ATMOSPHERIC ENVIRONMENT: X 31 (2026) 100492

Contents lists available at ScienceDirect

Atmospheric Environment: X

journal homepage: www.journals.elsevier.com/atmospheric-environment-x

Air pollution exposure inequality across U.S. income and racial groups, 2000–2023: A hybrid monitor–satellite analysis with two new shape diagnostics

Tom Regpala[1], Eric Palafox[1], Anjana Yatawara[*]

Department of Mathematics, California State University, Bakersfield, 9001 Stockdale Highway, Bakersfield, 93311, CA, USA

ARTICLE INFO

Dataset link: https://github.com/anjana-yatawara/air-pollution-exposure-inequality, https://aqs.epa.gov/aqsweb/airdata/download_files.html, https://registry.opendata.aws/surface-pm2-5-v6gl/, https://api.census.gov/data

Keywords:
Air pollution exposure
Health inequality metrics
Slope index of inequality
Middle-group residual
Within-stratum slope heterogeneity
Monitor-network selection bias

ABSTRACT

Estimates of exposure inequality depend both on where air pollution is measured and on how inequality is summarised from those measurements. Environmental-justice surveillance typically compresses a pollutant's socioeconomic gradient into one regression slope, the slope index of inequality (SII), defined as the population-weighted slope of group-mean concentration on income rank. A single slope assumes that the concentration–rank profile is a straight line and that exposure and income move together within every income group. The SII still returns a number when either assumption fails, but it no longer flags the misfit. We characterise county-level exposure inequality across the contiguous United States for four criteria pollutants ($PM_{2.5}$, NO_2, O_3, SO_2) over 2000–2023 (up to 44 states). We add two dimensionless shape diagnostics to the standard slope. The Middle-Group Residual (MGR) is the signed departure of the middle-income stratum from the SII line, and the Within-Stratum Slope Heterogeneity (WSSH) is the spread of within-stratum exposure–income correlations. $PM_{2.5}$ exposure uses a population-weighted hybrid of regulatory monitors and satellite-derived concentrations; the other pollutants use the monitor network. The mean income-stratified relative index of inequality (RII) is −0.048 ($PM_{2.5}$), +0.278 (NO_2), −0.018 (O_3), and +0.147 (SO_2). The $PM_{2.5}$ gradient reverses sign between the monitor-equipped county subset (RII +0.074, lower-income more exposed) and the population-representative full-county sample (RII −0.048). This reversal directly reflects the siting of the monitoring network. Non-linear profiles dominate; |MGR| exceeds 0.02 in 50%–92% of state-years and rises monotonically at finer partitions. The SO_2 gradient reversed direction over the period (HAC $p = 0.003$). Its remaining burden along the racial-composition axis falls on lower-income but predominantly White rural counties, so the income and racial axes diverge. A single slope cannot reveal this divergence. Across these four pollutants, the slope alone is therefore an incomplete summary. Its sign can hinge on which counties host monitors, and its straight-line form fails in most state-years. The MGR and WSSH recover the missing shape and coverage information from the same county data the SII already uses.

1. Introduction

Fine particulate matter and the reactive trace gases nitrogen dioxide (NO_2), ozone (O_3), and sulfur dioxide (SO_2) are not distributed evenly across a population. Exposure to these pollutants remains a leading driver of premature mortality even as mean concentrations fall (Cohen et al., 2017; Jbaily et al., 2022; Kerr et al., 2024). Air-quality science and regulation therefore routinely ask which communities breathe the most pollution and how that burden tracks income and race. The standard answer collapses a pollutant's socioeconomic gradient into a single regression slope and reports whether that slope is positive or negative. We show, for four U.S. criteria pollutants over 2000–2023, that this single number can hide two features of direct atmospheric-science relevance. The first is the *shape* of the gradient (whether middle-income communities lie on the straight line the slope implies). The second is the *coverage* of the measurement network (whether the counties that host monitors represent the country). We add two simple, unit-free diagnostics that recover the shape, and we use a population-weighted satellite–monitor hybrid to recover the coverage. Both change the conclusions drawn from the slope alone. The problem extends beyond the United States. Exposure to fine particulate matter tracks income worldwide, with the heaviest burden in lower-income countries and a profile across the global income distribution that bends rather than

* Corresponding author.
E-mail addresses: sregpala@csub.edu (T. Regpala), epalafox1@csub.edu (E. Palafox), ayatawara@csub.edu (A. Yatawara).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.aeaoa.2026.100492
Received 12 May 2026; Received in revised form 13 July 2026; Accepted 20 July 2026
Available online 28 July 2026
2590-1621/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

rises in a straight line (Rentschler and Leonova, 2023; Sager, 2025). A diagnostic that flags when one slope misrepresents the gradient is therefore useful wherever exposure inequality is monitored.

We study the four criteria pollutants with continuous, county-resolvable monitoring across the full 2000–2023 window (Section 2.2). These pollutants span distinct source and chemistry regimes, from regionally transported secondary aerosol ($PM_{2.5}$) and traffic-dominated primary NO_2 to photochemical O_3 and point-source SO_2. They therefore test a shape diagnostic against gradients of very different origin. Disparities in these pollutants are well documented and have not closed in step with the overall decline in concentrations (Colmer et al., 2020; Jbaily et al., 2022; Tessum et al., 2021; Kerr et al., 2024). The slope and relative indices of inequality (SII, RII) are the most widely reported summaries of such disparities (Moreno-Betancur et al., 2015). The SII is the population-weighted slope of group-mean concentration on socioeconomic rank, and the RII is that slope divided by the mean concentration. They are faithful when the group profile is a straight line. The slope is silent about the misfit when middle-income communities are more, or less, exposed than a line through the extremes would predict. The U.K. Office for Health Improvement and Disparities currently handles this pattern by transforming the outcome before fitting (Office for Health Improvement and Disparities, 2024). The transformation preserves one reportable slope but removes the evidence of the shape problem from the reported number. Flexible spline regression has already shown that the demographic–exposure relationship is genuinely curved rather than linear for $PM_{2.5}$ (Daouda et al., 2022). However, no compact statistic flags the curvature wherever the slope is reported without refitting a model. A second blind spot is the measurement network itself. Regulatory monitors are sited where population and infrastructure already concentrate, so the counties they cover are not a random sample of either exposure or income (Wang et al., 2024; Haskell-Craig et al., 2025).

Here we make three contributions. First, we define two named, dimensionless diagnostics that specialise textbook regression residual analysis (Cook and Weisberg, 1982) to the SII and are reported alongside it. The Middle-Group Residual (MGR) is the signed, mean-normalised residual of the middle-income stratum from the straight line the SII fits. It is zero exactly when the three stratum means are collinear. The Within-Stratum Slope Heterogeneity (WSSH) is the spread of the within-stratum exposure–income correlations. It indicates whether income and exposure trend the same way inside every stratum or flip direction across the income distribution. Each is computed from the same county-level inputs that yield the SII, and each is unit-free. The two formalise distinct ways a single slope can mislead (Section 2.1). Second, we apply the full EPA Air Quality System monitor-metadata filter chain (Completeness Indicator; Monitoring Objective; Sample-Duration tiebreaker; NAAQS-Primary collocation; Event-Type policy; U.S. Environmental Protection Agency, 2024) at continental scale and quantify how the resulting coverage gaps bias the measured gradient. Third, for $PM_{2.5}$ we replace the monitor-only county sample with a population-weighted hybrid that fills unmonitored counties with the satellite-derived WUSTL ACAG V6.GL.03 estimate (Hammer et al., 2020; van Donkelaar et al., 2021; Shen et al., 2024). We show that the cross-state mean $PM_{2.5}$ gradient reverses sign between the monitored subset and the full population. This monitor-network selection-bias signature has not previously been quantified in U.S. environmental-justice work. We find that non-linear stratum profiles predominate across 2292 state-pollutant-year panels. We also find a sign-reversing SO_2 gradient over the study period and a racial-axis divergence in the residual SO_2 burden that a single slope cannot reveal.

2. Methods

2.1. Rationale for two new diagnostic statistics

Six standard inequality metrics are computed routinely on stratified panels (see Section 2.3): the slope and relative indices of inequality (SII and RII), the relative concentration index (RCI), the between-group variance (BGV), the Theil index, and the mean log deviation (MLD). All six summarise the between-stratum dispersion of mean concentrations or its slope on socioeconomic rank. Two structural blind spots remain. First, none of the standard slope metrics flags a stratum profile that departs from a straight line. The SII reports the best-fitting line regardless of whether the middle-income stratum sits on that line, above it, or below it. Flexible spline regression has already established that the demographic–exposure relationship for $PM_{2.5}$ is curved (Daouda et al., 2022). The spline recovers the curvature by fitting a more elaborate model. The literature has lacked a single, named number that flags the same curvature wherever the slope is reported, without refitting. The U.K. Office for Health Improvement and Disparities technical guidance (Office for Health Improvement and Disparities, 2024) acknowledges the curvature and handles it by log- or logit-transforming the concentration before fitting the SII. The transformation preserves one reportable slope, but it removes the evidence of the shape problem from the reported number and breaks comparability with the concentration-scale exposures used in dose–response work (Cohen et al., 2017; Burnett et al., 2018). Residual analysis of regression models is a textbook statistical technique (Cook and Weisberg, 1982). However, no named, dimensionless residual statistic adapts it to the SII for routine reporting. Second, none of the standard between-stratum metrics is sensitive to within-stratum gradient heterogeneity. Exposure can rise with income inside the lowest stratum yet fall with income inside the highest stratum. Such a state-year is summarised identically to one in which all strata trend the same way, even though the underlying exposure patterns differ substantively.

We introduce two named diagnostics to fill these two gaps without sacrificing the SII's concentration-scale interpretability. They are SII-anchored, single-number statistics computable alongside the slope itself. They do not offer a new way to detect non-linearity, which spline regression already provides. The Middle-Group Residual (MGR) scales the SII residual at the middle stratum by the population-weighted mean concentration; the Within-Stratum Slope Heterogeneity (WSSH) is the spread of the within-stratum exposure–rank correlations across the partition. Each is dimensionless, is computable from the same county-level inputs that yield the SII, and reduces to zero when the implicit straight-line, uniform-gradient assumption underlying the SII holds. Section 3.3 demonstrates that they carry information substantively distinct from the six standard metrics across our 2292-panel sample.

2.2. Data sources

We study four criteria pollutants ($PM_{2.5}$, NO_2, O_3, and SO_2). They are the criteria species with continuous, county-resolvable EPA monitoring across the full 2000–2023 window, and they span distinct source and chemistry regimes. $PM_{2.5}$ is a regionally transported secondary aerosol, and NO_2 is a primary, traffic-dominated gas. O_3 is a secondary photochemical product whose peaks are rural, and SO_2 is a point-source gas dominated by combustion and industrial emitters. A shape diagnostic that holds across all four therefore faces gradients of very different physical origin. Lead (Pb) and CO are excluded because their post-2000 county-resolvable coverage is too sparse for a continuous 24-year panel.

EPA Air Quality System annual_conc_by_monitor files (U.S. Environmental Protection Agency, 2024) for 2000–2023 were downloaded for parameter codes 88101 ($PM_{2.5}$ FRM/FEM Mass), 44201 (O_3 8-hour daily maximum), 42602 (NO_2), and 42401 (SO_2). Records were filtered to per-pollutant Metric-Used and Sample-Duration combinations specified in the EPA AQS Data Mart manuals, then inner-joined to the EPA `aqs_monitors.csv` metadata.

Five quality filters were then applied. Filter (1) retained records with Completeness Indicator equal to Y. Filter (2) screened the Monitoring Objective field for population-exposure or peak-concentration keywords. Panels with at least one of POPULATION EXPOSURE, HIGHEST

CONCENTRATION, MAX OZONE CONCENTRATION, SOURCE ORIENTED, or UNKNOWN were retained; panels with only non-exposure background or quality-assurance designations were dropped. Filter (3) applied a Sample-Duration tiebreaker preferring FRM 24 HOUR over FEM 24-HR BLK AVG. Filter (4) applied a NAAQS-Primary collocation tiebreaker for collocated POCs at the same site. Filter (5) applied an Event-Type policy preferring Events Included over No Events. The retained monitor-level rows were aggregated to the county-pollutant-year by population-weighted mean. O_3 concentrations in EPA AQS are reported in ppm; values are converted to ppb (factor of 1000) for unit consistency with NO_2 and SO_2.

Coverage as a missingness protocol. A regulatory monitor is sited to judge area-wide compliance. It is not sited to sample the population at random. The counties a network covers are therefore not a random draw with respect to either income or concentration. Roughly one in five U.S. counties hosts any regulatory $PM_{2.5}$ monitor (Haskell-Craig et al., 2025), and the national $PM_{2.5}$ network leaves about 44% of nonattainment urban areas uncaptured under the current standard (Wang et al., 2024). We therefore treat the filter chain above as an explicit missingness protocol whose retained set is demographically structured. We do not treat it as a data-cleaning step that silently discards rows. The filter chain retains 11,155 of 14,668 candidate county-years for $PM_{2.5}$, 13,723 of 18,315 for O_3, 4413 of 5976 for NO_2, and 5743 of 7576 for SO_2 when applied to the raw annual files. The losses fall disproportionately on rural and lower-population counties. We handle the resulting gap in two ways. The unmonitored $PM_{2.5}$ county-years are filled with a satellite-derived estimate (below), so the analysis covers the full population rather than only the monitored subset. The consequence of that choice is quantified directly in Section 3.4. No comparable satellite-fused product is used here for NO_2, O_3, and SO_2. The unmonitored counties for these gases remain a transparent sample restriction, and we characterise their demographic tilt in Section 4 rather than impute it away. Two checks bound the effect of this missingness on the gas results. First, the Completeness Indicator filter alone removed 13.7% of the metric-selected NO_2 monitor-year records, 5.3% of O_3, and 16.1% of SO_2 as within-year-incomplete. These records constitute the missing data present in the downloaded files and were excluded before any coverage consideration. Second, the gas gradients are robust to which states are covered. We drop the lowest-coverage states (the bottom quartile by state-years, then all states with fewer than five state-years). This exclusion shifts the cross-state mean income-axis RII from +0.278 to +0.314 (NO_2), from −0.018 to −0.018 (O_3), and from +0.147 to +0.136 (SO_2), preserving sign and direction in every case.

We additionally require each state-year in the primary inferential sample to contain enough counties to define three strata robustly. The thresholds are at least 6 quality-passed monitor counties per state-year for NO_2, O_3, and SO_2, and at least 10 counties for the $PM_{2.5}$ panel. State-years below these thresholds are excluded from the primary panels (and from the bootstrap of Section 2.7) but retained for descriptive coverage maps.

A hybrid monitor–satellite county dataset was constructed for $PM_{2.5}$ by pairing each contiguous-state county-year with the best available concentration estimate. Quality-passed AQS monitor measurements were used where available, covering approximately 15% of county-years. The remaining county-years were filled with the WUSTL ACAG V6.GL.03 ConvNN $PM_{2.5}$ satellite-fused estimate (van Donkelaar et al., 2021; Shen et al., 2024), a product that fuses TROPOMI/MODIS/MISR satellite columns, GEOS-Chem chemical-transport-model output, and ground-based AQS surface measurements at 0.10° × 0.10° resolution. The satellite product was aggregated to each county by a population-weighted spatial average. The gridded 0.10° concentration field was intersected with U.S. Census TIGER/Line county polygons, and each grid cell within a county was weighted by its gridded population before averaging. This population-weighted aggregation replaces the earlier county-centroid windowing scheme and ties the county estimate to the spatial distribution of population within the county. The resulting dataset comprises 74,568 county-years across 3107 contiguous U.S. counties for 2000–2023.

Validation of the satellite estimates against the monitor-derived county $PM_{2.5}$ for the 11,155 monitor-equipped county-years yielded Pearson $r = 0.889$ and mean absolute difference 1.04 $\mu g\, m^{-3}$ (monitor-derived mean 9.77, satellite-derived mean 9.24), consistent with published validation of the V6.GL.03 product.

County-level demographic estimates come from U.S. Census ACS 5-year tabulations (U.S. Census Bureau, 2024). We use B19013 direct median household income for the income axis, B01003_001E total population as the stratum weight, and B03002 race-by-Hispanic-origin counts for race composition. ACS 5-year vintages for 2009 to 2023 are used directly; pre-2009 county-years carry the closest available (2009–2013) ACS estimate. Six U.S. territories and the District of Columbia (FIPS 60, 66, 69, 72, 78, 11) are excluded.

2.3. Stratification and standard inequality metrics

Counties are stratified by their state's median-household-income distribution at $J-1$ cut points. The primary partition is $J = 3$ at 30/70. The low-income stratum contains the bottom 30% of state population by cumulative population share of median household income, the middle stratum the next 40%, and the high-income stratum the top 30%. Robustness analyses use $J = 4$ (25/50/75) and $J = 5$ (20/40/60/80) partitions reported in Supplementary Table S1. The county is the natural analytic unit here for two reasons. It is the finest spatial scale at which EPA monitor exposures and ACS income estimates are jointly resolved nationally. In addition, income rank is defined relative to each state's own county distribution, which makes the gradient comparable across states. The spatial-scale implications of this choice are taken up in Section 4.

For a state-year with stratum population shares π_g ($g = 1, \ldots, J$, indexed in descending order of median household income, so that stratum rank increases with socio-economic disadvantage) and population-weighted stratum mean concentrations μ_g, the rank-midpoint position of stratum g is

$$R_g = \sum_{h<g} \pi_h + \frac{1}{2}\pi_g, \tag{1}$$

the cumulative population-share midpoint of stratum g. The slope index of inequality (SII) is the population-weighted least-squares slope of stratum mean concentration on stratum rank, and the relative index of inequality (RII) is that slope divided by the mean concentration:

$$\mathrm{SII} = \frac{\sum_g \pi_g (R_g - \bar{R})(\mu_g - \bar{Y})}{\sum_g \pi_g (R_g - \bar{R})^2}, \qquad \mathrm{RII} = \frac{\mathrm{SII}}{\bar{Y}}, \tag{2}$$

where $\bar{R} = \sum_g \pi_g R_g$ and $\bar{Y} = \sum_g \pi_g \mu_g$ are the population-weighted mean rank and concentration. A positive RII in our income-axis sign convention means that the low-income stratum is more exposed. The remaining standard metrics are the relative concentration index (RCI), the between-group variance (BGV), the Theil index, and the mean log deviation (MLD). The RCI is a rank-weighted analogue of the Gini index, and the BGV is the population-weighted variance of the stratum means. The Theil index and the MLD are the two standard generalised-entropy summaries of dispersion. Each is computed following standard definitions (Moreno-Betancur et al., 2015; Wagstaff, 2005; Erreygers, 2009; Theil, 1967).

2.4. The Middle-Group Residual (MGR)

The SII is the population-weighted ordinary-least-squares slope of μ_g on R_g. The linear fit has one residual degree of freedom when there are three strata and one regression slope. The entire deviation of the stratum profile from a straight line is therefore captured by a single signed quantity, the residual at the middle stratum. (A quadratic-versus-linear F-test is degenerate at $J = 3$, since a three-parameter quadratic fits three points exactly, leaving zero residual degrees of freedom.) The Middle-Group Residual indicates, in one number, whether the middle-income stratum lies off the straight line fitted by the SII. We define it at $J = 3$ as that residual, normalised by the population-weighted mean concentration:

Definition 1. For a state-year with $J = 3$ strata indexed in increasing order of R_g,

$$\mathrm{MGR} = \frac{\mu_M - \left[\bar{Y} + \mathrm{SII}\,(R_M - \bar{R})\right]}{\bar{Y}}, \tag{3}$$

where μ_M and R_M are the population-weighted mean concentration and rank-midpoint position of the middle ($g = 2$) stratum.

Appendix A lists and proves four properties of the MGR: (i) dimensionlessness and scale invariance, (ii) zero if and only if the three stratum means lie on a line, (iii) sign convention, and (iv) a positive-scalar-multiple relation between the residual variance from the SII linear fit and $\mathrm{MGR}^2\bar{Y}^2$. Property (iv) formalises the sense in which the MGR captures the entirety of the linear-fit residual at $J = 3$. Population-level prevalence of $|\mathrm{MGR}| > 0.02$ across the primary sample is reported in Section 3.

Worked example. We illustrate with two cases computed from the satellite-augmented $PM_{2.5}$ panel. The population-weighted means for North Carolina 2005 (Fig. 5a) at the 30/70 partition are $\mu_L = 12.78$, $\mu_M = 13.56$, $\mu_H = 14.35\ \mu\mathrm{g\,m^{-3}}$, with $\bar{Y} = 13.56\ \mu\mathrm{g\,m^{-3}}$ and a near-symmetric partition so $\bar{R} \approx 0.5$. Direct computation of Eq. (2) gives $\mathrm{SII} = -2.26\ \mu\mathrm{g\,m^{-3}}$ in our income-axis convention (a mild pro-high-income gradient, meaning the high-income stratum is more exposed). The partition is symmetric, so the middle stratum sits at the mean rank, $R_M - \bar{R} \approx 0$. The SII-predicted middle-stratum mean is therefore $\bar{Y} + \mathrm{SII}\,(R_M - \bar{R}) \approx \bar{Y} = 13.56$. The actual middle-stratum mean is $\mu_M = 13.56$, so the MGR is essentially $(13.56 - 13.56)/13.56 = 0.000$. The three stratum means are collinear, and the SII alone is a faithful summary of the North Carolina 2005 profile. Colorado 2020 (Fig. 5c) shows a different profile. The analogous computation yields $\mathrm{SII} = -2.63\ \mu\mathrm{g\,m^{-3}}$ and $\mathrm{MGR} = +0.080$. The slope is again mildly towards the high-income stratum ($\mu_L = 5.90$, $\mu_M = 7.95$, $\mu_H = 7.76\ \mu\mathrm{g\,m^{-3}}$, $\bar{Y} = 7.33\ \mu\mathrm{g\,m^{-3}}$). However, the middle stratum sits about 8% above the straight-line prediction, indicating a curved profile that the SII alone cannot detect.

The threshold $|\mathrm{MGR}| > 0.02$ flags state-years in which the middle-stratum residual exceeds 2% of the state-year population-weighted mean concentration. The choice of 0.02 is conservative; Supplementary Table S2 reports sensitivity at thresholds 0.01, 0.02, 0.03, and 0.05.

For $J \geq 4$, the regression has $J - 2$ residual degrees of freedom, so a single scalar cannot capture the entire profile-from-linearity deviation. The natural generalisation is the mean absolute interior residual (MAIR):

$$\mathrm{MAIR} = \frac{1}{J-2}\sum_{g=2}^{J-1}\left|\frac{\mu_g - [\bar{Y} + \mathrm{SII}\,(R_g - \bar{R})]}{\bar{Y}}\right|. \tag{4}$$

At $J = 3$ the sum reduces to a single term and $\mathrm{MAIR} = |\mathrm{MGR}|$; Supplementary Table S1 reports MAIR at $J = 4$ and $J = 5$.

2.5. The Within-Stratum Slope Heterogeneity (WSSH)

The MGR captures the residual of the stratum-aggregate means from the SII straight-line fit; it operates on $\{\mu_g\}_{g=1}^{J}$, the population-weighted stratum means, and discards within-stratum variation. The WSSH answers the complementary question of whether income and exposure trend in the same direction inside every stratum or flip direction across the income distribution. Let S_g denote the set of counties assigned to stratum g in a given state-year, y_i county i's annual mean concentration, and m_i its estimated median household income. The within-stratum income–exposure correlation is the rank correlation between county concentration and county income inside the stratum:

$$\rho_g = \mathrm{Spearman}(y_i, m_i)_{i\in S_g}, \tag{5}$$

and the Within-Stratum Slope Heterogeneity (WSSH) is its standard deviation across strata:

Definition 2.

$$\mathrm{WSSH} = \mathrm{sd}_{g\in\{1,\ldots,J\}}(\rho_g), \tag{6}$$

with $\mathrm{WSSH} = 0$ when every stratum has the same internal income–exposure correlation, and increasing as the within-stratum gradient varies in sign or magnitude across strata. We flag a state-year as having a within-stratum sign flip when $\min_g \rho_g < 0$ and $\max_g \rho_g > 0$.

WSSH and MGR together characterise the two ways a single SII number can mislead. MGR flags state-years in which the stratum-aggregate profile departs from a straight line (a between-stratum diagnostic). WSSH flags state-years in which within-stratum gradients vary in sign or magnitude across the income partition (a within-stratum diagnostic). Both are dimensionless and bounded ($|\rho_g| \leq 1$ implies $0 \leq \mathrm{WSSH} \leq 1$), are computable from the same county-level inputs that generate the SII, and reduce to zero when the data satisfy the straight-line, uniform-gradient assumption embedded in the SII.

2.6. Race-stratified replication

We replicate the analysis to assess whether the income-axis findings extend to race-axis disparities. Counties within each state are stratified by their share of non-White population, computed from the ACS B03002 race-by-Hispanic-origin tabulation as one minus the non-Hispanic White share of total population. Counties are ranked by non-White share within each state, and 30/70 cumulative-population-share cuts define low-rank (smallest non-White share, 30% of state population), middle (next 40%), and high-rank (largest non-White share, 30%) strata. The same SII, RII, MGR, and WSSH definitions then apply, with the regression rank ordered in descending non-White share, mirroring the income-axis construction. A positive race-stratified RII therefore indicates that lower-non-White-share counties have higher concentrations.

2.7. Bootstrap and temporal trend inference

We quantify the uncertainty in each state-pollutant-year's metrics with a non-parametric county-resample bootstrap. We draw $B = 1000$ replicates for each panel by resampling its counties with replacement *within each stratum*, holding the stratum membership and population weights fixed. The population-weighted stratum means and all inequality metrics are recomputed on each replicate. The 95% percentile intervals for the SII, RII, and MGR follow from the replicate distribution (Efron, 1979; Davison and Hinkley, 1997). The within-stratum resample requires at least three counties per stratum to avoid degeneracy, so a panel is bootstrap-valid only when every stratum meets that minimum. The bootstrap-valid fraction is 86.7% of state-years on the satellite-augmented $PM_{2.5}$ dataset and 23.6% (NO_2), 45.7% (O_3), and 25.8% (SO_2) on the monitor-only panels. The lower fractions for the gases reflect the sparser county coverage of the gas-phase networks. Among bootstrap-valid panels, the proportion of RII estimates whose 95% interval excludes zero is 32.3% ($PM_{2.5}$), 34.6% (O_3), 2.3% (NO_2), and 3.8% (SO_2). The combustion-driven gases yield noisy per-panel estimates, a point we return to in Section 4.

Temporal-trend regressions on the cross-state-mean annual series (24 observations per pollutant) are fit by ordinary least squares with Newey–West heteroskedasticity-and-autocorrelation-consistent (HAC) standard errors at lag 3 (Newey and West, 1987), which corrects the inference for the autocorrelation inherent in annual environmental series. All p-values reported as "HAC" use this correction.

Table 1
Summary of the primary analytic sample by pollutant, 30/70 partition, 2292 state-pollutant-year panels, 2000–2023.

Pollutant	N	States	$\bar{Y}$	Mean RII	% SII > 0	Mean |MGR|	% |MGR| > 0.02
$PM_{2.5}$ [a,d]	1033	44	8.98	−0.048	33.4	0.041	68.2
NO_2 [b,e]	182	17	11.94	+0.278	79.7	0.127	91.8
O_3 [b,c,e]	772	36	43.39	−0.018	31.7	0.026	49.6
SO_2 [b,e]	305	22	2.30	+0.147	52.5	0.225	88.5

[a] $\bar{Y}$ in μg m^{-3}.
[b] $\bar{Y}$ in ppb.
[c] O_3 is the 8-hour daily maximum annual mean (ppm × 1000).
[d] $PM_{2.5}$ hybrid county dataset: population-weighted aggregation of the V6.GL.03 satellite-fused product, using a quality-passed monitor measurement where available (~15% of county-years) and the satellite estimate elsewhere.
[e] Counties with quality-passed monitors only.
The 30/70 partition assigns counties to three income strata by cumulative population share of state median household income: low-income (bottom 30%), middle (middle 40%), high-income (top 30%).

3. Results

3.1. National patterns and temporal trends

The period-mean gradients divide the four pollutants into two groups. The combustion gases NO_2 and SO_2 show higher exposure in lower-income counties, while $PM_{2.5}$ and O_3 exposure is close to uniform across strata or slightly higher in higher-income counties. Across the primary sample of 2292 state-pollutant-year panels, the cross-state mean relative index of inequality (RII, the SII divided by the mean concentration, positive when the lower-income stratum is more exposed) is +0.278 for NO_2 (17 states) and +0.147 for SO_2 (22 states). The corresponding means are −0.018 for O_3 (36 states) and −0.048 for $PM_{2.5}$ on the population-weighted satellite–monitor hybrid county dataset (44 states). The monitor-equipped subset of the $PM_{2.5}$ panel yields a gradient of the opposite sign (+0.074). We trace that contrast to the spatial coverage of the monitoring network in Section 3.4. Table 1 summarises the analytic sample by pollutant. NO_2 carries the most consistent pro-low-income gradient. O_3 carries the most consistent inverted (pro-high-income) gradient, which reflects the rural geography of peak 8-hour-maximum ozone (Kerr et al., 2024). The $PM_{2.5}$ hybrid gradient is modest, and the next section shows that it depends on which counties enter the analysis.

Newey–West heteroskedasticity-and-autocorrelation-consistent (HAC) trend tests at lag 3 on the cross-state mean annual series reveal four distinct temporal signatures that the period means do not show (Fig. 1). SO_2 shows the largest change, a reversal in the sign of its gradient. The cross-state mean RII rose from −0.121 early in the record (higher-income counties more exposed) to +0.388 late in the record (lower-income counties more exposed). This trend of +0.033 per year is significant under HAC standard errors (HAC $p = 0.003$). Mean |MGR| for SO_2, the typical size of the middle-stratum departure from the SII line, also increased over the same window (HAC $p < 0.0001$). The gradient therefore became steeper and more curved as its sign reversed.

NO_2 shows divergent trends in slope and shape. The pro-low-income NO_2 gradient is narrowing significantly, with mean RII declining by −0.013 per year (HAC $p = 4.5 \times 10^{-5}$). However, the mean |MGR| for NO_2 rose significantly over the same window (+0.0033 per year, HAC $p = 5.1 \times 10^{-7}$). A slope-only summary would report that NO_2 disparities are shrinking. The shape diagnostic shows that the stratum profile became markedly more non-linear over the same period. The two summaries therefore carry independent temporal information.

O_3 shows the opposite combination, with a stable slope and a stratum profile that is becoming more linear. Mean O_3 RII is statistically flat (slope +0.00026 per year, HAC $p = 0.58$), but its mean |MGR| declines significantly (HAC $p < 0.001$). O_3 is therefore the one pollutant for which the linear approximation is becoming a better description over time.

$PM_{2.5}$ trends differ between the two coverage definitions. The cross-state mean RII is statistically flat on the full-county hybrid sample (slope −0.00005 per year, HAC $p = 0.92$). The same series declines significantly on the monitor-equipped subset alone (−0.00089 per year, HAC $p = 0.018$). In contrast, the shape diagnostic increases mildly, with the full-county mean |MGR| rising over the record (+0.00049 per year, HAC $p = 0.005$). We do not find a significant trend in the *signed* full-county $PM_{2.5}$ MGR (slope −0.00012 per year, HAC $p = 0.76$). The gradient is becoming more curved in magnitude without a consistent direction to that curvature.

3.2. State-level heterogeneity

The sign of the income–$PM_{2.5}$ gradient differs sharply from state to state. For most states the sign is a stable feature of local geography and not year-to-year noise. Fig. 2 ranks the 44 states in the hybrid $PM_{2.5}$ sample by their 2000–2023 mean RII. The strongest pro-low-income gradients among states with at least five primary-sample state-years are in California, New Mexico, Wyoming, and Oregon. In these states wildfire smoke, residential wood combustion, and agricultural-dust mixtures place the highest $PM_{2.5}$ concentrations in lower-income counties. Massachusetts is also among the strongest, with dense lower-income urban cores carrying the highest concentrations. The strongest inverted gradients are in Nebraska, Maine, Minnesota, South Dakota, and West Virginia. The urban-industrial centres that carry the highest concentrations in those states are also the highest-income counties.

The state-level direction is persistent. Just over half the states (23 of 44) are pro-high-income in at least 80% of their state-years on record. Only six states, led by California, are pro-low-income with the same consistency. The remaining 15 states near the middle of the ranking switch sign across years. Ten states (Arizona, Georgia, Iowa, Kentucky, Minnesota, Nebraska, North Carolina, Oklahoma, Tennessee, and West Virginia) show an inverted $PM_{2.5}$ gradient in every state-year on record. This uniform inversion reflects the urban-industrial high-income mechanism described above.

The state-level pattern differs across pollutants (Fig. 3). Non-linearity dominates the cell-by-cell map. Most state-pollutant cells with valid 30/70 panels (99 of 119) carry a state-pooled mean |MGR| above the 0.02 threshold. NO_2 (17 of 17) and SO_2 (22 of 22) cells are uniformly above the threshold, and O_3 cells are the least often above it (21 of 36). The income strata also differ in their racial composition, and the degree of overlap varies sharply by pollutant. The gap in non-White population share between the lowest- and highest-income strata is largest for NO_2 and several times smaller for $PM_{2.5}$. This overlap between the race and income axes motivates the direct race-axis analysis in Section 3.5.

3.3. Diagnostic orthogonality

The signed middle-group residual carries directional shape information that the six standard inequality metrics do not. The largest absolute Pearson correlation between the *signed* MGR and any of the six standard metrics (SII, RII, RCI, BGV, Theil index, mean log deviation) is 0.159, pooled across the primary-sample state-pollutant-years. On

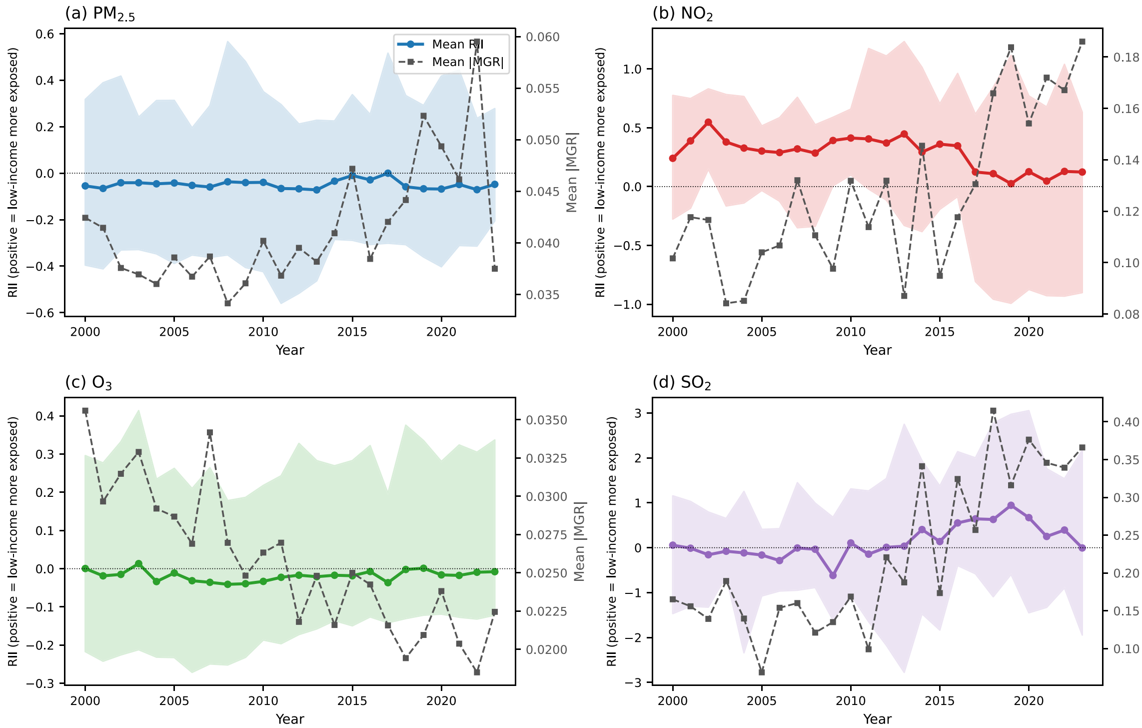

Fig. 1. National-mean air-pollution exposure inequality, 2000–2023, by pollutant. Solid coloured line: cross-state mean RII in each year among state-pollutant-years in the primary sample. Shaded band: 2.5th to 97.5th percentile of RII across states in each year. Dashed grey line (right axis): cross-state mean |MGR|. Horizontal dotted reference line at RII = 0. Panels: (a) $PM_{2.5}$, (b) NO_2, (c) O_3, (d) SO_2. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the satellite-$PM_{2.5}$ panel the corresponding value is 0.176 (Fig. 4). The signed |r| stays below 0.18 in both cases, so more than nine-tenths of the variance in the signed MGR is unexplained by even its most-correlated standard counterpart. The directional component of the MGR is therefore only weakly correlated with the standard slope and spread metrics.

This weak correlation holds only for the *signed* statistic. In absolute value, |MGR| co-moves in magnitude with the generalised-entropy indices (Theil, MLD) at $|r| \approx 0.65$–0.75, because both quantify how far the stratum profile departs from uniformity. The contribution of the MGR is therefore its direction. Its magnitude does track the entropy indices. The entropy indices, however, cannot indicate which side of the SII line the middle stratum falls on. The sign of the MGR shows whether middle-income communities are more or less exposed than a straight line through the extremes would predict.

The within-stratum slope heterogeneity (WSSH, the spread of within-stratum exposure–income correlations) is fully specified by the same county inputs and carries its own substantive signal. Its mean value differs systematically by pollutant (mean WSSH = 0.240 for $PM_{2.5}$, 0.506 for NO_2, 0.412 for O_3, and 0.413 for SO_2). The within-stratum gradient is therefore far from uniform across strata, and the departure from uniformity is strongest for the gas-phase pollutants. Its largest absolute correlation with any between-stratum metric is 0.172 (with the between-group variance), so WSSH carries information substantially distinct from the slope and spread metrics. The within-stratum sign flip is the configuration in which one stratum's internal income–exposure gradient runs opposite to another's. This configuration is common. It occurs in 65.3% of $PM_{2.5}$ panels, 70.9% of NO_2, 62.2% of O_3, and 51.8% of SO_2. WSSH and MGR therefore cover complementary failure modes of a single slope. WSSH detects the within-stratum sign flips; MGR detects the between-stratum curvature. Both failure modes are prevalent across all four pollutants.

The SII linear fit has mean $R^2 = 0.532$ across the primary sample, so on average roughly 47% of the between-stratum variance is left to the middle-stratum residual the MGR measures. The SII alone therefore captures only about half of the stratum profile. Non-linear stratum profiles are the norm. Values of |MGR| exceed the 0.02 threshold in 50%–92% of state-years across the four pollutants at the primary $J = 3$ partition. The exceedance rate rises monotonically at finer partitions (Supplementary Table S1), so the curvature is not an artefact of the three-stratum cut.

The four exemplar profiles in Fig. 5 illustrate the misfit (SII values in $\mu g\,m^{-3}$, income-axis convention; MGR is dimensionless). North Carolina 2005 (panel a) has a mild pro-high-income slope (SII = −2.26) and an effectively linear profile (MGR = 0.000); the three stratum means are collinear, so the SII alone describes it faithfully. California 2009 (panel b) has a strong pro-low-income slope (SII = +5.60) but a middle stratum sitting below the line (MGR = −0.061). Colorado 2020 (panel c) has a mild pro-high-income slope (SII = −2.63) with the opposite shape; the middle stratum sits well above the line (MGR = +0.080). Montana 2020 (panel d) pairs a pro-low-income slope (SII = +1.21) with the largest middle-stratum elevation in the set (MGR = +0.094). A slope-only report would obscure the shape pattern in three of these four cases. The SII, MGR, and WSSH together preserve the slope, the curvature, and the within-stratum consistency.

3.4. Monitor-network selection bias

The direction of the estimated inequality depends on the spatial coverage of the $PM_{2.5}$ monitoring network, not only on ambient concentrations. The cross-state mean $PM_{2.5}$ gradient reverses sign when unmonitored counties are added to the analysis. The monitor-equipped subset (the roughly 15% of county-years with quality-passed AQS monitors) gives a cross-state mean RII of +0.074 (at a population-weighted mean concentration $\bar{Y} = 10.29\ \mu g\,m^{-3}$), indicating that lower-income counties are more exposed. The population-representative hybrid value, obtained by filling the unmonitored counties with the satellite-derived estimate and population-weighting, is −0.048, indicating that higher-income counties are more exposed. Under a pure-satellite reference

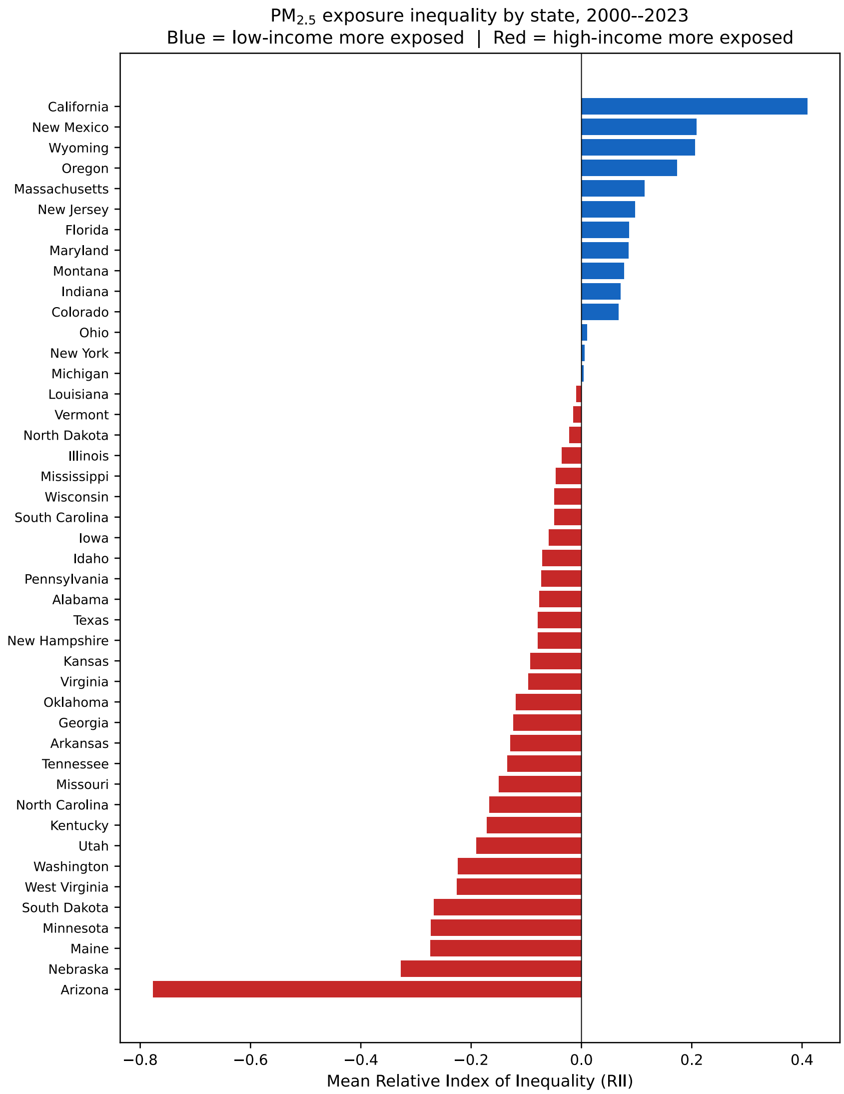

Fig. 2. Mean $PM_{2.5}$ RII by state, 2000–2023, hybrid monitor–satellite dataset. Blue bars: mean RII > 0 (low-income stratum exposed more). Red bars: mean RII < 0 (high-income stratum exposed more). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the value moves slightly further, to −0.082. The monitor-only and population-representative estimates differ by 0.122 RII units and span zero. The two coverage samples therefore yield opposite-signed gradients over the same 24-year period. Fig. 6 shows the state-level scatter (panel a), the state coverage by data source (panel b), and the persistent annual offset between the two cross-state mean trajectories (panel c).

The reversal reflects the structure of the measurement network. It is not a property of the pollutant. Regulatory $PM_{2.5}$ monitors are sited where population and infrastructure concentrate (Wang et al., 2024; Haskell-Craig et al., 2025), and within the monitored counties the highest concentrations occur in lower-income urban cores. The unmonitored counties are disproportionately rural, lower-income, and lower-$PM_{2.5}$. A monitor-only analysis must drop these counties, and dropping them biases the cross-state gradient towards pro-low-income. Monitor-only samples underlie most of the published U.S. literature. The discrepancy between the two estimates is not a matter of precision. The two coverage definitions describe inequality for different county populations, and a monitor-only analysis cannot recover the population-representative estimate without the satellite fill.

The temporal trends show the same split. The $PM_{2.5}$ RII on the monitor-equipped subset declines significantly (−0.00089 per year, HAC p = 0.018) and describes inequality dynamics within monitored counties. The slope across the full county universe is flat (−0.00005 per year, HAC p = 0.92). The signed full-county MGR shows no significant trend (slope −0.00012 per year, HAC p = 0.76), so we do not claim a directional shift in the full-county gradient shape. The magnitude of the curvature is mildly increasing (mean |MGR| slope +0.00049 per year, HAC p = 0.005), but its sign is not consistently moving one way.

3.5. Race-stratified gradients

We re-ran the analysis with counties ordered by racial composition instead of income. The race-stratified results confirm that higher-non-White-share counties bear higher $PM_{2.5}$ and NO_2 exposure, while SO_2 reveals a divergence between the two axes that a single slope cannot show. Counties are partitioned within each state by non-White share using the same 30/70 population-weighted cuts (Methods). Under the race-axis sign convention, a positive RII means that lower-non-White-share counties bear higher mean exposure. Mean race-stratified RII is −0.160 for $PM_{2.5}$ (N = 1034 panels, 44 states), −0.456 for NO_2 (N = 105, 6 states), +0.053 for O_3 (N = 566, 28 states), and +0.336 for SO_2 (N = 168, 13 states). Table 2 gives the full breakdown. The negative $PM_{2.5}$ and NO_2 values place higher exposure on higher-minority counties, consistent with prior satellite-fused $PM_{2.5}$ (Jbaily

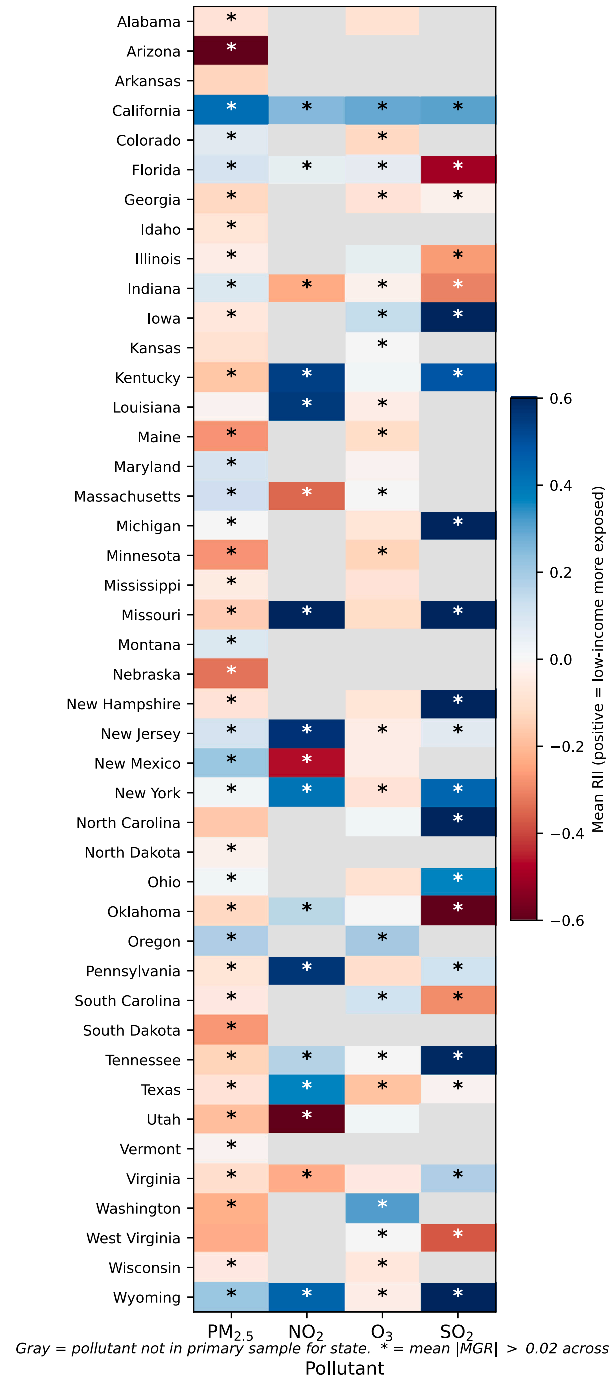

Fig. 3. State-by-pollutant mean RII, 2000–2023. Colour: signed RII magnitude (blue = pro-low-income; red = pro-high-income). Stars mark state-pollutant cells with state-pooled mean |MGR| > 0.02. Grey cells indicate state-pollutant pairs not in the monitor-equipped sample. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

et al., 2022; Colmer et al., 2020; Daouda et al., 2022) and NO_2 (Kerr et al., 2024) work. The slightly positive O_3 value reflects the rural ozone geography; predominantly White counties on average lie in higher-ozone rural regions.

SO_2 is the informative case for the cross-axis comparison. Its income-stratified RII is +0.147 (lower-income counties more exposed) and its race-stratified RII is +0.336 (lower-non-White-share counties more exposed). The residual SO_2 sources surviving the utility-coal phaseout therefore lie in lower-income but predominantly White rural counties. An analysis on either axis alone would miss that the income burden and the racial burden fall on different communities. The relationship between the two axes also differs by pollutant across states. The state-level correlation between race-stratified and income-stratified mean RII is small for $PM_{2.5}$, modest for O_3, moderately negative for NO_2, and moderately positive for SO_2. The negative correlation reflects a doubly-disadvantaged pattern; states where higher-minority counties bear more NO_2 tend also to be states where lower-income counties do. The positive correlation matches the income-versus-race divergence described above.

The shape diagnostics are informative on the race axis as well. The fraction of race-stratified panels with positive RII (lower-non-White-share counties more exposed) is 16.6% for $PM_{2.5}$, 21.9% for NO_2, 72.4% for O_3, and 64.9% for SO_2. These fractions are consistent in sign with the mean-RII direction in each case. Race-stratified |MGR| exceeds the 0.02 threshold in roughly half of $PM_{2.5}$ panels, four-fifths of NO_2 panels, and the large majority of SO_2 panels. Non-linear stratum profiles are therefore as common on the race axis as on the income axis. Within-stratum sign flips also recur on the race axis. The same two failure modes of a single slope, between-stratum curvature and within-stratum sign reversal, arise whether counties are ordered by income or by racial composition.

4. Discussion

Two empirical findings anchor this discussion. The first concerns *where* pollution is measured, and the second concerns *when* a gradient changes sign. First, the cross-state mean $PM_{2.5}$ income gradient reverses direction depending on which counties enter the analysis. The relative index of inequality (RII) is the SII slope divided by the mean concentration, and positive values mean the lower-income stratum is more exposed. The monitored subset, the ~15% of county-years that host a quality-passed regulatory monitor, gives an RII of +0.074, a pro-low-income gradient. The satellite-derived concentration fills the unmonitored counties, so every contiguous-state county is represented. This full county sample gives −0.048, a pro-high-income gradient. The two numbers are computed from the same state-years over the same 24-year window. Only the county coverage differs, and the sign of the estimated disparity flips between them. Second, the cross-state mean SO_2 gradient itself reversed direction over the study period, from pro-high-income in the early years to pro-low-income by the end (HAC $p = 0.003$). A single slope reported on a single county sample reveals neither finding. Each finding has a concrete physical cause in source and measurement geography, and neither reflects the choice of statistic.

The $PM_{2.5}$ coverage reversal isolates monitor-network selection bias as a structural feature of U.S. air-pollution inequality measurement. Monitor-only analyses dominate the published U.S. environmental-justice literature. The direction sensitivity we document therefore implies that monitor-based and satellite-augmented $PM_{2.5}$ inequality estimates answer different questions about exposure inequality. The two designs do not estimate the same quantity with greater or lesser precision. The cause is physical. Regulatory $PM_{2.5}$ monitors are sited where population and infrastructure concentrate, and only about one in five U.S. counties hosts one (Haskell-Craig et al., 2025; Wang et al., 2024). The highest concentrations in the monitored counties co-locate with lower-income urban cores, so the monitored sample yields a pro-low-income gradient. The counties the network misses lean rural, lower-income, and lower in $PM_{2.5}$. The satellite surface restores these counties and pulls the population-representative gradient towards pro-high-income. Haskell-Craig et al. (2025) find the urban monitor network itself reasonably equitable but document a rural-poverty coverage gap, which is precisely the gap our satellite fill closes. Population weighting of the satellite estimate attenuates the hybrid pro-high-income signal (the hybrid RII is −0.048, weaker than a pure satellite −0.082) but does not remove it. The monitored-versus-population gap

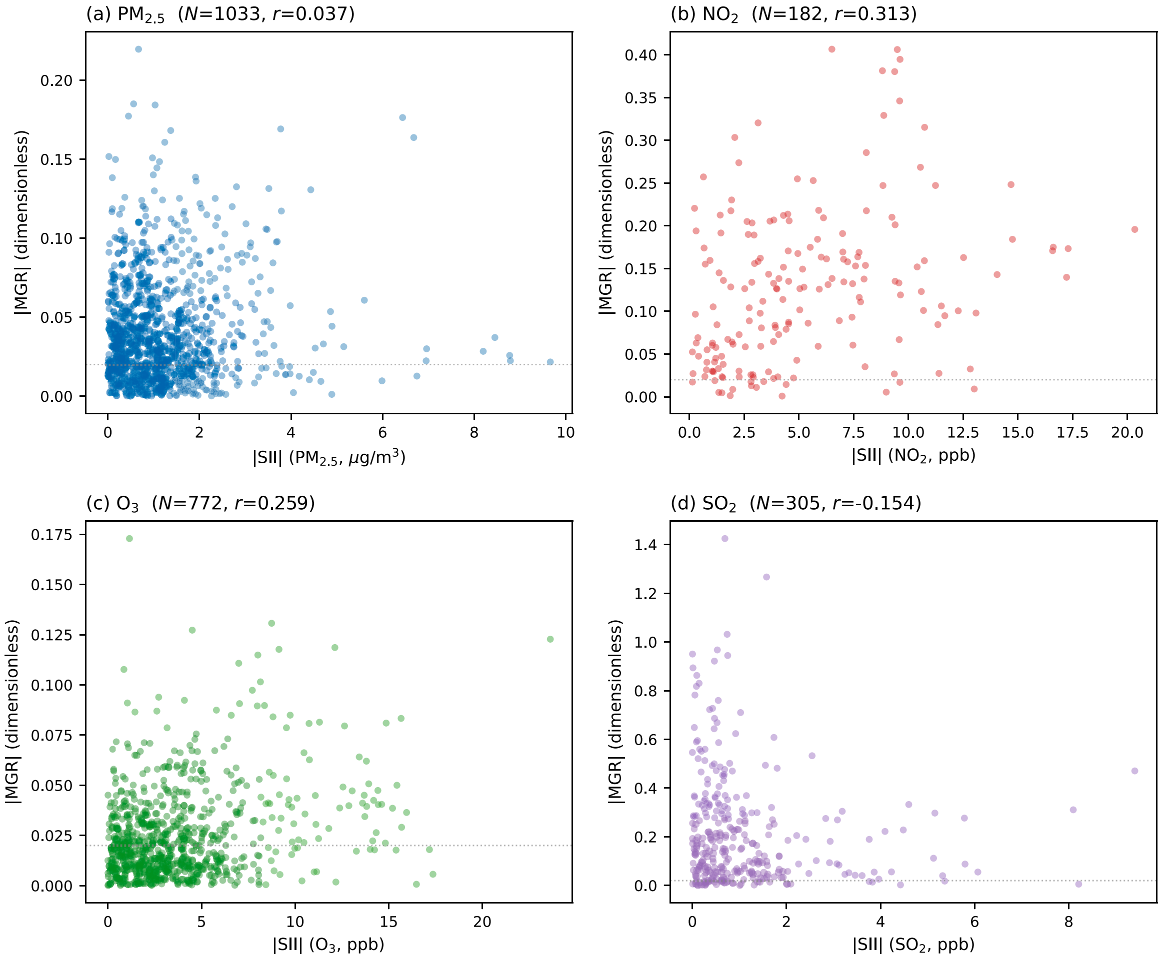

Fig. 4. Per-state-year joint distribution of |SII| and |MGR| for each pollutant. Dotted horizontal line: |MGR| = 0.02 threshold. Pearson correlations are shown in each panel title.

Table 2
Summary of the race-stratified analytic sample by pollutant, 30/70 partition on cumulative non-White population share, 1873 state-pollutant-year panels, 2000–2023.

Pollutant	N	States	Mean RII	% RII > 0	Mean |MGR|	% |MGR| > 0.02
$PM_{2.5}$[a]	1034	44	−0.160	16.6	0.032	54.7
NO_2[b]	105	6	−0.456	21.9	0.119	90.5
O_3[b]	566	28	+0.053	72.4	0.021	42.9
SO_2[b]	168	13	+0.336	64.9	0.235	97.0

[a] $PM_{2.5}$ hybrid (population-weighted monitor–satellite) county dataset.

[b] Counties with quality-passed monitors only.

The income- and race-axis $PM_{2.5}$ panels differ by one state-year (1033 vs. 1034). Arizona 2010 fills all three strata under the non-White-share ranking but leaves the highest-income stratum empty under the income ranking (a consequence of the 30/70 population-share cut and the concentration of the state's population in Maricopa and Pima counties). This state-year is therefore retained on the race axis only.

Sign convention: positive race-stratified RII indicates lower-non-White-share counties have higher concentrations. Negative RII indicates higher-non-White-share counties have higher concentrations. The 30/70 partition assigns counties to three race-rank strata by cumulative population share of state non-White share: low-rank (smallest non-White share, 30% of state population), middle (next 40%), and high-rank (largest non-White share, 30%).

of ~0.12 RII units still spans zero. The U.S. $PM_{2.5}$ income disparity is therefore modest and measurement-dependent. The sign of the reported disparity depends on which dataset the analyst treats as primary, not on the size of the effect.

The four pollutants split into two groups on the monitored sample. NO_2 and SO_2 show positive cross-state mean RII, with lower-income strata bearing higher mean exposure where monitors are present. O_3 shows the inverted pattern characteristic of rural-skewed peak ozone, where whiter, higher-income rural counties sit in higher-ozone regions (Kerr et al., 2024). $PM_{2.5}$ on the population-representative county dataset is pro-high-income (−0.048) and reverses to +0.074 on the monitored subset, as above.

These results extend the published U.S. literature rather than overturning it. Jbaily et al. (2022) reported a small national-mean $PM_{2.5}$ RII. This result is consistent with our finding that the monitor-only gradient direction is close to evenly split between positive and negative at the state-year level. Colmer et al. (2020) found that absolute $PM_{2.5}$ disparities fell over 1981–2016 while relative disparities persisted. Our cross-state mean $PM_{2.5}$ RII on the monitored subset declines modestly over the period (slope −0.00089 per year, HAC $p = 0.018$) but remains

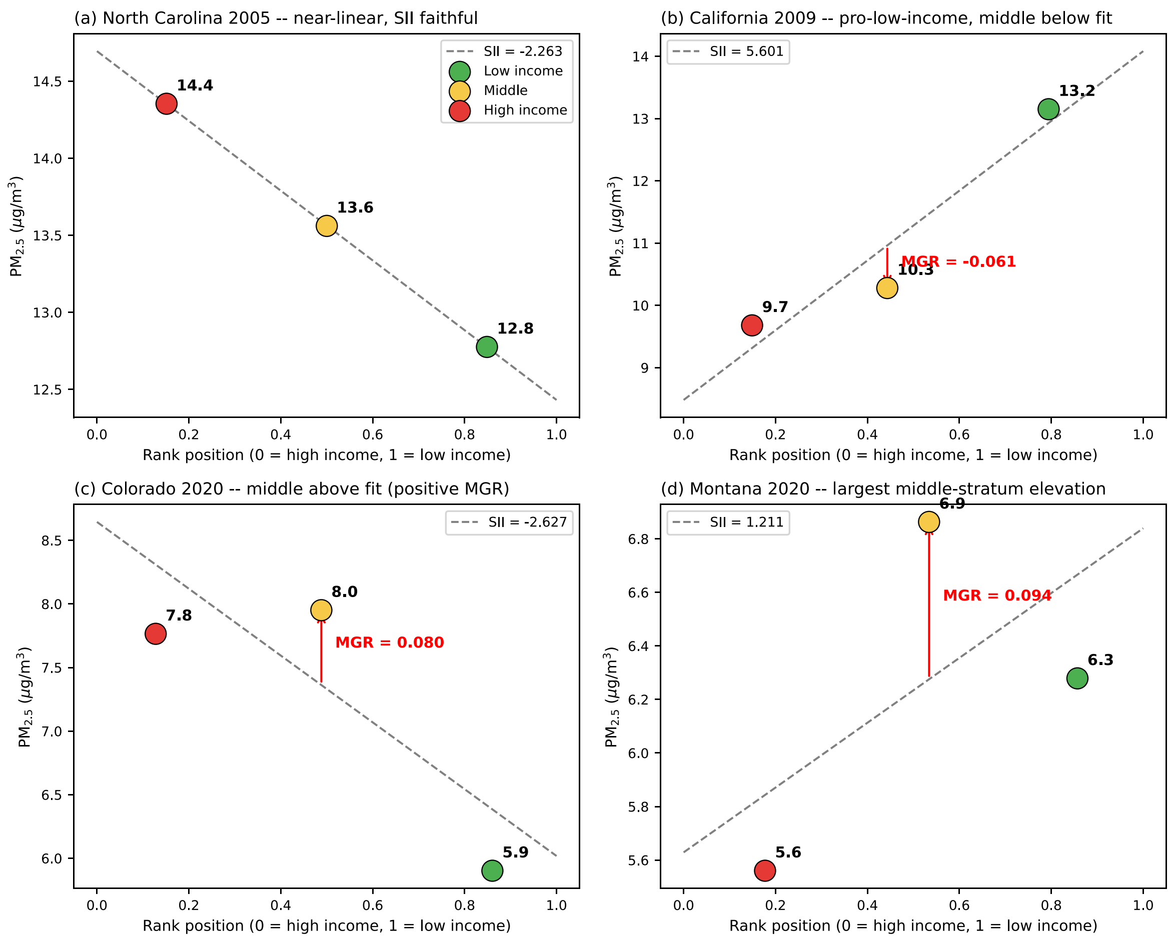

Fig. 5. Four $PM_{2.5}$ state-year exemplar stratum profiles drawn from the primary sample. Dashed grey line: SII linear fit. Three coloured points: low-income (green), middle-income (yellow), high-income (red) stratum means μ_L, μ_M, μ_H, plotted against stratum rank-midpoint R_g. Red arrow: signed MGR residual where the middle-stratum point lies off the SII line. (a) North Carolina 2005: near-linear stratum profile, MGR = 0.000, the three means collinear so the SII is a faithful summary. (b) California 2009: strong pro-low-income gradient with the middle stratum sitting *below* the SII line (MGR = −0.061). (c) Colorado 2020: pro-high-income slope with the middle stratum sitting *above* the SII line (MGR = +0.080). (d) Montana 2020: largest middle-stratum elevation in the set, with the middle stratum well above the line (MGR = +0.094). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

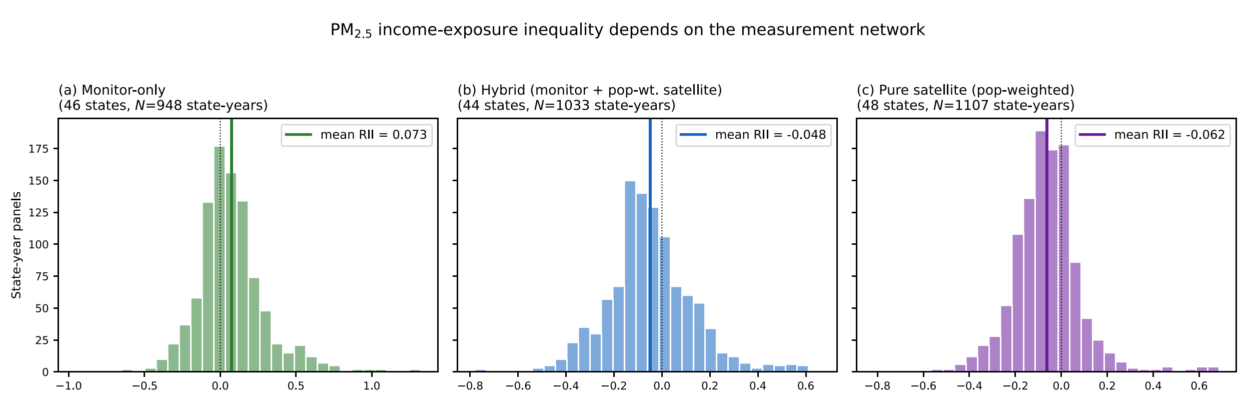

Fig. 6. $PM_{2.5}$ cross-panel comparison: monitor-only sample, satellite-augmented county dataset (primary), and pure-satellite reference. (a) State-level mean RII, monitor-only versus hybrid. (b) Cross-state mean RII for each panel: monitor-only (+0.074, pro-low-income), hybrid (population-weighted monitor–satellite; −0.048, the primary analytic dataset, pro-high-income), and pure population-weighted satellite (−0.082). The monitored subset and the population-representative sample give opposite-signed gradients. (c) Annual cross-state mean RII trajectory under each panel.

pro-low-income in every year of the window. The relative disparity therefore persists, narrowing slowly, against a falling absolute baseline. Kerr et al. (2024), using satellite-fused NO_2, reported widening racial NO_2-attributable health-burden disparities. Our monitor-based NO_2 results add a complementary income-axis finding. We find a 16.1 percentage-point minority-share gap at the income extremes and a cross-state mean RII that declines significantly over the study window (slope −0.013 per year, HAC $p < 0.0001$). The shape of the NO_2

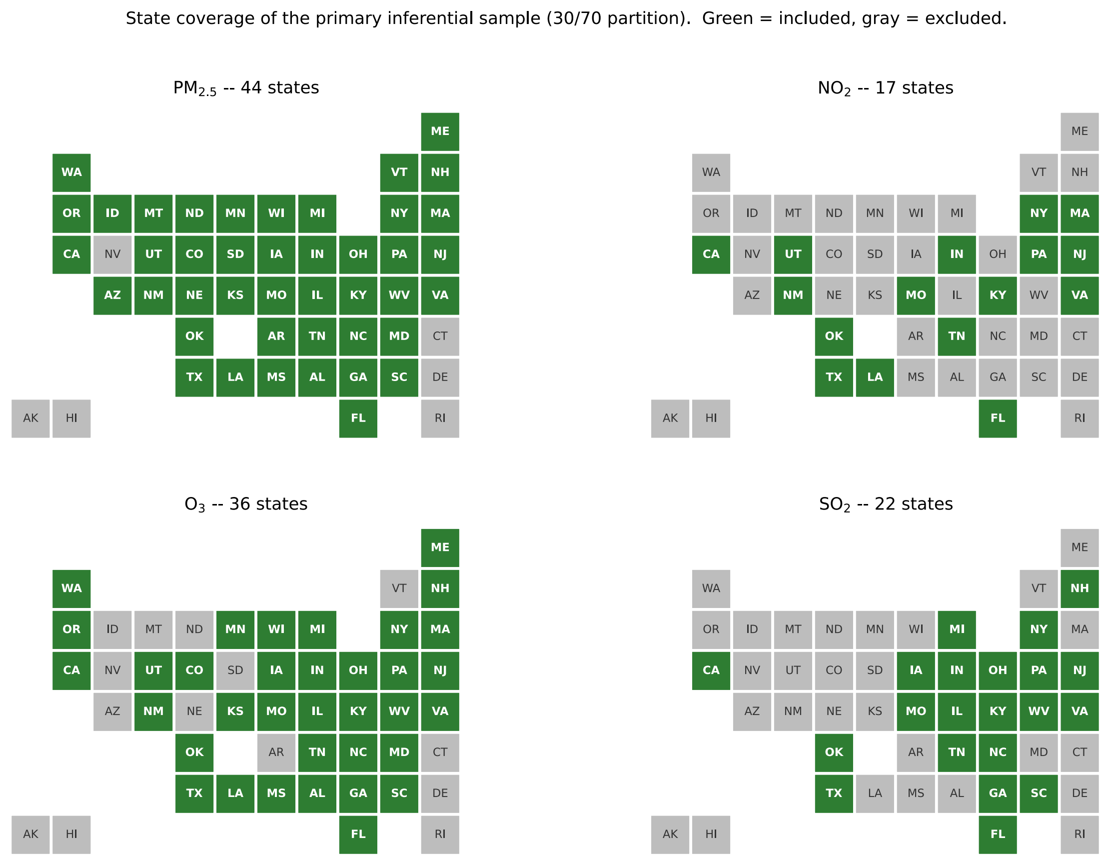

Fig. 7. Primary inferential sample coverage by pollutant (4-panel state tile-grid layout). Green: state included in the primary analytic panel. Grey: state excluded for insufficient quality-passed monitor coverage. $PM_{2.5}$ (top-left) is the hybrid monitor–satellite dataset and covers 44 states. NO_2 (top-right) covers 17 states, concentrated on the Eastern seaboard plus California, Texas, and Wyoming, and is demographically distinctive (47.2% non-White, 24.5% Hispanic population-weighted versus 34.2% / 12.0% in the excluded states). O_3 (bottom-left) covers 36 states with intermediate demographic tilt (about 42% non-White among included states). SO_2 (bottom-right) covers 22 states with an intermediate demographic tilt (about 44% non-White among included states versus about 35% excluded). The SO_2 sign-reversal finding is temporal (a within-state change over 2000–2023) and is not threatened by this cross-sectional coverage tilt. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

gradient moved in the opposite direction. Mean |MGR|, the size of the middle-income stratum's departure from the SII line, rose significantly over the same window (slope +0.0033 per year, HAC $p < 0.0001$). A slope-only summary would describe NO_2 disparities as a steadily shrinking pro-low-income gradient. The shape diagnostic shows the stratum profile becoming markedly more non-linear at the same time, so slope and shape carry independent temporal information here.

The shape gap the MGR fills has been handled elsewhere by changing the outcome scale. The OHID technical guidance (Office for Health Improvement and Disparities, 2024) acknowledges that the gradient often flattens or bends in the middle stratum. Its remedy is to log- or logit-transform the concentration before fitting the SII. This preserves one reportable slope but removes the evidence of the bend from the reported number. Our contribution is to keep the original concentration scale and report the middle-stratum misfit directly. The OHID log transform reduces middle-stratum residuals for $PM_{2.5}$, NO_2, and O_3 (mean |MGR| falls on the log scale in 97%–99% of panels). It amplifies them for SO_2, whose near-zero concentrations make log-transformation counterproductive. We did not discover that the demographic–exposure relationship is curved; flexible spline regression has already shown this for $PM_{2.5}$ (Daouda et al., 2022). Our contribution is a named, dimensionless statistic that flags the curvature wherever the slope is reported, without refitting a model.

The SO_2 temporal sign reversal is the largest single-pollutant change in the record. The cross-state mean RII shifted from pro-high-income in the early period to pro-low-income by the end (early-period mean −0.121, late-period mean +0.388; slope +0.033 per year, HAC p = 0.003). Mean |MGR| rose over the same window. The likely mechanism is physical and regulatory. The federal utility-coal phaseout under the Mercury and Air Toxics Standards (U.S. Environmental Protection Agency, 2012) and the 2010 SO_2 NAAQS revision (U.S. Environmental Protection Agency, 2010) cleared the highest-exposure urban-industrial sites disproportionately quickly (Henneman et al., 2023b). Henneman et al. (2023a) document that coal-power-plant $PM_{2.5}$ exposure has been inequitably distributed for over two decades, with the geographic incidence shifting as plants scrubbed or retired. The residual SO_2 geography we observe is consistent with that shift. Emissions now concentrate near remaining chemical-manufacturing, metals-smelting, and pulp-and-paper point sources whose host counties lean rural and lower-income. Koolik et al. (2025) argue that emission reductions alone cannot eliminate exposure disparities and can shift their demographic incidence. Geldsetzer et al. (2024) describe the analogous burden-shift in $PM_{2.5}$-attributable mortality, with the largest absolute improvements in already-high-pollution areas and relative disparities widening against a smaller absolute baseline. Our result aligns with both accounts. The race axis shows a further divergence. SO_2's income-stratified RII is +0.147 (lower-income counties more exposed), but its residual burden falls on lower-income but predominantly White rural counties (race-stratified RII +0.336, lower-non-White-share counties more exposed). Single-axis surveillance cannot reveal this divergence.

Several mechanisms contribute to the state-level heterogeneity in $PM_{2.5}$ direction. The states with the strongest pro-low-income gradients (California, New Mexico, Wyoming, Oregon, and Massachusetts) have geographies in which wildfire smoke, residential wood combustion, agricultural dust, and dense lower-income urban cores place peak $PM_{2.5}$ in lower-income counties. The rising wildfire burden over the study period (Burke et al., 2021) plausibly contributes to the upward $PM_{2.5}$ RII trend in the Western states among them. In contrast, ten states show pro-high-income gradients in every state-year on record. In these states, the urban and industrial centres that carry the highest $PM_{2.5}$ also tend to contain the highest-income counties. A formal demographic-geographic decomposition of this reversal is left for future work. NO_2's distinctive race-composition gap reflects its traffic-source origin and the legacy of twentieth-century urban policy. Lane et al. (2022) document that historical Home Owners' Loan Corporation neighbourhood grades remain associated with present-day air-pollution disparities through highway placement and housing-market sorting. Tessum et al. (2021) show that essentially all major $PM_{2.5}$ source sectors expose non-White populations above their population share.

Global applicability. The two blind spots we document are properties of the measurement method. They are not properties of U.S. geography, so they transfer to any jurisdiction that summarises exposure inequality with a regression slope. Disproportionate $PM_{2.5}$ exposure is not a uniquely American phenomenon. Across 211 countries the heaviest burdens fall on lower-income populations, and four in five over-exposed people live in low- and middle-income nations (Rentschler and Leonova, 2023). Global $PM_{2.5}$ inequality has also risen over 2000–2020 (Sager, 2025). Both blind spots are most consequential where that burden concentrates. The shape blind spot is acute because the global $PM_{2.5}$–income profile bends. Rentschler and Leonova (2023) find pollution peaking at *middle* income, which is a middle-elevated, non-linear stratum profile and a positive-MGR signature at the cross-country scale. Our middle-group diagnostic is built to detect exactly this pattern. The coverage blind spot is acute because ground-monitor networks are sparse across much of the low- and middle-income world. Satellite-fused surfaces therefore become the only practical exposure substrate in those settings, and the coverage-driven gradient reversal we document for U.S. counties becomes a central concern. The framing also transfers. Global inequality splits into a between-country and a within-country share (Sager, 2025), and a single gradient likewise splits into a slope and a shape. Current reporting discards the shape component. We use the same satellite $PM_{2.5}$ surface that underpins global exposure accounting (van Donkelaar et al., 2021). The shape and coverage diagnostics introduced here are therefore directly portable to any national surveillance program built on that data.

Spatial scale (MAUP). Our analytic unit is the county. Any statistic computed from spatially aggregated data depends on the size and boundaries of the reporting units. This dependence is the modifiable areal unit problem (MAUP), which comprises a scale effect (changing unit size by aggregation) and a zoning effect (redrawing boundaries at fixed unit count) (Parenteau and Sawada, 2011). We adopt the county because it is the resolution at which monitor-derived concentrations and ACS income and race tabulations are jointly available across all states and all study years. Income stratification is also defined relative to each state's distribution, and this definition requires a unit nested within states. The direction of the resulting bias is known. The bias runs towards attenuation, so our county-scale estimates are conservative. Aggregation chiefly compresses within-unit variance while leaving unit means largely intact, and it inflates measured between-area correlations (Parenteau and Sawada, 2011). The SII, RII, and MGR are functions of stratum means and do not depend on within-stratum spread, so they are the quantities least disturbed by coarsening. A finer tract- or ZIP-level analysis would therefore be expected to strengthen the gradients we report, not to weaken them. Empirically, county aggregation has been shown to understate tract-level $PM_{2.5}$ exposure disparities by roughly 20%, with tract-level estimates close to block-level (Clark et al., 2022). For the shape diagnostic, we expect that finer spatial units would raise the |MGR| exceedance rates, though we do not demonstrate this here. A sub-county analysis would require a continuous tract- or ZIP-level concentration surface for 2000–2023, which we do not assemble here. A formal scale-sensitivity analysis is left to future work.

Several further limitations bound the scope of these conclusions. The monitor-only panels for NO_2, O_3, and SO_2 restrict to counties meeting our quality-control filter chain (17–36 states; Fig. 7 maps state coverage by pollutant). The 17-state NO_2 sample covers 56.6% of the U.S. population but is demographically distinctive. Population-weighted, the included states are 47.2% non-White and 24.5% Hispanic, versus roughly 34% non-White in the 33 excluded states. The 16.1-percentage-point race-composition gap at the income extremes should therefore be read as a characterisation of states with adequate NO_2 monitor coverage. It is not a fully representative national estimate. The 22-state SO_2 sample carries an intermediate demographic tilt. Its included states are about 44% non-White versus about 35% in the excluded states, a roughly 9-percentage-point gap. That tilt, however, does not threaten the SO_2 finding. The SO_2 sign reversal is a *temporal* result, a within-state change in the income gradient over 2000–2023 among the covered states, identified from each state's own trajectory. A fixed cross-sectional difference in the covered states' demographics cannot generate a trend in the gradient over time. The 36-state O_3 sample is also intermediate, at about 42% non-White among included states. The race-stratified NO_2 panel is narrower still (6 states, 105 panels). Its −0.456 mean RII is suggestive and should not be read as a national point estimate. Equivalent satellite expansion for NO_2 (Kerr et al., 2024; Larkin et al., 2017), SO_2, and O_3 is left to future work because of access and spatial-coverage constraints on the relevant satellite-fused products. The per-state-year SII point estimate is also noisy for the combustion-driven pollutants. Only 2.3% of NO_2 and 3.8% of SO_2 county-resample bootstrap intervals exclude zero, so individual state-year SII values for these pollutants should not be over-interpreted in isolation.

A demographic-data limitation applies to the early years. ACS 5-year tabulations begin in 2009, so county-years from 2000 to 2008 carry the 2009-vintage income, population, and race estimates (Section 2.2). The exposure series is unaffected (the EPA and satellite concentrations are observed per year throughout), but the income and race *stratification labels and weights* are held constant over 2000–2008. The temporal trends we report are trends in exposure-based RII and |MGR| and remain valid. A reader should not interpret any apparent demographic-composition change before 2009 as real, since it is flat by construction. Finally, the 3-stratum partition is coarser than the decile construction of Jbaily et al. (2022) or the quintile construction of Kerr et al. (2024). The MGR has a clean closed form only at $J = 3$, with $J \geq 4$ handled by the mean absolute interior residual (Supplementary Table S1). A related coarseness operates within counties, whose mean exposures aggregate over sub-county heterogeneity. Sub-county mobile-monitoring work shows that within-block variability can rival between-county variability (Apte et al., 2017), so the classic ecological-fallacy caution (Morgenstern, 1995) applies as discussed above.

A health-burden translation puts the measurement dependence in mortality terms and indicates the policy stakes. We translate the income gradient via the GBD 2019 $PM_{2.5}$ integrated exposure–response function (Cohen et al., 2017; GBD 2019 Risk Factors Collaborators, 2020), against the 2021 WHO Air Quality Guideline annual mean of 5 $\mu g\, m^{-3}$ (World Health Organization, 2021). The +0.074 monitored-subset RII at $\bar{Y} = 10.29$ $\mu g\, m^{-3}$ implies an approximately 3% relative-risk excess for the low-income stratum within monitored counties. The population-representative dataset gives the opposite-signed −0.048 at $\bar{Y} = 8.98\ \mu g\,m^{-3}$, placing the income-stratified attributable burden on average on the high-income stratum. Both readings, applied to the roughly 200,000 annual $PM_{2.5}$-attributable adult deaths in the contiguous U.S. Cohen et al. (2017) and Geldsetzer et al. (2024), imply on the order of 5000–10,000 excess deaths per year[2] carried disproportionately by one income stratum. The direction of that disparity depends on which dataset the analyst selects as primary. A formal disparity-attribution analysis with explicit counterfactual specification and uncertainty propagation is left for future work.

5. Conclusions

A single slope summary of an exposure–income gradient misses two distinct features. We introduce two dimensionless shape diagnostics that extend the slope index of inequality (SII): the Middle-Group Residual (MGR) and the Within-Stratum Slope Heterogeneity (WSSH). The MGR flags state-years in which the middle-income stratum departs from the fitted SII line. The WSSH flags state-years in which the direction of the income–exposure gradient varies across the income partition. Both diagnostics are computable from the same county inputs that yield the SII, and both are largely orthogonal to the six standard between-stratum inequality metrics. Non-linear stratum profiles, flagged by $|\mathrm{MGR}| > 0.02$, occur in 50%–92% of state-years across the four pollutants at the 3-stratum partition. This exceedance rate rises monotonically to 68%–100% at the 5-stratum partition. The SII linear fit captures only about half of the between-stratum variance (mean $R^2 = 0.532$), leaving roughly 47% to the MGR. The WSSH criterion for within-stratum sign flips is met in 52%–71% of panels across the four pollutants.

Two empirical findings are independent of the diagnostic methodology. First, the cross-state mean $PM_{2.5}$ income gradient reverses sign between the monitor-equipped subset of U.S. counties (RII +0.074, pro-low-income) and the population-representative full county sample under satellite augmentation (RII −0.048, pro-high-income). This reversal is a direct signature of monitor-network selection bias and has not previously been quantified at continental scale in U.S. environmental-justice work. The population-weighted satellite–monitor hybrid yields the population-representative estimate by including the unmonitored counties in the analysis. Second, the cross-state mean SO_2 gradient reversed direction over 2000–2023, from pro-high-income to pro-low-income (HAC $p = 0.003$). This inversion is plausibly driven by the post-2010 utility-coal phaseout and by the geographic concentration of residual SO_2 near rural, lower-income point sources.

The shape blind spot and the coverage blind spot are properties of the measurement method, not of U.S. geography. They therefore apply to any jurisdiction running regression-based exposure-inequality surveillance. They apply most acutely to monitor-sparse low- and middle-income settings, where the global exposure burden concentrates. We recommend that continental-scale exposure-inequality surveillance report the middle-stratum residual and the within-stratum heterogeneity alongside the slope. We also recommend that such surveillance state the monitoring network's coverage explicitly. Together, these quantities document the profile curvature and the coverage dependence that a slope summary alone omits.

[2] ≈3% excess risk (from the +0.074 monitored-subset RII via the GBD 2019 $PM_{2.5}$ exposure–response function relative to the WHO 5 $\mu g\,m^{-3}$ guideline) applied to the ≈200,000 annual $PM_{2.5}$-attributable adult deaths in the contiguous U.S. (Cohen et al., 2017; Geldsetzer et al., 2024) gives ≈6000 deaths; a 2.5%–5% excess-risk range spans the 5000–10,000 band.

CRediT authorship contribution statement

Tom Regpala: Writing – original draft, Software, Formal analysis, Data curation. **Eric Palafox:** Writing – original draft, Formal analysis, Data curation. **Anjana Yatawara:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of Generative AI and AI-assisted technologies in the writing process

During the preparation of this work the authors used Claude (Anthropic) to improve the readability and language of the manuscript. After using this tool, the authors reviewed, verified, and edited the content as needed and take full responsibility for the content of the publication.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research was supported by the Center for Environmental Studies (CES) Mini-Grant at California State University, Bakersfield (2025 cycle). The funder had no role in study design, data analysis, interpretation, writing, or the decision to submit.

Appendix A. Mathematical properties of MGR and WSSH

A.1. Properties of the MGR at $J = 3$

Proposition 1. *At $J = 3$, with π_g, R_g, and μ_g defined as in Section 2.3 and the SII as the population-weighted OLS slope of μ_g on R_g (Eq. (2)), the MGR (Definition 1) satisfies:*

(i) MGR is dimensionless, invariant under multiplicative rescaling of all μ_g, and its sign is interpretable (Property (iii)). Its numerator, the middle-stratum residual e_M, is invariant under uniform additive shifts of μ_g; the MGR itself is not, because the normalising $\bar{Y}$ shifts with the data.

(ii) $\mathrm{MGR} = 0$ if and only if the three stratum means μ_g lie on a line in R_g.

(iii) $\mathrm{MGR} > 0$ indicates the middle-income stratum lies above the SII linear prediction (a convex departure when the gradient is pro-low-income); $\mathrm{MGR} < 0$ indicates the middle stratum lies below the SII line.

(iv) The residual sum of squares from the SII linear fit is determined entirely by MGR: $\sum_g \pi_g e_g^2 = c(\boldsymbol{\pi})\,\mathrm{MGR}^2\bar{Y}^2$, where $c(\boldsymbol{\pi}) > 0$ depends only on the partition population shares π_g and the corresponding rank-midpoints R_g. For a symmetric partition ($\pi_L = \pi_H$) the constant has the closed form $c(\boldsymbol{\pi}) = \pi_M/(1-\pi_M)$, which equals $2/3$ for the 30/40/30 partition used in the primary analysis.

Proof of (i). The numerator of Eq. (3) is the residual $e_M = \mu_M - \left[\bar{Y} + \mathrm{SII}\,(R_M - \bar{R})\right]$ at the middle stratum, which carries the same units as μ_g; dividing by $\bar{Y}$ removes those units, so the MGR is dimensionless. A proportional rescaling $\mu_g \to \lambda\,\mu_g$ (any $\lambda > 0$) scales both the fitted line and $\bar{Y}$ by λ, so e_M scales by λ and the ratio $e_M/\bar{Y}$ is unchanged. The MGR is therefore scale-invariant, and its sign carries the directional meaning established in (iii). Translation behaves differently. Under a uniform additive shift $\mu_g \to \mu_g + a$, the fitted line shifts by a as well, so the residual numerator e_M is unchanged (translation-invariant); but the denominator becomes $\bar{Y}+a$, so MGR $\to e_M/(\bar{Y}+a)$. The MGR is thus *not* translation-invariant. This is the intended behaviour for a concentration gradient on a ratio scale, where the zero is physical and uniform additive shifts are not a natural symmetry. We record the property explicitly so that its statement here matches its use elsewhere in the paper.

Proof of (ii) and (iii). Eq. (3) gives MGR $= e_M/\bar{Y}$. The MGR vanishes precisely when $e_M = 0$, that is, when $\mu_M = \bar{Y} + \mathrm{SII}(R_M - \bar{R})$, the SII-predicted middle-stratum mean; equivalently, when the three points (R_g, μ_g) are collinear. With $\bar{Y} > 0$ for any positive-concentration outcome, MGR inherits the sign of e_M, so a middle stratum sitting above (below) the SII line gives MGR > 0 (<0).

Proof of (iv). The OLS first-order conditions for the population-weighted regression of μ_g on R_g (intercept and slope) give $\sum_g \pi_g e_g = 0$ and $\sum_g \pi_g R_g e_g = 0$. At $J = 3$ this is a system of two linear equations in the three residuals $\{e_L, e_M, e_H\}$, leaving one degree of freedom. Solving for e_L and e_H in terms of e_M:

$$e_L = -\frac{\pi_M(R_M - R_H)}{\pi_L(R_L - R_H)}\, e_M, \qquad e_H = -\frac{\pi_M(R_L - R_M)}{\pi_H(R_L - R_H)}\, e_M. \tag{A.1}$$

For the symmetric 30/40/30 partition the population shares are $\boldsymbol{\pi} = (0.3, 0.4, 0.3)$ with equal-spaced rank-midpoints $R = (0.15, 0.50, 0.85)$, so $R_M - R_H = R_L - R_M$ and $R_L - R_H = -2(R_M - R_L)$. Substituting into Eq. (A.1) gives the symmetric solution $e_L = e_H = -(2/3)\,e_M$. The residual sum of squares is then

$$\sum_g \pi_g e_g^2 = \pi_L e_L^2 + \pi_M e_M^2 + \pi_H e_H^2 = \left[(0.3+0.3)\tfrac{4}{9} + 0.4\right] e_M^2 = \tfrac{2}{3}\, e_M^2. \tag{A.2}$$

Since $e_M = \mathrm{MGR}\cdot\bar{Y}$, this is $\sum_g \pi_g e_g^2 = \frac{2}{3}\,\mathrm{MGR}^2\bar{Y}^2$, i.e. $c(\boldsymbol{\pi}) = 2/3$. More generally, substitution of Eq. (A.1) into the residual variance yields a quadratic $c(\boldsymbol{\pi})\,e_M^2$, where $c(\boldsymbol{\pi})$ depends on the shares π_g and the implied R_g alone. The constant evaluates to the closed form $c(\boldsymbol{\pi}) = \pi_M/(1-\pi_M)$ for any symmetric partition ($\pi_L = \pi_H$, equal-spaced midpoints). We verified this closed form by direct derivation and numerically to better than 10^{-6} over 10^3 partitions: 30/40/30 $\to 2/3$, equal thirds $\to 1/2$, 25/50/25 $\to 1$, 20/60/20 $\to 3/2$. The MGR therefore equals the linearity misfit at $J = 3$ up to the fixed partition constant. This identity justifies interpreting the MGR as a curvature coordinate of the stratum profile rather than as an auxiliary residual.

A.2. The MAIR generalisation at $J \geq 4$

For $J \geq 4$, the SII linear fit has $J - 2$ residual degrees of freedom; a single signed scalar cannot capture the entire departure of the profile from linearity. The mean absolute interior residual MAIR (Eq. (4)) averages the absolute interior residuals normalised by $\bar{Y}$:

$$\mathrm{MAIR} = \frac{1}{J-2}\sum_{g=2}^{J-1}\left|\frac{e_g}{\bar{Y}}\right|. \tag{A.3}$$

At $J = 3$ the sum reduces to its single interior term and MAIR = |MGR|, so MAIR is a continuous extension of the MGR magnitude. MAIR exceedance rates rise monotonically with J for all four pollutants in our analysis (Supplementary Table S1). This pattern supports reading the $J = 3$ MGR exceedance rate as a lower bound on partition-resolution sensitivity rather than an artefact of coarse stratification.

A.3. Boundedness of WSSH

The Within-Stratum Slope Heterogeneity (Definition 2) is the standard deviation across strata of ρ_g, the within-stratum Spearman correlation between county-level concentration and county-level median household income. Since each $\rho_g \in [-1, 1]$, and the standard deviation of a set of values confined to an interval of width 2 is at most 1, WSSH $\in [0, 1]$. WSSH equals zero exactly when every stratum has the same internal exposure–income correlation. It attains its maximum when the strata split into two subsets with $\rho_g = +1$ and $\rho_g = -1$. In that configuration the association between income and exposure has opposite signs in different parts of the distribution, and a single cross-stratum slope cannot register this pattern.

Appendix B. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.aeaoa.2026.100492.

Data and code availability

Analysis code, derived panel data, and figure regeneration scripts are openly available at https://github.com/anjana-yatawara/air-pollution-exposure-inequality. EPA AQS `annual_conc_by_monitor` files are publicly downloadable from https://aqs.epa.gov/aqsweb/airdata/download_files.html. The WUSTL ACAG V6.GL.03 $PM_{2.5}$ satellite-fused product is publicly available via the NASA AWS Open Data registry at https://registry.opendata.aws/surface-pm2-5-v6gl/. U.S. Census American Community Survey 5-year estimates are publicly available via the Census Data API at https://api.census.gov/data.

References

Apte, J.S., Messier, K.P., Gani, S., Brauer, M., Kirchstetter, T.W., Lunden, M.M., Marshall, J.D., Portier, C.J., Vermeulen, R.C., Hamburg, S.P., 2017. High-resolution air pollution mapping with Google Street View cars: exploiting big data. Environ. Sci. Technol. 51, 6999–7008. http://dx.doi.org/10.1021/acs.est.7b00891.

Burke, M., Driscoll, A., Heft-Neal, S., Xue, J., Burney, J., Wara, M., 2021. The changing risk and burden of wildfire in the United States. Proc. Natl. Acad. Sci. 118, e2011048118. http://dx.doi.org/10.1073/pnas.2011048118.

Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope, C.A., Apte, J.S., Brauer, M., Cohen, A., Weichenthal, S., Coggins, J., Di, Q., Brunekreef, B., Frostad, J., Lim, S.S., Kan, H., Walker, K.D., Thurston, G.D., Hayes, R.B., Lim, C.C., Turner, M.C., Jerrett, M., Krewski, D., Gapstur, S.M., Diver, W.R., Ostro, B., Goldberg, D., Crouse, D.L., Martin, R.V., Peters, P., Pinault, L., Tjepkema, M., van Donkelaar, A., Villeneuve, P.J., Miller, A.B., Yin, P., Zhou, M., Wang, L., Janssen, N.A.H., Marra, M., Atkinson, R.W., Tsang, H., Quoc Thach, T., Cannon, J.B., Allen, R.T., Hart, J.E., Laden, F., Cesaroni, G., Forastiere, F., Weinmayr, G., Jaensch, A., Nagel, G., Concin, H., Spadaro, J.V., 2018. Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. Proc. Natl. Acad. Sci. 115, 9592–9597. http://dx.doi.org/10.1073/pnas.1803222115.

Clark, L.P., Harris, M.H., Apte, J.S., Marshall, J.D., 2022. National and intraurban air pollution exposure disparity estimates in the United States: impact of data-aggregation spatial scale. Environ. Sci. Technol. Lett. 9, 786–791. http://dx.doi.org/10.1021/acs.estlett.2c00403.

Cohen, A.J., Brauer, M., Burnett, R., Anderson, H.R., Frostad, J., Estep, K., Balakrishnan, K., Brunekreef, B., Dandona, L., Dandona, R., Feigin, V., Freedman, G., Hubbell, B., Jobling, A., Kan, H., Knibbs, L., Liu, Y., Martin, R., Morawska, L., Pope, C.A., Shin, H., Straif, K., Shaddick, G., Thomas, M., van Dingenen, R., van Donkelaar, A., Vos, T., Murray, C.J.L., Forouzanfar, M.H., 2017. Estimates and 25-year trends of the global burden of disease attributable to ambient air pollution: an analysis of data from the Global Burden of Diseases Study 2015. Lancet 389, 1907–1918. http://dx.doi.org/10.1016/S0140-6736(17)30505-6.

Colmer, J., Hardman, I., Shimshack, J., Voorheis, J., 2020. Disparities in $PM_{2.5}$ air pollution in the United States. Science 369, 575–578. http://dx.doi.org/10.1126/science.aaz9353.

Cook, R.D., Weisberg, S., 1982. Residuals and influence in regression. Chapman and Hall, New York.

Daouda, M., Henneman, L., Goldsmith, J., Kioumourtzoglou, M.A., Casey, J.A., 2022. Racial/ethnic disparities in nationwide $PM_{2.5}$ concentrations: perils of assuming a linear relationship. Environ. Health Perspect. 130, 077701. http://dx.doi.org/10.1289/EHP11048.

Davison, A.C., Hinkley, D.V., 1997. Bootstrap methods and their application. Cambridge University Press, Cambridge.

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., Martin, R.V., 2021. Monthly global estimates of fine particulate matter and their uncertainty. Environ. Sci. Technol. 55, 15287–15300. http://dx.doi.org/10.1021/acs.est.1c05309.

Efron, B., 1979. Bootstrap methods: another look at the jackknife. Ann. Statist. 7, 1–26. http://dx.doi.org/10.1214/aos/1176344552.

Erreygers, G., 2009. Correcting the concentration index. J. Health Econ. 28, 504–515. http://dx.doi.org/10.1016/j.jhealeco.2008.02.003.

GBD 2019 Risk Factors Collaborators, 2020. Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet 396, 1223–1249. http://dx.doi.org/10.1016/S0140-6736(20)30752-2.

Geldsetzer, P., Fridljand, D., Kiang, M.V., Bendavid, E., Heft-Neal, S., Burke, M., Thieme, A.H., Benmarhnia, T., 2024. Disparities in air pollution attributable mortality in the US population by race/ethnicity and sociodemographic factors. Nature Med. 30, 2821–2829. http://dx.doi.org/10.1038/s41591-024-03117-0.

Hammer, M.S., van Donkelaar, A., Li, C., Lyapustin, A., Sayer, A.M., Hsu, N.C., Levy, R.C., Garay, M.J., Kalashnikova, O.V., Kahn, R.A., Brauer, M., Apte, J.S., Henze, D.K., Zhang, L., Zhang, Q., Ford, B., Pierce, J.R., Martin, R.V., 2020. Global estimates and long-term trends of fine particulate matter concentrations (1998–2018). Environ. Sci. Technol. 54, 7879–7890. http://dx.doi.org/10.1021/acs.est.0c01764.

Haskell-Craig, Z., Josey, K.P., Kinney, P.L., deSouza, P., 2025. Equity in the distribution of regulatory $PM_{2.5}$ monitors. Environ. Sci. Technol. 59, 15843–15852. http://dx.doi.org/10.1021/acs.est.4c12915.

Henneman, L.R., Choirat, C., Dedoussi, I., Dominici, F., Roberts, J., Zigler, C., 2023a. Inequitable exposures to U.S. coal power plant-related $PM_{2.5}$: 22 years and counting. Environ. Health Perspect. 131, 037005. http://dx.doi.org/10.1289/EHP11605.

Henneman, L.R., Choirat, C., Dedoussi, I., Dominici, F., Roberts, J., Zigler, C., 2023b. Mortality risk from United States coal electricity generation. Science 382, 941–946. http://dx.doi.org/10.1126/science.adf4915.

Jbaily, A., Zhou, X., Liu, J., Lee, T.H., Kamareddine, L., Verguet, S., Dominici, F., 2022. Air pollution exposure disparities across US population and income groups. Nature 601, 228–233. http://dx.doi.org/10.1038/s41586-021-04190-y.

Kerr, G.H., van Donkelaar, A., Martin, R.V., Brauer, M., Burkart, K., Wozniak, S., Goldberg, D.L., Anenberg, S.C., 2024. Increasing racial and ethnic disparities in ambient air pollution-attributable morbidity and mortality in the United States. Environ. Health Perspect. 132, 037002. http://dx.doi.org/10.1289/EHP11900.

Koolik, L.H., Bullard, R.D., Min, E., Morello-Frosch, R., Patterson, R., Salgado, M., Wedekind, N., Marshall, J.D., Apte, J.S., 2025. Eliminating air pollution disparities requires more than emission reduction. Proc. Natl. Acad. Sci. 122, e2505888122. http://dx.doi.org/10.1073/pnas.2505888122.

Lane, H.M., Morello-Frosch, R., Marshall, J.D., Apte, J.S., 2022. Historical redlining is associated with present-day air pollution disparities in U.S. cities. Environ. Sci. Technol. Lett. 9, 345–350. http://dx.doi.org/10.1021/acs.estlett.1c01012.

Larkin, A., Geddes, J.A., Martin, R.V., Xiao, Q., Liu, Y., Marshall, J.D., Brauer, M., Hystad, P., 2017. Global land use regression model for nitrogen dioxide air pollution. Environ. Sci. Technol. 51, 6957–6964. http://dx.doi.org/10.1021/acs.est.7b01148.

Moreno-Betancur, M., Latouche, A., Menvielle, G., Kunst, A.E., Rey, G., 2015. Relative index of inequality and slope index of inequality: a structured regression framework for estimation. Epidemiology 26, 518–527. http://dx.doi.org/10.1097/EDE.0000000000000311.

Morgenstern, H., 1995. Ecologic studies in epidemiology: concepts, principles, and methods. Annu. Rev. Public. Health 16, 61–81. http://dx.doi.org/10.1146/annurev.pu.16.050195.000425.

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55, 703–708. http://dx.doi.org/10.2307/1913610.

Office for Health Improvement and Disparities, 2024. Method for the slope index of inequality indicators. In: Fingertips Public Health Technical Guidance. URL: https://fingertips.phe.org.uk/static-reports/public-health-technical-guidance/Inequality/SII.html. (Accessed April 2026).

Parenteau, M.P., Sawada, M.C., 2011. The modifiable areal unit problem (MAUP) in the relationship between exposure to NO_2 and respiratory health. Int. J. Health Geogr. 10, 58. http://dx.doi.org/10.1186/1476-072X-10-58.

Rentschler, J., Leonova, N., 2023. Global air pollution exposure and poverty. Nat. Commun. 14, 4432. http://dx.doi.org/10.1038/s41467-023-39797-4.

Sager, L., 2025. Global air quality inequality over 2000–2020. J. Environ. Econ. Manag. 130, 103112. http://dx.doi.org/10.1016/j.jeem.2024.103112.

Shen, S., Li, C., van Donkelaar, A., Jacobs, N., Wang, C., Martin, R.V., 2024. Enhancing global estimation of fine particulate matter concentrations by including geophysical a priori information in deep learning. ACS ES T Air 1, 332–345. http://dx.doi.org/10.1021/acsestair.3c00054.

Tessum, C.W., Paolella, D.A., Chambliss, S.E., Apte, J.S., Hill, J.D., Marshall, J.D., 2021. $PM_{2.5}$ polluters disproportionately and systemically affect people of color in the United States. Sci. Adv. 7, eabf4491. http://dx.doi.org/10.1126/sciadv.abf4491.

Theil, H., 1967. Economics and Information Theory. North-Holland, Amsterdam.

U.S. Census Bureau, 2024. American Community Survey 5-year estimates, 2005–2009 through 2019–2023. Census Bureau data platform, URL: https://www.census.gov/programs-surveys/acs/data.html. (Accessed April 2026).

U.S. Environmental Protection Agency, 2010. Primary National Ambient Air Quality Standard for sulfur dioxide; final rule. Fed. Regist. 75 (119), 35520–35603, 22 June 2010.

U.S. Environmental Protection Agency, 2012. National emission standards for hazardous air pollutants from coal- and oil-fired electric utility steam generating units (Mercury and Air Toxics Standards); final rule. Fed. Regist. 77 (32), 9304–9513, 16 February 2012; promulgated 21 December 2011.

U.S. Environmental Protection Agency, 2024. Air Quality System (AQS) annual summary data. Pre-generated data files, annual summary, URL: https://aqs.epa.gov/aqsweb/airdata/download_files.html. (Accessed April 2026).

Wagstaff, A., 2005. The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. Health Econ. 14, 429–432. http://dx.doi.org/10.1002/hec.953.

Wang, Y., Marshall, J.D., Apte, J.S., 2024. U.S. ambient air monitoring network has inadequate coverage under new $PM_{2.5}$ standard. Environ. Sci. Technol. Lett. 11, 1220–1226. http://dx.doi.org/10.1021/acs.estlett.4c00605.

World Health Organization, 2021. WHO Global Air Quality Guidelines: Particulate Matter ($PM_{2.5}$ and PM_{10}), Ozone, Nitrogen Dioxide, Sulfur Dioxide and Carbon Monoxide. Technical Report, World Health Organization, Geneva, URL: https://www.who.int/publications/i/item/9789240034228. (Accessed April 2026).