichael, A. J., Wigg, N. R., Vimpani, G. V., Robertson, E. F., Roberts, R. J., and Tong, S.-L. (1992). Environmental Exposure to Lead and Children's Intelligence at the Age of Seven Years. New England Journal of Medicine, 327(18):1279–1284. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJM199210293271805.

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

Billings, S. B. and Schnepel, K. T. (2018). Life after Lead: Effects of Early Interventions for Children Exposed to Lead. American Economic Journal: Applied Economics, 10(3):315–344.

Blackowicz, M. J., Hryhorczuk, D. O., Rankin, K. M., Lewis, D. A., Haider, D., Lanphear, B. P., and Evens, A. (2016). The Impact of Low-Level Lead Toxicity on School Performance among Hispanic Subgroups in the Chicago Public Schools. International Journal of Environmental Research and Public Health, 13(8):774.

Braun, J. M., Hoffman, E., Schwartz, J., Sanchez, B., Schnaas, L., Mercado-Garcia, A., Solano-Gonzalez, M., Bellinger, D. C., Lanphear, B. P., Hu, H., Tellez-Rojo, M. M., Wright, R. O., and Hernandez-Avila, M. (2012). Assessing windows of susceptibility to lead-induced cognitive deficits in Mexican children. NeuroToxicology, 33(5):1040–1047.

Cai, Q.-l., Peng, D.-j., Lin-Zhao, Chen, J.-w., Yong-Li, Luo, H.-l., Ou, S.-y., Huang, M.-l., and Jiang, Y.-m. (2021). Impact of Lead Exposure on Thyroid Status and IQ Performance among School-age Children Living Nearby a Lead-Zinc Mine in China. NeuroToxicology, 82:177–185.

Canfield, R. L., Henderson, C. R., Cory-Slechta, D. A., Cox, C., Jusko, T. A., and Lanphear, B. P. (2003). Intellectual Impairment in Children with Blood Lead Concentrations below 10 g per Deciliter. New England Journal of Medicine, 348(16):1517–1526. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMoa022848.

Chen, A., Cai, B., Dietrich, K. N., Radcliffe, J., and Rogan, W. J. (2007). Lead exposure, IQ, and behavior in urban 5-7 year olds: Does lead affect behavior only by lowering IQ? Pediatrics, 119(3):e650–e658. asterix.

Chiodo, L. M., Covington, C., Sokol, R. J., Hannigan, J. H., Jannise, J., Ager, J., Greenwald, M., and Delaney-Black, V. (2007). Blood lead levels and specific attention effects in young children. Neurotoxicology and Teratology, 29(5):538–546.

Chowdhury, K. I. A., Nurunnahar, S., Kabir, M. L., Islam, M. T., Baker, M., Islam, M. S., Rahman, M., Hasan, M. A., Sikder, A., Kwong, L. H., Binkhorst, G. K.,

Nash, E., Keith, J., McCartor, A., Luby, S. P., and Forsyth, J. E. (2021). Child lead exposure near abandoned lead acid battery recycling sites in a residential community in Bangladesh: Risk factors and the impact of soil remediation on blood lead levels. Environmental Research, 194:110689.

Clay, K., Portnykh, M., and Severnini, E. (2019). The legacy lead deposition in soils and its impact on cognitive function in preschool-aged children in the United States. Economics & Human Biology, 33:181–192.

Cooney, G. H., Bell, A., McBride, W., and Carter, C. (1989). Low-Level Exposures to Lead: The Sydney Lead Study. Developmental Medicine & Child Neurology, 31(5):640–649.

Copas, J. B. and Shi, J. Q. (2001). A sensitivity analysis for publication bias in systematic reviews. Statistical Methods in Medical Research, 10(4):251–265.

Counter, S. A., Buchanan, L. H., and Ortega, F. (2005). Neurocognitive impairment in lead-exposed children of Andean lead-glazing workers. Journal of Occupational and Environmental Medicine, 47(3):306–312.

Dantzer, J., Ryan, P., Yolton, K., Parsons, P. J., Palmer, C. D., Cecil, K., and Unrine, J. M. (2020). A comparison of blood and toenails as biomarkers of children's exposure to lead and their correlation with cognitive function. The Science of the total environment, 700:134519.

Desrochers-Couture, M., Oulhote, Y., Arbuckle, T. E., Fraser, W. D., Séguin, J. R., Ouellet, E., Forget-Dubois, N., Ayotte, P., Boivin, M., Lanphear, B. P., and Muckle, G. (2018). Prenatal, concurrent, and sex-specific associations between blood lead concentrations and IQ in preschool Canadian children. Environment International, 121(Pt 2):1235–1242.

Dietrich, K. N., Berger, O. G., Succop, P. A., Hammond, P. B., and Bornschein, R. L. (1993). The developmental consequences of low to moderate prenatal and postnatal lead exposure: Intellectual attainment in the cincinnati lead study cohort following school entry. Neurotoxicology and Teratology, 15(1):37–44.

Dobrescu, A.-I., Ebenberger, A., Harlfinger, J., Griebler, U., Klerings, I., Nußbaumer-Streit, B., Chapman, A., Affengruber, L., and Gartlehner, G. (2022). Effectiveness of interventions for the remediation of lead-contaminated soil to prevent or reduce lead exposure - A systematic review. Science of The Total Environment, 806:150480.

Duval, S. and Tweedie, R. (2000). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. Biometrics, 56(2):455–463.

Dzierlenga, M. W., Crawford, L., and Longnecker, M. P. (2020). Birth weight and perfluorooctane sulfonic acid: a random-effects meta-regression analysis. Environmental Epidemiology, 4(3):e095.

Earl, R., Burns, N., Nettelbeck, T., and Baghurst, P. (2016). Low-level environmental lead exposure still negatively associated with children's cognitive abilities. Australian Journal of Psychology, 68(2):98–106. Publisher: Routledge _eprint: https://doi.org/10.1111/ajpy.12096.

Edwards, S., Anthopolos, R., and Miranda, M. L. (2013). The Impact of Early Childhood Lead Exposure on Educational Test Performance among Connecticut Schoolchildren: Phase II Report. Technical report, Children's Environmental Health Initiative.

Egger, M., Davey Smith, G., Schneider, M., and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. BMJ (Clinical research ed.), 315(7109):629–634.

Ericson, B., Caravanos, J., Depratt, C., Santos, C., Cabral, M. G., Fuller, R., and Taylor, M. P. (2018). Cost Effectiveness of Environmental Lead Risk Mitigation in Low- and Middle-Income Countries. GeoHealth, 2(2):87–101.

Ericson, B., Hu, H., Nash, E., Ferraro, G., Sinitsky, J., and Taylor, M. P. (2021). Blood lead levels in low-income and middle-income countries: a systematic review. The Lancet Planetary Health, 5(3):e145–e153.

Ettinger, A. S., Lamadrid-Figueroa, H., Téllez-Rojo, M. M., Mercado-García, A., Peterson, K. E., Schwartz, J., Hu, H., and Hernández-Avila, M. (2009). Effect of Calcium Supplementation on Blood Lead Levels in Pregnancy: A Randomized Placebo-Controlled Trial. Environmental Health Perspectives, 117(1):26–31.

Evans, D. K. and Yuan, F. (2022). How Big Are Effect Sizes in International Education Studies? Educational Evaluation and Policy Analysis, 44(3):532–540.

Evens, A., Hryhorczuk, D., Lanphear, B. P., Rankin, K. M., Lewis, D. A., Forst, L., and Rosenberg, D. (2015). The impact of low-level lead toxicity on school performance among children in the Chicago Public Schools: a population-based retrospective cohort study. Environmental Health, 14(1):21.

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

Fisher, Z. and Tipton, E. (2015). robumeta: An R-package for robust variance estimation in meta-analysis. arXiv:1503.02220 [stat].

Fulton, M., Thomson, G., Hunter, R., Raab, G., Laxen, D., and Hepburn, W. (1987). Influence of Blood Lead on the Ability and Attainment of Children in Edinburgh. The Lancet, 329(8544):1221–1226. Publisher: Elsevier.

Galiciolli, M. E. A., Lima, L. S., da Costa, N. d. S., de Andrade, D. P., Irioda, A. C., and Oliveira, C. S. (2022). IQ alteration induced by lead in developed and underdeveloped/developing countries: A systematic review and a meta-analysis. Environmental Pollution (Barking, Essex: 1987), 292(Pt A):118316.

Gilbert, S. G. and Rice, D. C. (1987). Low-level lifetime lead exposure produces behavioral toxicity (Spatial discrimination reversal) in adult monkeys. Toxicology and Applied Pharmacology, 91(3):484–490.

Grönqvist, H., Nilsson, J. P., and 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. Publisher: The University of Chicago Press.

Haryanto, B., Sutrisna, B., and Djaja, I. M. (2015). Effect of Calcium Supplementation on School Children's Blood Lead Levels in Indonesia. International Journal of Science and Research, 2015(1):5005.

Hedges, L. V., Tipton, E., and Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. Research Synthesis Methods, 1(1):39–65.

Heidari, S., Mostafaei, S., Razazian, N., Rajati, M., Saeedi, A., and Rajati, F. (2022). The effect of lead exposure on IQ test scores in children under 12 years: a systematic review and meta-analysis of case-control studies. Systematic Reviews, 11(1):106.

Heng, Y. Y., Asad, I., Coleman, B., Menard, L., Benki-Nugent, S., Were, F. H., Karr, C. J., and McHenry, M. S. (2022). Heavy metals and neurodevelopment of children in low and middle-income countries: A systematic review. PLOS ONE, 17(3):e0265536.

Higgins, J. P. T. and Green, S., editors (2008). Cochrane Handbook for Systematic Reviews of Interventions. Wiley–Blackwell, Chichester, England ; Hoboken, NJ, 1st edition edition.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., and Altman, D. G. (2003). Measuring inconsistency in meta-analyses. BMJ, 327(7414):557–560. Publisher: British Medical Journal Publishing Group Section: Education and debate.

Higney, A., Hanley, N., and Moro, M. (2022). The Impact of Lead Pollution on Human Capital Formation: Size of Dose Matters. Technical report.

Hollingsworth, A., Huang, M., Rudik, I., and Sanders, N. J. (2022). A Thousand Cuts: Cumulative Lead Exposure Reduces Academic Achievement. Working Paper.

Hong, S.-B., Im, M.-H., Kim, J.-W., Park, E.-J., Shin, M.-S., Kim, B.-N., Yoo, H.-J., Cho, I.-H., Bhang, S.-Y., Hong, Y.-C., and Cho, S.-C. (2015). Environmental lead exposure and attention deficit/hyperactivity disorder symptom domains in a community sample of South Korean school-age children. Environmental Health Perspectives, 123(3):271–276.

Huang, P.-C., Su, P.-H., Chen, H.-Y., Huang, H.-B., Tsai, J.-L., Huang, H.-I., and Wang, S.-L. (2012). Childhood blood lead levels and intellectual development after ban of leaded gasoline in Taiwan: A 9-year prospective study. Environment International, 40:88–96.

Jahir, T., Pitchik, H. O., Rahman, M., Sultana, J., Shoab, A. K. M., Nurul Huda, T. M., Byrd, K. A., Islam, M. S., Yeasmin, F., Baker, M., Yeasmin, D., Nurunnahar, S., Luby, S. P., Winch, P. J., and Forsyth, J. E. (2021). Making the invisible visible: Developing and evaluating an intervention to raise awareness and reduce lead exposure among children and their caregivers in rural Bangladesh. Environmental Research, 199:111292.

Jusko, T. A., Henderson, C. R., Lanphear, B. P., Cory-Slechta, D. A., Parsons, P. J., and Canfield, R. L. (2008). Blood Lead Concentrations < 10 g/dL and Child Intelligence at 6 Years of Age. Environmental Health Perspectives, 116(2):243–248.

Kamel, N. M., Ramadan, A. M., Kamel, M. I., Mostafa, Y. A. E. G., Abo el Naga, R. M., and Ali, A. M. (2003). Impact of lead exposure on health status and

scholastic achievement of school pupils in Alexandria. The Journal of the Egyptian Public Health Association, 78(1-2):1–28.

Keating, E. M., Fischer, P. R., Pettifor, J. M., Pfitzner, M., Isichei, C. O., and Thacher, T. D. (2011). The Effect of Calcium Supplementation on Blood Lead Levels in Nigerian Children. The Journal of Pediatrics, 159(5):845–850.e1.

Kim, D.-S., Yu, S.-D., and Lee, E.-H. (2010). Effects of blood lead concentration on intelligence and personality in school children. Molecular & Cellular Toxicology, 6(1):19–23.

Kim, Y., Kim, B.-N., Hong, Y.-C., Shin, M.-S., Yoo, H.-J., Kim, J.-W., Bhang, S.-Y., and Cho, S.-C. (2009). Co-exposure to environmental lead and manganese affects the intelligence of school-aged children. NeuroToxicology, 30(4):564–571.

Kordas, K., Canfield, R. L., López, P., Rosado, J. L., Vargas, G. G., Cebrián, M. E., Rico, J. A., Ronquillo, D., and Stoltzfus, R. J. (2006). Deficits in cognitive function and achievement in Mexican first-graders with low blood lead concentrations. Environmental Research, 100(3):371–386.

Lanphear, B. P., Hornung, R., Khoury, J., Yolton, K., Baghurst, P., Bellinger, D. C., Canfield, R. L., Dietrich, K. N., Bornschein, R., Greene, T., Rothenberg, S. J., Needleman, H. L., Schnaas, L., Wasserman, G., Graziano, J., and Roberts, R. (2005). Low-level environmental lead exposure and children's intellectual function: an international pooled analysis. Environmental Health Perspectives, 113(7):894–899. 1449 citations (Crossref) [2022-11-07].

Leighton, J., Klitzman, S., Sedlar, S., Matte, T., and Cohen, N. L. (2003). The effect of lead-based paint hazard remediation on blood lead levels of lead poisoned children in New York City. Environmental Research, 92(3):182–190.

Linakis, M. W., Landingham, C. V., and Longnecker, M. P. (2021). Re-expressing coefficients from regression models for inclusion in a meta-analysis. Pages: 2021.11.02.466931 Section: New Results.

Liu, J., Li, L., Wang, Y., Yan, C., and Liu, X. (2013). Impact of Low Blood Lead Concentrations on IQ and School Performance in Chinese Children. PLoS ONE, 8(5):e65230.

Lucchini, R. G., Guazzetti, S., Renzetti, S., Conversano, M., Cagna, G., Fedrighi, C., Giorgino, A., Peli, M., Placidi, D., Zoni, S., Forte, G., Majorani, C., Pino,

A., Senofonte, O., Petrucci, F., and Alimonti, A. (2019). Neurocognitive impact of metal exposure and social stressors among schoolchildren in Taranto, Italy. Environmental Health: A Global Access Science Source, 18(1):67.

Lucchini, R. G., Zoni, S., Guazzetti, S., Bontempi, E., Micheletti, S., Broberg, K., Parrinello, G., and Smith, D. R. (2012). Inverse association of intellectual function with very low blood lead but not with manganese exposure in Italian adolescents. Environmental Research, 118:65–71.

Manser, W. W., Lalani, R., Haider, S., and Khan, M. A. (1990). Trace element studies on Karachi populations. Part V: Blood lead levels in normal healthy adults and grammar school children. JPMA The Journal of the Pakistan Medical Association, 40(7):150–154.

Markowitz, M. E., Sinnett, M., and Rosen, J. F. (2004). A randomized trial of calcium supplementation for childhood lead poisoning. Pediatrics, 113(1 Pt 1):e34–39.

Mathee, A., Röllin, H., von Schirnding, Y., Levin, J., and Naik, I. (2006). Reductions in blood lead levels among school children following the introduction of unleaded petrol in South Africa. Environmental Research, 100(3):319–322.

McLaine, P., Navas-Acien, A., Lee, R., Simon, P., Diener-West, M., and Agnew, J. (2013). Elevated blood lead levels and reading readiness at the start of kindergarten. Pediatrics, 131(6):1081–1089.

Menezes-Filho, J. A., Carvalho, C. F., Rodrigues, J. L. G., Araújo, C. F. S., dos Santos, N. R., Lima, C. S., Bandeira, M. J., Marques, B. L. d. S., Anjos, A. L. S., Bah, H. A. F., Abreu, N., Philibert, A., and Mergler, D. (2018). Environmental Co-Exposure to Lead and Manganese and Intellectual Deficit in School-Aged Children. International Journal of Environmental Research and Public Health, 15(11):2418.

Min, M. O., Singer, L. T., Kirchner, H. L., Minnes, S., Short, E., Hussain, Z., and Nelson, S. (2009). Cognitive Development and Low-Level Lead Exposure in Poly-Drug Exposed Children. Neurotoxicology and teratology, 31(4):225–231.

Nwobi, N. L., Adedapo, S. K., Olukolade, O., Oyinlade, O. A., Lagunju, I. A., Atulomah, N. O., Nwazuoke, I. A., and Anetor, J. I. (2019). Positive and inverse correlation of blood lead level with erythrocyte acetylcholinesterase and intelligence quotient in children: implications for neurotoxicity. Interdisciplinary Toxicology, 12(3):136–142.

Oster, E. (2016). Unobservable Selection and Coefficient Stability: Theory and Validation. Journal of Business & Economic Statistics.

O'Brien, R. M., Phelan, T. J., Smith, N. M., and Smits, K. M. (2021). Remediation in developing countries: A review of previously implemented projects and analysis of stakeholder participation efforts. Critical Reviews in Environmental Science and Technology, 51(12):1259–1280.

Pan, S., Lin, L., Zeng, F., Zhang, J., Dong, G., Yang, B., Jing, Y., Chen, S., Zhang, G., Yu, Z., Sheng, G., and Ma, H. (2018). Effects of lead, cadmium, arsenic, and mercury co-exposure on children's intelligence quotient in an industrialized area of southern China. Environmental Pollution, 235:47–54.

Pfadenhauer, L. M., Burns, J., Rohwer, A., and Rehfuess, E. A. (2016). Effectiveness of interventions to reduce exposure to lead through consumer products and drinking water: A systematic review. Environmental Research, 147:525–536.

Pocock, S. J., Smith, M., and Baghurst, P. (1994). Environmental lead and children's intelligence: a systematic review of the epidemiological evidence. BMJ (Clinical research ed.), 309(6963):1189–1197.

Rahbar, M. H., White, F., Agboatwalla, M., Hozhabri, S., and Luby, S. (2002). Factors associated with elevated blood lead concentrations in children in Karachi, Pakistan. Bulletin of the World Health Organization, 80:769–775. Publisher: World Health Organization.

Rahman, A., Maqbool, E., and Zuberi, H. S. (2002). Lead-associated deficits in stature, mental ability and behaviour in children in Karachi. Annals of Tropical Paediatrics, 22(4):301–311.

Ramírez Ortega, D., González Esquivel, D. F., Blanco Ayala, T., Pineda, B., Gómez Manzo, S., Marcial Quino, J., Carrillo Mora, P., and Pérez de la Cruz, V. (2021). Cognitive Impairment Induced by Lead Exposure during Lifespan: Mechanisms of Lead Neurotoxicity. Toxics, 9(2):23.

Rasoul, G. A., Al-Batanony, M., Mahrous, O., Abo-Salem, M., and Gabr, H. (2012). Environmental Lead Exposure among Primary School Children in Shebin El-Kom District, Menoufiya Governorate, Egypt. 3(4).

Rau, T., Urzúa, S., and Reyes, L. (2015). Early Exposure to Hazardous Waste and Academic Achievement: Evidence from a Case of Environmental Negligence.

Journal of the Association of Environmental and Resource Economists, 2(4):527–563.

Rees, N. and Fuller, R. (2020). The Toxic Truth: Children's Exposure to Lead Pollution Undermines a Generation of Future Potential. Technical report, UNICEF, Pure Earth.

Reuben, A., Caspi, A., Belsky, D. W., Broadbent, J., Harrington, H., Sugden, K., Houts, R. M., Ramrakha, S., Poulton, R., and Moffitt, T. E. (2017). Association of Childhood Blood Lead Levels With Cognitive Function and Socioeconomic Status at Age 38 Years and With IQ Change and Socioeconomic Mobility Between Childhood and Adulthood. JAMA, 317(12):1244.

Rodríguez-Barranco, M., Tobías, A., Redondo, D., Molina-Portillo, E., and Sánchez, M. J. (2017). Standardizing effect size from linear regression models with log-transformed variables for meta-analysis. BMC Medical Research Methodology, 17(1):44.

Rothenberg, S. J. and Rothenberg, J. C. (2005). Testing the Dose–Response Specification in Epidemiology: Public Health and Policy Consequences for Lead. Environmental Health Perspectives, 113(9):1190–1195.

Roy, A., Ettinger, A. S., Hu, H., Bellinger, D., Schwartz, J., Modali, R., Wright, R. O., Palaniappan, K., and Balakrishnan, K. (2013). Effect modification by Transferrin C2 polymorphism on lead exposure, hemoglobin levels, and IQ. Neurotoxicology, 38:17–22.

Ruadze, E., Leonardi, G. S., Saei, A., Khonelidze, I., Sturua, L., Getia, V., Crabbe, H., Marczylo, T., Lauriola, P., and Gamkrelidze, A. (2021). Reduction in Blood Lead Concentration in Children across the Republic of Georgia following Interventions to Address Widespread Exceedance of Reference Value in 2019. International Journal of Environmental Research and Public Health, 18(22):11903.

Ruebner, R. L., Hooper, S. R., Parrish, C., Furth, S. L., and Fadrowski, J. J. (2019). Environmental lead exposure is associated with neurocognitive dysfunction in children with chronic kidney disease. Pediatric nephrology (Berlin, Germany), 34(11):2371–2379.

Schnaas, L., Rothenberg, S. J., Flores, M.-F., Martinez, S., Hernandez, C., Osorio, E., Velasco, S. R., and Perroni, E. (2006). Reduced Intellectual Development

in Children with Prenatal Lead Exposure. Environmental Health Perspectives, 114(5):791–797.

Shadbegian, R., Guignet, D., Klemick, H., and Bui, L. (2019). Early childhood lead exposure and the persistence of educational consequences into adolescence. Environmental Research, 178:108643.

Shen, X.-m., Yan, C.-h., Wu, S.-h., and Shi, R. (2004). Parental education to reduce blood lead levels in children with mild and moderate lead poisoning: a randomized controlled study. Zhonghua Er Ke Za Zhi = Chinese Journal of Pediatrics, 42(12):892–897.

Silbergeld, E. K. (1997). Preventing Lead Poisoning in Children. Annual Review of Public Health, 18(1):187–210. _eprint: https://doi.org/10.1146/annurev.publhealth.18.1.187.

Simonsohn, U., Nelson, L. D., and Simmons, J. P. (2014). p-Curve and Effect Size: Correcting for Publication Bias Using Only Significant Results. Perspectives on Psychological Science, 9(6):666–681. Publisher: SAGE Publications Inc.

Singh, A. K. and Singh, M. (2006). Lead decline in the Indian environment resulting from the petrol-lead phase-out programme. The Science of the Total Environment, 368(2-3):686–694.

Solon, O., Riddell, T. J., Quimbo, S. A., Butrick, E., Aylward, G. P., Lou Bacate, M., and Peabody, J. W. (2008). Associations between cognitive function, blood lead concentration, and nutrition among children in the central Philippines. The Journal of Pediatrics, 152(2):237–243.

Sorensen, L. C., Fox, A. M., Jung, H., and Martin, E. G. (2019). Lead exposure and academic achievement: evidence from childhood lead poisoning prevention efforts. Journal of Population Economics, 32(1):179–218.

Staes, C., Matte, T., Copley, C. G., Flanders, D., and Binder, S. (1994). Retrospective Study of the Impact of Lead-based Paint Hazard Remediation on Children's Blood Lead Levels in St. Louis, Missouri. American Journal of Epidemiology, 139(10):1016–1026.

Stanley, T. D. and Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. Research Synthesis Methods, 5(1):60–78.

Surkan, P. J., Zhang, A., Trachtenberg, F., Daniel, D. B., McKinlay, S., and Bellinger, D. C. (2007). Neuropsychological function in children with blood lead levels <10 microg/dL. Neurotoxicology, 28(6):1170–1177.

Syofyan, S. S., Wahyuni, A. S., Rusmil, K., and Lelo, A. (2020). The Effects of Calcium Supplementation on Blood Lead Levels and Short-term Memory of Chronically Exposed Children: A Clinical Trial Study. Open Access Macedonian Journal of Medical Sciences, 8(B):1144–1151.

Taylor, C. M., Kordas, K., Golding, J., and Emond, A. M. (2017). Effects of low-level prenatal lead exposure on child IQ at 4 and 8 years in a UK birth cohort study. NeuroToxicology, 62:162–169.

Tena, A., Peru, E., Martinetti, L. E., Cano, J. C., Loyola Baltazar, C. D., Wagler, A. E., Skouta, R., and Fenelon, K. (2019). Long-term consequences of early postnatal lead exposure on hippocampal synaptic activity in adult mice. Brain and Behavior, 9(8):e01307. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/brb3.1307.

Tirima, S., Bartrem, C., von Lindern, I., von Braun, M., Lind, D., Anka, S. M., and Abdullahi, A. (2016). Environmental Remediation to Address Childhood Lead Poisoning Epidemic due to Artisanal Gold Mining in Zamfara, Nigeria. Environmental Health Perspectives, 124(9):1471–1478.

UNEP (2020). Update on the Global Status of Legal Limits on Lead in Paint. Technical report.

United Nations, Department of Economic and Social Affairs, Population Division (2019). World Population Prospects 2019, Online Edition. Rev. 1.

United Nations Environment Programme (UNEP) (2014). The Impact of Phasing Out Leaded Petrol in Kenya: Based on blood lead tests carried out by Njoroge Kimani before and after leaded petrol phase-out. Technical report.

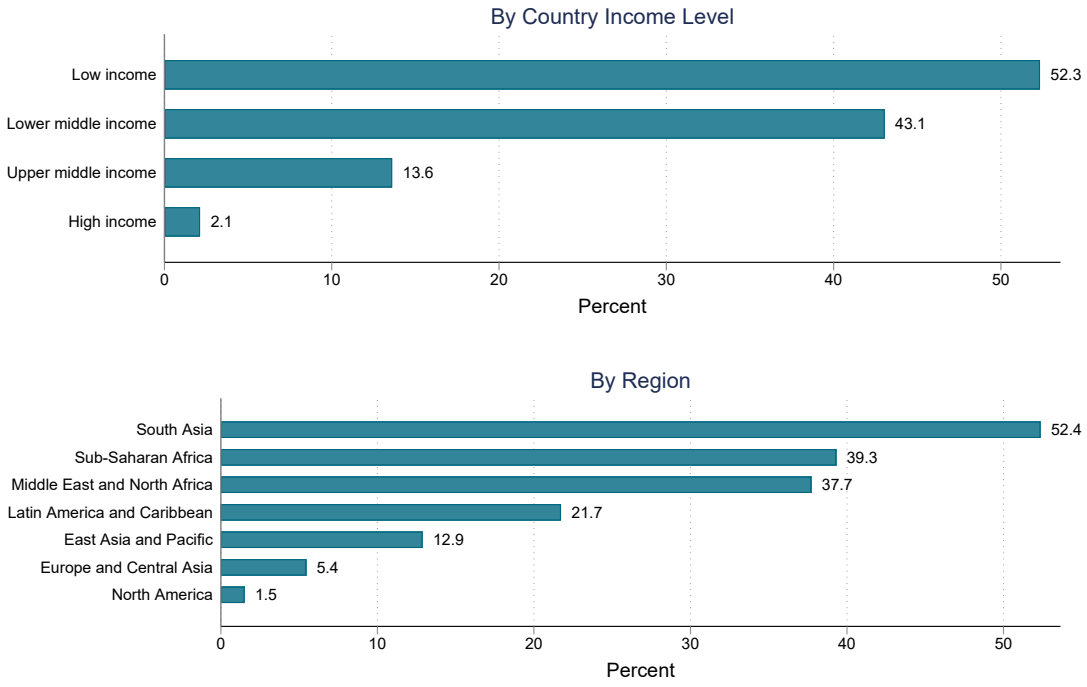
US EPA, O. (2015). Data Tables - Biomonitoring - Lead.

Vega-Dienstmaier, J. M., Salinas-Piélago, J. E., Gutiérrez-Campos, M. d. R., Mandamiento-Ayquipa, R. D., Yara-Hokama, M. d. C., Ponce-Canchihuamán, J., and Castro-Morales, J. (2006). Lead levels and cognitive abilities in Peruvian children. Revista Brasileira De Psiquiatria (Sao Paulo, Brazil: 1999), 28(1):33–39.

Wan, X., Wang, W., Liu, J., and Tong, T. (2014). Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC medical research methodology, 14:135.

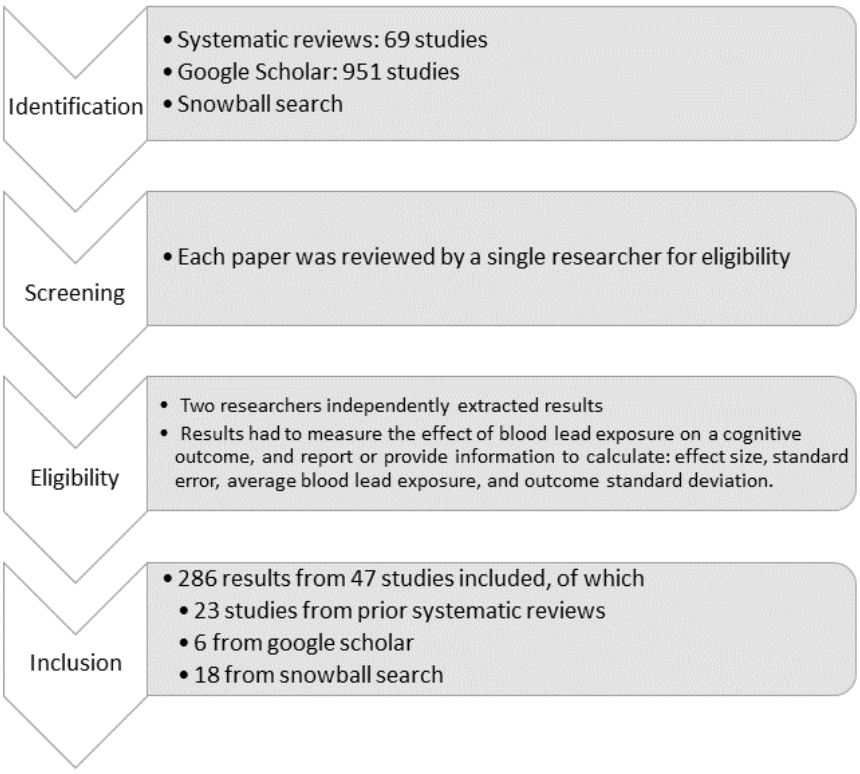
Appendix

Figure A1: Estimates of overall population exposure, by region



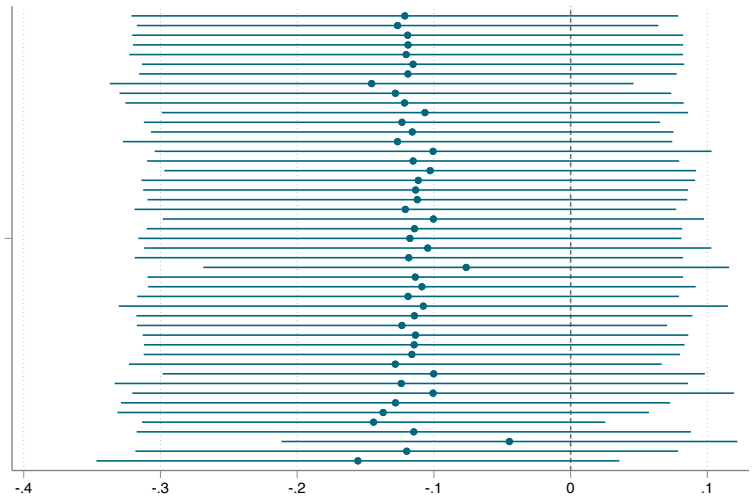
Note: This figure shows the share of children with an elevated blood lead level, defined as exceeding the US CDC reference level of 5 micrograms per deciliter. These estimates are based on our analysis of raw numbers produced by IHME/UNICEF (Rees and Fuller, 2020), combined with population estimates from the UN (United Nations, Department of Economic and Social Affairs, Population Division, 2019).

Figure A2: Systematic review process



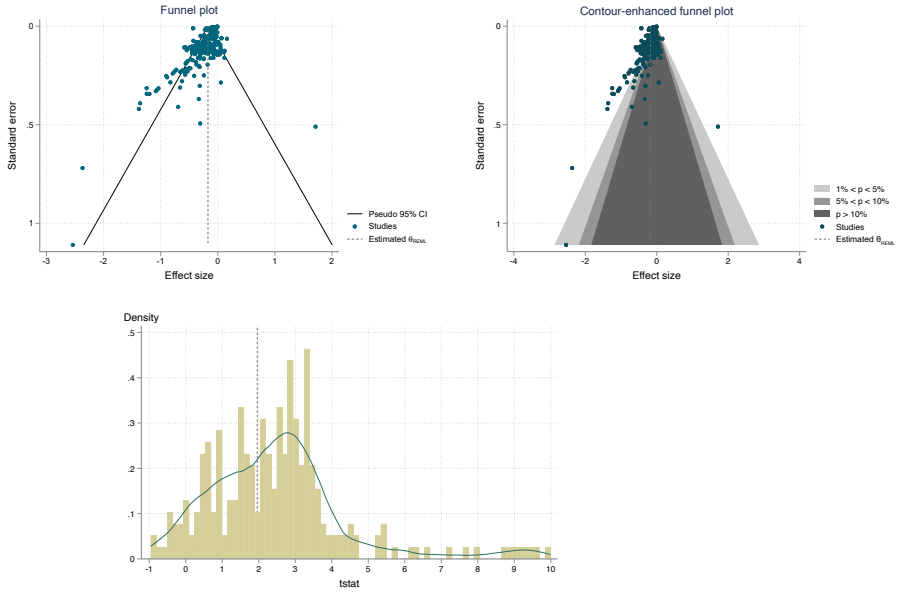
Note: This flowchart shows the process used to identify relevant studies and extract eligible results.

Figure A3: Sensitivity to exclusion of individual studies



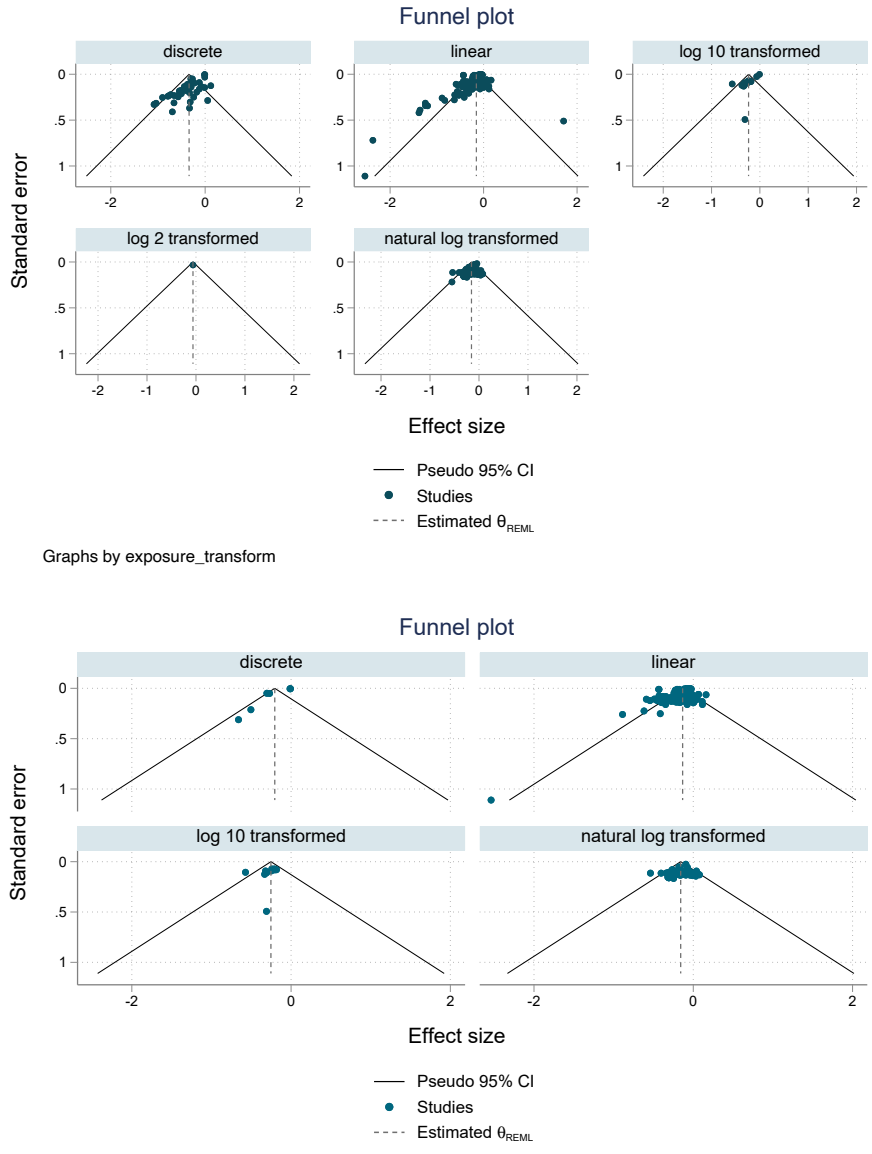
Note: This figure shows the robustness of our primary specification result result to leaving out individual studies. The study which substantially weakens the effect when excluded is Taylor et al. (2017).

Figure A4: Funnel Plot for Potential Publication Bias



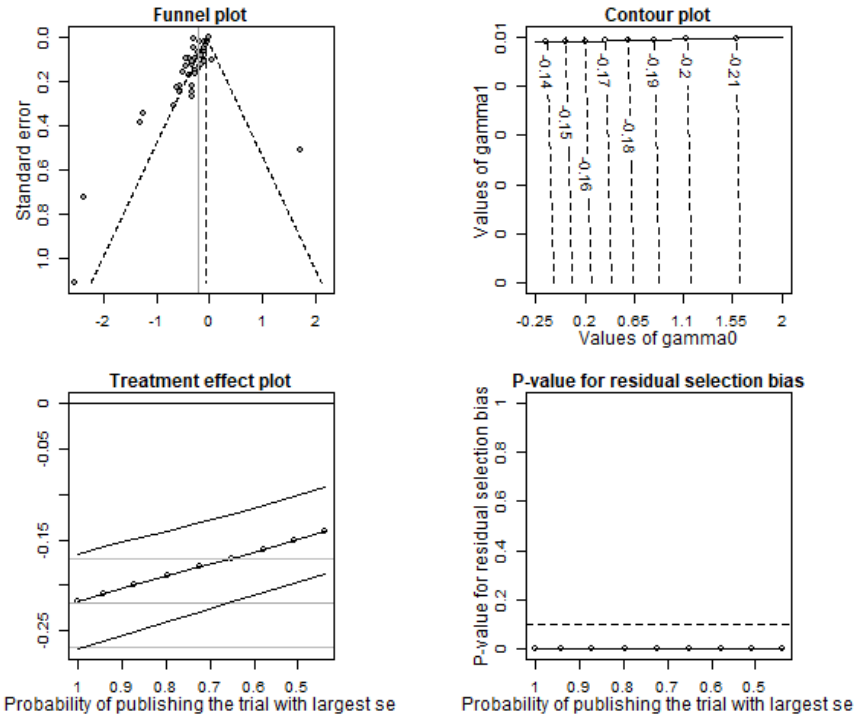
Note: This two top figures shows the distribution of study effect sizes and standard errors. An outlier result (Solon et al., 2008) was excluded to allow for visual inspection. Asymmetry around the vertical dashed line indicates that there may be publication bias present. Contours show that "missing" studies are primarily in the region of statistical insignificance - consistent with a statistical significance filter in publishing leading to this bias. The third figure shows the distribution of z-stats. The spike just right of the vertical dashed line separating significant and non-significant results provides further evidence for selection for significant results.

Figure A5: Funnel Plot by original exposure transformation



Note: This figure shows funnel plots for all results, faceted by the original functional form used for the relationship between blood lead and the outcome. An outlier result (Solon et al., 2008) was excluded to allow for visual inspection. The bottom plot also excludes results which required additional conversions or imputation. The error resulting from the re-expression process has the potential to generate a spurious correlation between effect sizes and standard errors, as both would be biased in the same direction. However, these plots illustrate that this correlation exists even when results are separated by their original functional form, suggesting that small-study effects - which in turn can indicate publication bias- are genuine.

Figure A6: Copas plot



Note: This figure shows Copas method plots where we have collapsed our results, assuming an intra-study correlation of 1. An outlier result (Solon et al., 2008) was excluded to allow for visual inspection. The funnel plot (top left) is equivalent to figure A4. The contour plot (top right) shows results from simulations assuming different selection probabilities for a study with a given precision. The treatment effect plot (bottom left) shows that the estimated treatment effect is lower if we assume these selection probabilities to be lower, but even at low selection probabilities, there is still a substantial negative effect of lead.

Table A1: Studies on association of lead with IQ

Authors	Year	Country code	Country income group
Alvarez-Ortega et al	2017	COL	M
Baghurst et al	1992	USA	H
Bellinger et al	1992	USA	H
Braun et al	2012	MEX	M
Cai et al	2020	CHN	M
Canfield et al	2003	USA	H
Chen et al	2007	USA	H
Chiodo et al	2007	USA	H
Cooney et al	1989	AUS	H
Counter et al	2005	ECU	M
Dantzer et al	2020	USA	H
Desrochers-Couture et al	2018	CAN	H
Dietrich et al	1993	USA	H
Earl et al	2016	AUS	H
Fulton et al	1987	GBR	H
Hong et al	2015	KOR	H
Huang et al	2012	TWN	H
Jusko et al	2008	USA	H
Kamel et al	2003	EGY	M
Kim et al	2009	KOR	H
Kim, Yu, and Lee	2010	KOR	H
Liu et al	2013	CHN	M
Lucchini et al	2012	ITA	H
Lucchini et al	2019	ITA	H
Menezes-Filho et al	2018	BRA	M
Min et al	2009	USA	H
Nwobi et al	2019	NGA	M
Pan et al	2018	CHN	M
Rahman et al	2002	PAK	L
Rasoul et al	2012	EGY	M
Reuben et al	2017	NZL	H
Roy et al	2013	IND	M
Ruebner et al	2019	USA	H
Schnaas et al	2006	MEX	M
Schnaas et al	2000	MEX	M
Solon et al	2008	PHL	M
Surkan et al	2007	USA	H
Taylor et al	2017	GBR	H
Vega-Dienstmaier et al	2006	PER	M
Zailina et al	2011	MYS	M

Note: We include here some three cognitive assessments that are similar but not identical to IQ tests - Vega-Dienstmaier et al. (2006) use the "Graphic Test of Reasoning" and the "Kohs Block Design Test", while Cooney et al. (1989) and Braun et al. (2012) use the McCarthy Scales of Children's Abilities, General Cognitive Index.

Table A2: Studies on association of lead with reading and mathematics scores

Authors	Year	Country code	Country Income Group
Reading			
Aizer et al	2016	USA	H
Bellinger et al	1992	USA	H
Blackowicz et al	2016	USA	H
Chiodo et al	2007	USA	H
Evens et al	2015	USA	H
Fulton et al	1987	GBR	H
Kamel et al	2003	EGY	M
Kim, Yu, and Lee	2010	KOR	H
Kordas et al	2006	MEX	M
Lanphear et al	2000	USA	H
Liu et al	2013	CHN	M
McLaine et al	2013	USA	H
Min et al	2009	USA	H
Shadbegian et al	2019	USA	H
Surkan et al	2007	USA	H
Maths			
Aizer et al	2016	USA	H
Bellinger et al	1992	USA	H
Blackowicz et al	2016	USA	H
Chiodo et al	2007	USA	H
Evens et al	2015	USA	H
Fulton et al	1987	GBR	H
Kamel et al	2003	EGY	M
Kim, Yu, and Lee	2010	KOR	H
Kordas et al	2006	MEX	M
Lanphear et al	2000	USA	H
Liu et al	2013	CHN	M
Min et al	2009	USA	H
Shadbegian et al	2019	USA	H
Surkan et al	2007	USA	H

Table A3: Alternative re-expression algorithms

	Mean BLL	SD BLL	True effect (log)	RB	L	D
NHANES 99-2000	3.060	3.042	-0.261	-0.244	-0.148	-0.105
NHANES 2001-02	2.629	2.205	-0.194	-0.263	-0.159	-0.121
NHANES 2011-12	1.953	1.657	-0.219	-0.247	-0.150	-0.113
Canfield et al 2003	9.015	5.488	-0.489	-0.845	-0.512	-0.427
Vega-Dienstmaier et al 2006	10.330	7.360	-0.672	-2.371	-0.750	-0.522
Kordas et al maths	11.793	6.400	-0.405	-0.573	-0.348	-0.299
Kordas et al reading	11.787	6.393	-0.463	-0.674	-0.409	-0.351
Crump et al 2013	14.607	14.230	-0.119	-0.205	-0.125	-0.089

Note: In this table we assess the performance of three different algorithms used to re-express effects estimated with linear models in terms of log units. We first use microdata to estimate directly the effect of a log unit increase in BLL on standardised test score outcomes (column 2). We then estimate the effect of a linear unit increase in BLL, and apply the algorithms to re-express the effect estimated with a linear model in terms of log units (columns 3-5). RB indicates results from the Rodríguez-Barranco et al. (2017) algorithm, L from the Linakis et al. (2021) algorithm, and D from the Dzierlenga et al. (2020) algorithm. The root mean squared error indicated that the Linakis et al. (2021) algorithm clearly performed best.

Table A4: Alternative adjustments for publication-bias

	(1) None	(2) Egger	(3) Non-linear	(4) Trim-and-fill	(5) P-uniform*
Effect Standard Error		-1.470***			
		(0.290)			
Effect Variance			-0.252		
			(0.188)		
Constant	-0.212***	-0.085***	-0.207***	-0.157***	-0.178***
	(0.024)	(0.029)	(0.023)	(0.026)	(0.031)
N (Studies)	47	47	47	47	47

Note: * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses. This table shows four alternative standard approaches for adjusting for publication bias, applied to our results when unadjusted and averaged by study (column 1) (we don't use the adjusted result, as the trim-and-fill and p-uniform methods would not be applicable): in column 2, the Egger intercept (adjusting for the study standard error) (Egger et al., 1997); in column 3, the PET-PEESE nonlinear intercept (adjusting for the variance) (Stanley and Doucouliagos, 2014), also used in our main specification; in column 4, the trim-and-fill method (attempting to 'fill in' hypothetical results given no publication bias) (Duval and Tweedie, 2000); and in column 5, the p-uniform* method (Aert and Assen, 2018).

Table A5: Bivariate meta-regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Not logarithmic		-0.042						
		(0.050)						
Mean blood lead			-0.008**					
			(0.004)					
No control for parent ed/IQ				-0.173***				
				(0.036)				
No control for family income					-0.123***			
					(0.046)			
Exposure: not average or lag						-0.036		
						(0.049)		
Outcome: Math							0.021	
							(0.055)	
Outcome: Reading							-0.015	
							(0.050)	
High income country								0.035
								(0.057)
Constant	-0.227***	-0.204***	-0.175***	-0.155***	-0.169***	-0.212***	-0.233***	-0.255***
	(0.025)	(0.033)	(0.033)	(0.021)	(0.023)	(0.036)	(0.032)	(0.047)
N (Estimates)	286	286	286	286	286	286	286	286
N (Studies)	47	47	47	47	47	47	47	47

Note: * p < 0.1, ** p < 0.05, *** p < 0.01. Standard errors in parentheses. We use the Hedges et al. (2010) estimator to account for dependence between multiple estimates from the same study. The dependent variable in each case is the standardized effect size of a natural log unit increase in lead exposure on the cognitive outcome.

Table A6: Sensitivity

	(1)	(2)	(3)
Imputed Outcome SD		-0.041	
		(0.066)	
Constant	-0.115	-0.094	-0.159
	(0.100)	(0.110)	(0.115)
N (Estimates)	286	286	212
N (Studies)	47	47	30

Note: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. Standard errors in parentheses. We use the Hedges et al. (2010) estimator to account for dependence between multiple estimates from the same study. This table shows our adjusted meta-regression estimate (column 1), the estimate after additionally including a dummy for whether a population rather than sample standard deviation was used to calculate an effect size (column 2), and the estimate in a sub-sample that excludes those in which various imputations had to be made (column 3). These exclusions include: results in which we impute mean blood lead levels, results in which we calculate the standard error using a reported p-value, and results where the effect size is converted from a reported correlation coefficient.