CENTER FOR GLOBAL DEVELOPMENT

Schools in the Shadow of Toxic Sites: Pollution Proximity in Low- and Middle-Income Countries

Lee Crawfurd

Abstract

How many children in low- and middle-income countries attend school near sources of pollution? We match the locations of 2.6 million schools across 17 countries to 11,301 documented toxic sites and find that 9.7 percent of schools in our sample lie within 5 km of a site. Weighting by enrollment, 12.7 percent of students in the 7 countries with enrollment data attend a school within this distance. These remain lower bounds: the available data capture only a fraction of actual contaminated sites. Proximity is overwhelmingly an urban phenomenon—urban schools are 4 to 28 times more likely than rural schools to be near a site, depending on the country. Within countries, schools in the wealthiest neighborhood quintile are about 14 times more likely to be near a site than schools in the poorest (34.5 percent versus 2.4 percent), reflecting the spatial concentration of industry in wealthier urban areas of LMICs. Where data on school management are available, private schools are also more likely than public schools to be near sites in all eight such countries.

KEYWORDS
pollution, schools, environmental justice, lead, low- and middle-income countries

JEL CODES
Q53, I25, O15, R14

WORKING PAPER 749 · JUNE 2026

Schools in the Shadow of Toxic Sites: Pollution Proximity in Low- and Middle-Income Countries

Lee Crawfurd
Center for Global Development

Lee Crawfurd. 2026. "Schools in the Shadow of Toxic Sites: Pollution Proximity in Low- and Middle-Income Countries." CGD Working Paper 749. Washington, DC: Center for Global Development. https://www.cgdev.org/publication/schools-shadow-toxic-sites-pollution-proximity-low-and-middle-income-countries

I'm grateful to Chris Kinally, Maria Jose Talayero, and Yi Ning Wong for helpful comments, and to Claude Code (Opus 4.7 and 4.8) for research assistance. I take full responsibility for all results and any errors.

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

Center for Global Development. 2026.

1 Introduction

A central question in environmental health is who bears the burden of pollution. In high-income countries, the answer is well-established: decades of research document that exposure falls disproportionately on poorer communities and racial minorities, through a land-market mechanism in which polluting facilities locate where opposition is weakest (Pastor et al., 2002). This pattern extends directly to schools in the United States, where 44% of public schools are within one mile of a government-tracked hazard site (Malik et al., 2026). Exposure to air pollution is disproportionate for students of color (Cheeseman et al., 2022). Whether the same pattern holds in low- and middle-income countries is less clear. Across 16 of 17 countries, schools near documented contaminated sites are in *wealthier* neighborhoods — though the same is true of the general population in most of these countries (Section 4).

We document this pattern using geocoded data on the best available documented contaminated sites in each country—combining Pure Earth's Toxic Sites Identification Program (TSIP, 2,840 sites across our 17 study countries) with national contaminated-site registers in India, Brazil São Paulo, Mexico, and Peru (11,301 sites in total). We overlay this with the locations of 2,601,218 schools. Data for eleven countries comes from national school censuses or government education registers, and for six countries from Overture Maps. 252,809 of these schools (9.7% of all schools) lie within 5 km of a documented site. Enrollment-weighted figures are higher, with 42,783,664 students (12.7% of students in the 7 countries with enrollment data) attending schools in proximity. These figures are lower bounds: the toxic site data captures only a fraction of actual sites (Dowling et al., 2016). Children near such sites face real risks. Doubling blood lead is associated with about a 0.08 SD reduction in test scores (Crawfurd et al., 2024). Mercury, arsenic, cadmium, and organophosphate pesticides cause IQ losses of similar magnitude (1–7 points in heavily exposed children; Appendix Table A1). Causal evidence from six of our study countries shows that pollution exposure lowers test scores or cognitive outcomes: air pollution in India (Balakrishnan and Tsaneva, 2021), Brazil (Carneiro et al., 2021), and Colombia (Villalobos and Blackman, 2025); and lead exposure in Mexico (Litzow et al., 2024), Kenya (Ipapa, 2023), and Indonesia (Berkhout et al., 2025). Similar findings appear in US schools near hazard sites (Persico and Venator, 2021).

We measure *proximity*, not *exposure* or *health impacts*. Proximity is a necessary but not sufficient condition for exposure, which also depends on transport pathways, dose, and individual susceptibility; our estimates identify schools that warrant further environmental investigation, not confirmed cases of harm. The economic stakes are nonetheless large—childhood lead exposure alone costs LMICs nearly $1 trillion annually in lost productivity (Attina and Trasande, 2013)—and children are particularly vulnerable because developing physiology amplifies toxicant effects (Grandjean and Landrigan, 2014).

Despite this, there is almost no systematic information on how close LMIC schools are to known pollution sources. Proximity rates vary widely—from over a quarter of schools within 5 km in Armenia, Mongolia, and Ghana to under 5% in India—reflecting differences in data coverage as much as actual pollution prevalence.

This paper makes three descriptive contributions. First, we provide the first multi-country, multi-pollutant assessment of school proximity to known contaminated sites in LMICs, extending the single-pollutant cross-country analysis of Crawfurd et al. (2025) across 17 countries covering close to half the population of the developing world. Second, proximity is overwhelmingly an urban phenomenon: urban schools are 4 to 28 times more likely than rural schools to be near a documented contaminated site, and schools in the wealthiest within-country quintile are about 14 times more likely than schools in the poorest (34.5% versus 2.4%). This wealth gradient mirrors recent district-level evidence (Behrer and Heft-Neal, 2024; Ascencio et al., 2025) and reverses the US pattern (Malik et al., 2026; Cheeseman et al., 2022)—reflecting how industry and schools concentrate in wealthier urban cores in LMICs rather than the price-driven sorting of high-income settings. Third, in the eight countries where management type is available, private schools are more likely than public schools to be near contaminated sites. We treat this finding as preliminary given the limited country coverage and the well-known heterogeneity of LMIC private schools.

Section 2 describes the contaminated site and school location data. Section 3 outlines the proximity computation and wealth matching. Section 4 reports proximity rates by country, school type, and student wealth quintile. Section 5 discusses mechanisms, limitations, and policy implications.

2 Data

2.1 Polluted sites: the Toxic Sites Identification Program

Data on the location of polluted sites are available from the Toxic Sites Identification Program (TSIP) run by the NGO Pure Earth (Caravanos et al., 2014; Ericson et al., 2013). This is a major global effort to identify the most dangerous polluted sites around the world. There is no single standardized approach to identifying sites, as each country establishes its own priorities and methods. Sources of information include requests or concerns from local authorities or national government agencies, reports of incidents or alerts from affected communities, findings from government agencies or academic research, and news or media coverage. The database includes both active and inactive legacy sites. We use the version of the database downloaded in March 2026 via the contaminatedsites.org API. This is an updated version of the same source used in previous modeling efforts (Ericson et al., 2016; Crawfurd et al., 2025).

We use the full TSIP database including all industry types and pollutants. After dropping sites with missing or imprecise geocoordinates (fewer than 4 decimal places) and sites outside the expected bounding box for their country, our sample contains 4,596 sites across 89 countries. Lead is the most common key pollutant (1,582 sites globally); mercury, arsenic, chromium and pesticides are also represented. Across the 2,840 sites in our 17 study countries, the leading source industries are mining and smelting, battery recycling, manufacturing, and waste handling (Appendix Table A2). Source industries matter for interpretation because contamination radii vary: small-scale used-lead-acid-battery operations and ceramics workshops typically affect a few hundred metres to roughly one kilometre, while primary smelters can disperse airborne lead over 2–5 km. Section 3 discusses our 1 km and 5 km thresholds in light of this dispersion evidence.

Several features of the TSIP source data warrant flagging. Pure Earth describes the Initial Site Screening (ISS) protocol as a rapid hazard-screening tool "not designed as the basis for epidemiological studies" (Pure Earth, 2017), and 1,259 of our 2,840 study-country sites (44%) have an ISS status below "approved." We therefore treat TSIP as a register of sites warranting further investigation, not as confirmed exposure measurements. TSIP conventions also allow several small operations clustered in one village to be recorded as a single "local" site, so the count understates the number of physical pollution sources. TSIP is also not comprehensive: it reflects where Pure Earth has had funding and trained investigators, with sampling targeted at suspected "hotspots," and Dowling et al. (2016) estimate that it captures only around one-seventh to one-ninth of contaminated sites in Ghana. Coverage is concentrated in five countries—India (668 sites), Bangladesh (359), Indonesia (265), the Philippines (228), and Ghana (158)—which together hold 36% of all sites. Finally, TSIP covers point-source heavy-metals contamination but not traffic-related air pollution or agricultural pesticide use.

Our analysis focuses on the 17 countries where TSIP has at least 25 sites and we have either national school census or Overture Maps school-location coverage, yielding 2,840 sites. Table 1 reports the country-by-pollutant distribution.

2.2 Government contaminated site registers

Because TSIP is known to undercount contaminated sites and the magnitude of undercounting varies dramatically across countries, we supplement TSIP with national government contaminated-site registers for the four countries where they exist with publicly available geocoordinates: India's Central Pollution Control Board (CPCB) State-Wise List of Contaminated Sites (105 sites covering heavy-metals contamination at industrial areas, landfills, steel plants, and chemical facilities); the São Paulo state CETESB *Cadastro de Áreas Contaminadas* (Samlani et al., 2024) (1,270 metals-contaminated sites, restricted from the full 6,434-site registry that is dominated by gas-station hydrocarbon spills); Peru's Ministry of Energy and Mines *Pasivos Ambientales Mineros* inventory (6,122 mining liabilities—mine openings, waste dumps, tailings deposits—a major source of heavy-metal contamination); and Mexico's *Registro de Emisiones y Transferencia de Contaminantes* (RETC), which records 964 industrial facilities reporting lead emissions in 2021 across chemicals, food processing, automotive, and metallurgy

Table 1. Documented contaminated sites by country, source, and key pollutant

Country	Lead	Arsenic	Cadmium	Other	Total
Panel A: TSIP (Pure Earth Toxic Sites Identification Program)					
India	321	10	30	307	668
Bangladesh	293	31	1	34	359
Indonesia	158	4	3	100	265
Philippines	44	40	16	128	228
Colombia	28	6	1	127	162
Ghana	95	16	3	44	158
Brazil	48	5	1	101	155
Armenia	23	45	2	62	132
Kenya	74	23	4	19	120
Vietnam	7	1	2	88	98
Peru	25	49	1	18	93
Mongolia	20	31	0	40	91
Mexico	51	25	0	12	88
Argentina	47	6	0	33	86
Kyrgyzstan	6	4	0	44	54
Cambodia	3	6	0	39	48
Kazakhstan	10	2	1	22	35
TSIP subtotal	1,253	304	65	1,218	2,840
Panel B: National government contaminated-site registers[a]					
India (CPCB)	—	—	—	105	105
Brazil (CETESB, São Paulo state)[b]	—	—	—	1,270	1,270
Mexico (RETC)	964	—	—	—	964
Peru (PAM)[c]	—	—	—	6,122	6,122
Register subtotal	964	—	—	7,497	8,461
Combined total (17 countries)	2,217	304	65	8,715	11,301

Note: Number of documented contaminated sites by source, country, and key pollutant. Panel A is TSIP after dropping sites with imprecise coordinates (<4 decimal places) or outside expected country bounding boxes. TSIP sites may be associated with multiple pollutants; the key pollutant is the primary contaminant identified during site investigation. "Other" for TSIP includes mercury, chromium, pesticides, various organic and inorganic compounds, and unspecified contaminants. Mercury (620 sites globally) and chromium (404 sites globally) are present in the full TSIP database but none of the sites in the 17 study countries list them as the key pollutant.

[a] Government registers cover only four of the 17 study countries. Each register has a different scope and pollutant focus, so a unified pollutant decomposition is not meaningful; counts are reported under "Other" except for RETC, which specifically catalogs lead emitters.

[b] CETESB *Cadastro de Áreas Contaminadas* (2024) covers only São Paulo state. The 1,270 count is the metals-contamination subset of the full 6,434-site register, excluding sites where the primary contaminant is hydrocarbons (gas-station spills).

[c] MINEM *Pasivos Ambientales Mineros* (2024) catalogs mining liabilities (mine openings, waste dumps, tailings deposits), which are a primary source of heavy-metal contamination in Peru.

sectors.[1]

For each of these four countries we use the union of TSIP and the national register as the headline measure of "schools near documented heavy-metal or industrial contamination"—a school is recorded as within 5 km of a documented site if it falls within 5 km of either source. The thirteen other countries in our sample use TSIP only. Section 4.6 reports robustness using TSIP only across all 17 countries. The contrast between TSIP and the government registers is large: the CETESB registry identifies eight times more metals-contaminated sites in São Paulo state alone than TSIP catalogs in all of Brazil; Peru's mining liability inventory contains 66 times more sites than TSIP's 93 Peruvian sites; Mexico's RETC identifies 964 active lead-emitting facilities nationwide compared to TSIP's 88. This broadly matches Dowling et al.'s (2016) estimate that TSIP captures only one-seventh to one-ninth of contaminated sites in Ghana, and suggests the undercount may be larger in countries with extensive mining or industrial activity. We acknowledge that the four government registers measure different concepts (CETESB: registered contaminated land; RETC: current lead emitters; PAM: mining liabilities; CPCB: regulatory-inspection-flagged sites) and we treat "schools near documented contamination" inclusively rather than imposing a single uniform site definition.

2.3 National school census data

For eleven countries we use national school census or Education Management Information System (EMIS) data: India (UDISE+, 2022–23), Bangladesh (LGED, 2023), Indonesia (Dapodik, 2024), Kenya (NEMIS, 2023), Mexico (SEP, 2025), Argentina (Mapa Educativo Nacional, 2023), Brazil (INEP Censo Escolar, 2023), Philippines (DepEd, 2023), Colombia (SISE, 2024), Peru (MINEDU Padrón, 2024), and Ghana (ESRAG EMIS, 2020/21). Sources, vintages, and available variables (management type, urban/rural classification, enrollment, school level, gender) are documented in Appendix Table A3; school-type breakdowns by country are reported in Tables A9 and A8. Enrollment data are available for seven countries (Brazil, Ghana, India, Indonesia, Kenya, Peru, Philippines); urban/rural classifications for seven (Argentina, Brazil, Colombia, Ghana, India, Mexico, Peru); and management type for eight (Argentina, Brazil, Ghana, India, Indonesia, Kenya, Mexico, Peru), with Colombia public-only.

The Ghana EMIS requires a geocoding step: it records schools by GhanaPostGPS digital address codes rather than coordinates. Of the 34,464 schools, we resolve coordinates for 22,581 (65%) via the GhanaPostGPS API. Non-geocoded schools are disproportionately rural (34.3% urban versus 49.6% urban among geocoded schools, concentrated in northern Savanna, Bono East, and North East regions); management-type composition is similar across the two groups. The urban skew in the geocoded sample means Ghana's headline proximity rate (26.0%) likely overstates the rate from a

[1] Sources: https://cpcb.nic.in/uploads/hwmd/StateWise-list-of-ContaminatedSites.pdf; Samlani et al. (2024); https://geocatmin.ingemmet.gob.pe/geocatmin/; https://www.gob.mx/semarnat/acciones-y-programas/registro-de-emisiones-y-transferencia-de-contaminantes-retc. RETC facility locations are geocoded from postal codes, which may be offset from actual facility locations by up to several kilometres, particularly in rural areas. The SEMARNAT Inventory of Contaminated Sites was also examined but is dominated by hydrocarbon spills (91%) with only 39 metals-contaminated sites; Indonesia's PROPER industrial pollution ratings cover 2,000–3,000 facilities but publish ratings only at the regency level without precise facility coordinates.

complete census.

2.4 School locations: Overture Maps

For six countries where national school census data are not publicly available with geocoordinates, school locations come from the Overture Maps Foundation, which aggregates and standardizes geospatial data from OpenStreetMap and other open sources. We download all places classified under school-related categories (school, education, elementary school, high school, middle school, and preschool) from the February 2026 release. Table A4 in the Appendix compares Overture school counts to official statistics for each country. We include only countries where Overture coverage exceeds about 80% of the official school count: Armenia ($1{,}379$ schools), Cambodia ($17{,}148$), Kazakhstan ($6{,}365$), Kyrgyzstan ($1{,}935$), Mongolia (795), and Vietnam ($67{,}087$). Table A5 in the Appendix lists all countries with 10 or more TSIP sites that are not included in the main analysis, together with the reason for exclusion.

The Overture data have several limitations. First, completeness varies across countries and within countries, with rural and remote areas typically less well-mapped. Second, the data lack information on school characteristics such as enrollment, grade levels served, or demographic composition. Third, classification of school types is inconsistent across countries. To assess comparability, we computed proximity rates using both Overture and EMIS data for all eleven EMIS countries (Table A6 in the Appendix). In 10 of 11 countries, the Overture-derived proximity rate exceeds the EMIS rate, with a mean difference of +10.7 pp; Argentina is the lone exception (Overture 11.7% versus EMIS 11.9%, -0.1 pp). The cross-country ranking is broadly preserved (Pearson $r = 0.87$, $p < 0.001$). The systematic upward bias reflects the urban concentration of OSM mapping: where Overture coverage is sparse, the schools that get mapped tend to be urban, and urban schools are mechanically closer to TSIP sites. Argentina escapes this because its school sample is overwhelmingly urban regardless of source—Argentina is roughly 93% urbanised. The bias is largest where Overture under-covers rural areas: Ghana (Overture captures 12% of the EMIS count, +29 pp bias), Kenya (20%, +27 pp), and India (25%, +15 pp).

Total Overture coverage alone therefore does not bound the bias—what matters is the urban–rural composition of the schools that are mapped. Argentina sits at 74% Overture coverage (below the 80% threshold we use for the Overture-only countries) and shows near-zero bias, while Bangladesh at 49% coverage shows a +11 pp bias. The six Overture-only countries in our sample (Armenia, Cambodia, Kazakhstan, Kyrgyzstan, Mongolia, Vietnam) span urbanisation rates from 24% (Cambodia) to 69% (Mongolia), so the bias likely varies substantially across them. We treat their proximity rates as potentially upward-biased point estimates rather than strict upper bounds, with the bias likely largest for the less urbanised countries in the group.

2.5 Neighborhood wealth data

To examine the relationship between neighborhood wealth and TSIP proximity, we use the Meta Relative Wealth Index (RWI) of Chi et al. (2022), which provides gridded estimates of relative household wealth at 2.4 km resolution for all low- and middle-income countries, derived from machine learning on satellite imagery and connectivity data. RWI data are available for all 17 countries.

2.6 Satellite $PM_{2.5}$ data

We use annual mean $PM_{2.5}$ for 2023 from the satellite-derived global surface $PM_{2.5}$ dataset of van Donkelaar et al. (2021), at 0.05° (≈5 km) resolution.[2]

2.7 Sample construction

Table A7 in the Appendix summarizes the number of observations at each stage of the data processing pipeline, from the raw TSIP download (5,016 sites) through coordinate cleaning (4,596 sites) and country selection (2,840 sites across 17 countries), and from the two school data sources (2,628,805 EMIS schools and 94,709 Overture schools) to the final analysis sample of 2,723,514 schools. Figure 1 shows the distribution of distance to the nearest TSIP site across the about 900,000 Overture-mapped schools in all 17 countries; the country-level analysis in Table 2 uses EMIS data for the eleven EMIS countries and Overture data only for the remaining six.

3 Methods

3.1 Proximity computation

For each school, we compute the distance to the nearest TSIP site using a two-step procedure. First, we project all school and site coordinates onto an approximate Cartesian plane using a local equirectangular projection and construct a k-d tree over TSIP site locations. We query this tree to identify candidate nearest TSIP sites for each school, querying $k = 3$ candidates to guard against the approximation error inherent in using a flat projection for countries spanning large latitude ranges. Second, we compute the exact great-circle (geodesic) distance between each school and all three candidate sites using the haversine formula, selecting the site with the minimum haversine distance as the true nearest TSIP site. This measures straight-line distance across the Earth's surface, without accounting for terrain, elevation, or road networks. Geodesic distance is preferable to road distance for assessing pollution exposure potential because contaminants disperse through air, water, and soil

[2] Specifically, version V5GL0502, product GWRPM25c_0p05. The van Donkelaar group also produces a higher-resolution 0.01° product; we use the 0.05° version to reduce memory requirements and because 5 km resolution is well-matched to the spatial scale of our proximity analysis.

rather than along transportation networks. At the 1–5 km thresholds we use, differences between geodesic distance and terrain-adjusted measures are small relative to the inherent imprecision of using proximity as a proxy for exposure. We construct binary indicators for whether each school falls within 1 and 5 kilometers of a TSIP site.

3.2 Distance thresholds

The choice of distance thresholds—1 and 5 km—is motivated by evidence on pollution dispersion from the types of sources that dominate the TSIP database. For lead contamination from battery recycling sites, which account for roughly 20% of our sample (see Section 2.1), soil contamination is typically most severe within a few hundred meters of the source and can remain elevated further out, as documented at individual sites in Vietnam (Ericson et al., 2018); legacy heavy-metal contamination persists for decades after operations cease (Mielke et al., 2016). However, Crawfurd et al. (2025) document much wider dispersion from used lead-acid battery (ULAB) recycling sites, with elevated blood lead levels detected in populations living several kilometers from active operations, suggesting that the conventional "hotspot" framing may understate the affected radius. Airborne lead and particulate matter from smelting and industrial operations can disperse further, with elevated concentrations detected at distances of 2–5 km depending on stack height, meteorology, and terrain. Our two thresholds bracket these ranges: 1 km captures the zone of highest soil contamination around point sources and 5 km captures the broader dispersion zone relevant for airborne pollutants. These thresholds align with US environmental health studies of school proximity to hazards (Malik et al., 2026; Cheeseman et al., 2022).

We emphasize that proximity is a crude proxy for actual exposure. The health consequences depend on the type of contaminant (soil-borne versus airborne versus waterborne), site activity and remediation status, local meteorology and hydrology, and school building characteristics. A school 1 km from an active battery recycling operation may face considerable lead exposure through soil and dust pathways, while a school the same distance from a remediated mining site may face negligible risk. Without site-specific exposure assessment, we cannot distinguish these cases.

3.3 Neighborhood wealth matching

We match each school to the nearest RWI grid cell (Section 2.5) using a KD-tree, excluding matches beyond 5 km. For the enrollment-weighted quintile analysis, quintiles are defined so that each represents about 20% of enrolled students within each country, sorted by neighborhood RWI. We also standardise RWI to a within-country z-score for use as a continuous control in the regression specification below.

3.4 Satellite $PM_{2.5}$ measurement

For each school, we extract the annual mean $PM_{2.5}$ value from the satellite-derived dataset described in Section 2.6 by matching the school's coordinates to the underlying $0.01°$ raster cell. We report the share of schools whose $PM_{2.5}$ value exceeds the WHO 2021 annual guideline (5 μg/m^3), the WHO 2005 interim target (15 μg/m^3), and the US EPA non-attainment standard (35 μg/m^3), and document the $PM_{2.5}$ gradient by wealth and urban location. Ambient $PM_{2.5}$ captures combustion-driven air pollution from traffic, household fuels, biomass burning, and regional dispersion—a different pollution phenomenon from the point-source heavy-metal contamination cataloged by TSIP and the government registers. We therefore present $PM_{2.5}$ as a complementary measure of environmental quality.

3.5 Regression specification

We estimate OLS regressions of the form:

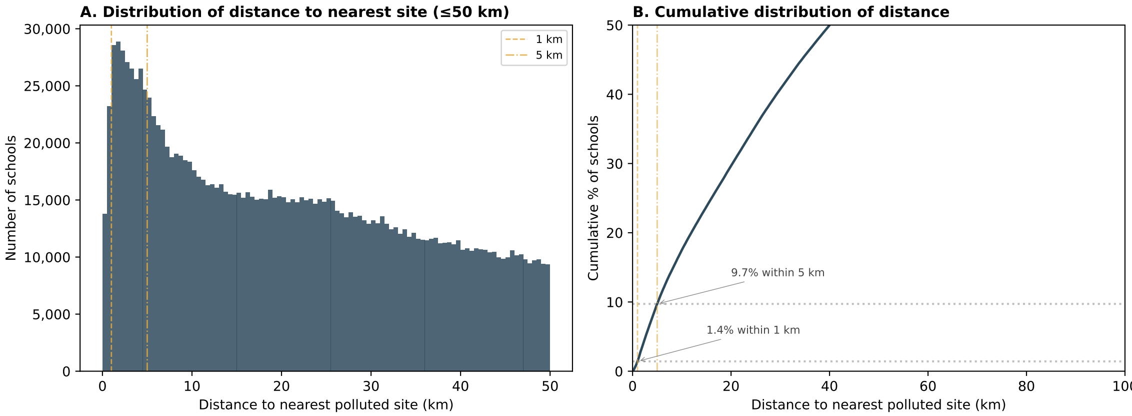
$$\text{Within5km}_{ic} = \beta\, X_{ic} + \gamma_1\, \text{Urban}_{ic} + \gamma_2\, \text{RWI}_{ic} + \alpha_{c(s)} + \varepsilon_{ic}$$

where Within5km_{ic} is an indicator for school i in country c being within 5 km of the nearest documented contaminated site (best-available source per country: TSIP combined with national government registers for India, Brazil São Paulo, Mexico, and Peru; TSIP only elsewhere); X_{ic} is the school characteristic of interest; Urban_{ic} is an indicator for urban location; RWI_{ic} is standardised within-country neighborhood RWI; and $\alpha_{c(s)}$ are fixed effects. Column (1) of Table 4 reports the baseline specification with country fixed effects only (α_c) and no controls. Column (2) replaces country with sub-national (state, province, or region) fixed effects (α_s) and adds the urban and RWI controls; coefficients on the controls are not displayed. We use a linear probability model for ease of interpretation. Standard errors are heteroskedasticity-robust.

4 Results

Our country-level analysis is organised around two data tiers (Table 2). Panel A reports proximity rates for the eleven countries with national school census or government education-register data—India, Bangladesh, Indonesia, Kenya, Mexico, Argentina, Philippines, Brazil, Colombia, Peru, and Ghana—using the full school population for each country. Panel B reports proximity rates for six additional countries (Armenia, Cambodia, Kazakhstan, Kyrgyzstan, Mongolia, Vietnam) for which only Overture Maps school locations are available. As documented in Section 2.3, OSM tends to undersample rural schools in countries that are not heavily urbanised, which produces an upward bias in proximity rates; the Panel B rates are therefore best read as potentially upward-biased point

Figure 1. Distribution of school distance to nearest documented contaminated site



Note: Both panels use the headline analysis sample of 2,601,218 schools across 17 countries (national school censuses where available; Overture Maps for the remaining six countries). Distance is to the nearest documented contaminated site, using the best-available combined source per country (TSIP combined with the national government register for India, Brazil São Paulo, Mexico, and Peru; TSIP only for the other 13 countries). Panel A shows the distance distribution for schools within 50 km, with vertical lines at the 1 km and 5 km thresholds. Panel B shows the cumulative distribution function. Appendix Figure 5 maps schools and documented contaminated sites separately for each country.

estimates rather than precise measures. We restrict the within-country breakdowns in Sections 4.3–4.7 to the Panel A countries, where comparable population coverage is available.

Across all 17 countries (using best-available contamination data per country: TSIP combined with national government registers for India, Brazil São Paulo, Mexico, and Peru; TSIP only elsewhere), 252,809 of 2,601,218 schools (9.7%) lie within 5 km of a documented contaminated site.[3] For cross-country comparability we also report TSIP-only rates, which give a substantially lower aggregate of 6.6% — a difference driven by the fact that government registers in Mexico, Peru, and São Paulo state capture far more sites than TSIP cataloged in the same areas. Proximity rates vary across countries. The most exposed countries are Brazil São Paulo (66.4% combined), Armenia (64.6% TSIP only) and Mongolia (58.4%) — both Overture-mapped countries where the OSM urban bias inflates the rate — and Mexico (26.0% combined) and Ghana (26.0% TSIP). The least exposed in our sample are India (4.7% combined) and Kazakhstan (6.5%). Cross-country variation reflects differences in source coverage (TSIP investigation intensity, register comprehensiveness) as much as actual differences in pollution prevalence: the national-Brazil TSIP-only rate (5.6%), for example, reflects only 155 cataloged sites relative to Brazil's vast geography, while CETESB's 1,270 metals-contaminated sites in São Paulo state alone put 66% of São Paulo schools within 5 km of a documented contaminated site. The within-country comparisons by school type in the following sections hold TSIP coverage constant within each country and are therefore more informative than cross-country comparisons of

[3]These figures pool active and legacy sites; we cannot distinguish schools near currently active contamination from those near sites that have been remediated or are no longer operating.

raw proximity rates. Figure 5 in the Appendix maps schools and TSIP sites separately for each country.

4.1 Enrollment-weighted proximity

School-count proximity rates understate the number of children potentially affected, because schools near contaminated sites tend to be larger urban institutions. The last two columns of Table 2 report the number of children (in millions) attending schools within 1 km and 5 km of a documented site, for the 7 EMIS countries with enrollment data. In every country the share of students near a site exceeds the share of schools, often by a lot: the Philippines (27.1% of students versus 8.9% of schools, threefold, driven by large schools in the National Capital Region), Indonesia (11.5% versus 8.0%), and Ghana (29.4% versus 26.0%) show the largest gaps.[4] Across all 7 countries with enrollment data, 42,783,664 students (12.7% of total enrollment) attend schools within 5 km of a documented contaminated site.

4.2 Proximity to lead and heavy-metals sites

Because lead is the most extensively documented developmental neurotoxicant and accounts for the largest share of TSIP sites, we separately examine proximity to sites where lead is the dominant contaminant. Table 3 repeats Table 2 for lead / heavy-metals sites only, so the lead-specific rates can be read directly against the all-contaminant rates in Table 2. For the 13 countries without a national government register, the lead column is restricted to TSIP sites where lead is the key pollutant (1,253 sites, 44% of TSIP sites in our study countries); for the four countries with a register (India, Brazil São Paulo, Mexico, Peru), it adds the register, which catalogs heavy-metals contamination predominantly involving lead—RETC (Mexico) is strictly lead-specific, while CPCB (India), CETESB metals (Brazil São Paulo), and PAM (Peru) record heavy-metals sites in which lead is the most prevalent contaminant. Across all 17 countries, 7.5% of schools are within 5 km of a lead / heavy-metals site, compared to 9.7% for all sites. Proximity rates to lead-relevant sites are highest in countries where lead sites dominate the inventory: in Bangladesh, where lead accounts for 82% of TSIP sites (primarily from battery recycling), the lead rate (10.6%) is close to the all-site rate (11.3%); in Mexico, where RETC is lead-specific and contributes most of the register sites, the gap between lead (25.7%) and all (26.0%) is tiny. In Ghana (21.1% versus 26.0%) lead also accounts for most of the observed proximity. By contrast, in countries with more diverse pollution profiles—such as Armenia (3.9% lead versus 64.6% all) and Mongolia (28.2% versus 58.4%)—proximity to lead-specific sites is much lower than proximity to all sites.

4.3 Proximity by school characteristics

The within-country variation in proximity is dominated by the urban–rural divide, with capital cities driving the urban concentration further still (Figure 2). Pooled across the seven countries with

[4] Ghana proximity rates are likely overstated due to the urban concentration of schools with resolvable GhanaPostGPS codes; see Section 2.3.

Table 2. School proximity to documented contaminated sites (all contaminants)

				Within 1 km		Within 5 km		Children (millions)	
Country	Schools	Enroll. (M)	Sites	Schools	%	Schools	%	<1 km	<5 km
Panel A: National school census (EMIS) and government register countries									
Argentina	64,638	—	86	564	0.9	7,660	11.9	—	—
Bangladesh	78,129	—	359	887	1.1	8,819	11.3	—	—
Brazil (national)	150,131	42.6	155	831	0.6	8,460	5.6	0.3	3.2
Brazil (São Paulo)†	27,835	9.7	155 / 1,270	6,097	21.9	18,493	66.4	2.4	7.2
Colombia	56,650	—	162	443	0.8	4,943	8.7	—	—
Ghana	22,581	6.7	158	1,094	4.8	5,861	26.0	0.4	2.0
India†	1,372,195	237.8	668 / 105	7,911	0.6	65,103	4.7	2.8	21.8
Indonesia	396,280	56.4	265	2,094	0.5	31,680	8.0	0.4	6.5
Kenya	26,197	8.3	120	576	2.2	3,732	14.2	0.2	1.2
Mexico†	270,228	—	88 / 964	13,453	5.0	70,139	26.0	—	—
Peru†	175,754	9.4	93 / 6,122	2,549	1.5	25,547	14.5	0.2	1.7
Philippines	16,022	9.3	228	271	1.7	1,432	8.9	0.5	2.5
Panel A subtotal	2,506,509	337.5		35,939	1.4	243,409	9.7	6.9	42.8
Panel B: Overture Maps / OpenStreetMap countries (upper bound; see Section 2.3)									
Armenia	1,379	—	132	184	13.3	891	64.6	—	—
Cambodia	17,148	—	48	257	1.5	1,531	8.9	—	—
Kazakhstan	6,365	—	35	44	0.7	411	6.5	—	—
Kyrgyzstan	1,935	—	54	3	0.2	189	9.8	—	—
Mongolia	795	—	91	48	6.0	464	58.4	—	—
Vietnam	67,087	—	98	507	0.8	5,914	8.8	—	—
Panel B subtotal	94,709	—		1,043	1.1	9,400	9.9	—	—
Total	2,601,218	337.5		36,982	1.4	252,809	**9.7**	6.9	**42.8**

† Best-available contaminated-site source: TSIP combined with the national government register (India CPCB; Brazil CETESB *Cadastro de Áreas Contaminadas* metals subset, São Paulo state only; Mexico RETC lead emitters; Peru *Pasivos Ambientales Mineros*). For these four countries, "Sites" reports TSIP sites / government-register sites; "Within X km" columns reflect proximity to either source. Appendix Table A17 reports a strictly-comparable TSIP-only specification across all 17 countries.

Brazil appears in two rows: "Brazil (national)" uses TSIP only across all states (CETESB covers only São Paulo state), while "Brazil (São Paulo)" restricts to São Paulo schools and reports the best-available combined rate. To avoid double-counting, the panel subtotal and overall total exclude the Brazil (national) row.

Distance is haversine distance from each school to the nearest contaminated site. Panel A sources: India (UDISE+ 2022–23), Bangladesh (LGED 2023), Indonesia (Dapodik 2024), Kenya (NEMIS 2023), Mexico (SEP Catálogo 2025), Argentina (Mapa Educativo Nacional 2023), Philippines (DepEd 2023), Brazil (INEP Censo Escolar 2023), Colombia (SISE 2024), Peru (MINEDU Padrón 2024), and Ghana (EMIS 2020/21, geocoded via GhanaPostGPS). Panel B uses Overture Maps / OpenStreetMap school locations (February 2026), which disproportionately map urban schools and therefore yield proximity rates that are upper bounds (see Section 2.3). Enrollment is reported in millions of children. The "Children (millions)" columns give the number of enrolled children (in millions) attending schools within 1 km and within 5 km of a documented site, available only for the countries that record enrollment; the Panel A subtotal and overall total sum these counts across those countries. Countries with insufficient Overture school coverage (Georgia, Nepal, Pakistan, Senegal) are excluded (see Table A4). TSIP sites from the March 2026 database, excluding sites with missing or imprecise coordinates.

Table 3. School proximity to documented contaminated sites (lead / heavy-metals only)

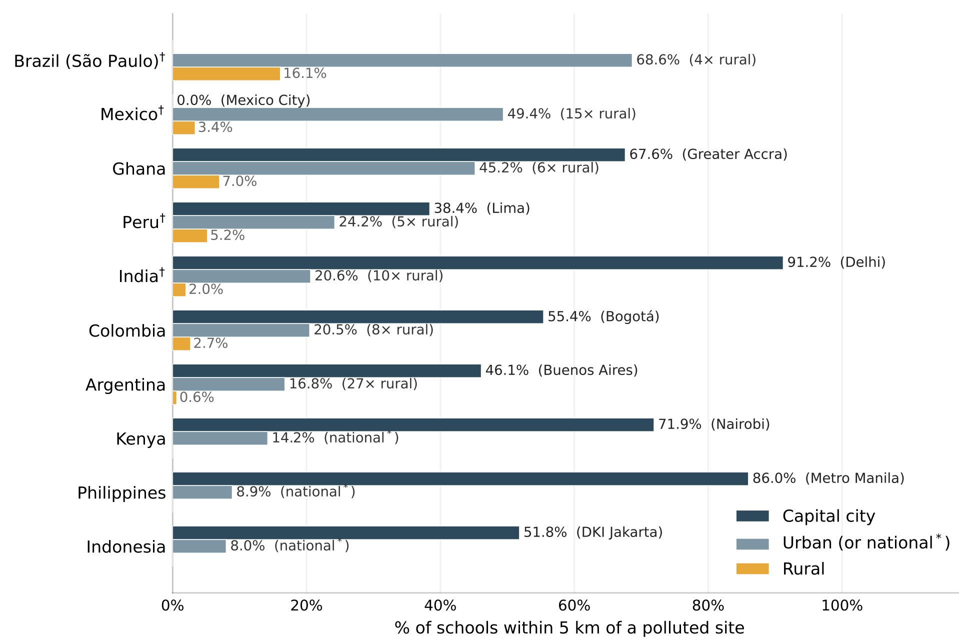
				Within 1 km		Within 5 km		Children (millions)	
Country	Schools	Enroll. (M)	Lead sites	Schools	%	Schools	%	<1 km	<5 km
Panel A: National school census (EMIS) and government register countries									
Argentina	64,638	—	47	313	0.5	4,644	7.2	—	—
Bangladesh	78,129	—	293	748	1.0	8,266	10.6	—	—
Brazil (national)	150,131	42.6	48	261	0.2	3,917	2.6	0.1	1.5
Brazil (São Paulo)†	27,835	9.7	48 / 1,270	5,966	21.4	18,461	66.3	2.4	7.2
Colombia	56,650	—	28	196	0.3	2,792	4.9	—	—
Ghana	22,581	6.7	95	660	2.9	4,770	21.1	0.2	1.6
India†	1,372,195	237.8	321 / 105	4,774	0.3	39,101	2.8	1.8	13.6
Indonesia	396,280	56.4	158	1,567	0.4	24,058	6.1	0.3	4.9
Kenya	26,197	8.3	74	499	1.9	2,943	11.2	0.2	0.9
Mexico†	270,228	—	51 / 964	13,239	4.9	69,433	25.7	—	—
Peru†	175,754	9.4	25 / 6,122	1,397	0.8	18,619	10.6	0.1	1.2
Philippines	16,022	9.3	44	114	0.7	954	6.0	0.2	2.1
Panel A subtotal	2,506,509	337.5		29,473	1.2	194,041	7.7	5.2	31.5
Panel B: Overture Maps / OpenStreetMap countries (upper bound; see Section 2.3)									
Armenia	1,379	—	23	1	0.1	54	3.9	—	—
Cambodia	17,148	—	3	96	0.6	971	5.7	—	—
Kazakhstan	6,365	—	10	0	0.0	39	0.6	—	—
Kyrgyzstan	1,935	—	6	3	0.2	154	8.0	—	—
Mongolia	795	—	20	4	0.5	224	28.2	—	—
Vietnam	67,087	—	7	27	0.0	565	0.8	—	—
Panel B subtotal	94,709	—		131	0.1	2,007	2.1	—	—
Total	2,601,218	337.5		29,604	1.1	196,048	**7.5**	5.2	**31.5**

† Uses the best-available combined lead source (TSIP lead sites combined with the national government register: India CPCB; Brazil São Paulo CETESB metals; Mexico RETC lead emitters; Peru PAM mining liabilities). "Lead sites" reports TSIP-lead sites / government-register sites for these countries.

Note: This table repeats Table 2 restricted to lead / heavy-metals contamination. Proximity is to (a) TSIP sites whose key pollutant is lead and (b) for the four † countries, the national heavy-metals register, in which lead is the predominant contaminant (RETC is strictly lead-specific). For the other 13 countries the columns reflect TSIP-lead sites only. Enrollment is in millions of children; the "Children (millions)" columns give the number of enrolled children attending schools within 1 km and within 5 km of a lead site, available only where enrollment is recorded. Brazil appears in two rows; the subtotal and total exclude Brazil (national) to avoid double-counting São Paulo. Distance is haversine distance. TSIP sites from the March 2026 database (contaminatedsites.org).

school-level urban/rural classifications, 31.1% of urban schools are within 5 km of a documented contaminated site compared to 2.4% of rural schools—about a thirteen-to-one ratio. The gap is most extreme in Argentina (28:1) and Mexico (15:1), smaller where government registers bring rural sites into view: Brazil São Paulo (4:1, with CETESB blanketing the state) and Peru (5:1, with PAM mining liabilities in highland regions). Capital cities concentrate exposure dramatically: nine of the ten capitals shown in Figure 2 have proximity rates well above their national averages, led by NCT Delhi (91%), Metro Manila (86%), Nairobi (72%), Greater Accra (68%), Bogotá (55%), DKI Jakarta (52%), Buenos Aires (46%), and Lima Metropolitana (38%). Mexico City is the striking exception—the cataloged contaminated sites in Mexico cluster in industrial corridors in Estado de México, Hidalgo, and the Bajío, so essentially no school within CDMX sits within 5 km of one.

Figure 2. School proximity to documented contaminated sites, by country: capital city, urban, and rural



Note: Percentage of schools within 5 km of the nearest documented contaminated site, separately for capital city, urban, and rural schools, across ten EMIS countries. For the seven countries with school-level urban/rural indicators (Argentina, Brazil São Paulo, Colombia, Ghana, India, Mexico, Peru), the three bars show capital city, all urban, and rural rates, with multipliers giving the urban-to-rural ratio; Brazil São Paulo has no separate capital bar because the panel already restricts to São Paulo state. For the remaining three countries (Indonesia, Kenya, Philippines), where the EMIS files do not include a school-level urban/rural classification, only two bars are shown—the capital-city rate and the country-wide rate (marked *). Capital cities: Buenos Aires (Argentina), Bogotá (Colombia), Greater Accra (Ghana), NCT Delhi (India), Mexico City (Mexico), Lima Metropolitana (Peru), DKI Jakarta (Indonesia), Nairobi (Kenya), Metro Manila (Philippines). † uses the best-available combined source (TSIP combined with the national government register for India CPCB, Brazil São Paulo CETESB, Mexico RETC, and Peru PAM). Countries sorted by overall urban / national rate (ascending). See Appendix Table A9 Panel A and Appendix Table A10 for underlying counts.

The urban–rural divide is the proximate explanation for most of the cross-cutting patterns we document for other school characteristics. Table 4 reports OLS regressions of the within-5 km indicator on five school characteristics: each row is a separate regression with country fixed effects (column 1) or sub-national fixed effects plus urban and standardized RWI controls (column 2). Per-country descriptive rates are in Appendix Tables A9 and A8.

Private schools (eight countries with management-type data) are 11.6 percentage points more likely than public schools to be within 5 km of a documented contaminated site (17.0% versus 6.9%, pooled). Magnitudes range from 7 pp gaps in India and Argentina to over 20 pp in Brazil São Paulo and Mexico, where government registers add sites in urban industrial corridors. Once we partial out sub-national fixed effects, urban location, and neighborhood wealth, the private coefficient shrinks to about 2.5 pp—small but still highly significant. Restricting to urban schools narrows but does not eliminate the gap (Appendix Table A11), and it is robust to enrollment weighting (Appendix Table A19). LMIC private schools are heterogeneous—spanning elite urban institutions, low-fee peri-urban schools (Tooley et al., 2007; Andrabi et al., 2008; Crawfurd et al., 2024), and religious schools—and our data do not separate these tiers; the RWI gradient (reported below in §4.6) is cleaner in this respect because it identifies the neighborhood the school sits in, not the families it serves.

A further test of the wealth-and-minority sorting story is possible for India, where we link UDISE schools to Census 2011 district-level Scheduled Caste (SC) and Muslim population shares. Asher et al. (2026) document that Indian government schools are located further from SC and Muslim neighbourhoods than private schools, implying minority communities face a double disadvantage in school access. If the same dynamic operates for pollution exposure, the private–public proximity gap should widen where SC or Muslim shares are higher. We find the opposite: the gap is, if anything, slightly smaller in high-SC and high-Muslim districts (interaction coefficients -0.25 and -0.05, both $p < 0.001$, across $1{,}046{,}675$ schools in 477 districts). The sorting of private schools toward industrial urban areas in India operates primarily through economic geography rather than through residential segregation by caste or religion.

The remaining four characteristics follow a similar pattern of substantial baseline gaps that compress under the column-2 controls. *Single-sex schools* (India, Kenya, Peru) are 6 pp closer to documented sites than co-educational schools at baseline; the gap disappears once urban concentration is partialled out. *Special education schools* (Argentina, Colombia, Kenya, Mexico, Peru) show a 6 pp raw gap that flips to -4 pp under controls. *Indigenous and minority schools*—Argentine intercultural/bilingual, Indian Madrasas, Mexican indigenous schools—are 12 pp *further* from documented sites at baseline, reflecting their rural and remote locations; the gap persists at -4 pp under controls, with Indian Madrasas a partial exception because of their urban concentration. *Secondary schools* (Colombia, India, Indonesia, Mexico, Peru) are 4 pp closer to documented sites than primary schools at baseline; the gap vanishes once urban location is partialled out. The secondary–school result is the most policy–relevant of these, given causal evidence that pollution reduces cognitive function and test performance in several of our study countries—India (Balakrishnan and Tsaneva, 2021), Brazil (Carneiro et al., 2021), Colombia (Villalobos and Blackman, 2025)—and elsewhere (Ebenstein et al., 2016; Heissel et al., 2022; Persico and Venator, 2021).

Table 4. OLS regressions of school proximity on school characteristics

	(1) Baseline	(2) + Controls	Countries	Observations
Private	0.1321^{***}	0.0252^{***}	6	1,913,176
	(0.0005)	(0.0005)		
Secondary	0.0386^{***}	-0.0074^{***}	4	1,326,206
	(0.0005)	(0.0005)		
Single-sex	0.0605^{***}	-0.0056^{***}	2	1,547,949
	(0.0012)	(0.0011)		
Special education	0.0607^{***}	-0.0397^{***}	4	566,861
	(0.0051)	(0.0043)		
Indigenous/minority	-0.1154^{***}	-0.0442^{***}	3	1,705,221
	(0.0013)	(0.0012)		
Wealth (RWI, std)	0.1016^{***}	0.0627^{***}	7	1,976,134
	(0.0002)	(0.0002)		
Country FE	Yes	Yes		
Sub-national FE	No	Yes		
Urban control	No	Yes		
Wealth (RWI) control	No	Yes[a]		

Note: Each row reports a separate OLS regression of an indicator for school location within 5 km of the nearest documented contaminated site (best-available combined source per country: TSIP combined with the national government register for India, Brazil São Paulo, Mexico, and Peru; TSIP only elsewhere) on the named school characteristic. The within-row sample is the same across columns (1) and (2). Column (1) is the baseline with country fixed effects only. Column (2) replaces country with sub-national (state, province, or region) fixed effects and adds urban and standardized neighborhood RWI (Chi et al., 2022) as controls; the control coefficients are not displayed. The wealth row regresses the within-5 km indicator on within-country standardized RWI, so its coefficient is the percentage-point change per 1 SD of RWI; [a] in this row the wealth control is the regressor itself, not an additional control. Standard errors are heteroskedasticity-robust. Country-level descriptive proximity rates by these characteristics are in Appendix Tables A9 and A8; the wealth-quintile descriptive table is Table 5. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

4.4 Neighborhood wealth and proximity

A central question in the US environmental-justice literature is whether pollution exposure falls disproportionately on poorer communities. We examine this using the RWI neighborhood-wealth matching described in Section 3.5. In 16 of 17 countries, schools near documented contaminated sites are in wealthier neighborhoods, with an unweighted cross-country mean RWI gap of 0.577 points

(Appendix Table A12). Kazakhstan is the lone exception, where the small TSIP sample (35 sites) is dominated by Soviet-era mining and metallurgy operations in resource regions—East Kazakhstan, Karaganda, and Aktyubinsk—spatially separated from the main wealth concentrations in Almaty and Astana.

Table 5 reports enrollment-weighted proximity by within-country RWI quintile for the 7 EMIS countries with enrollment data. The gradient is steep and monotonic in every country: pooling within-country quintiles, 2.4% of students in the poorest quintile attend a school within 5 km of a documented contaminated site, compared to 34.5% in the wealthiest—a ratio of about 14 to 1. The gradient is steepest in Indonesia ($49\times$), India ($39\times$, combined), and the Philippines ($31\times$), and most compressed where government registers densely cover all wealth tiers: Brazil São Paulo ($1.9\times$ combined; Q1 at 49%, Q5 at 93%) and Peru ($6.0\times$, with PAM mining liabilities raising Q1 exposure in rural mining regions). Urban co-location explains a large share of the gradient but does not eliminate it: within urban areas in the four EMIS countries with both indicators (India, Brazil, Peru, Ghana), the Q5/Q1 ratio remains about 5:1. The wealth row of Table 4 confirms this in a pooled regression: a 1 SD increase in within-country RWI is associated with a 10 pp increase in within-5 km proximity at baseline, falling to 6 pp after sub-national fixed effects and urban controls.

A natural question is whether the school proximity rates we document reflect a school-specific exposure pattern or simply where people live: schools, like other facilities, are sited where the population is. To assess this, we compare each country's school proximity rate to the population proximity rate—the share of the population within 5 km of a documented contaminated site, computed by summing WorldPop 2020 1 km population cells within 5 km of any documented site (Appendix Table A14). In 8 of the 11 EMIS countries with WorldPop data, the general population is at least as likely as schools to be within 5 km of a documented site, sometimes substantially more so: 24% of Bangladesh's population versus 11% of its schools, 20% in the Philippines versus 9%, 19% in Colombia versus 9%, and 9% in India versus 5%. The three exceptions are Brazil São Paulo, Mexico, and Peru—all of which combine TSIP with a government register that adds many urban industrial sites. The within-country wealth gradient documented above is therefore best read as a within-school stratification (wealthier-neighborhood schools are more exposed than poorer-neighborhood schools) rather than a school-specific over-exposure claim relative to the general population.

Three potential mechanisms could in principle generate this co-location of industry, schools, and wealth: *(i)* urban co-location of industry and private schools while public schools are more evenly distributed; *(ii)* within-city land-market sorting, in which private schools locate on cheaper industrial-adjacent land (Tooley et al., 2007; Andrabi et al., 2008; Crawfurd et al., 2024); and *(iii)* weak environmental regulation and land-use enforcement in LMIC cities more generally (Greenstone and Hanna, 2014; Duflo et al., 2018). The within-urban regression results above are consistent with mechanism *(i)* and inconsistent with mechanism *(ii)*: if private schools near contaminated sites located there because industrial-adjacent land is cheaper, they should sit in *poorer* micro-neighborhoods than unexposed urban private schools. Appendix Table A16 shows the opposite—in all six countries with urban indicators, urban private schools near documented sites are in *wealthier* neighborhoods than unexposed urban private schools, with the same pattern holding for urban public schools. Within cities, documented contaminated sites, private schools, and public schools alike cluster in the wealthier sub-areas; the zoning mechanism likely contributes but cannot be separated from urban co-location in these data.

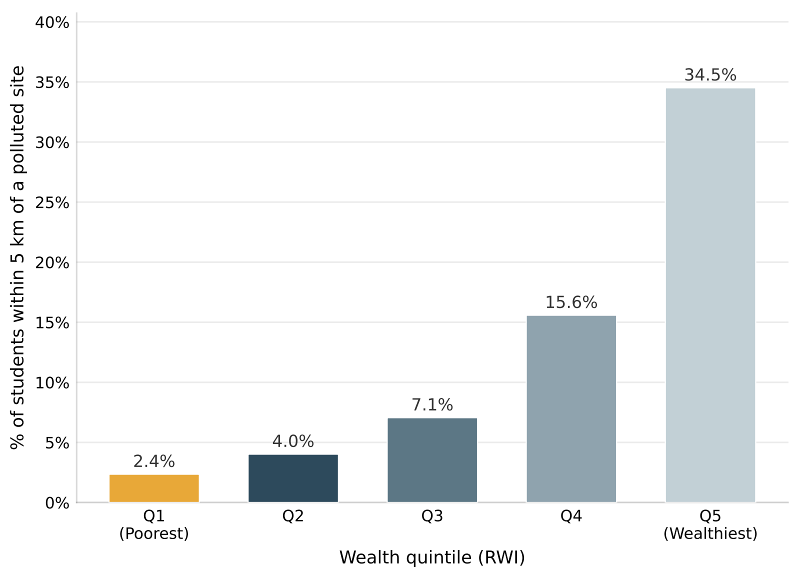
A natural concern is that the wealth gradient reflects TSIP's investigation pattern (NGO investigators may concentrate in wealthier urban areas) rather than the underlying spatial distribution of contamination. The four government registers provide a direct test, because they are administrative records of regulatory- or compliance-flagged sites and are not subject to NGO investigation bias. Appendix Table A13 compares the wealth Q5/Q1 ratio under TSIP-only versus register-only data, both within and across countries. Within the four register countries, the average Q5/Q1 ratio is, if anything, *steeper* under the government registers than under TSIP ($46\times$ versus $34\times$). The within-country contrast varies: the gradient is much steeper under the register in India (CPCB: $60\times$ vs. TSIP $36\times$) and Mexico (RETC: $118\times$ vs. TSIP $36\times$), and shallower under the register in Brazil São Paulo (CETESB: $3\times$ vs. TSIP $6\times$) and Peru (PAM: $2\times$ vs. TSIP $59\times$). The Peru and São Paulo cases reflect the specific nature of those registers—PAM's mining liabilities are in remote highland regions and CETESB blankets all wealth tiers across São Paulo state—rather than evidence of TSIP investigation bias. Across countries, the mean Q5/Q1 ratio is $46\times$ under government registers (4 countries) versus $34\times$ under TSIP-only (7 countries with RWI but no register). The pooled wealth gradient is therefore not an artefact of which sites Pure Earth happens to investigate.

Table 5. School proximity to documented contaminated sites by wealth quintile

	Q1 (Poorest)	Q2	Q3	Q4	Q5 (Wealthiest)	Q5/Q1 ratio
India†	0.8	1.3	3.1	10.1	30.5	39.2
Indonesia	0.7	2.7	5.9	15.3	33.3	49.2
Kenya	1.9	4.1	7.6	12.9	45.7	24.6
Philippines	1.8	5.7	27.5	47.0	53.7	30.6
Brazil (São Paulo)†	49.1	68.3	75.0	85.8	93.1	1.9
Peru†	4.4	10.6	17.7	30.8	26.1	6.0
Ghana	3.2	7.2	14.9	50.3	71.7	22.1
Pooled	2.4	4.0	7.1	15.6	34.5	14.7

Note: Percentage of students attending schools within 5 km of the nearest documented contaminated site, by enrollment-weighted within-country Meta Relative Wealth Index (RWI) quintile. Quintiles are defined within each country so each contains ≈20% of enrolled students sorted by neighborhood RWI; results are then pooled across countries. † uses the best-available combined source (TSIP combined with the national government register); Brazil restricts to São Paulo schools because CETESB covers only São Paulo state. Q5/Q1 ratio compares the wealthiest to the poorest quintile.

Figure 3. Proximity to documented contaminated sites by wealth quintile



Note: Enrollment-weighted percentage of students attending schools within 5 km of the nearest documented contaminated site (best-available source per country: TSIP combined with the national government register for India, Brazil São Paulo, and Peru; TSIP only for Indonesia, Kenya, Philippines, and Ghana), by Meta Relative Wealth Index (RWI) quintile, pooled across the 7 EMIS countries with enrollment data. Quintiles are defined so each represents approximately 20% of enrolled students within each country, sorted by neighborhood RWI. Appendix Figure 4 adds a second panel showing the analogous ambient $PM_{2.5}$ gradient.

4.5 Air pollution at the schoolyard

TSIP catalogs point-source contamination—heavy metals, industrial chemicals, pesticides—but most children's exposure to airborne pollutants comes from ambient air pollution from combustion sources not captured by TSIP. As an independent measure of environmental quality at school locations, we extract satellite-derived annual mean $PM_{2.5}$ (van Donkelaar et al., 2021) for each school in our sample and report the share above three air-quality thresholds: the WHO 2021 annual guideline of 5 µg/m^3, the WHO 2005 interim target of 15 µg/m^3, and the US EPA non-attainment annual standard of 35 µg/m^3 (Table 6).

Across the eleven Panel A countries, virtually every school (99.9%) sits in air that exceeds the WHO 2021 guideline, 85% sit above the WHO 2005 interim target, and 41% are in air that would trigger non-attainment status under US EPA standards. The picture is starkest in South Asia and South-east Asia: in Bangladesh and India, essentially every school sits above the WHO 2005 interim target (15 µg/m^3) and 56–69% sit above the US EPA threshold; in the Philippines, the Peruvian highlands, and Ghana, the WHO 2005 target is exceeded for nearly all schools while exceedance of the US EPA standard is in the low single digits. In Mongolia and Armenia, where coal-burning concentrates in primate cities, 57–60% of mapped schools sit above 35 µg/m^3. Even in the cleanest country in our sample (Argentina), the WHO 2021 guideline is exceeded everywhere. Ambient $PM_{2.5}$ in LMIC school neighbourhoods is therefore not a tail-of-the-distribution problem; it is the modal exposure.

The wealth gradient in ambient $PM_{2.5}$ is much weaker than the wealth gradient in proximity to documented contaminated sites, and reverses direction once we weight by enrolment. Pooling within-country RWI quintiles across the 11 countries with both school location and RWI data, mean $PM_{2.5}$ at school locations is only 9% higher in the wealthiest quintile than in the poorest (35.5 versus 32.6 µg/m^3; Q5/Q1 ratio of $1.09\times$). Weighting by enrolment across the seven countries with enrolment data, the gradient reverses: students in the poorest quintile attend schools with mean $PM_{2.5}$ of 41.8 µg/m^3 versus 32.5 µg/m^3 for the wealthiest quintile (Q5/Q1 ratio of $0.78\times$; Appendix Table A15). The enrolment-weighted reversal is driven by India and Indonesia, where large numbers of low-RWI students attend schools in high-$PM_{2.5}$ regions. The contrast with the $14\times$ wealth gradient in documented-site proximity is informative: $PM_{2.5}$ and point-source heavy-metal contamination are different pollution phenomena with different drivers, so we should not expect their spatial gradients to align. At the district level, Behrer and Heft-Neal (2024) report a positive $PM_{2.5}$-wealth gradient in most LMICs; our school-level results are directionally consistent at the school-weighted level but flip under enrolment weighting.

4.6 Robustness check: TSIP-only comparability across all 17 countries

The headline rates above use the best-available contamination data per country (Section 2.2), which means India, Brazil São Paulo, Mexico, and Peru use national government registers combined with

Table 6. Share of schools in $PM_{2.5}$ above air-quality thresholds

		% of schools with mean annual $PM_{2.5}$ above:			Mean
	Schools	5 μg/m^3	15 μg/m^3	35 μg/m^3	$PM_{2.5}$
Panel A: National school census data					
Bangladesh	77,688	100	100	100	65.3
India	1,366,331	100	100	70	45.1
Ghana	22,236	100	100	2	26.0
Peru	169,982	100	97	7	25.4
Philippines	14,375	100	90	2	21.7
Kenya	26,028	100	89	0	20.1
Indonesia	372,206	100	71	4	19.4
Colombia	56,370	100	85	0	18.1
Mexico	268,598	99	56	0	16.1
Argentina	64,528	100	41	0	15.3
Brazil	149,212	100	35	2	14.1
Pooled	2,587,554	100	85	41	
Panel B: Overture Maps / OpenStreetMap					
Mongolia	793	100	66	60	46.1
Armenia	1,379	100	100	57	32.0
Cambodia	17,039	100	100	7	29.2
Vietnam	66,356	100	100	20	28.9
Kyrgyzstan	1,934	100	99	33	28.7
Kazakhstan	6,295	100	41	12	17.8
Pooled	93,796	100	96	18	

Note: Share of schools whose location falls in a $PM_{2.5}$ pixel above the indicated threshold. $PM_{2.5}$ from van Donkelaar et al. V5GL0502 (annual mean 2023; 0.05° resolution). Thresholds: 5 μg/m^3 is the WHO 2021 annual air-quality guideline; 15 μg/m^3 is the WHO 2005 interim target (and roughly the global population-weighted mean); 35 μg/m^3 is the US EPA non-attainment annual standard. Panel A uses national school census data; Panel B uses Overture Maps / OpenStreetMap school locations. Pooled rates are weighted by school counts within each panel. Countries are sorted by descending mean $PM_{2.5}$.

TSIP, while the other 13 countries use TSIP alone. To check that the headline patterns are not driven by this asymmetric data hybridization, we re-estimate all results using TSIP only across all 17 countries.

The TSIP-only headline within-5 km rate is 6.6% (171,950 schools) compared to 9.7% with combined data (Appendix Table A17). The 3-percentage-point gap is concentrated in Brazil São Paulo (where CETESB metals coverage transforms the rate from 16% to 66%) and Mexico (4.8% to 26%); Appendix Table A18 decomposes the source contribution for each of the four government-register countries. India and Peru, where the government registers add either few sites (CPCB) or different sites (PAM), show much smaller shifts. Importantly, the qualitative within-country patterns are robust to source: in all four combined-source countries, the urban-rural gradient, wealth-quintile gradient, and private-public gap survive when computed using TSIP only (see appendix tables). The key shifts in moving from TSIP-only to combined data are: (a) the headline rate rises by 3 percentage points; (b) the pooled wealth Q5/Q1 ratio compresses from 34× to 14× because PAM mining registers add Q1-quintile rural schools in Peru and CETESB densely covers all wealth tiers in São Paulo; and (c) the private-public regression coefficient strengthens from 7.7 to 11.6 percentage points. The qualitative patterns are also robust to changes in the proximity threshold (Appendix Table A20) and to spatial clustering of standard errors at the contaminated-site level (Appendix Table A21). State-level decompositions for India and Brazil are reported in Appendix Tables A22 and A23; per-country private/public detail is in Appendix Table A19.

These shifts reflect a real and important phenomenon: TSIP's investigation pattern is correlated with both urbanization and Pure Earth's country-by-country programmatic choices. The TSIP-only specification understates true contaminated-site exposure in countries with comprehensive industrial-pollution registers (Mexico, Brazil), while the combined specification gives weight to different site definitions across countries. Neither specification is unambiguously the right one, and we present both. The substantive findings—urban concentration, wealth gradient, and private-public gap—hold under both. By way of contrast with a far more comprehensive data system: in the United States, where the EPA tracks a broad set of hazard sites, 44% of public schools sit within one mile of one (Malik et al., 2026); at the comparable 1 km threshold our combined-source rate is just 0.7%. The 13 study countries without any national contaminated-site register may have true prevalences higher than their current TSIP-only headlines suggest, though by how much depends on country-specific industrial geography and data-collection histories that we cannot adjudicate here.

5 Conclusion

We provide the first multi-country descriptive assessment of school proximity to documented industrially contaminated sites in low- and middle-income countries. Combining the Pure Earth Toxic Sites Identification Program with national government contaminated-site registers in four countries

(11,301 sites across our 17 LMICs), and the locations of 2,601,218 schools, we find that 252,809 schools—9.7% of all schools in our sample (or 6.6% under a TSIP-only comparability check)—lie within 5 km of a documented contaminated site. Weighting by enrollment, 12.7% of students in the 7 countries with enrollment data attend school within this distance. Because TSIP and government registers each capture only a fraction of actual contaminated sites (Dowling et al., 2016; Crawfurd et al., 2025), these remain lower bounds on the true prevalence.

Proximity rates vary widely across countries: from over half of schools within 5 km in Brazil São Paulo (66%), Armenia (65%), and Mongolia (58%), to under 5% in India (4.7%). The cross-country variation reflects differences in source coverage (TSIP investigation intensity, government-register comprehensiveness) as much as actual differences in pollution prevalence. Within countries, the dominant pattern is urban concentration: urban schools are 4 to 28 times more likely than rural schools to be near a documented contaminated site, depending on the country. Reflecting this urban concentration, schools in the wealthiest within-country quintile are about 14 times more likely to be near a site than schools in the poorest. This reverses the US environmental justice pattern but matches recent LMIC evidence on the broader wealth gradient in pollution exposure (Behrer and Heft-Neal, 2024; Ascencio et al., 2025). Where school management type is available, private schools are also more likely than public schools to be near polluted sites in all eight such countries; we treat this finding as preliminary given the limited country coverage and the well-known heterogeneity of LMIC private school provision (from elite urban institutions to low-fee peri-urban schools), and we flag it as an important area for further research.

Several limitations merit flagging. First, school and contaminated-site data are not contemporaneous (school censuses range from 2020/21 to 2025; TSIP was downloaded March 2026), introducing measurement error from interim openings, closures, and remediation. Only 3% of TSIP sites in our 17 study countries are identifiable as "former" or "abandoned" from their names, though legacy heavy-metal contamination can persist for decades (Mielke et al., 2016), and the six Overture-only countries should be interpreted with particular caution despite the 80% coverage threshold. Second, the within-country wealth gradient we document does not imply that all forms of environmental injustice in LMICs run opposite to the US pattern: within India, Scheduled Castes and Tribes remain more exposed than the national average to $PM_{2.5}$ (deSouza et al., 2023) and to coal-fired power plant emissions (Kopas et al., 2020), so the wealth gradient and the caste/ethnicity gradient can run in opposite directions within the same country—a tension we cannot adjudicate with our school-level data.

While our proximity estimates cannot determine whether specific schools face actual contamination, they suggest three directions for policy. First, countries could consider environmental screening as part of the process for constructing or approving new school buildings, in urban and peri-urban areas where proximity to industrial activity is common. Second, targeted environmental monitoring—including soil, water, and air sampling—around existing schools in proximity to known contaminated sites could identify cases where remediation or other protective measures are warranted. Third, and perhaps most fundamentally, the analysis highlights the value of building more comprehensive pollution inventories in LMICs. The TSIP represents an important start, but the substantial undercounting of actual sites limits its utility for environmental health surveillance. Investments in systematic pollution mapping, potentially leveraging remote sensing and satellite data alongside ground-based investigation, would enable more precise identification of at-risk school populations.

Future work could extend this analysis in several directions. Linking school proximity data to education outcome measures—such as standardized test scores available through EMIS systems in India (UDISE+) and other countries—would enable direct estimation of whether proximity to polluted sites is associated with lower student achievement. More comprehensive pollution data, including satellite-derived measures of air quality ($PM_{2.5}$, NO_2) and industrial land use, could capture the many pollution sources not covered by the TSIP. Within-city analyses using more granular spatial data could better characterize the micro-geography of school-pollution co-location. Finally, extending the coverage of geocoded school data in LMICs—through continued expansion of OpenStreetMap mapping and open publication of government school census databases—would improve the precision and coverage of analyses like ours.

Data Availability Statement

The replication package for this paper—including all Python analysis code, Stata and R cross-language replication scripts, raw school census data, TSIP pollution site data, government contaminated site registers, and generated tables and figures—is archived at:

https://doi.org/10.5281/zenodo.19359187

The TSIP contaminated sites database was downloaded in March 2026 via the contaminatedsites.org API (https://www.contaminatedsites.org). National school census data are obtained from government open-data portals; sources and access conditions for each country are documented in the README file accompanying the replication package and in Table A3 in the Appendix. School location data for six countries are derived from Overture Maps (https://overturemaps.org), available under the CDLA Permissive 2.0 licence. Meta Relative Wealth Index data are available via the Humanitarian Data Exchange (https://data.humdata.org). Satellite $PM_{2.5}$ data are from the global annual mean surface $PM_{2.5}$ dataset available from NASA SEDAC (https://sedac.ciesin.columbia.edu). Government contaminated site registers for India (CPCB), Brazil (CETESB), Mexico (RETC), and Peru (MINEM) are publicly available from the respective agency websites, as documented in Section 2.2.

References

Andrabi, T., Das, J. and Khwaja, A. I. (2008), 'A dime a day: The possibilities and limits of private schooling in Pakistan', *Comparative Education Review* **52**(3), 329–355.

Ascencio, E. J., Barja, A., Montes-Alvis, J., Kephart, J. L., Gouveia, N., Rodriguez, D. A., Benmarhnia, T., Diez Roux, A. V., Bilal, U., Miranda, J. J. and Carrasco-Escobar, G. (2025), 'Urban and socioeconomic disparities in $PM_{2.5}$ exposure across 340 Latin American cities', *Environmental Research Letters* **20**(12), 124044. 3, 24

Asher, S., Jha, K., Novosad, P., Adukia, A. and Tan, B. (2026), Residential segregation and unequal access to local public services in India: Evidence from 1.5m neighborhoods, Technical report, Center for Global Development.
URL: *https://www.cgdev.org/publication/residential-segregation-and-unequal-access-local-public-services-india-evidence-15m* 16

Attina, T. M. and Trasande, L. (2013), 'Economic costs of childhood lead exposure in low- and middle-income countries', *Environmental Health Perspectives* **121**(9), 1097–1102. 2

Axelrad, D. A., Bellinger, D. C., Ryan, L. M. and Woodruff, T. J. (2007), 'Dose–response relationship of prenatal mercury exposure and IQ: an integrative analysis of epidemiologic data', *Environmental Health Perspectives* **115**(4), 609–615. 30

Balakrishnan, U. and Tsaneva, M. (2021), 'Air pollution and academic performance: Evidence from India', *World Development* **146**, 105553. 2, 17

Behrer, A. P. and Heft-Neal, S. (2024), 'Higher air pollution in wealthy districts of most low- and middle-income countries', *Nature Sustainability* **7**(2), 203–212. 3, 21, 24

Berkhout, E., Maulana, S., Molato-Gayares, R., Park, A. and Suryadarma, D. (2025), 'Lead Exposure and Cognitive Skills in a Developing Country: Evidence from Toxic Sites in Indonesia'. 2

Bouchard, M. F., Chevrier, J., Harley, K. G., Kogut, K., Vedar, M., Calderon, N., Trujillo, C., Johnson, C., Bradman, A., Barr, D. B. and Eskenazi, B. (2011), 'Prenatal exposure to organophosphate pesticides and IQ in 7-year-old children', *Environmental Health Perspectives* **119**(8), 1189–1195. 30

Caravanos, J., Gualtero, S., Dowling, R., Ericson, B., Keith, J., Hanrahan, D. and Fuller, R. (2014), 'A simplified risk-ranking system for prioritizing toxic pollution sites in low- and middle-income countries', *Annals of Global Health* **80**(4), 278–285. 3

Carneiro, J., Cole, M. A. and Strobl, E. (2021), 'The effects of air pollution on students' cognitive performance: Evidence from Brazilian university entrance tests', *Journal of the Association of Environmental and Resource Economists* **8**(6), 1051–1077. 2, 17

Cheeseman, M., Ford, B., Volckens, J., Lyapustin, A. and Pierce, J. R. (2022), 'Disparities in air pollutants across racial, ethnic, and poverty groups at US public schools', *GeoHealth* **6**(12), e2022GH000672. 2, 3, 9

Chi, G., Fang, H., Chatterjee, S. and Blumenstock, J. E. (2022), 'Microestimates of wealth for all low- and middle-income countries', *Proceedings of the National Academy of Sciences* **119**(3), e2113658119. 8, 17

Crawfurd, L., Hares, S. and Todd, R. (2024), 'The impact of private schools, school chains and PPPs in developing countries', *The World Bank Research Observer* **39**(1), 97–123.

Crawfurd, L., Hu, J. and Mitchell, T. (2025), 'Beyond hotspots: estimating population lead exposure from battery recycling', *Working Paper* . 3, 9, 24

Crawfurd, L., Todd, R., Hares, S., Sandefur, J. and Bonnifield, R. S. (2024), 'The Effect of Lead Exposure on Children's Learning in the Developing World: A Meta-Analysis', *The World Bank Research Observer* p. lkae010.

deSouza, P. N., Chaudhary, E., Dey, S., Ko, S., Németh, J., Guttikunda, S., Chowdhury, S., Kinney, P., Subramanian, S. V., Bell, M. L. and Kim, R. (2023), 'An environmental justice analysis of air pollution in India', *Scientific Reports* **13**, 16690. 24

Dowling, R., Caravanos, J., Grigsby, P., Rivera, A., Ericson, B., Amoyaw-Osei, Y., Akuffo, B. and Fuller, R. (2016), 'Estimating the Prevalence of Toxic Waste Sites in Low- and Middle-Income Countries', *Annals of Global Health* **82**(5), 700–710. 2, 4, 6, 24

Duflo, E., Greenstone, M., Pande, R. and Ryan, N. (2018), 'The value of regulatory discretion: Estimates from environmental inspections in India', *Econometrica* **86**(6), 2123–2160. 18

Ebenstein, A., Lavy, V. and Roth, S. (2016), 'The long-run economic consequences of high-stakes examinations: Evidence from transitory variation in pollution', *American Economic Journal: Applied Economics* **8**(4), 36–65. 17

Ericson, B., Caravanos, J., Chatham-Stephens, K., Landrigan, P. and Fuller, R. (2013), 'Approaches to systematic assessment of environmental exposures posed at hazardous waste sites in the developing world: The Toxic Sites Identification Program', *Environmental Monitoring and Assessment* **185**(2), 1755–1766. 3

Ericson, B., Duong, T. T., Keith, J., Nguyen, T. C., Havens, D., Daniell, W., Karr, C. J., Ngoc Hai, D., Van Tung, L., Thi Nhi Ha, T., Wilson, B., Hanrahan, D., Croteau, G. and Patrick Taylor, M. (2018), 'Improving human health outcomes with a low-cost intervention to reduce exposures from lead acid battery recycling: Dong Mai, Vietnam', *Environmental Research* **161**, 181–187. 9

Ericson, B., Landrigan, P., Taylor, M. P., Frostad, J., Caravanos, J., Keith, J. and Fuller, R. (2016), 'The Global Burden of Lead Toxicity Attributable to Informal Used Lead-Acid Battery Sites', *Annals of Global Health* **82**(5), 686–699. 3

Grandjean, P. and Landrigan, P. J. (2014), 'Neurobehavioural effects of developmental toxicity', *The Lancet Neurology* **13**(3), 330–338. 2

Greenstone, M. and Hanna, R. (2014), 'Environmental regulations, air and water pollution, and infant mortality in India', *American Economic Review* **104**(10), 3038–3072. 18

Heissel, J. A., Persico, C. and Simon, D. (2022), 'Does pollution drive achievement? the effect of traffic pollution on academic performance', *Journal of Human Resources* **57**(3), 747–776. 17

Ipapa, G. (2023), The Hidden Costs of Recycling: Lead Exposure and Student Learning, Working Paper. 2

Kippler, M., Tofail, F., Hamadani, J. D., Gardner, R. M., Grantham-McGregor, S. M., Bottai, M. and Vahter, M. (2012), 'Early-life cadmium exposure and child development in 5-year-old girls and boys: a cohort study in rural Bangladesh', *Environmental Health Perspectives* **120**(10), 1462–1468. 30

Kopas, J., York, E., Jin, X., Harish, S. P., Kennedy, R., Shen, S. V. and Urpelainen, J. (2020), 'Environmental justice in India: Incidence of air pollution from coal-fired power plants', *Ecological Economics* **176**, 106711. 24

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. 30

Litzow, E., Cecato, B., Zarate-Barrera, T. and Romero, M. (2024), Toxic Recycling: The Cost of Used Lead-Acid Battery Processing in Mexico, Working Paper. 2

Malik, S., Kraft, M. A. and Falken, G. T. (2026), U.S. schools' proximity to environmental hazard sites: A national analysis, EdWorkingPaper 26-1384, Annenberg Institute at Brown University.
URL: *https://edworkingpapers.com/ai26-1384* 2, 3, 9, 23

Mielke, H. W., Gonzales, C. R., Powell, E. T. and Mielke, P. W. (2016), 'Spatiotemporal dynamic transformations of soil lead and children's blood lead ten years after Hurricane Katrina: New grounds for primary prevention', *Environment International* **94**, 567–575. 9, 24

Pastor, M., Sadd, J. L. and Morello-Frosch, R. (2002), 'Who's minding the kids? Pollucion, public schools, and environmental justice in Los Angeles', *Social Science Quarterly* **83**(1), 263–280. 2

Persico, C. L. and Venator, J. (2021), 'The effects of local industrial pollution on students and schools', *Journal of Human Resources* **56**(2), 406–445. 2, 17

Pure Earth (2017), Toxic sites identification program (tsip) investigator handbook, Technical report, Pure Earth, New York, NY. Revised March 2019.
URL: *https://www.pureearth.org/wp-content/uploads/2023/11/TSIP-Investigator-Handbook-2017.pdf* 4, 31

Samlani, N., Pino, D. S., Bertolo, R. and Pak, T. (2024), 'A comprehensive dataset of environmentally contaminated sites in the state of São Paulo in Brazil', *Scientific Data* **11**, 263. 4, 6

Tooley, J., Dixon, P. and Gomathi, S. V. (2007), 'Private schools and the Millennium Development Goal of universal primary education: A census and comparative survey in Hyderabad, India', *Oxford Review of Education* **33**(5), 539–560.

van Donkelaar, A., Hammer, M. S., Bindle, L., Brauer, M., Brook, J. R., Garay, M. J., Hsu, N. C., Kalashnikova, O. V., Kahn, R. A., Lee, C., Levy, R. C., Lyapustin, A., Sayer, A. M. and Martin, R. V. (2021), 'Monthly global estimates of fine particulate matter and their uncertainty', *Environmental Science & Technology* **55**(22), 15287–15300. 8, 21

Villalobos, L. and Blackman, A. (2025), 'Air pollution and education in the Global South: Fine particulates reduce high-stakes test scores in Bogotá', *Applied Economics Letters* . 2, 17

Wasserman, G. A., Liu, X., Parvez, F., Ahsan, H., Factor-Litvak, P., van Geen, A., Slavkovich, V., LoIacono, N. J., Cheng, Z., Hussain, I., Momotaj, H. and Graziano, J. H. (2004), 'Water arsenic exposure and children's intellectual function in Araihazar, Bangladesh', *Environmental Health Perspectives* **112**(13), 1329–1333. 30

A Appendix

Table A1. Key neurotoxicants at TSIP sites and their effects on child cognition

Contaminant	TSIP sites (17 countries)	Exposure metric	IQ loss (points) per doubling	IQ loss (points) heavily exposed	Key source
Lead	1,253	Blood lead (μg/dL)	3–5	5–15	Lanphear et al. (2005)
Methylmercury	324	Maternal hair Hg (ppm; prenatal)	—[a]	1–10	Axelrad et al. (2007)
Arsenic	304	Drinking water / urinary As	$\sim$0.8	3–7	Wasserman et al. (2004)
Cadmium	65	Urinary Cd	0.5–3	1–5	Kippler et al. (2012)
Organophosphate pesticides	61[b]	Maternal urinary DAP metabolites	$\sim$1.7	5–7	Bouchard et al. (2011)

Notes: TSIP site counts are for the 2,840 sites in our 17 study countries with the indicated contaminant recorded as the key pollutant (March 2026 database). "Per doubling" is the IQ-point loss associated with a doubling of the exposure metric, derived from the cited dose-response estimates; comparisons across rows are approximate because exposure metrics differ. "Heavily exposed" is the typical IQ-point deficit observed in children in cohorts with high exposure relative to background populations (e.g., for lead, blood lead 5–15 μg/dL compared to <2 μg/dL; for arsenic, drinking water above the WHO guideline of 10 μg/L). Lead is the most potent per unit of exposure and has the most consistently estimated dose-response across cohorts. The other contaminants produce IQ losses of the same order of magnitude (typically 1–7 points) at exposure levels common at heavily polluted sites in low- and middle-income countries.

[a] Methylmercury dose-response is conventionally expressed as a linear effect per ppm of maternal hair mercury ($\sim$0.18 IQ points per ppm in the Axelrad et al. (2007) meta-analysis of the Faroe Islands, Seychelles, and New Zealand cohorts), not as a per-doubling effect. [b] Includes 61 TSIP sites with organochlorine pesticides (DDT, Aldrin, Lindane, and other unspecified pesticides) recorded as the key pollutant. The TSIP database records few sites with organophosphate pesticides specifically as the key pollutant, but pesticide manufacturing facilities (49 sites in study countries) commonly handle both classes.

Table A2. Composition of TSIP sites by industry category

	All TSIP sites		Lead sites only	
Source industry	N	%	N	%
Lead - Battery Recycling	472	16.6	464	37.0
Mining and Ore Processing	296	10.4	63	5.0
Artisanal Mining (hand mining)	260	9.2	22	1.8
Unknown	233	8.2	58	4.6
Lead Smelting (with ingot production)	177	6.2	170	13.6
Industrial/Municipal Dump Site	171	6.0	64	5.1
Industrial Estate (mixed industries)	102	3.6	41	3.3
Agriculture	99	3.5	8	0.6
Chemical Manufacturing (acids, organics, base chemicals)	88	3.1	14	1.1
Lead-acid battery manufacturing or repair	83	2.9	82	6.5
Multiple Diverse Industries	83	2.9	35	2.8
Tannery Operations	81	2.9	6	0.5
Recycling / Recyclers (including salvage yards)	71	2.5	39	3.1
Product Manufacturing (electronics, equipment, clothing)	64	2.3	8	0.6
Mechanic/Garage (Vehicle Repair)	53	1.9	12	1.0
Pesticide Manufacturing	49	1.7	1	0.1
Dye Industry	44	1.5	7	0.6
Power Plants (coal or oil)	43	1.5	8	0.6
Smelting (everything except Lead)	42	1.5	15	1.2
Transportation (bus stations, rail yards)	39	1.4	9	0.7
Heavy Industry (casting, rolling, stamping)	39	1.4	16	1.3
Ceramics (lead glaze)	38	1.3	38	3.0
Petrochemical Industries (refineries)	37	1.3	16	1.3
E-waste recycling	26	0.9	15	1.2
Lead Mines	21	0.7	14	1.1
Other (rare categories)	129	4.5	28	2.2
Total	2,840	100.0	1,253	100.0

Notes: Industry composition of the 2,840 TSIP sites in the 17 study countries (downloaded March 2026 via the contaminatedsites.org API), and of the 1,253 sites where lead is the recorded key pollutant. "Source industry" is recorded by TSIP investigators from a controlled vocabulary of approximately 40 categories (Pure Earth, 2017). Categories accounting for less than 1% of both the all-sites and lead-sites totals are aggregated into the "Other (rare categories)" row. The Lead–Battery Recycling category combines informal backyard battery breakers, scrap dealers, and larger industrial recyclers; the TSIP vocabulary does not distinguish further within this category.

Table A3. School location data sources by country

Country	Source	Schools	Available variables: Enroll.	Mgmt	Urban	Level
National school census (EMIS)						
Argentina	Mapa Educativo Nacional (2023)	64,638		✓	✓	
Bangladesh	LGED / HDX (2023)	78,129				
Brazil	INEP Censo Escolar (2023)	150,131	✓	✓	✓	
Colombia	SISE / DANE (2024)	56,650			✓	✓
Ghana	EMIS / ESRAG (2020/21)	22,581	✓	✓	✓	
India	UDISE+ (2022-23)	1,372,195	✓	✓	✓	✓
Indonesia	Dapodik (2024)	396,280	✓	✓		✓
Kenya	NEMIS (2023)	26,197	✓	✓		
Mexico	SEP Catálogo (2025)	270,228		✓	✓	✓
Peru	MINEDU Padrón (2024)	175,754	✓	✓	✓	✓
Philippines	DepEd Master List (2023)	16,022	✓			
Overture Maps / OpenStreetMap (February 2026)						
Armenia	Overture Maps	1,379				
Cambodia	Overture Maps	17,148				
Kazakhstan	Overture Maps	6,365				
Kyrgyzstan	Overture Maps	1,935				
Mongolia	Overture Maps	795				
Vietnam	Overture Maps	67,087				
Total		2,723,514				

Note: Summary of school location data sources. Years in parentheses indicate the data vintage. All EMIS datasets are the most recent publicly available versions at the time of analysis (early 2026). Available variables: Enroll. = student enrollment counts; Mgmt = management type (public/private); Urban = urban/rural classification; Level = school level (primary/secondary). Colombia covers public schools only. Ghana schools geocoded via GhanaPostGPS digital address codes. See Table A4 for Overture Maps coverage assessment.

Table A4. Overture Maps school coverage assessment

Country	Overture schools	Actual schools	Coverage (%)	Source
Armenia	1,379	1,372	>100	Ministry of Education (2023)
Cambodia	17,148	18,830	91	Ministry of Education (2022/23)
Kazakhstan	6,365	7,859	81	Ministry of Education (2024/25)
Kyrgyzstan	1,935	2,266	85	National Statistics Committee (2023)
Mongolia	795	859	93	Ministry of Education (2023)
Vietnam	67,087	26,300	>100	General Statistics Office (2021/22)

Note: Overture Maps schools counted from the February 2026 Overture Places release, filtered to education-related categories (school, elementary_school, high_school, middle_school, preschool, education). Actual school counts from the most recent available government statistics. Vietnam's Overture count exceeds the official school count because Overture includes preschools, kindergartens, and other education places not counted in the official school total (which covers primary through upper secondary only). We include only countries where Overture coverage exceeds approximately 80% of the official school count. Nepal (18,587 Overture schools, 63% coverage) and Pakistan (82,166 Overture schools, 31% coverage) are excluded due to insufficient coverage; other countries assessed but excluded due to both low coverage and small absolute school counts include Azerbaijan, Bolivia, Nigeria, Tajikistan, and Tanzania.

Table A5. TSIP countries with 10 or more sites excluded from the analysis

Country	Income group	TSIP sites	Reason for exclusion
Insufficient school data coverage			
Tanzania	LMIC	256	Overture Maps coverage ≈25% of actual schools; no EMIS data obtained
Tajikistan	LMIC	220	Overture Maps returns <2% of actual schools
Azerbaijan	UMIC	127	Overture Maps coverage ≈56% of actual schools; no EMIS data obtained
Senegal	LMIC	119	OSM/Overture data lacks coordinates; no EMIS data obtained
Nepal	LMIC	103	Overture Maps coverage 63% of actual schools; OSM coordinates unavailable
Georgia	UMIC	78	OSM data lacks coordinates; no EMIS data obtained
Bolivia	LMIC	47	Overture Maps coverage ≈35% of actual schools; no EMIS data obtained
Nigeria	LMIC	31	Overture Maps coverage ≈16% of actual schools; no EMIS data obtained
No geocoded school data available			
China	UMIC	268	No geocoded school data obtained
Russia	UMIC	175	No geocoded school data obtained
Ukraine	LMIC	75	No geocoded school data obtained
Uganda	LIC	24	No geocoded school data obtained
Zimbabwe	LMIC	18	No geocoded school data obtained
Myanmar	LMIC	17	No geocoded school data obtained
Belarus	UMIC	16	No geocoded school data obtained
Outside paper scope			
Chile	HIC	24	High-income country

Note: Table shows all countries with 10 or more TSIP contaminated sites that are not included in the main analysis. Income group classifications follow World Bank FY2025 definitions: LIC = low income, LMIC = lower middle income, UMIC = upper middle income, HIC = high income. Overture coverage percentages compare Overture Maps school counts to the most recent available government school census. For countries excluded due to insufficient Overture coverage, applying an urban-bias correction derived from countries with complete EMIS data suggests true proximity rates are approximately 1.3–4.3× lower than the raw Overture estimates, consistent with proximity rates observed in the main sample.

Table A6. Robustness check: EMIS versus Overture Maps school proximity rates

	EMIS		Overture Maps		
Country	Schools	%<5 km	Schools	%<5 km	Diff. (pp)
Argentina	64,638	12.4	47,609	12.1	−0.3
Bangladesh	78,129	11.4	38,104	21.9	+10.5
Brazil	150,131	5.7	168,388	8.8	+3.1
Colombia	56,650	8.7	22,150	20.5	+11.7
Ghana	22,581	26.0	2,657	57.8	+31.9
India	1,372,195	4.6	337,030	20.8	+16.3
Indonesia	396,280	8.4	184,521	14.9	+6.5
Kenya	26,197	14.6	5,241	41.6	+27.1
Mexico	270,228	4.8	118,617	7.2	+2.4
Peru	175,754	10.4	23,945	15.1	+4.6
Philippines	16,022	9.0	61,553	18.3	+9.3

Note: Percentage of schools within 5 km of the nearest TSIP pollution site, computed separately using national school census (EMIS) data and Overture Maps / OpenStreetMap data for the same country. EMIS school counts correspond to the geocoded school census databases described in Section 2.2. Overture school counts are from the February 2026 release, filtered to education-related place categories. In 10 of 11 countries, the Overture-derived proximity rate exceeds the EMIS rate (mean difference: +11.2 percentage points); Argentina is the lone exception (Overture 12.1% vs EMIS 12.4%, -0.3 pp), likely because Argentina's high urbanisation rate (≈93%) means the urban–rural composition of Overture-mapped schools is similar to the EMIS census even at moderate coverage. The Pearson correlation between EMIS and Overture proximity rates across countries is $r = 0.88$ ($p < 0.001$). The upward bias reflects the urban concentration of OpenStreetMap mapping: where Overture coverage is sparse, the mapped schools skew urban, and urban schools are mechanically closer to TSIP sites. Total Overture coverage does not bound the bias: Argentina at 74% coverage shows near-zero bias while Bangladesh at 49% coverage shows +11 pp—the urban–rural composition of mapped schools matters more than the total count. Proximity rates for the six Overture-only countries in Table 2 should therefore be read as potentially upward-biased point estimates rather than strict upper bounds, with the bias likely largest for the less urbanised countries in that group.

Table A7. Data construction workflow

Stage	Description	Observations
Panel A: TSIP pollution sites		
Raw download	Sites from contaminatedsites.org API (March 2026)	5,016 sites
Coordinate cleaning	Drop sites with <4 decimal places precision or outside country bounding box	4,596 sites
Study sample	Retain countries with ≥25 sites and school data	2,840 sites (17 countries)
Panel B: Government contaminated-site registers		
India (CPCB)	Central Pollution Control Board contaminated-site inventory (2021)	105 sites
Brazil São Paulo (CETESB)	State environmental agency metals/heavy-metal sites (2024)	1,270 sites
Mexico (RETC)	Federal lead/heavy-metal release registry (PRTR; 2022)	964 sites
Peru (PAM)	MINEM Pasivos Ambientales Mineros (mining liabilities; 2024)	6,122 sites
All registers combined		**8,461 sites**
Panel C: School locations		
EMIS school censuses	11 countries with geocoded national data	2,628,805 schools
Overture Maps / OSM	6 countries with ≥80% coverage	94,709 schools
Total analysis sample	**17 countries**	**2,723,514 schools**
Panel D: Proximity results (combined sources)		
Within 5 km (combined)	Schools within 5 km of any TSIP or government-register site	252,809 (9.7%)
Within 5 km (TSIP only)	Schools within 5 km of nearest TSIP site	171,950 (6.6%)
Panel E: Additional datasets		
Overture (all countries)	School locations for map (Figure 1)	900,194 schools
Enrollment data	7 EMIS countries with non-zero enrollment	337,509,321 students
	Students within 5 km of nearest site (combined)	42,783,664 (12.7%)

Note: This table tracks observations through each stage of the data construction pipeline. Panel A shows the TSIP pollution site sample, starting from the raw API download and progressing through coordinate quality filtering and country selection. Panel B shows the national government contaminated-site registers we combine with TSIP in four countries: India (CPCB), Brazil São Paulo state (CETESB), Mexico (RETC), and Peru (MINEM PAM). Panel C shows school location data from two sources: national school censuses (EMIS) for 11 countries and Overture Maps / OpenStreetMap for 6 countries where geocoded EMIS data are not publicly available. The total analysis sample of 2,723,514 schools reflects the combination of these two sources (Table 2). Panel D summarizes the headline proximity results using best-available combined sources (TSIP combined with the national register where available; TSIP only elsewhere) and reports the TSIP-only comparison as a reference. Panel E notes two additional datasets: Overture Maps school locations are downloaded for all 17 countries (including those with EMIS data) and used for the map in Figure 1, which is why Figure 1 shows approximately 900,000 schools rather than the full 2,723,514; and enrollment data are available for 7 of the 11 EMIS countries (Brazil, Ghana, India, Indonesia, Kenya, Peru, Philippines).

Table A8. School proximity to documented contaminated sites by other school characteristics

	Schools		<5 km (%)	
Panel A: School gender type				
	Co-educational	Single-sex	Co-educational	Single-sex
India†	1,335,674	36,521	4.6	10.3
Kenya	25,637	269	13.8	24.2
Peru†	173,689	2,065	14.4	27.3
Total	1,535,000	38,855	5.9	11.3
Panel B: Special education status				
	Regular	Special ed.	Regular	Special ed.
Argentina	62,281	1,953	11.8	12.6
Colombia	55,590	1,060	8.7	10.6
Kenya	25,654	543	14.1	18.8
Mexico†	268,490	1,733	25.9	41.2
Peru†	174,557	1,197	14.5	19.2
Total	586,572	6,486	18.8	21.6
Panel C: Indigenous or minority status				
	Other	Indigenous/minority	Other	Indigenous/minority
Argentina	62,058	740	11.7	0.3
India†	1,352,775	19,420	4.7	8.9
Mexico†	247,492	22,736	28.1	2.1
Total	1,662,325	42,896	8.4	5.2
Panel D: School level				
	Primary	Secondary	Primary	Secondary
Colombia	26,425	6,085	5.2	6.8
India†	710,368	255,760	3.3	7.8
Indonesia	160,047	376,023	6.5	7.9
Mexico†	99,171	61,669	23.5	25.1
Peru†	147,123	19,605	13.9	16.5
Total	1,143,134	719,142	6.9	9.6

Note: Percentage of schools within 5 km of the nearest documented contaminated site, using the best-available source per country: † indicates countries where the source is TSIP combined with the national government register (India CPCB, Mexico RETC, Peru PAM; Brazil São Paulo CETESB is not represented in this table because school-level gender/special-ed/indigenous classifications are not recorded in CETESB). Panel A: Single-sex includes boys-only and girls-only schools. 3 countries: India (UDISE+), Kenya (NEMIS), Peru (MINEDU). Panel B: Special education as classified in each country's school census. 5 countries: Argentina, Colombia, Kenya, Mexico, Peru. Panel C: Argentina (intercultural/bilingual schools), India (Madarsa), Mexico (indigenous schools). Panel D: Primary includes primary and upper-primary schools; secondary includes lower-secondary, upper-secondary, and equivalent levels. Pre-primary, vocational, and higher education are excluded. 5 countries: Colombia, India, Indonesia, Mexico, Peru. Distance is haversine distance.

Table A9. School proximity to documented contaminated sites by location type and management type

	Schools		<5 km (%)	
Panel A: Location type				
	Urban	Rural	Urban	Rural
Argentina	44,981	19,657	16.8	0.6
Brazil (São Paulo)†	26,667	1,168	68.6	16.1
Colombia	19,271	37,379	20.5	2.7
Ghana	11,086	11,250	45.2	7.0
India†	203,864	1,168,330	20.6	2.0
Mexico†	132,642	137,586	49.4	3.4
Peru†	86,312	89,442	24.2	5.2
Subtotal	524,823	1,464,812	31.1	2.4
Panel B: Management type				
	Public	Private	Public	Private
Argentina	51,925	12,078	9.8	21.2
Brazil (São Paulo)†	18,427	9,408	59.0	81.0
Ghana	15,003	7,578	18.4	40.9
India†	1,053,265	299,510	3.0	10.5
Indonesia	172,622	223,658	4.8	10.4
Kenya	19,397	6,800	8.6	30.4
Mexico†	224,705	45,523	18.6	62.5
Peru†	128,665	47,089	10.3	26.0
Subtotal	1,684,009	651,644	6.9	17.0

Note: Percentage of schools within 5 km of the nearest documented contaminated site, using the best-available source per country (Table 2). † indicates countries where the best-available source combines TSIP with a national government register (India CPCB, Brazil CETESB metals (São Paulo only), Mexico RETC lead emitters, Peru *Pasivos Ambientales Mineros*). Panel A: urban/rural classification from national school census; Colombia includes public schools only. Panel B: Public includes government and government-aided schools; private includes unaided and unrecognised. Distance is haversine distance. Brazil rows use the São Paulo subset since CETESB covers only São Paulo state.

Table A10. Schools near TSIP sites: capital city breakdown

Country	Capital city	Schools	Enrollment	Within 1 km Schools	Within 1 km %	Within 5 km Schools	Within 5 km %
Indonesia	DKI Jakarta	7,895	1,733,671	369	4.7	4,088	51.8
Kenya	Nairobi	1,288	346,053	190	14.8	926	71.9
Philippines	Metro Manila	762	1,933,185	128	16.8	655	86.0
Colombia	Bogotá	2,595	—	30	1.2	1,438	55.4
Peru	Lima[1]	29,376	2,456,008	404	1.4	11,271	38.4
Mexico	Ciudad de México[2]	8,897	—	0	0.0	1	0.0
Argentina	Buenos Aires[3]	9,128	—	307	3.4	4,211	46.1
Ghana	Greater Accra	3,416	1,091,313	415	12.1	2,309	67.6

Note: Capital city breakdown for the eight EMIS countries with separately identifiable capital-region schools (India and Brazil are shown in Appendix Tables A22–A23). School data from country EMIS sources (see Section 2.1). Distance is haversine distance from each school to the nearest TSIP site. Enrollment is total enrolled students in capital-city schools; "—" indicates enrollment data are unavailable.

[1] Lima Metropolitana only; excludes Lima Provincias.

[2] TSIP database contains no sites within 5 km of any school in Mexico City.

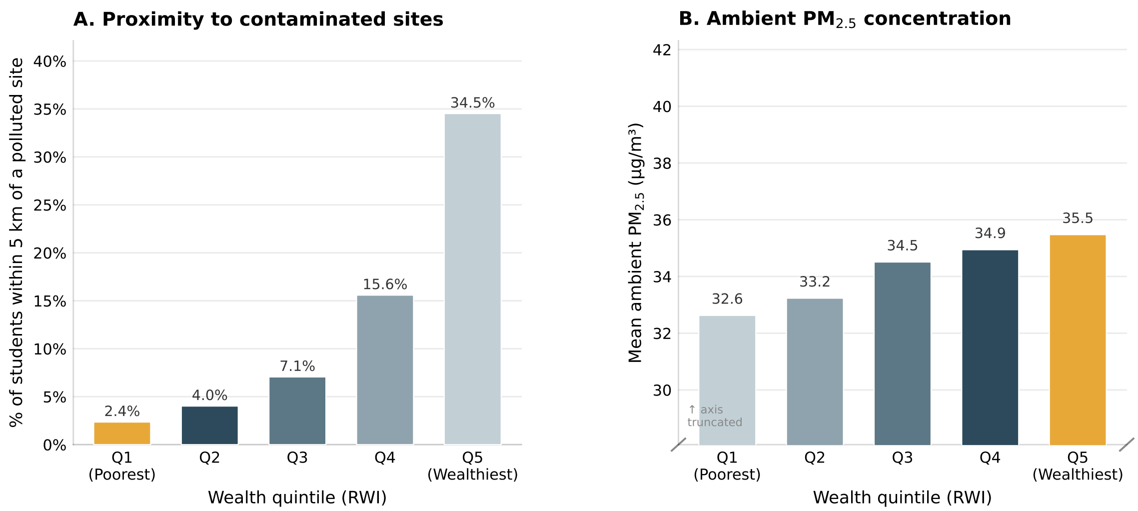
[3] Defined by geographic bounding box (lat -34.8 to -34.3, lon -58.7 to -58.2).

Table A11. Within-urban private–public proximity gap

Country	Urban private N	Urban private % <5 km	Urban public N	Urban public % <5 km	Urban gap (pp)	Overall gap (pp)
Argentina	11,725	21.7	32,745	15.2	+6.6	+11.4
Brazil (São Paulo)†	9,354	81.3	17,313	61.8	+19.5	+22.1
Ghana	5,663	50.7	5,423	39.4	+11.4	+22.4
India†	107,730	22.3	90,073	18.4	+4.0	+7.5
Mexico†	44,366	63.3	88,276	42.4	+20.9	+43.9
Peru†	44,876	27.0	41,436	21.2	+5.8	+15.7

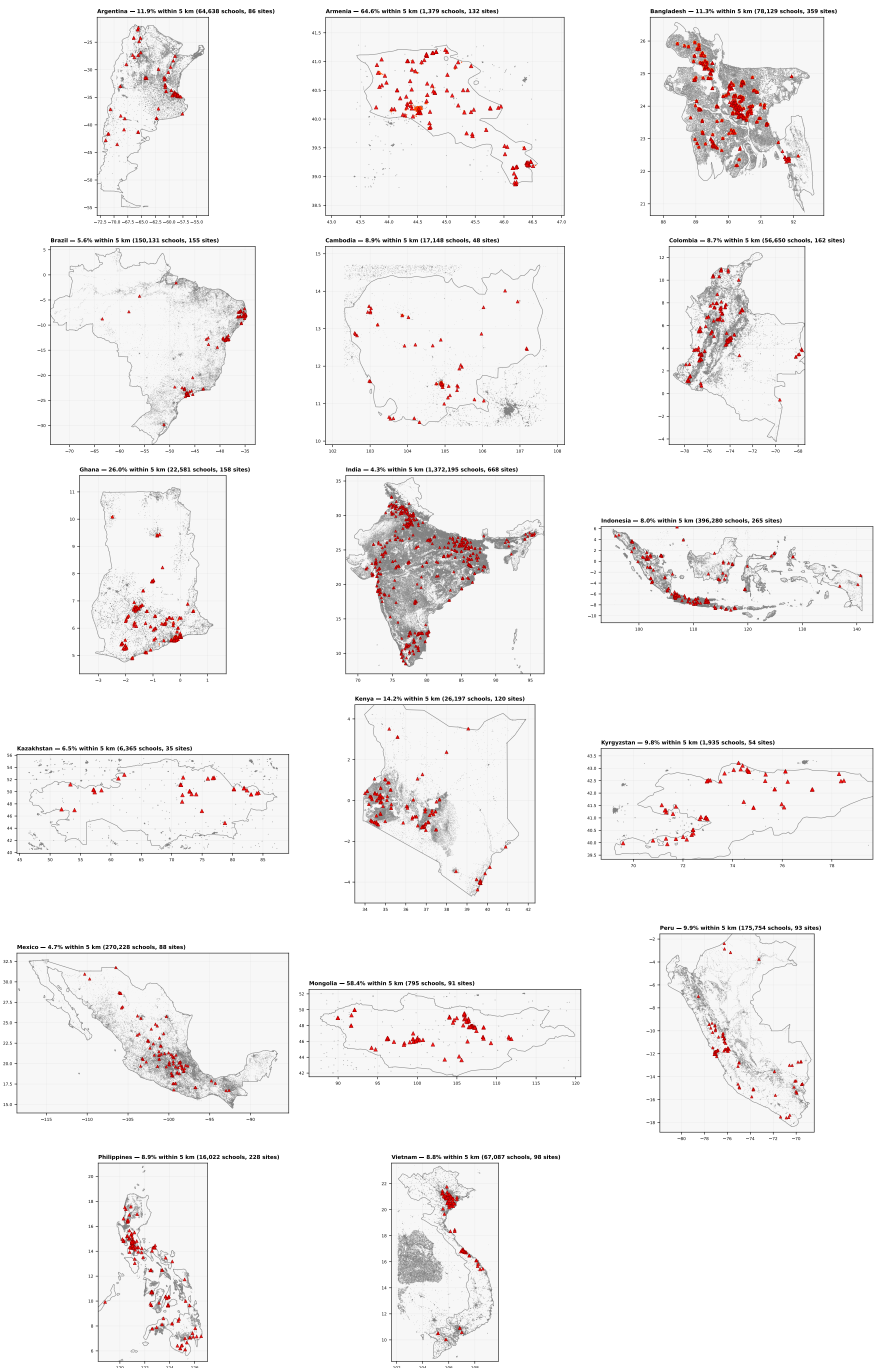
Note: Percentage of urban private and urban public schools within 5 km of a documented contaminated site (best-available source per country). Urban gap = within-urban private − within-urban public; overall gap = unrestricted private − public. † uses combined source.

Figure 4. School-level wealth gradient in TSIP proximity and ambient $PM_{2.5}$



Note: Both panels show a measure of school-environment pollution by Meta Relative Wealth Index (RWI) quintile, from poorest (Q1) to wealthiest (Q5). Panel A reports the enrollment-weighted percentage of students attending schools within 5 km of a documented contaminated site (TSIP ∪ government register where available), pooled across the 7 EMIS countries with enrollment data, with quintiles defined so each represents about 20% of enrolled students within each country. Panel B reports mean ambient $PM_{2.5}$ (annual average 2023; van Donkelaar et al. V5GL0502 at 0.05° resolution) at school locations, pooling within-country quintiles across the 11 countries with both EMIS school data and RWI coverage. The two panels measure different pollution phenomena: Panel A captures proximity to point-source heavy-metal contamination, while Panel B captures ambient combustion-driven air pollution. Within-country quintiles avoid the cross-country composition confound from differences in average $PM_{2.5}$ and wealth levels. The Panel B y-axis is truncated to make the modest gradient visible. The enrolment-weighted $PM_{2.5}$ gradient (not shown) reverses direction (Appendix Table A15, Panel B).

Figure 5. School and TSIP site locations by country



Note: Each panel maps all schools (grey dots) and TSIP polluted sites (red triangles) for one of the 17 study countries. Orange dots indicate schools located within 5 km of the nearest TSIP site. School locations are from national school censuses for eleven countries and Overture Maps for the remaining six (Armenia, Cambodia, Kazakhstan, Kyrgyzstan, Mongolia, Vietnam). Panel titles report the percentage of schools within 5 km, the total number of schools, and the number of TSIP sites in each country.

Table A12. Mean RWI for schools near and far from documented contaminated sites

	Within 5 km		Beyond 5 km		
Country	N	Mean RWI	N	Mean RWI	Diff.
Argentina	7,660	1.089	54,106	0.649	+0.440
Bangladesh	8,819	0.305	69,233	0.018	+0.287
Brazil (national)	8,460	1.193	139,431	0.643	+0.550
Brazil (São Paulo)†	18,493	1.166	9,319	0.826	+0.340
Colombia	4,890	0.923	48,189	0.095	+0.828
Ghana	5,861	1.034	16,683	0.228	+0.806
India†	65,072	0.908	1,305,282	0.111	+0.797
Indonesia	31,667	1.078	358,595	0.342	+0.737
Kenya	3,732	0.808	22,450	0.059	+0.749
Mexico†	70,134	1.104	198,197	0.206	+0.898
Peru†	25,500	0.951	146,748	0.259	+0.692
Philippines	1,432	1.037	14,570	0.137	+0.901
Mean across countries					+0.669

Note: Mean Meta Relative Wealth Index (RWI) for schools within 5 km of a documented contaminated site (best-available source per country) vs. schools beyond 5 km. Positive Diff. means schools near sites are in wealthier areas. 12 of 12 countries show schools near sites in wealthier neighborhoods. † uses best-available combined source.

Table A13. Wealth-quintile proximity ratios under TSIP-only vs government register: testing investigation-bias

	TSIP only			Government register		
Country	Q1 %	Q5 %	Q5/Q1	Q1 %	Q5 %	Q5/Q1
Panel A: Register countries (within-country comparison)						
India[a]	0.5	17.2	35.9×	0.1	4.4	60.1×
Brazil (São Paulo)[a]	4.8	29.9	6.2×	33.5	90.3	2.7×
Mexico[a]	0.4	13.4	35.6×	0.5	61.4	118.3×
Peru[a]	0.5	31.4	59.2×	3.1	5.3	1.7×
Panel A mean			*34.2×*			*45.7×*
Panel B: TSIP-only countries (no government register available)						
Bangladesh	6.7	25.0	3.7×	—	—	—
Indonesia	0.4	27.4	68.6×	—	—	—
Kenya	1.4	47.6	32.9×	—	—	—
Argentina	0.5	23.7	48.9×	—	—	—
Philippines	0.9	34.2	37.9×	—	—	—
Colombia	1.2	26.7	22.7×	—	—	—
Ghana	3.1	68.6	22.0×	—	—	—
Panel B mean			*33.8×*			—

Note: School-weighted within-country wealth quintile (Q1 = poorest, Q5 = wealthiest neighborhood RWI) proximity rates, reported separately under two contaminated-site source definitions. The TSIP-only column uses Pure Earth's investigation-based database; the government register column uses the relevant national administrative register (India CPCB, Brazil São Paulo CETESB metals, Mexico RETC, Peru PAM). [a] marks countries in Panel A where both sources are available. The within-country comparison (Panel A) holds school identity constant, isolating the source-definition effect on the gradient steepness: in 2 of 4 register countries, the TSIP-only gradient is steeper than the register-only gradient. Comparison across panels: mean Q5/Q1 is 45.7× in register countries (using register-only) versus 33.8× across the 7 TSIP-only countries. If TSIP's investigation pattern systematically amplifies the apparent wealth gradient (because investigators concentrate in wealthier urban areas), we expect both the within-country contrast (steeper under TSIP than register) and the cross-country contrast (steeper in TSIP-only than register-only countries) to be positive.

Table A14. School proximity versus population proximity to documented contaminated sites

	Schools		Population (WorldPop 2020)		School / pop.
Country	N	% <5 km	Millions	% <5 km	ratio
India†	1,372,195	4.7	1380.0	9.2	0.51×
Brazil (São Paulo)†	27,835	66.4	62.1	14.2	4.67×
Mexico†	270,228	26.0	128.9	6.2	4.19×
Peru†	175,754	14.5	33.0	13.6	1.07×
Bangladesh	78,129	11.3	164.7	24.4	0.46×
Indonesia	396,280	8.0	273.5	12.2	0.65×
Kenya	26,197	14.2	53.8	18.0	0.79×
Argentina	64,638	11.9	45.2	17.1	0.69×
Philippines	16,022	8.9	109.6	19.5	0.46×
Colombia	56,650	8.7	50.9	19.2	0.45×
Ghana	22,581	26.0	31.1	28.8	0.90×
Unweighted mean		18.2		16.6	1.35×

Note: For each of the eleven EMIS countries, this table compares the percentage of schools within 5 km of a documented contaminated site (best-available source per country) with the percentage of the country's population within 5 km of the same set of sites. The population baseline uses WorldPop's 2020 UN-adjusted 1 km population raster, summing population over cells whose centroid is within 5 km of any documented site and dividing by the country's total mapped population. The school/population ratio expresses school proximity relative to the population baseline: a ratio of 1.0 means schools sit where people sit; a ratio above 1.0 means schools are disproportionately concentrated near documented contaminated sites beyond what the geographic distribution of the population would imply. † uses the combined TSIP + government register; other countries use TSIP only. Brazil São Paulo restricts WorldPop, TSIP, and the CETESB register to a bounding box around São Paulo state.

Table A15. Ambient $PM_{2.5}$ by Wealth Quintile

Wealth quintile	Schools	Mean $PM_{2.5}$ ($\mu g/m^3$)	Mean RWI
Panel A: Within-country quintiles, pooled (11 countries with EMIS + RWI data)			
Q1 *(Poorest)*	513,843	32.6	-0.491
Q2	513,696	33.2	-0.094
Q3	512,931	34.5	0.242
Q4	513,901	34.9	0.607
Q5 *(Wealthiest)*	512,043	35.5	1.116
Q5/Q1 ratio		1.09	
Panel B: Enrollment-weighted (7 countries with enrollment data)			
Q1 *(Poorest)*	653,307	41.8	-0.373
Q2	483,515	43.3	0.082
Q3	381,938	38.9	0.468
Q4	284,691	33.8	0.860
Q5 *(Wealthiest)*	231,991	32.5	1.302
Q5/Q1 ratio		0.78	

Notes: Mean ambient $PM_{2.5}$ (annual average, 2023; van Donkelaar et al. V5GL0502 at 0.05° resolution) at school locations, stratified by Meta Relative Wealth Index (RWI) quintile. Panel A assigns quintiles within each country (equal school counts per quintile), then pools results across countries, weighting each school equally. This within-country approach avoids the composition confound from cross-country differences in average $PM_{2.5}$ and wealth levels. Panel B weights by student enrollment and is restricted to countries with enrollment data; quintiles are defined over the cross-country cumulative-enrollment-weighted RWI distribution. Q1 = poorest 20% of schools within their country; Q5 = wealthiest 20%. A Q5/Q1 ratio >1 indicates that students in wealthier neighborhoods attend higher-$PM_{2.5}$ schools.

Table A16. Mean RWI of urban private schools near vs. far from documented contaminated sites

	Urban private schools					Urban public schools				
Country	N near	RWI near	N far	RWI far	Diff.	N near	RWI near	N far	RWI far	Diff.
India†	24,053	1.112	83,633	0.749	+0.363	16,522	1.113	73,508	0.757	+0.355
Brazil (SP)†	7,604	1.235	1,750	1.011	+0.224	10,701	1.128	6,611	0.873	+0.255
Mexico†	28,093	1.193	16,239	1.057	+0.136	37,417	1.120	50,805	0.763	+0.357
Peru†	12,104	1.243	32,771	1.059	+0.184	8,791	1.118	32,623	0.872	+0.246
Ghana	2,873	1.142	2,784	0.639	+0.503	2,135	1.076	3,270	0.470	+0.606
Argentina	2,550	1.123	9,169	1.040	+0.083	4,972	1.086	27,735	0.889	+0.198

Note: Mean Meta Relative Wealth Index (RWI) of urban schools, split by management type and proximity to the nearest documented contaminated site (best-available source per country). Diff. is (Near − Far) in RWI units. Positive Diff. indicates schools near contamination are in *wealthier* neighborhoods than unexposed schools of the same management type — the opposite of what land-market sorting toward cheaper polluted land would predict. Restricted to countries with both urban/rural indicators and management-type data; Brazil restricts to São Paulo schools.

Table A17. School proximity to TSIP sites only (comparability robustness for Table 2)

		TSIP within 5 km		Combined
Country	Schools	N	%	% (Table 2)
Panel A: National school census (EMIS) and government register countries				
Argentina	64,638	7,660	11.9	11.9
Bangladesh	78,129	8,819	11.3	11.3
Brazil (national)	150,131	8,460	5.6	5.6
Brazil (São Paulo)†	27,835	4,548	16.3	66.4
Colombia	56,650	4,943	8.7	8.7
Ghana	22,581	5,861	26.0	26.0
India†	1,372,195	62,484	4.6	4.7
Indonesia	396,280	31,680	8.0	8.0
Kenya	26,197	3,732	14.2	14.2
Mexico†	270,228	13,054	4.8	26.0
Peru†	175,754	18,337	10.4	14.5
Philippines	16,022	1,432	8.9	8.9
Panel A subtotal	2,506,509	162,549	6.5	9.7
Panel B: Overture Maps / OpenStreetMap countries (TSIP only by construction)				
Armenia	1,379	891	64.6	64.6
Cambodia	17,148	1,531	8.9	8.9
Kazakhstan	6,365	411	6.5	6.5
Kyrgyzstan	1,935	189	9.8	9.8
Mongolia	795	464	58.4	58.4
Vietnam	67,087	5,914	8.8	8.8
Panel B subtotal	94,709	9,400	9.9	9.9
Total	2,601,218	171,949	**6.6**	9.7

Note: TSIP-only within-5 km rates across all 17 countries, reported for cross-country comparability against Table 2's headline rates (which use combined TSIP ∪ government-register data for the four countries marked †). The last column repeats the Table 2 combined rate; for the 13 countries without government registers and Brazil (national), the two rates coincide. The gap between TSIP-only and combined is concentrated in Brazil São Paulo and Mexico, reflecting the fact that CETESB and RETC catalog substantially more contaminated sites than TSIP investigators have visited.

Table A18. Source contributions to the combined within-5 km rate, four countries with government registers

	% of schools within 5 km of:			Schools within 5 km of:		
Country	TSIP only	Govt only	Combined	TSIP only	Govt only	Both
India	4.6	1.0	4.7	51,277	2,619	11,207
Brazil (SP)	16.3	66.3	66.4	34	13,945	4,514
Mexico	4.8	23.8	26.0	5,829	57,085	7,225
Peru	10.4	4.6	14.5	17,419	7,210	918

Note: For each country with a national government contaminated-site register, decomposition of the combined within-5 km rate into schools near TSIP only, schools near the government register only, and schools near both. "Combined" equals the union: a school is within 5 km of either source. Registers: India CPCB (105 sites); Brazil São Paulo CETESB metals (1,270 sites); Mexico RETC lead emitters (964); Peru PAM mining liabilities (6,122). For the other 13 study countries, no government register is available; combined and TSIP-only rates coincide.

Table A19. Country-by-country private–public proximity regressions

	Baseline		+ Urban		
Country	Private (pp)	SE	Private (pp)	SE	N
Argentina	11.36^{***}	(0.39)	6.34^{***}	(0.33)	64,003
Brazil (São Paulo)†	22.07^{***}	(0.54)	19.52^{***}	(0.58)	27,835
Ghana	22.43^{***}	(0.65)	9.06^{***}	(0.60)	22,581
India†	7.49^{***}	(0.06)	2.68^{***}	(0.04)	1,352,775
Indonesia	5.58^{***}	(0.08)	—	—	396,280
Kenya	21.83^{***}	(0.59)	—	—	26,197
Mexico†	43.95^{***}	(0.24)	21.19^{***}	(0.21)	270,228
Peru†	15.69^{***}	(0.22)	5.38^{***}	(0.22)	175,754

Note: Country-by-country OLS regressions of an indicator for school location within 5 km of a documented contaminated site on a private-school indicator. Baseline is the difference in means between private and public schools (in percentage points); "+ Urban" adds an urban location indicator. † uses combined source. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A20. Private-school proximity coefficient: threshold sensitivity

	Baseline		+ Urban	
Threshold	Private (pp)	SE	Private (pp)	SE
Within 1 km	0.0196^{***}	(0.0002)	0.0107^{***}	(0.0002)
Within 5 km	0.1158^{***}	(0.0004)	0.0514^{***}	(0.0005)

Note: OLS regressions of an indicator for school location within X km of the nearest documented contaminated site (best-available source per country) on a private-school indicator, pooled across the 8 countries with public/private data, with country fixed effects. Baseline = country FE only; "+ Urban" adds an urban location indicator (restricted to 6 countries with urban/rural data). Heteroskedasticity-robust SE. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A21. Spatial clustering robustness: private school proximity coefficient

Clustering level	(1) Baseline	(2) + Urban
Site-level clustering (baseline)	0.0765^{***}	0.0339^{***}
	(0.0060)	(0.0045)
≈5 km geographic blocks	0.0765^{***}	0.0339^{***}
	(0.0043)	(0.0033)
≈11 km geographic blocks	0.0765^{***}	0.0339^{***}
	(0.0058)	(0.0044)

Note: Each row reports the private school coefficient from the pooled OLS regression of school proximity (within 5 km) on an indicator for private management and country fixed effects, varying only the clustering of standard errors. The baseline is the nearest-TSIP-site cluster (same as Table 4). Geographic block clustering assigns schools to a country $\times$ grid-cell cluster, grouping together all schools in the same spatial neighborhood regardless of which TSIP site they are nearest to. This directly addresses the concern that sites less than 5–10 km apart share unobserved local characteristics. Block cells measure approximately $0.05° \times 0.05°$ (≈5 km) and $0.1° \times 0.1°$ (≈11 km). Number of clusters: site-level 1,570, 5 km blocks 226,752, 11 km blocks 88,224. Column (1): all countries with private data ($N = 2,478,004$). Column (2): countries with urban/rural data ($N = 2,055,527$). Standard errors in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$.

Table A22. India: school proximity to documented contaminated sites by state

State	Schools	Combined (TSIP ∪ CPCB)		TSIP only	
		N <5 km	%	N <5 km	%
Delhi	2,154	1,964	91.2	1,963	91.1
Chandigarh	228	172	75.4	172	75.4
Punjab	19,354	2,116	10.9	2,017	10.4
Gujarat	52,229	5,585	10.7	5,206	10.0
Haryana	23,009	2,256	9.8	2,197	9.5
Tamil Nadu	58,559	5,587	9.5	5,461	9.3
Bihar	84,352	7,288	8.6	7,288	8.6
Jharkhand	44,146	3,022	6.8	3,007	6.8
Uttarakhand	20,311	1,256	6.2	1,207	5.9
Maharashtra	104,903	5,870	5.6	5,592	5.3
Telangana	42,158	2,202	5.2	1,759	4.2
Uttar Pradesh	231,499	11,994	5.2	11,380	4.9
Kerala	15,440	697	4.5	685	4.4
Goa	1,440	52	3.6	0	0.0
Madhya Pradesh	121,326	4,051	3.3	4,044	3.3
Other states	551,087	10,991	2.0	10,506	1.9
All India	1,372,195	65,103	4.7	62,484	4.6

Note: Top 15 Indian states by combined within-5 km percentage; remaining states aggregated. "Combined" uses TSIP ∪ CPCB; "TSIP only" is the strictly-comparable subset. Schools from UDISE+ (2022–23).

Table A23. Brazil: school proximity to TSIP sites by state, with CETESB combined for São Paulo

State	TSIP only		Combined (where available)	
	Schools	% <5 km	N <5 km	% <5 km
PE	6,671	28.1	—	—
São Paulo (SP)†	27,835	16.3	18,493	66.4
AL	2,229	13.0	—	—
PB	3,923	8.7	—	—
BA	10,065	8.2	—	—
RJ	9,133	4.5	—	—
RS	6,999	1.7	—	—
PA	8,965	0.3	—	—
MG	14,328	0.3	—	—
PI	3,388	0.0	—	—
Other states	56,595	0.0	—	—
All Brazil	150,131	5.6	—	—

Note: Top 10 Brazilian states by TSIP-only within-5 km percentage. The CETESB *Cadastro de Áreas Contaminadas* covers only São Paulo state; for SP, "Combined" adds CETESB's 1,270 metals-contaminated sites to TSIP. † uses best-available combined source. Schools from INEP Censo Escolar (2023).