Supplementary Information

Spatial decomposition of air pollution concentrations highlights historical causes for current exposure-disparities in the United States

Jiawen Liu[1], Julian D. Marshall[1]*

[1] *Department of Civil & Environmental Engineering, University of Washington, Seattle, Washington, 98125, United States of America.*
**Correspondence to: jdmarsh@uw.edu*

Summary

Literature reviews

Demographic data

Methods of spatial decomposition on administration boundaries

Methods of spatial decomposition on length scales

Figure S1-S5

Table S1-S5

Literature reviews

Findings from previous environmental justice (EJ) literature include that (i) race-ethnicity and income both correlate with exposures, but race/ethnicity is generally more important than income as a predictor, and racial/ethnic disparities hold even after correcting for income8,9. (ii) Improvements in air pollution in recent decades have benefited nearly all Americans; those improvements have reduced disparities if measured in absolute terms but less so if disparity is measured in relative terms9–12. (iii) Although people of color are, on average, exposed to higher-than-average concentrations of PM2.5, per capita they consume fewer-than-average goods and services that cause the pollution and so are less "responsible" for the pollution (i.e., a smaller-than-average fraction of pollution is attributable to their consumption)13. (iv) EJ movements benefit white folks as well14. (v) Historically, in many cases, pollution sources located near communities of color, rather than the other way around15–19. Communities experiencing higher-than-average air pollution exposure often experience additional stressors too (e.g., psychosocial, lack of greenspace)20–23. (vi) To address disparities24 may requiring location-specific emission reductions25; often, steps that help address disparities also help overall average exposures26; there might not be an "equity-efficiency tradeoff"27. (vii) The spatial scale of analysis can impact results; county-level data may be too spatially-coarse to identity disparities clearly28–32.

Recognized causes of contemporary environmental disparities include the following. (i) Starting in the early/mid-1900s, racial covenants and redlining reinforced segregation among neighborhoods33. Racial covenants were deed restrictions, reflecting actions by builders and individuals, disallowing subsequent home sale to someone of a specific nationality or racial/ethnic/religious group. Redlining was a set of rules and maps, reflecting actions by banks and the government, whereby home-loan conditions offered to buyers or government support to banks depended in part on the racial/ethnic/religious makeup of a community. A recent study found a gradient in present-day air pollution levels, from most-polluted in neighborhoods with the lowest grade (grade "D") to least-polluted highest-grade (grade "A") neighborhoods34. (ii) When polluting infrastructure (e.g., factories; roads) was constructed, on average it was more likely to locate in communities of color. Because of unequal power and access, those communities were less able to fight or reverse those decisions. (iii) "White flight," which refers to the movement of urban White people to suburbs, further strengthened and entrenched urban racial segregation, especially in the 1950s – 1970s. (iv) Racist planning combined and interacted with those trends. For example, the impact of an urban highway might become a physical barrier to reinforce segregation, and facilitated White flight to suburbs where residents could commute to urban jobs if they own a car. These historical aspects underscore concepts such as cumulative risk (i.e., considering only one emission source at a time may miss the multiple exposures impacting a community; it is important to look too at risks from all sources) and "sacrifice zones".

Demographic data

NHGIS provides easy access to US census data. Here, we used population estimates for 7 census racial groups and two ethnic groups for each census block: (i) White alone, (ii) Black or African American alone, (iii) American Indian and Alaska Native alone, (iv)

Asian alone, (v) Native Hawaiian and Other Pacific Islander alone, (vi) some other race alone, and (vii) two or more races. For each racial group, NHGIS reports population estimates for two ethnic groups: (i) Hispanic or Latino and (ii) not Hispanic nor Latino. Thus, there are 14 racial-ethnic groups in NHGIS (7 racial groups; 2 ethnic groups).

Disparity decomposition on administration boundaries

National total disparity between a racial-ethnic group and whole population is calculated using equation (1)

$$\text{Disparity} = \frac{\sum_{b=1}^{n} C_b P_{r,b}}{\sum_{b=1}^{n} p_{r,b}} - \frac{\sum_{b=1}^{n} C_b P_{t,b}}{\sum_{b=1}^{n} P_{t,b}} \quad (1)$$

Where n is the number of census blocks in the contiguous United States; C_b is pollution level for census block b; $P_{r,b}$ is population for each racial-ethnic group for census block b; $P_{t,b}$ is total population for census block b.

Total disparity is decomposed into between-state disparity and within-state disparity. Between-state disparity aims to remove the effect of within-state exposure difference. Thus, we assume all population in the same state expose to the same pollution level, regardless of race-ethnicity. Between-state disparity between a racial-ethnic group and whole population is calculated using equation (2).

$$\text{Between-state disparity} = \frac{\sum_{s=1}^{48} C_s P_{r,s}}{\sum_{s=1}^{48} P_{r,s}} - \frac{\sum_{s=1}^{48} C_s P_{t,s}}{\sum_{s=1}^{48} P_{t,s}} \quad (2)$$

Where C_s is population-weighted average pollution level for state, $P_{r,s}$ is population for each racial-ethnic group in state s; $P_{t,s}$ is total population in state s.

Within-state disparity aims to find disparity between different racial-ethnic groups and whole population in each state. To do this, we first calculate state-standardized pollution levels using equation (3).

$$C_{ss} = C_b - C_s \quad (3)$$

Where C_{ss} is the state-standardized pollution level, C_b is the block pollution level, and C_s is population-weighted pollution level for state s. Then we calculate within-state disparity using equation (1) where C_b is standardized pollution level for census block b, i.e. C_{ss}.

Within-state disparity is further divided into between-urban/rural and within-urban/rural disparity. Within-urban/rural disparity includes within-urban blocks disparity and within-rural blocks disparity. Disparity between urban and rural blocks aims to remove disparity within urban blocks and within rural blocks in the same state and determine how much does residence block urbanicity contribute to disparity after accounting difference between states. Between-urban/rural blocks disparity within states is also calculated using equation (1). To remove the effect of within-urban/rural blocks disparity, we assume all population reside in urban blocks in the

same state have the same pollution level, which is the population-weighted average (C_{ur}) of standardized blocks' pollution levels (C_{ss}) for all urban blocks in state *s*. We then used equation (1), substitute C_b with C_{ur}. Same procedure applied to all rural blocks.

To calculate within-urban blocks disparity, we standardized pollution levels for all urban blocks in state *s* by subtracting population-weighted average pollution level for all urban blocks in the state, equation (4).

$$C_{sus} = C_{ss} - C_{ur} \qquad (4)$$

Where C_{sus} is the state urban blocks-standardized pollution level, C_{ss} is the state-standardized pollution level, and C_{ur} is population-weighted pollution level of standardized blocks' pollution levels (C_{ss}) for all urban blocks in state *s*. We then used equation (1), substitute C_b with C_{sus} for all urban blocks to get the unweighted within-urban disparity. We applied the same procedure for all rural blocks in the state *s* to calculate the unweighted within-rural blocks disparity. Then we calculate population-weighted within-urban disparity using equation (5). Same procedure is applied for population-weighted within-rural disparity.

$$\text{Weighted within} - \text{urban disparity} = \frac{p_{urban}}{p_{urban} + p_{rural}} \times \text{unweighted within} - \text{urban disparity} \qquad (5)$$

Within-urban blocks disparity can be further divided into between-urban area and within-urban area disparity. Between-urban area disparity aims to reveal disparity between different urban areas, after accounting state locations and urbanicity. To do this, we assume all population reside in the same urban area have the same pollution level, which is the population-weighted average exposure (C_{ua}) of standardized blocks' pollution level (C_{sus}) for all urban blocks in the same urban area. We then used equation (1), substitute C_b with C_{ua}.

To calculate within-urban area disparity, we standardized pollution levels for all urban blocks in the same urban area by subtracting population-weighted average pollution level for all urban blocks in the urban area, equation (6).

$$C_{suas} = C_{sus} - C_{ua} \qquad (6)$$

Where C_{suas} is the state urban area-standardized pollution level, C_b is the block pollution level, and C_{ua} is population-weighted pollution level of standardized blocks' pollution level (C_{sus}) for all urban blocks in urban area *a*. We then used equation (1), substitute C_b with C_{suas} for all blocks.

Disparity decomposition on length scales

States' and blocks' boundaries are defined administratively. Besides administrative boundary, we also study natural length scale based on distance. Disparities at the four different length scales are calculated using equation (7)

$$\text{Distance-based disparity} = \frac{\sum_{d=1}^{m} C_d P_{r,d}}{\sum_{d=1}^{m} P_{r,d}} - \frac{\sum_{d=1}^{m} C_d P_{t,d}}{\sum_{d=1}^{m} P_{t,d}} \quad (7)$$

Where m is the number of census blocks; C_d is pollution level for decomposed spatial components in census block; $P_{r,d}$ is population for census blocks for each racial-ethnic group; $P_{t,d}$ is total population for census blocks.

There are four categories included in the study: *long*, *mid-long*, *mid-short*, and *short*, corresponding to length scales of, respectively, >100 km, 10-100 km, 1-10 km, and <1 km. Briefly, for a given census block, the long (>100 km) pollution level is the minimum concentration for all census block centroids within a 100 km radius of the given centroid. The mid-long (10-100 km) pollution level is the difference between the minimum concentration within 10 km and the long pollution level. Similarly, mid-short is the concentration difference between the 1-km minimum and the 10-km minimum. Lastly, short is the concentration difference between the predicted concentration and the 1-km minimum. National exposure disparities are then calculated separately for each of the four categories.

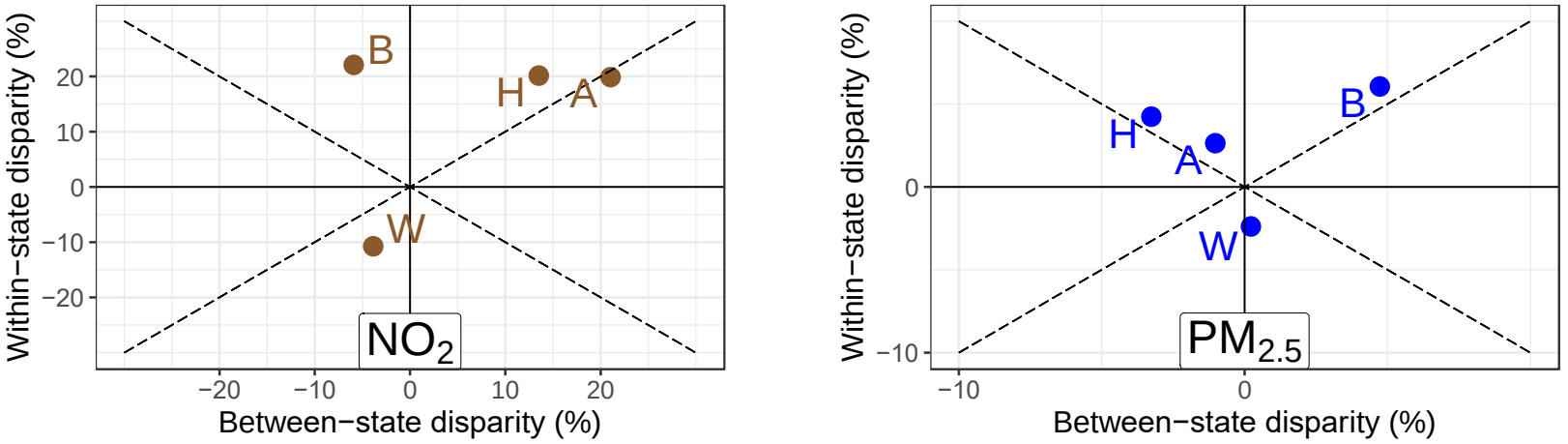


Figure S1. Scatterplot for within-state disparity and between-state disparity for the four racial-ethnic groups. For all panels, each point represents the racial-ethnic group. “W” refers to non-Hispanic White people. “B” refers to non-Hispanic Black people. “H” refers to Hispanic people of any race(s). “A” refers to non-Hispanic Asian people. Dash line is the 1:1 ratio reference line.

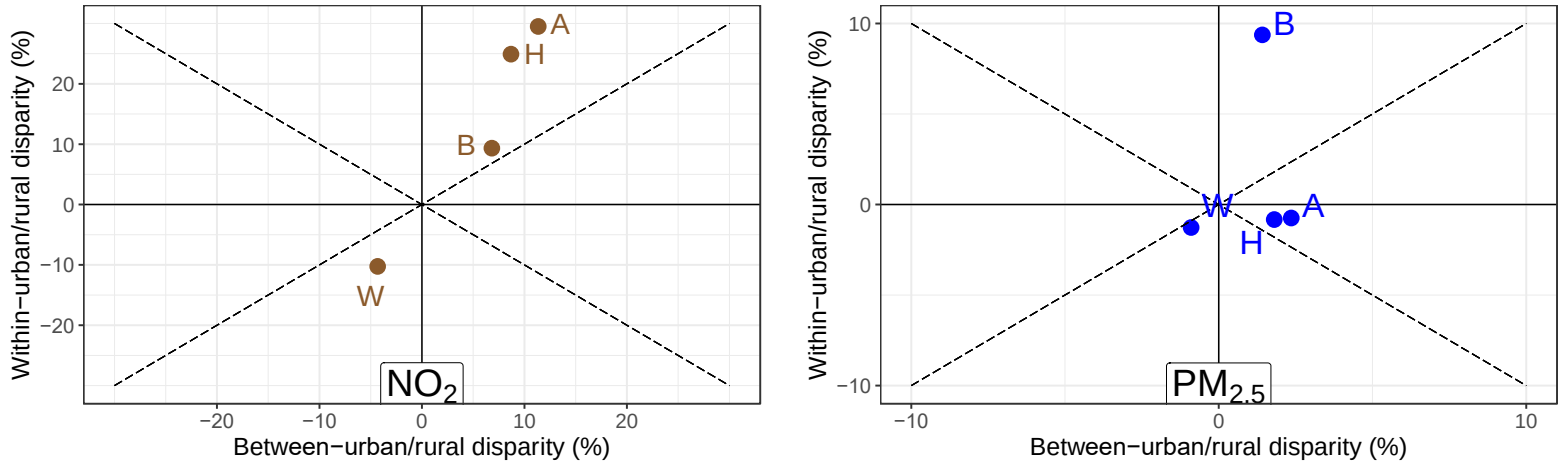


Figure S2. Scatterplot for within-urban/rural disparity and between-urban/rural disparity for the four racial-ethnic groups. For all panels, each point represents the racial-ethnic group. "W" refers to non-Hispanic White people. "B" refers to non-Hispanic Black people. "H" refers to Hispanic people of any race(s). "A" refers to non-Hispanic Asian people. Dash line is the 1:1 ratio reference line.

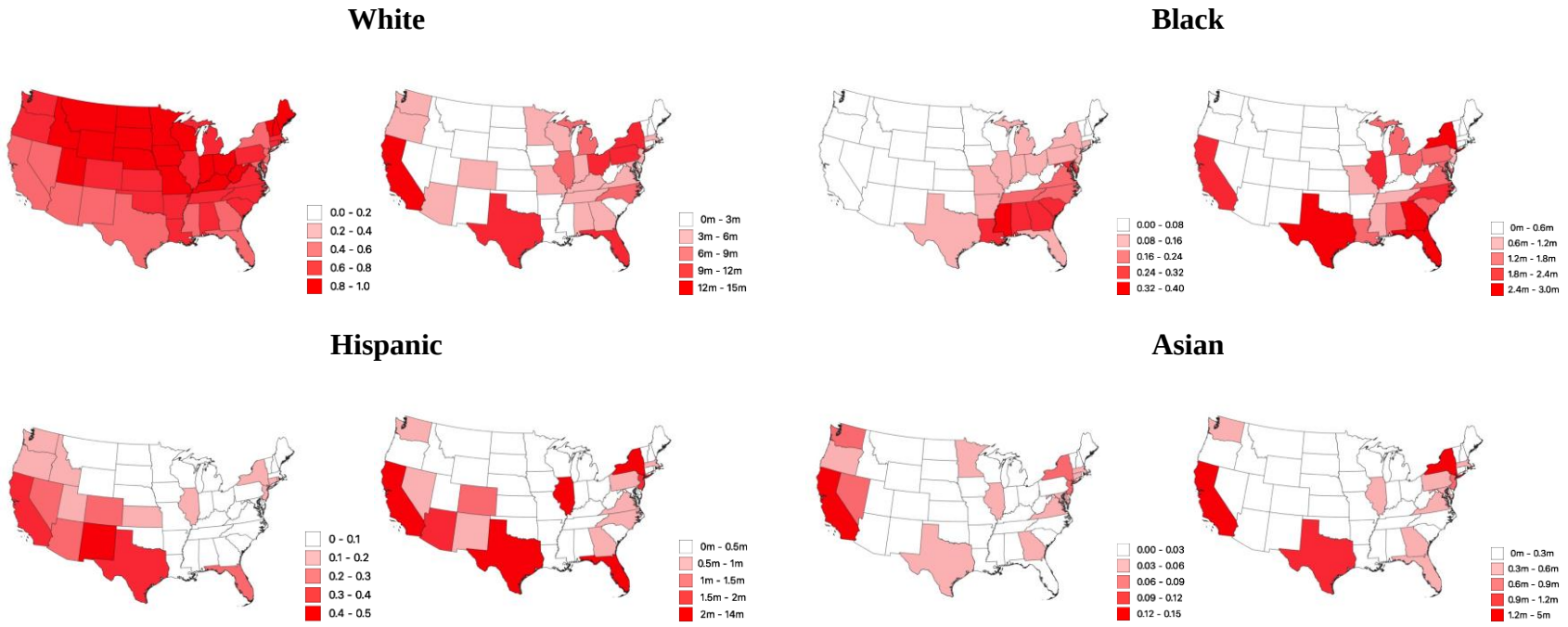


Figure S3. Population density map for US for the four racial-ethnic groups. For each group, the left map stands for proportion of the subpopulation in the state; the right map stands for actual subpopulation in the state.

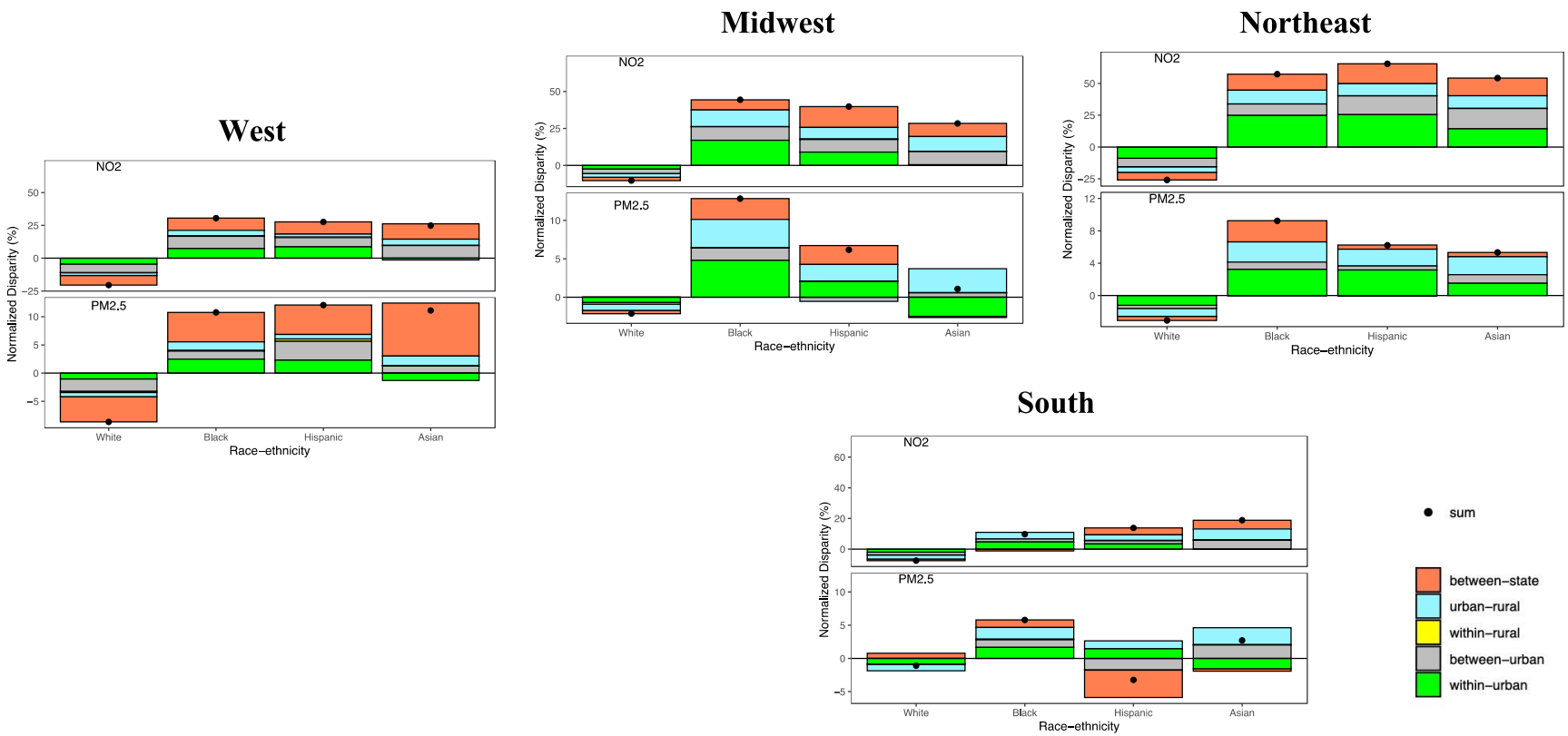
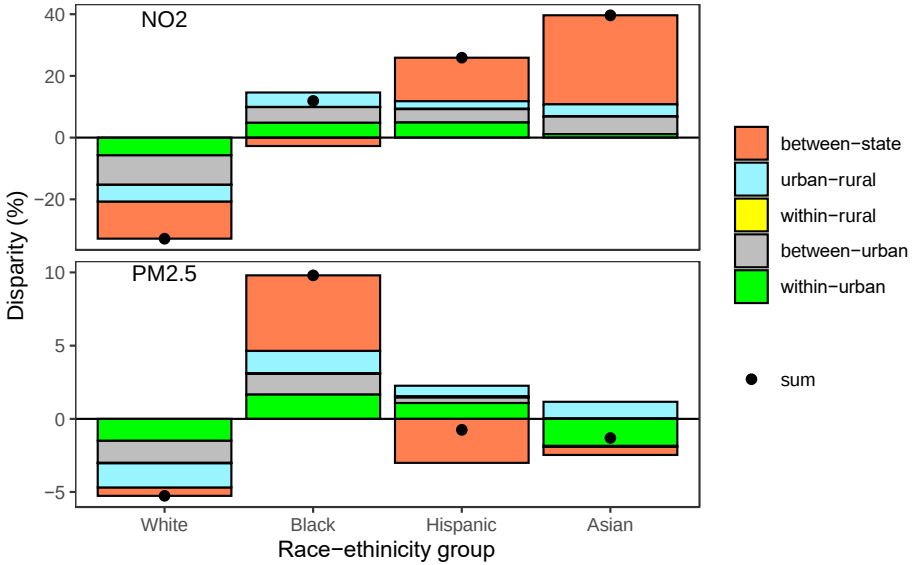
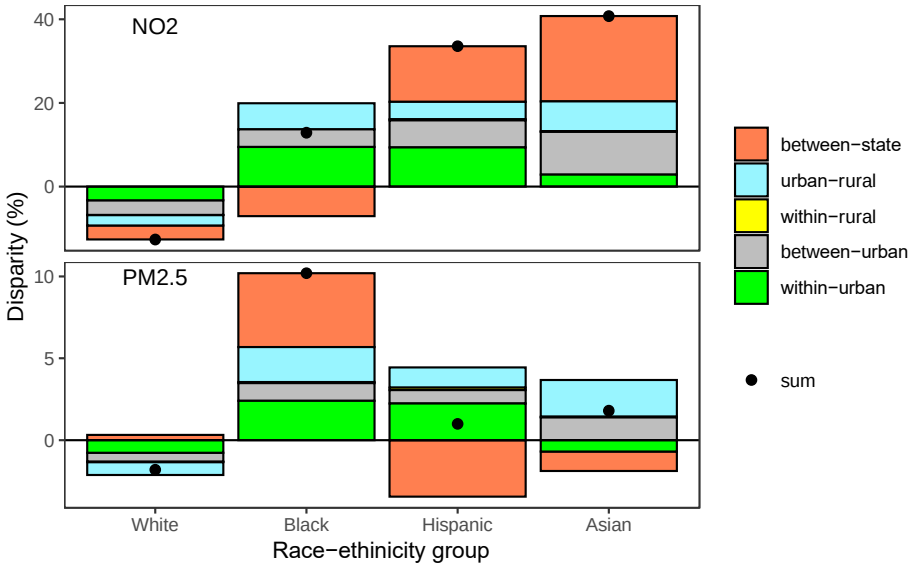


Figure S4. Normalized decomposed disparities for administrative boundaries for the four main racial-ethnic groups for four regions in the US. Point stands for total disparity for the racial-ethnic group.

G* 0.041-0.21



G* 0.21-0.47



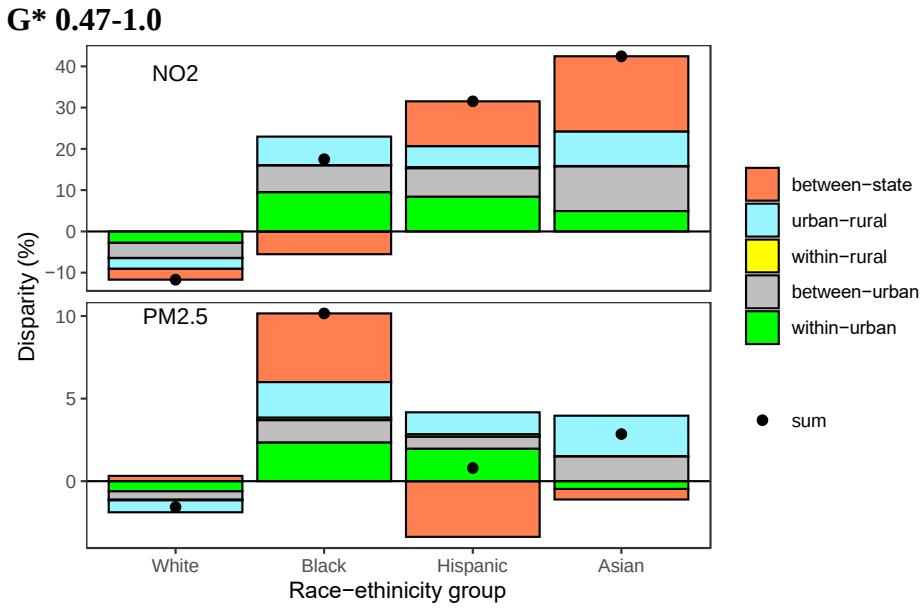
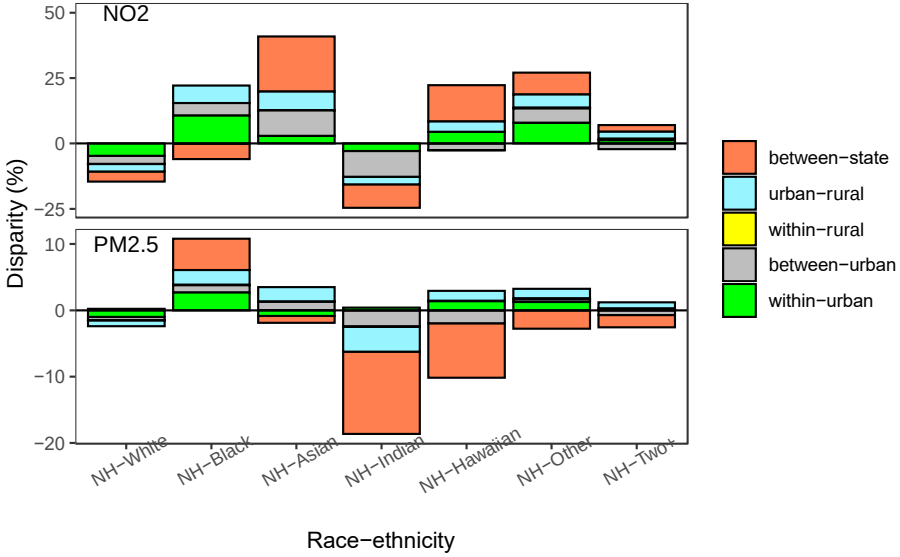
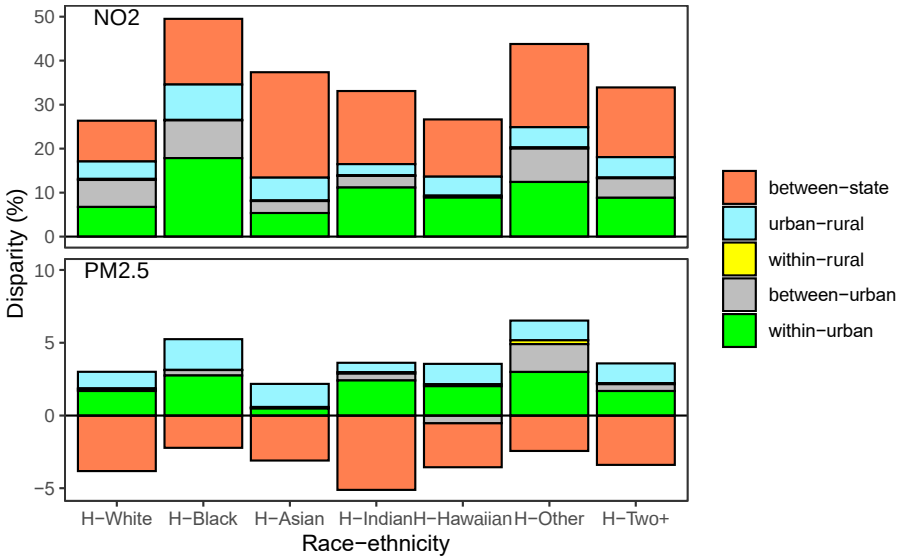


Figure S5. Normalized decomposed disparities for administrative boundaries for the four main racial-ethnic groups for three different ranges of G* values in the US. 10th percentile of G* is 0.21 and 90th percentile of G* is 0.47. Point stands for total disparity for the racial-ethnic group. The category with G* 0.041-0.21 is the 10% least segregated; 0.21-0.47 is the 80% in the middle; 0.47-1.0 is the 10% most segregated.

Non-Hispanic



Hispanic



Non-Hispanic + Hispanic

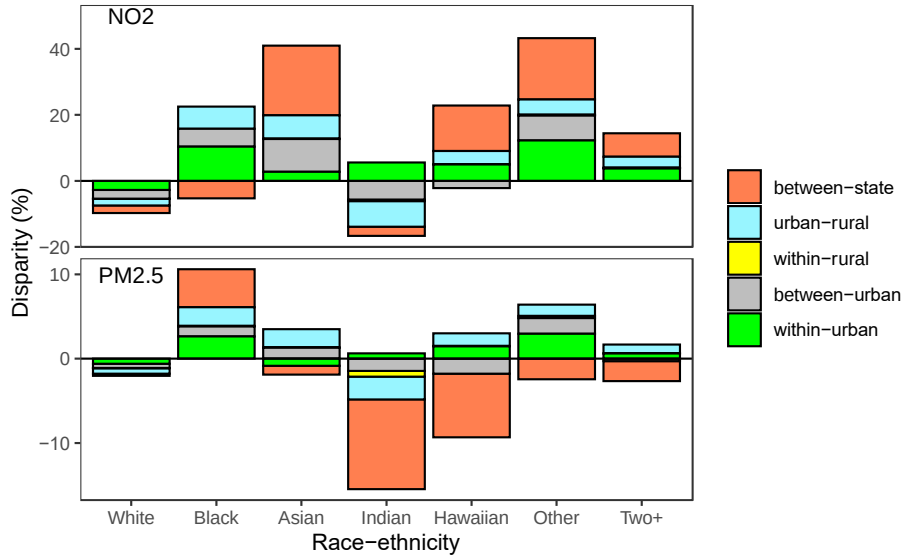


Figure S6. Normalized decomposed disparities for administrative boundaries for 7 racial-ethnic groups in the US for Hispanic, non-Hispanic, and any ethnicity. Non-Hispanic-White (population: 64%); Non-Hispanic-Black (12%); Non-Hispanic-Asian (4.5%); Non-Hispanic-Indian (0.7%); Non-Hispanic-Hawaiian (0.11%); Non-Hispanic-Other (0.20%); Non-Hispanic-Two+ (1.8%); Hispanic-White (8.7%); Hispanic-Black (0.40%); Hispanic-Asian (0.064%); Hispanic-Indian (0.22%); Hispanic-Hawaiian (0.017%); Hispanic-Other (6.0%); Hispanic-Two+ (0.97%); White (73%); Black (13%); Asian (4.6%); Indian (0.92%); Hawaiian (0.13%); Other (6.2%); Two+ (2.8%).

Table S1. Decomposed disparities normalized to pollutants' national population-weighted average exposure at five spatial levels on administration boundaries for the four main racial-ethnic groups.

NO_2	White	Black	Hispanic	Asian
Between-state	-4%	-6%	14%	21%
Urban-rural	-3%	7%	4%	7%
Within-rural	0%	0%	0%	0%
Between-urban	-3%	5%	6%	10%
Within-urban	-5%	11%	9%	3%
Total	-15%	17%	34%	41%
$PM_{2.5}$				
Between-state	0%	5%	-3%	-1%
Urban-rural	-1%	2%	1%	2%
Within-rural	0%	0%	0%	0%
Between-urban	-0%	1%	1%	1%
Within-urban	-1%	3%	2%	-1%
Total	-2%	11%	1%	2%

Table S2. Absolute contribution of normalized decomposed disparity [a] at five spatial levels on administration boundaries to total disparity.

NO_2	White	Black	Hispanic	Asian
Between-state	27%	21%	41%	51%
Urban-rural	20%	24%	13%	17%
Within-rural	0%	0%	1%	0%
Between-urban	21%	17%	18%	24%
Within-urban	32%	39%	28%	7%
$PM_{2.5}$				
Between-state	8%	44%	44%	19%
Urban-rural	34%	20%	16%	41%
Within-rural	2%	1%	2%	1%
Between-urban	18%	10%	10%	24%
Within-urban	38%	25%	28%	16%

[a] contribution of normalized decomposed disparity = abs(disparity at each spatial level) / sum[abs(disparity at each spatial level)] *100%

Table S3. Decomposed disparities normalized to pollutants' national population-weighted average exposure at five spatial levels on length scales for the four main racial-ethnic groups.

NO_2	White	Black	Hispanic	Asian
Short	-2%	2%	4%	6%
Short_mid	-7%	8%	15%	19%
Mid_long	-6%	6%	14%	16%
Long	-0%	0%	1%	0%
Total	-15%	16%	34%	41%
$PM_{2.5}$				
Short	0%	1%	0%	0%
Short_mid	-1%	2%	3%	4%
Mid_long	-2%	2%	6%	6%
Long	2%	7%	-8%	-8%
Total	-2%	11%	1%	2%

Table S4. Absolute contribution of normalized decomposed disparity [a] at five spatial levels on length scales to total disparity based on absolute values.

NO_2	White	Black	Hispanic	Asian
Short	14%	15%	13%	14%
Short_mid	45%	46%	44%	46%
Mid_long	40%	36%	41%	39%
Long	1%	3%	2%	1%
$PM_{2.5}$				
Short	3%	6%	1%	1%
Short_mid	24%	15%	16%	20%
Mid_long	43%	17%	35%	35%
Long	30%	62%	47%	45%

[a] contribution of normalized decomposed disparity = abs(disparity at each spatial level) / sum[abs(disparity at each spatial level)] *100%

Table S5. Total disparity percentage for four racial-ethnic groups for three categories of G* statistics

G* 0.041-0.21[a]	White	Black	Hispanic	Asian
NO_2	-33%	12%	26%	40%
$PM_{2.5}$	-5.3%	9.8%	-0.75%	-1.3%
G* 0.21-0.47[b]				
NO_2	-13%	13%	34%	41%
$PM_{2.5}$	-1.8%	10%	1.0%	1.8%
G* 0.47-1.0[c]				
NO_2	-12%	17%	32%	42%
$PM_{2.5}$	-1.6%	10%	0.79%	2.9%

[a] the 10% least segregated category

[b] the 80% in the middle category

[c] the 10% most segregated category

total disparity percentage = (population-weighted exposure for subpopulation in the category – national population-weighted exposure)/ national population-weighted exposure*100%