Supporting Information

Historical redlining is associated with present-day air pollution disparities in U.S. cities

Haley M. Lane[a,*], Rachel Morello-Frosch,[b,c] , Julian D. Marshall,[d] and Joshua S. Apte[a,b]

[a] Department of Civil and Environmental Engineering, University of California, Berkeley, CA 94720 USA

[b] School of Public Health, University of California, Berkeley, CA 94720 USA

[c] Department of Environmental Science, Policy, and Management, University of California, Berkeley, CA 94720 USA

[d] Department of Civil and Environmental Engineering, University of Washington, Seattle, WA USA

* Corresponding author. Email: apte@berkeley.edu

Contents:

S1 Materials and Methods

S1.1 Data Preparation

We used GIS software to merge U.S. Census Bureau TIGER/Line state census block shape files and HOLC map shape files created by the Mapping Inequality project.[1,2] Mapped census blocks outside of the HOLC map areas were extended to encompass the city's associated census urban area (CUA) using TIGER/Line CUA boundaries. A map of all HOLC map cities and their accompanying CUAs is included in Figure S1. Block groups which had partial overlap with a HOLC area were labeled as 100% the overlapping grade because HOLC map negative space often reflects non-residential areas. In cases where HOLC neighborhoods with different grades overlapped with a single block, the clipped areas of the overlapping HOLC sections were used to determine population portioning for each grade. For example, a block with 3,000 m^2 designated as 'B' and 7,000 m^2 designated as 'C' would be identified as 30% 'B' and 70% 'C' regardless of whether the total block was 10,000 m^2 or 15,000 m^2 in area. This approach assumes that, for blocks with multiple HOLC categories, racial/ethnic groups would be evenly distributed between A-D categories based solely on area portioning. Note that 88% of the blocks which overlap with HOLC maps only overlap with 1 grade, and 96% of the overlapping blocks have a single grade attribution of at least 80%.

We characterized ambient outside annual $PM_{2.5}$ and NO_2 pollution levels on a block level using v1 empirical models developed by the Center for Air, Climate and Energy Solutions (CACES).[3] CACES provides annual ambient levels for 1979-2015, and year-2010 levels were used to be contemporaneous with the most recently available complete U.S. Census. Figure S7 shows a sensitivity comparison using year-2015 estimates with the same year-2010 population

data. While overall estimated pollution levels decreased markedly over the 5-year period, intraurban HOLC disparities were largely unaffected by the overall trends.

The 2010 U.S. Census was used to characterize mapped area demographics. Total population was pulled from Table P1 while Hispanic and Latino population and non-Hispanic Asian, non-Hispanic Black, and non-Hispanic White populations were pulled from Table P5.[4] Population, demographics data, and air pollution levels were linked to the merged HOLC census block maps by fips code. A summary of the overall dataset populations is included in Table S1 and illustrated in Figure S3, and examples of the different dataset inputs for Atlanta, GA are shown in Figure S2.

The HOLC road and rail analyses were developed using the TIGER/Line datasets for rail lines and primary roads across the US.[1] Buffer polygons were created for both datasets with buffer distances 150 m and 300 m. The 150 m buffer was chosen in order to reflect the point at which there is typically an ~50% reduction from peak road concentrations,[5] while the 300 m buffer was selected to reflect the likely extent of near-road impact based on the analysis of Su et al (2015).[6] Each buffer was then used to clip the HOLC-census block shapefile to determine which blocks, and what percentage of each block, was within each buffer distance of primary roads or railroads.

The industrial facility proximity analysis used similar methods with the 2017 National Emissions Inventory (NEI) facility-level emissions dataset.[7] Circular buffers with radii of 500 m and 1000 m were created for each source facility location within 1000 m of HOLC mapped areas. Larger buffers were chosen to account for the fact that a single coordinate for a facility may not fully reflect the locations of individual emissions sources within the facility. The number of source buffers overlapping with a given census block-HOLC grade polygon was

counted, and the annual NO_x and primary $PM_{2.5}$ emissions for the overlapping sources were summed.

S1.2 Intraurban Analysis Methods

The size of HOLC mapped areas can vary widely between different regions and cities. Large, older cities such as New York, San Francisco, and Chicago, which were well developed in the 1930s, include mapped areas similar to the city's current footprint. However, there are notable variations in how the surrounding areas of these cities are represented given the significant increase in suburban sprawl since the 1930s. Conversely, some cities, particularly in the west, have grown significantly since HOLC maps were developed. For example, Phoenix, AZ has a population of more than 1 million people, but only approximately 75,000 people lived in the HOLC mapped areas as of 2010. The HOLC map does not represent modern-day Phoenix well because the city population has grown by a factor of 30 since 1930. There are also variations that reflect choices made by the mapping team which obfuscate intra-city comparisons. For example, the Los Angeles, CA shapefile includes several maps from different adjacent municipalities all labeled as Los Angeles while the Boston, MA suburbs are each included separately. However, we chose not to adjust these because not all cities within an individual CUA were contiguous, as was the case for Chicago, where closer suburbs were included as "Chicago" while further cities were separate. Additionally, the largest metropolitan region, New York, NY, includes each NYC borough as a separate city while incorporating the surrounding New Jersey regions, including the city of Newark, as county maps.

Overall city characteristics were characterized on a HOLC map basis to focus on differences within redlining areas. HOLC mapped areas exposures were higher than those in the broader urbanized areas given the HOLC focus on older, more developed sections of cities (see

Figure S4). Overall distributions were evaluated on a per-person basis, and statistics such as means and standard deviations were weighted by block population. Equation 1 illustrates how overall population-weighted means were calculated,

$$\overline{C_{g,r}} = \frac{\sum_{i=1}^{N} C_i * p_{i,g,r}}{\sum_{i=1}^{N} p_{i,g,r}} \quad \textbf{Eq. 1}$$

$N = \#\ of\ census\ blocks$
$n = \#\ of\ census\ blocks\ within\ a\ city$
$C = pollution\ level$
$p = population$
$i = block$
$g = grade$
$[A, B, C, D, All]$
$r = race/ethnicity$
$[Asian, Black, Hispanic, White, Other, Total]$

A city's average exposure level was characterized using the population-weighted mean pollution level for the total population within HOLC mapped areas (See Eq. 2). Intraurban disparities were determined by subtracting the city average exposure from individual block pollution estimates (See Eq. 3), and in-grade disparities were similarly ascertained by subtracting a city's average exposure within areas of the same HOLC grade (See Eq. 4).

$$\overline{C_{city,g}} = \frac{\sum_{i=1}^{n} C_i * p_{i,g,total}}{\sum_{i=1}^{n} p_{i,g,total}} \quad \textbf{Eq. 2}$$

$$\overline{dC_{IU,g,r}}(Intraurban) = \frac{\sum_{i=1}^{n}(C_i - \overline{C_{city,all}}) * p_{i,g,r}}{\sum_{i=1}^{N} p_{i,g,r}} \quad \textbf{Eq. 3}$$

$$\overline{dC_{IG,g,r}}(Intragrade) = \frac{\sum_{i=1}^{n}(C_i - \overline{C_{city,g}}) * p_{i,g,r}}{\sum_{i=1}^{n} p_{i,g,r}} \quad \textbf{Eq. 4}$$

All pollution analyses, with the exception of the work included in Figure S11, are presented on a population-weighted basis.

S1.3 Segregation Analysis Methods

We conducted a stylized sensitivity analysis to explore how overall racial-ethnic disparity arises from factors both related to and independent of segregation by historic HOLC group. We created two alternative exposure scenarios to evaluate the sensitivity of racial-ethnic disparities in HOLC mapped areas. These scenarios were as follows: (1) all between grade segregation is controlled for while preserving overall population density (Eq. 5) and (2) all within grade segregation is controlled for (Eq.6). Populations were adjusted as shown below.

$$p_{i,r}(1) = p_{i,total} * \frac{p_{City,r}}{p_{City,total}} * \frac{p_{g,r}}{p_{g,total}} \quad \textbf{Eq. 5}$$

$$p_{i,r}(2) = p_{i,total} * \frac{p_{g,r}}{p_{g,total}} \quad \textbf{Eq.6}$$

An overall exposure analysis was conducted, consistent with the methods outlined in S1.2 using these adjusted block population numbers ($p_{i,r}(1)$ and $p_{i,r}(2)$), and results are summarized in Figure S10.

References

(1) U.S. Census Bureau. 2010 TIGER/Line Shapefiles (Machine Readable Datafiles). 2010.

(2) Nelson, R. K.; Winling, L.; Marciano, R.; Connolly, N. Mapping Inequality. *Am. Panor.*

(3) Kim, S.-Y.; Bechle, M.; Hankey, S.; Sheppard, L.; Szpiro, A. A.; Marshall, J. D. Concentrations of Criteria Pollutants in the Contiguous U.S., 1979 – 2015: Role of Prediction Model Parsimony in Integrated Empirical Geographic Regression. *PLoS ONE* **2020**, *15* (2). https://doi.org/10.1371/journal.pone.0228535.

(4) U.S. Census Bureau. *2010 Census Summary File 2 - United States*; 2011.

(5) Karner, A. A.; Eisinger, D. S.; Niemeier, D. A. Near-Roadway Air Quality: Synthesizing the Findings from Real-World Data. *Environ. Sci. Technol.* **2010**, *44* (14), 5334–5344. https://doi.org/10.1021/es100008x.

(6) Su, J. G.; Apte, J. S.; Lipsitt, J.; Garcia-Gonzales, D. A.; Beckerman, B. S.; de Nazelle, A.; Texcalac-Sangrador, J. L.; Jerrett, M. Populations Potentially Exposed to Traffic-Related Air Pollution in Seven World Cities. *Environ. Int.* **2015**, *78*, 82–89. https://doi.org/10.1016/j.envint.2014.12.007.

(7) US EPA, OAR. 2017 National Emissions Inventory (NEI) Data https://www.epa.gov/air-emissions-inventories/2017-national-emissions-inventory-nei-data (accessed 2022 -02 -02).

S2 Supporting Information Tables

Table S1. Year-2010 U.S. Census population distribution by historic redlining status and overall population (includes 202 cities in 146 CUAs)

HOLC grade / spatial unit	Asian	Black	Hispanic	White	Other	Total
A	4%	3%	2%	10%	5%	6%
B	19%	18%	15%	27%	21%	22%
C	52%	45%	50%	43%	49%	46%
D	25%	34%	33%	19%	24%	26%
HOLC-Mapped Areas (10^6)	3.15	10.4	10.7	19.3	1.29	44.8
Census Urban Areas (10^6)	10.8	24.8	31.0	90.0	4.33	161
US Urban Population (10^6)						249

S3 Supporting Information Figures

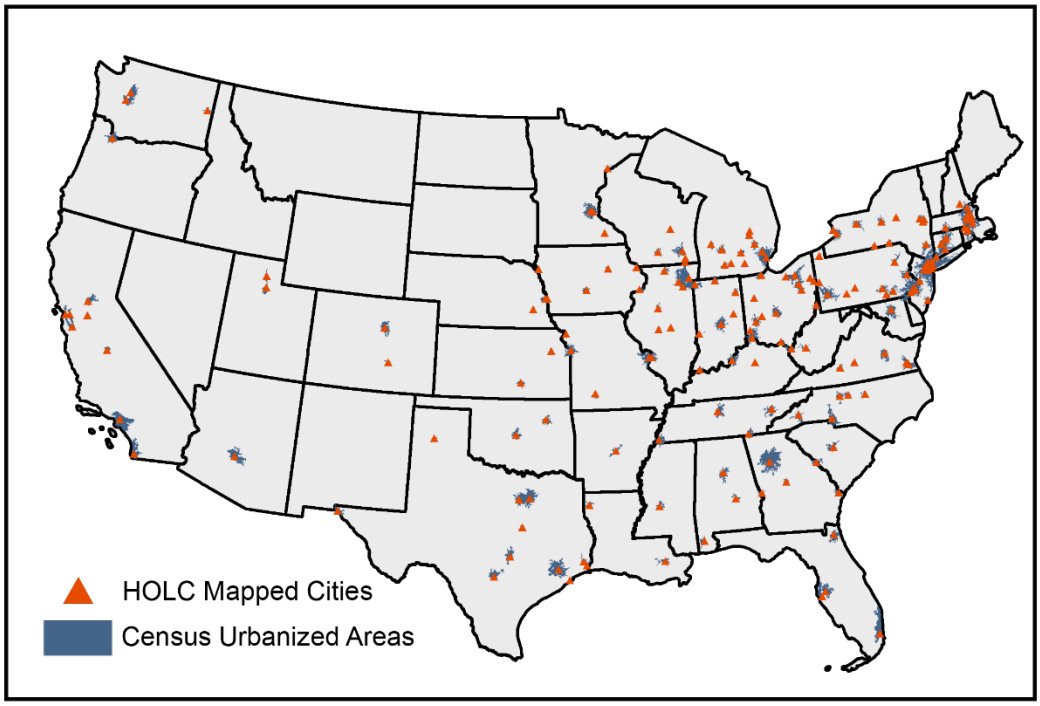

Figure S1. Locations of cities for which HOLC maps were drawn in the 20th century and the current census urbanized areas associated with each city. All cities shown here were included in this analysis.

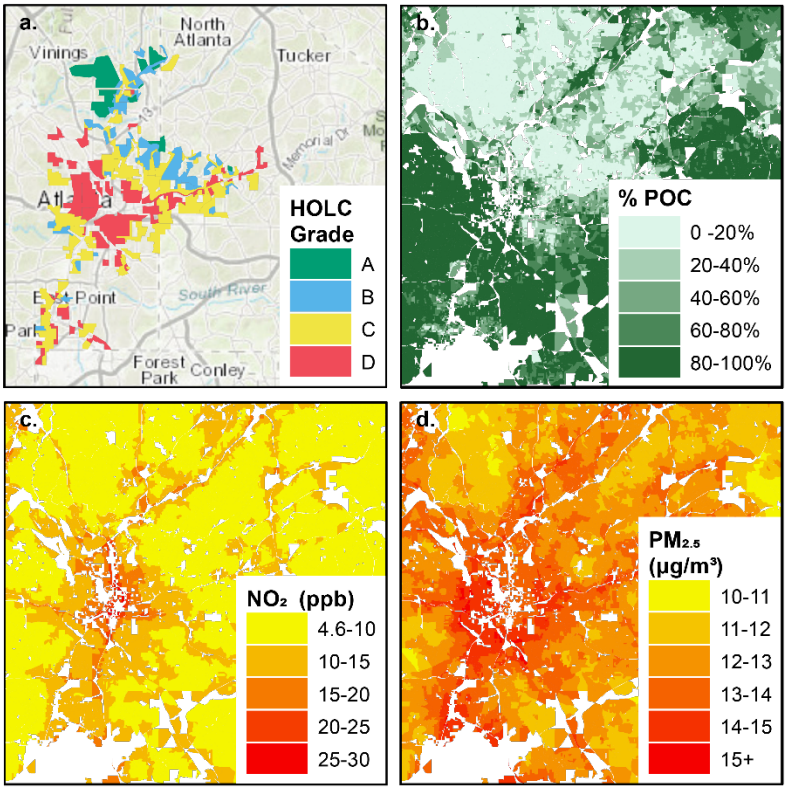

Figure S2. Illustration of input data sources in Atlanta, GA. (a) 1930's HOLC map for the City of Atlanta which classifies neighborhoods on an A-D scale based on the perceived favorability for home lending. (b) Prevalence of people of color in Atlanta, GA in year-2010 on a census block level. (c) Annual CACES modelled NO_2 levels for year-2010 on a census block level. (d) Annual CACES modelled $PM_{2.5}$ concentrations for year-2010 on a census block level.

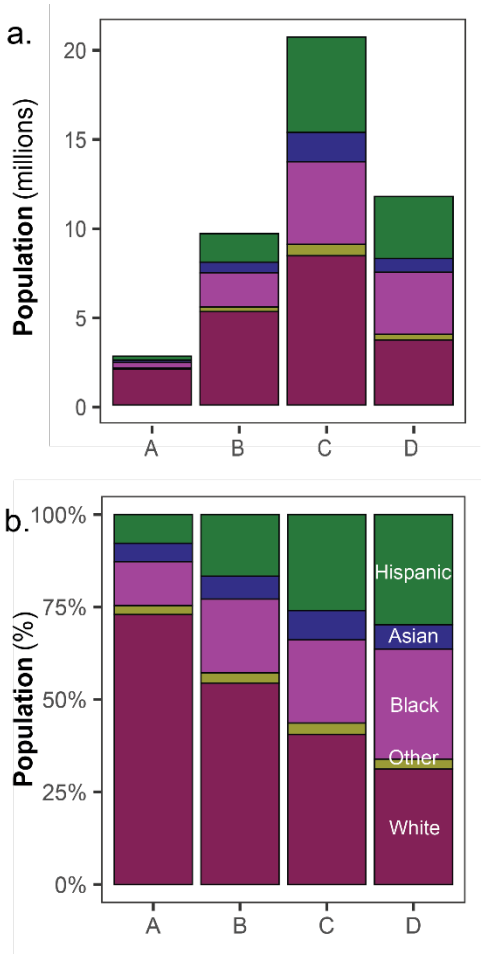

Figure S3. Demographics of population living within HOLC mapped areas in 2010 by (a) millions of people and (b) portion of the population within each grade. The A group population is small relative to other grades while the C grade contains the most people. Portion of population that is White decreases from A to D while the portion of people of color increases.

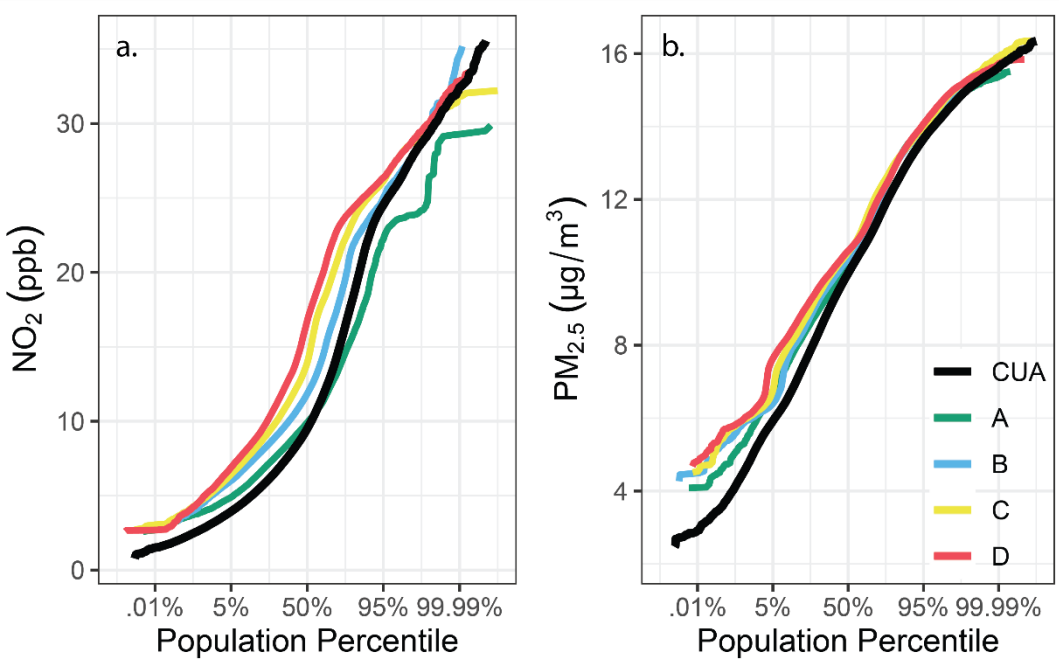

Figure S4. Overall cumulative exposure distribution for all HOLC mapped areas separated by HOLC grade. (a) NO_2 cumulative distribution. (b) $PM_{2.5}$ cumulative distribution. Both pollutants demonstrate a relatively monotonic relationship between worsening HOLC grade and an increased pollution level. However, NO_2 levels exhibit both a wider overall range and more between-grade difference than is observed for $PM_{2.5}$. Overall CUA pollution levels, particularly at lower percentiles, tend to be less than within the HOLC mapped areas, particularly for groups B-D, which reflects how HOLC mapped areas tend to already emphasize the more developed, more polluted areas of cities.

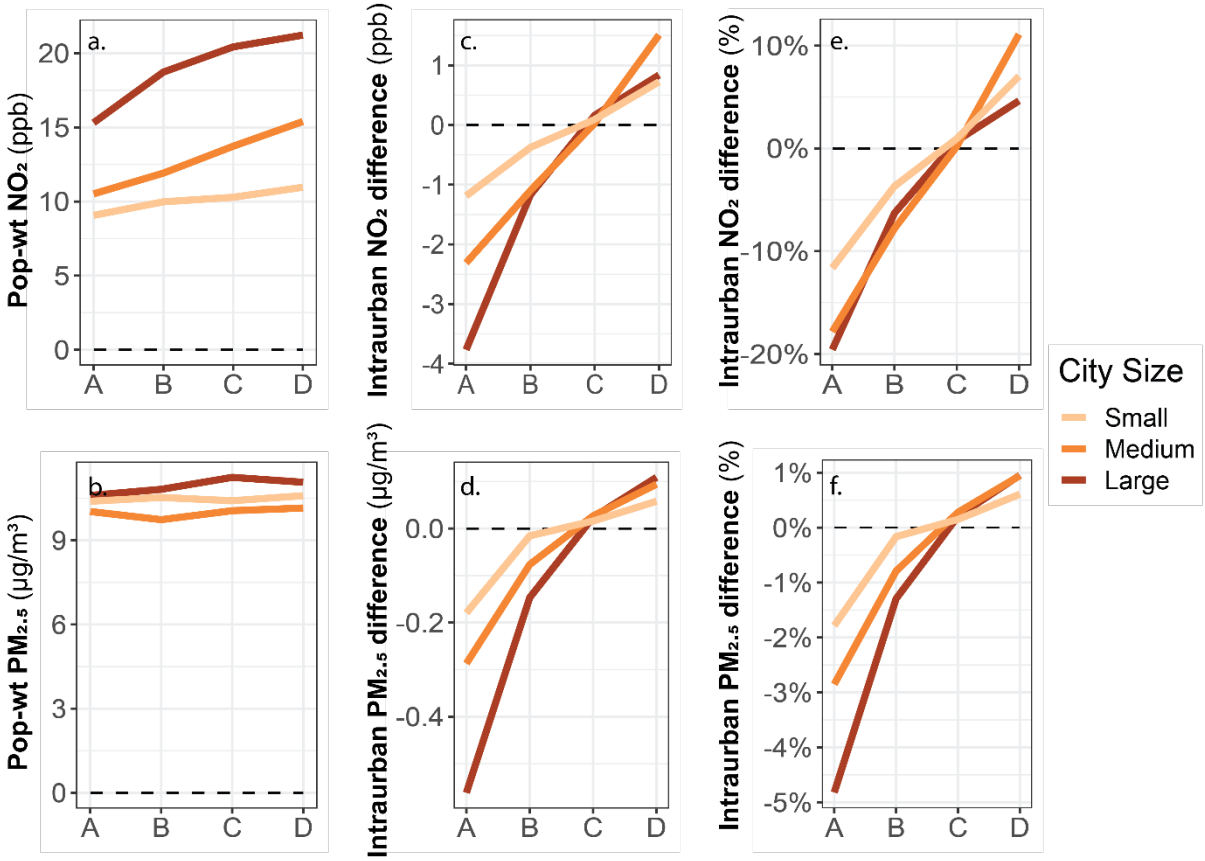

Figure S5. Comparison of population weighted mean pollution levels by city size. (a) Unadjusted NO_2 pop-weighted mean. (b) Unadjusted $PM_{2.5}$ pop-weighted mean. (c) Intraurban NO_2 pop-weighted mean. (d) Intraurban $PM_{2.5}$ pop-weighted mean. (e) Relative intraurban NO_2 pop-weighted mean. (f) Relative intraurban $PM_{2.5}$ pop-weighted mean. City size designations were determined to create groups of approximately equal overall populations within HOLC mapped areas (Small: n=162 pop=16 mill.; Medium: n=29 pop=11 mill.; Large: n=11 pop=18 mill.). This was chosen over using total city populations because, while unadjusted levels show some relationships with city size, the overall HOLC map size and population tended to correlate more with overall observed pollution level range. Larger cities tend to have substantially higher unadjusted PWM (a & b), but a monotonic A-D relationship is found across both pollutants and all 3 sizes of cities (c &d). NO_2 relative adjusted PWM also demonstrate less variation between city sizes than absolute adjusted PWM (e & f).

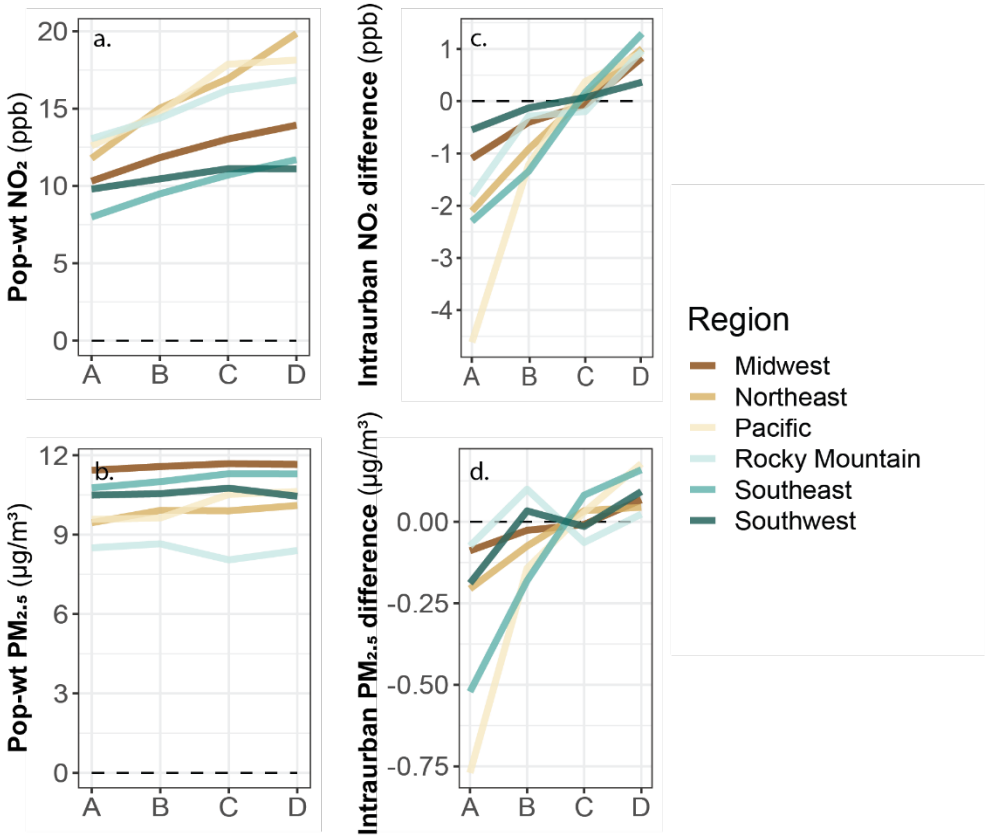

Figure S6. Regional comparison of population weighted mean pollution levels. (a) Unadjusted NO_2 pop-weighted mean. (b) Unadjusted $PM_{2.5}$ pop-weighted mean. (c) Intraurban NO_2 pop-weighted mean. (d) Intraurban $PM_{2.5}$ pop-weighted mean. Unadjusted results indicate regional trends for both pollutants. However, $PM_{2.5}$ results suggest that regional trends are a significant determiner of overall concentrations, whereas within city differences (A-D) are still substantial for NO_2**,** consistent with a wide body of literature. Adjusted intraurban results illustrate that regional trends are removed when pollution levels are adjusted relative to city means. NO_2 demonstrates a more consistent monotonic trend than $PM_{2.5}$, although trends are present for both, and the magnitude of HOLC trends varies by region. Note that these results can be heavily skewed by individual cities in the Pacific, Rocky Mountain, and Southwest due to the small numbers of mapped cities in these regions. The substantial A group differences observed for the Pacific region for both pollutants are driven by the particularly large A group disparities within only two cities: Oakland, CA and Los Angeles, CA.

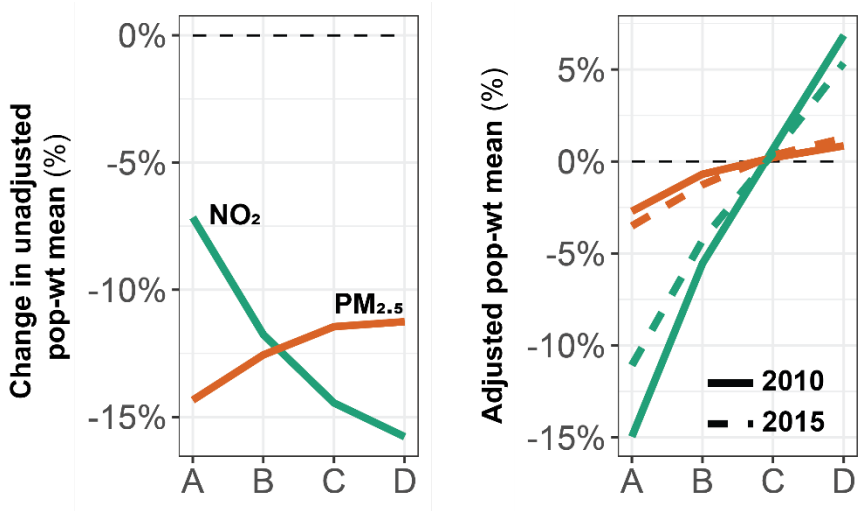

Figure S7. Model PWM pollution level interannual comparison. (a) Change in overall unadjusted pop-weighted mean between 2010 and 2015. (b) Comparison of overall interurban pop-weighted mean between 2010 and 2015. Overall populations decreased substantially between 2010 and 2015. However, changes in interurban HOLC, related disparities were minimal. This suggest that, at the summary level, populations within all grades experienced similar reductions in annual levels but that within-city HOLC related disparities were not substantially changed.

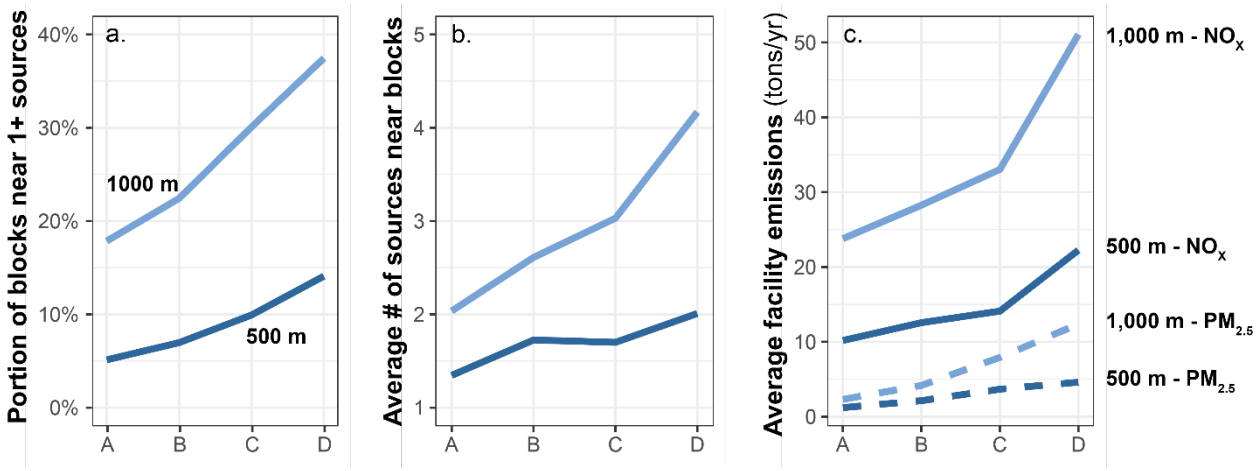

Figure S8. Distribution of industrial emissions sources by HOLC grade. National Emissions Inventory (NEI) facility prevalence in HOLC mapped areas by (a) % census blocks of a given grade within buffer distance of at least 1 source, by (b) average number of sources within buffer distance of a census block of a given grade for blocks near at least 1 source, and by (c) average annual emissions (tons/year) of sources within buffer distance of a census block of a given grade for blocks near at least 1 source. A strong, positive association was found between NEI facility prevalence and less-favorable HOLC grades.

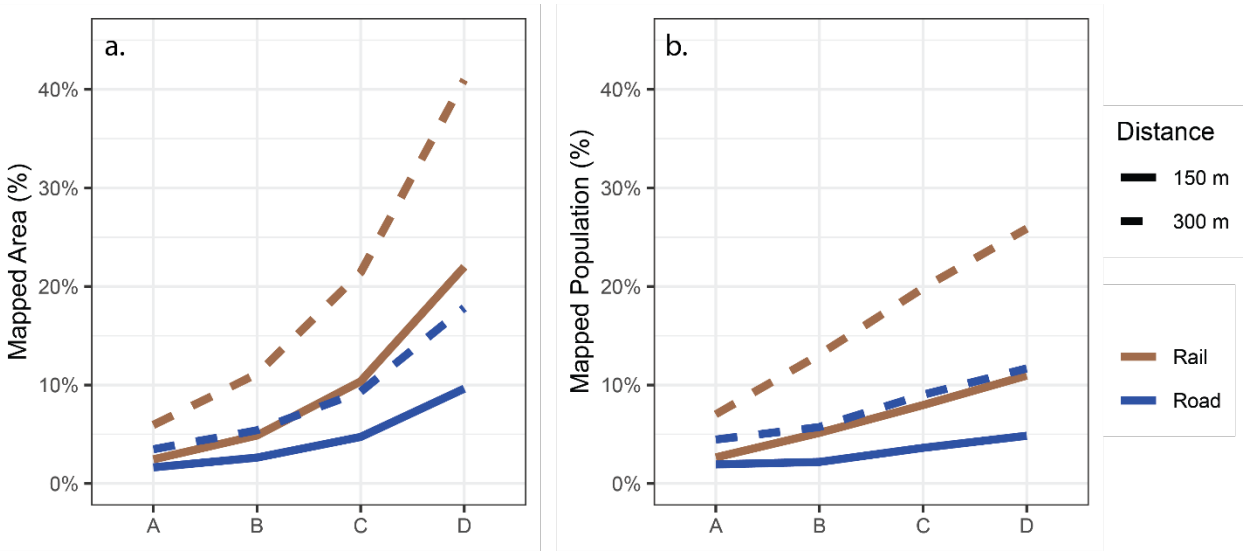

Figure S9. Transportation infrastructure by HOLC grade. Primary road and rail line prevalence in HOLC mapped areas by (a) % area within buffer distance of roads and rail and by (b) population living near roads and rail. A strong, positive association was found between rail line and road prevalence and worsening HOLC grades.

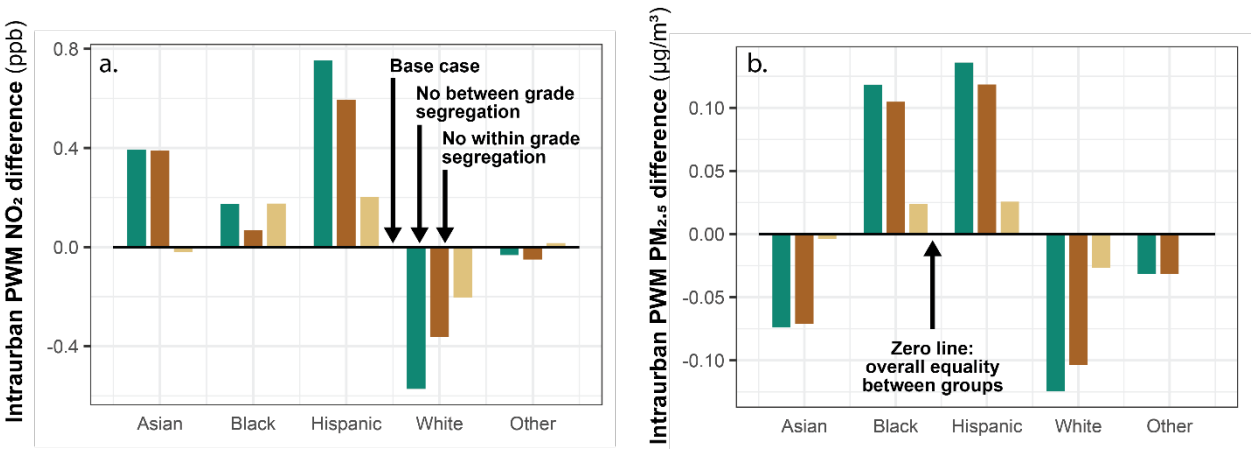

Figure S10. Segregation scenarios simulation. (a) Intraurban pop-weighted mean NO_2 levels by race/ethnicity. (b) Intraurban pop-weighted mean $PM_{2.5}$ concentrations by race/ethnicity. For $PM_{2.5}$, the aggregate intraurban disparities faced by each race-ethnicity are only weakly reduced (4-17%) by eliminating between-grade segregation, but would be nearly eliminated (79-95% reduction) by removing within-grade segregation. For NO_2, the results are more complex. Removing between-grade segregation would only minimally reduce the disparity for Asian and Hispanic populations (respectively 1 and 21%), and would moderately reduce the benefit that accrues to white populations (37%); in each case, within-grade disparity has a correspondingly larger effect. However, in contrast, removing between-HOLC-grade segregation would substantially reduce the intraurban NO_2 disparities that are faced by the Black population. This exception to the overall pattern is notable, especially in the context of the history of segregation and redlining: historic HOLC maps often made explicit and derogatory mention of Black communities as cause for an adverse classification.

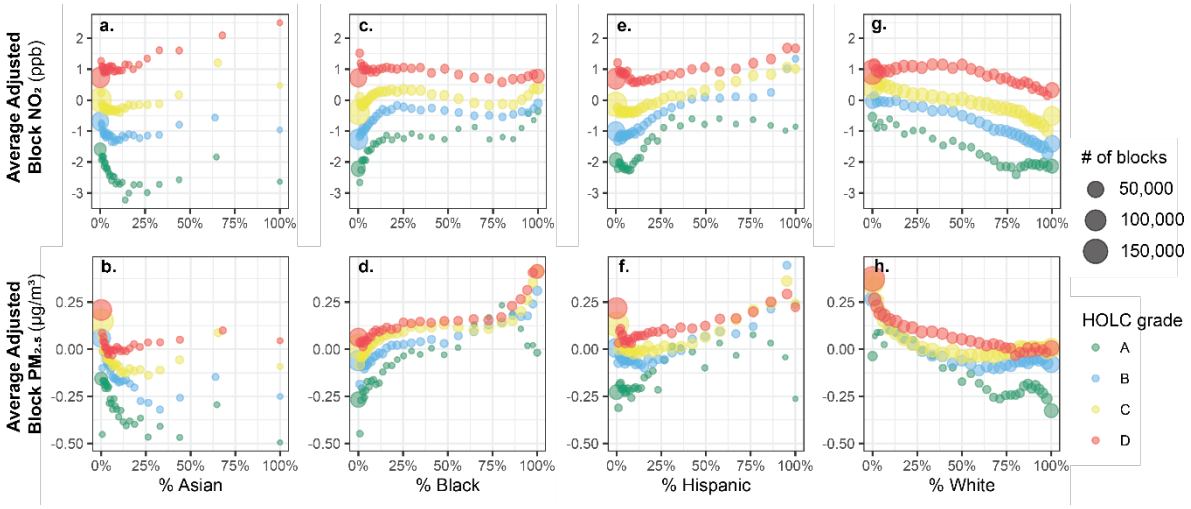

Figure S11. Average intraurban NO_2 levels by race/ethnicity prevalence and HOLC grade for (a) Asian, (c) Black, (e) Hispanic, and (g) White populations. Average intraurban $PM_{2.5}$ concentrations by race/ethnicity prevalence and HOLC grade for (b) Asian, (d) Black, (f) Hispanic, and (h) White populations. Census blocks which had a single HOLC grade areal apportionment of at least 80% were grouped into 31 equal number subsets based on increasing % race/ethnicity. Each subset was then divided by HOLC grade, and adjusted pollution level means are shown above, with point size indicating the number of census blocks included in each subset-grade group. These results illustrate that there are unique associations between intraurban pollution variation and race/ethnicity prevalence for each race/ethnicity group. Pollution tends to increase as Black and Hispanic prevalence increases, while levels decrease as the prevalence of White people increases. The relationships for Asian populations were less distinct, possibly in part due to the small number of census blocks with prevalence greater than 25%. Despite these overarching demographics related trends, distinct differences are observed by HOLC grade which, particularly for NO_2, seem to be somewhat stepwise with each worsening HOLC grade. This suggests that the HOLC effect on air pollution disparities is at least partially distinct from race/ethnicity related disparities, as is explored in more depth in Figure S10.