

SUPPLEMENTAL INFORMATION

Climate justice and California's methane super-emitters: An environmental equity assessment of community proximity and exposure intensity

Joan A. Casey^{1*}, Lara Cushing², Nicholas Depsky³, Rachel Morello-Frosch^{4*}

¹ Columbia University Mailman School of Public Health, Department of Environmental Health Sciences, New York, NY 10034, USA; jac2250@cumc.columbia.edu

² University of California, Los Angeles Fielding School of Public Health, Department of Environmental Health Sciences, Los Angeles, CA 90095, USA

³ University of California, Berkeley, Energy and Resources Group, Berkeley, CA 94720, USA

⁴ University of California, Berkeley, Department of Environmental Science, Policy and Management and School of Public Health, Berkeley, CA 94720, USA; rmf@berkeley.edu

*Co-corresponding authors

The supplemental information is 18 pages and contains 3 tables and 14 figures.

Supplemental Table 1: Super-emitters excluded and included in analyses.

Category	No populated area within 2km (excluded)	Populated area within 2km (included)
Landfill/compost	3 (9%)	29 (91%)
Power plant	0	7 (100%)
Refinery	0	37 (100%)
Wastewater treatment	0	14 (100%)
Oil/gas distribution	20 (26%)	56 (74%)
Oil/gas production	56 (31%)	127 (69%)
Dairy/manure	2 (1%)	213 (99%)
Total	81 (14%)	483 (86%)

Supplemental Table 2: Residential Parcel Classifications. Residential parcels were defined as any parcel classified by one of the following land-use codes in the statewide tax parcel database. Large parcels for any use low-density use code with areas greater than 1-acre (4,047 m²) were assumed to contain unpopulated, open space and were excluded. High-density residential parcels that tend to be larger (e.g., an entire apartment complex) were allowed to have areas of up to 50-acres before being excluded and are indicated in the list below by (*).

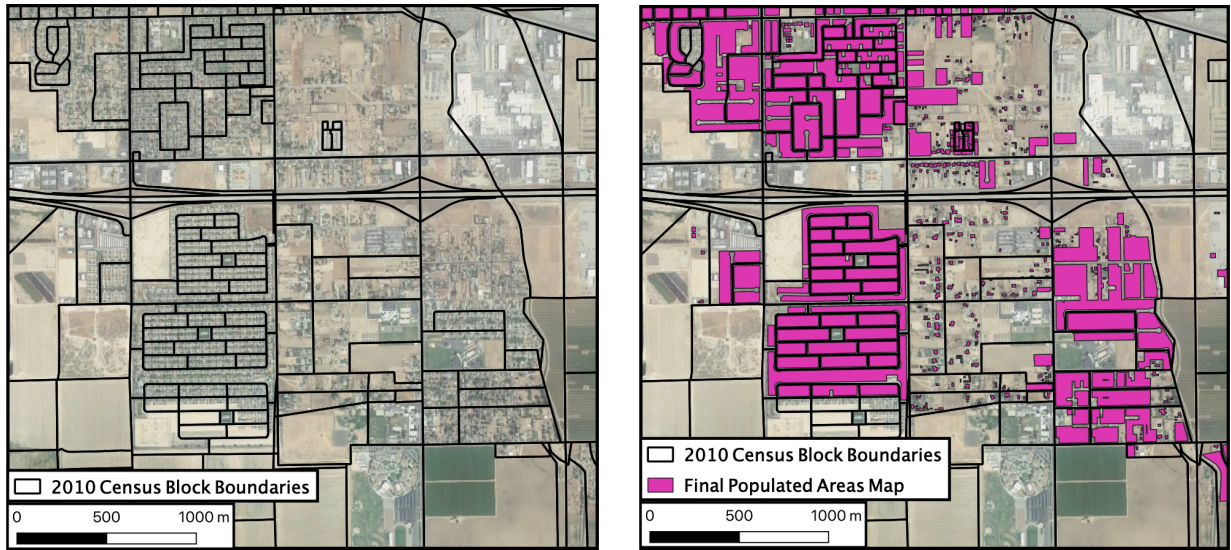
Parcel Land-Use Code
APARTMENT HOUSE (100+ UNITS)*
APARTMENT HOUSE (5+ UNITS)*
APARTMENTS (GENERIC)*
CLUSTER HOME (RESIDENTIAL)
COMM/OFC/RES MIXED USE
CONDOMINIUM (RESIDENTIAL)*
COOPERATIVE (RESIDENTIAL)*
DORMITORY, GROUP QUARTERS (RESIDENTIAL)
DUPLEX (2 UNITS, ANY COMBINATION)
FRATERNITY HOUSE, SORORITY HOUSE
GARDEN APT, COURT APT (5+ UNITS)*
HIGHRISE APARTMENTS*
HOMES (RETIRED; HANDICAP, REST; CONVALESCENT; NURSING)
MANUFACTURED, MODULAR, PRE-FABRICATED HOMES
MISC RESIDENTIAL IMPROVEMENT
MOBILE HOME
MOBILE HOME PARK, TRAILER PARK
MULTI-FAMILY DWELLINGS (GENERIC, ANY COMBINATION 2+)
PLANNED UNIT DEVELOPMENT (PUD) (RESIDENTIAL)
QUADRUPLEX (4 UNITS, ANY COMBINATION)
RESIDENTIAL (GENERAL) (SINGLE)
RESIDENTIAL COMMON AREA (CONDO/PUD/ETC.)
RESIDENTIAL INCOME (GENERAL) (MULTI-FAMILY)
RURAL RESIDENCE (AGRICULTURAL)
SINGLE FAMILY RESIDENTIAL
STORES & APARTMENTS
TIMESHARE (RESIDENTIAL)
TOWNHOUSE (RESIDENTIAL)
TRIPLEX (3 UNITS, ANY COMBINATION)
ZERO LOT LINE (RESIDENTIAL)

Supplemental Table 3: Distribution of sociodemographic variables by exposed versus unexposed groups and scaling factors used for regression analyses.

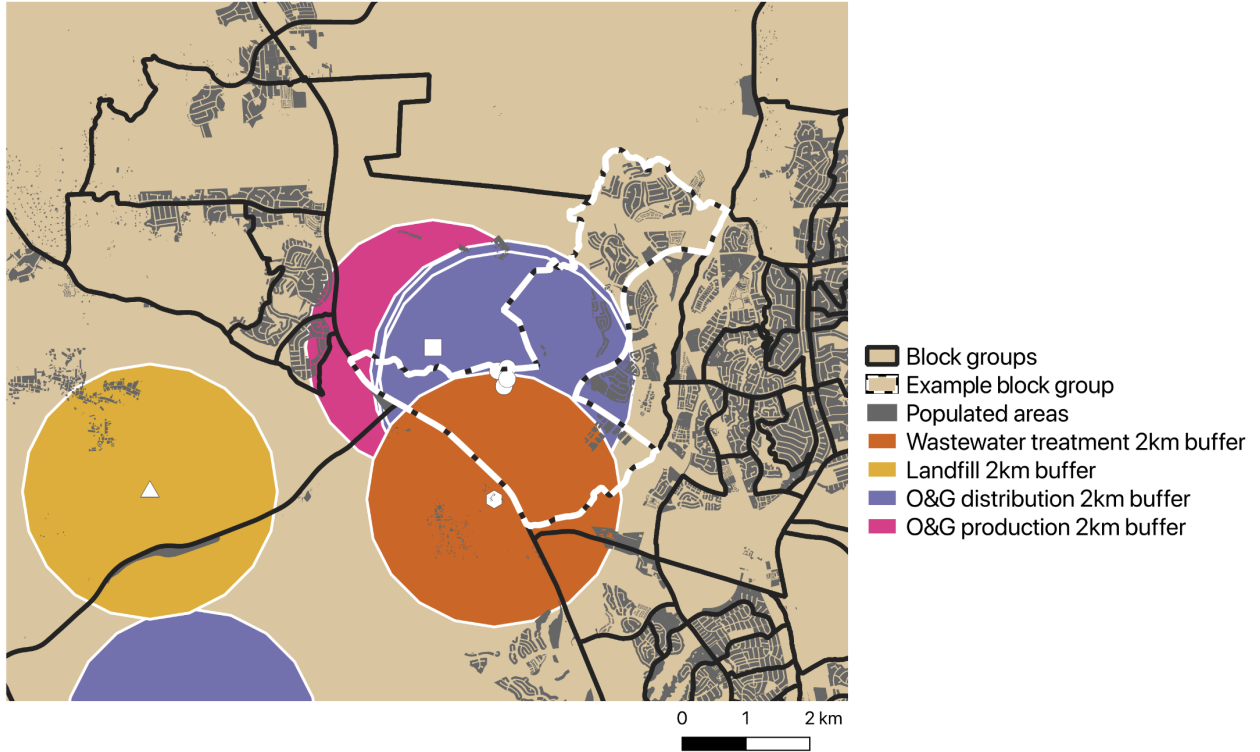
	Exposed block groups (within 2km of a super-emitter) N = 951	Unexposed block groups (5-10km from a super-emitter) N = 8,722	Linear β interpretation
Category	Median (25 th , 75 th percentiles)		
Race/ethnicity, %			
Hispanic	33 (16, 61)	37 (17, 67)	Per 10% increase
Non-Hispanic			
Native American ^a	0.4 (1.7)	0.2 (0.9)	Per 1% increase
Asian	6 (1, 16)	7 (2, 18)	Per 10% increase
Black	2 (0, 8)	2 (0, 8)	Per 10% increase
White	35 (11, 61)	28 (8, 56)	N/A
Poverty, %	15 (13, 23)	17 (14, 25)	Per 10% increase
Renters, %	41 (23, 62)	47 (25, 71)	Per 10% increase
Limited English-speaking households, %	6 (2, 14)	8 (3, 17)	Per 1% increase
Voters, %	71 (62, 76)	67 (59, 75)	Per 10% increase
Uninsured, %	10 (5, 18)	12 (6, 20)	Per 1% increase
Median household income, \$	\$64,700 (45,000, 92,200)	\$60,900 (41,700, 86,300)	Per \$10,000 increase
Less than a high school diploma, %	14.2 (5.3, 31.4)	16.2 (6.0, 33.5)	Per 10% increase
Unemployed, %	8 (5, 12)	8 (5, 12)	Per 1% increase
SNAP, %	7 (1, 15)	7 (2, 17)	Per 1% increase
Population density, individuals per km ²	3100 (2750, 4540)	4280 (3820, 5510)	Per 500 individuals per km ² increase

^a Mean (SD)

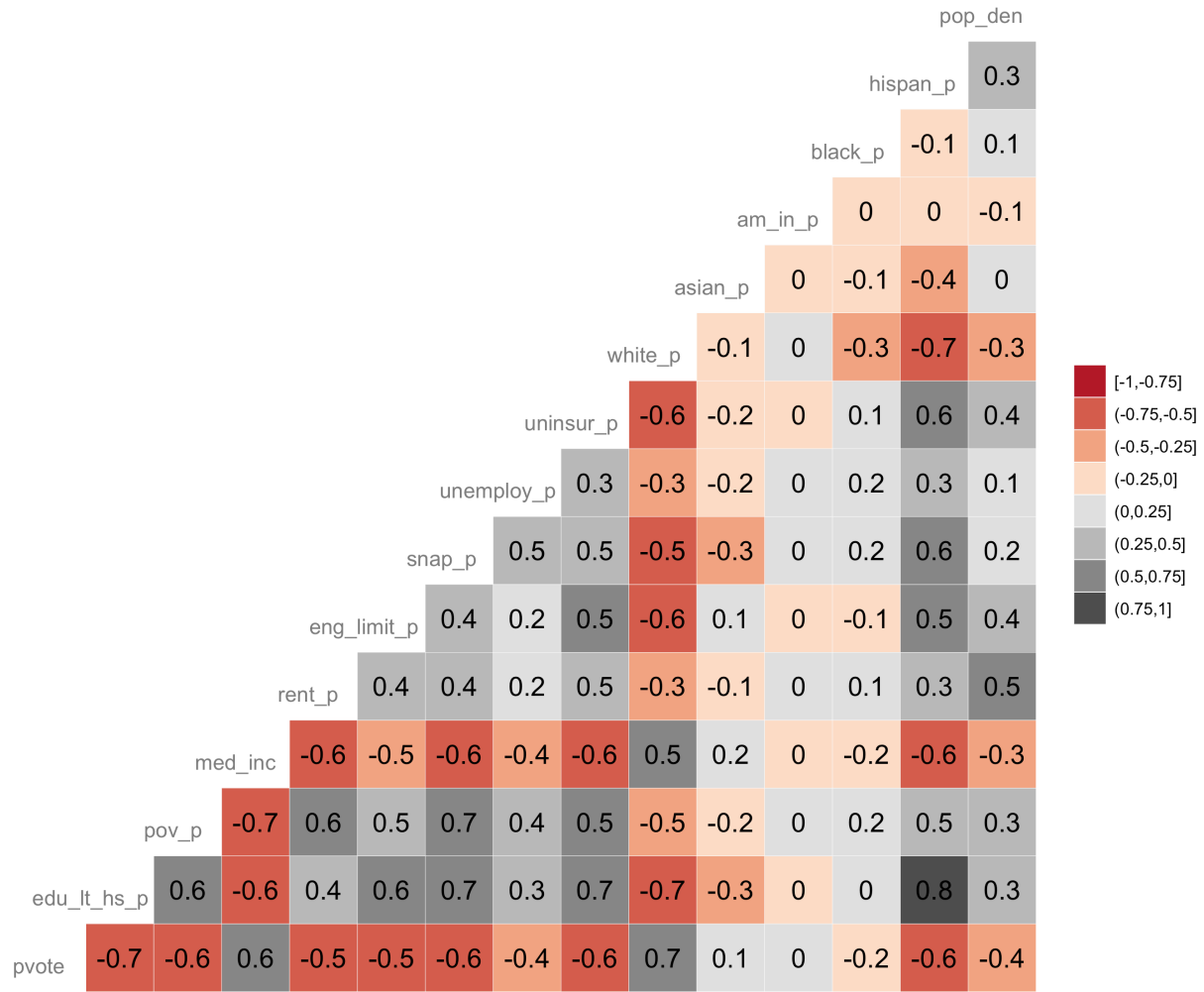
Supplemental Figure 1. Example of the creation of populated areas layer from parcel, building footprint and block boundary data. Image shown is in eastern Bakersfield, CA.



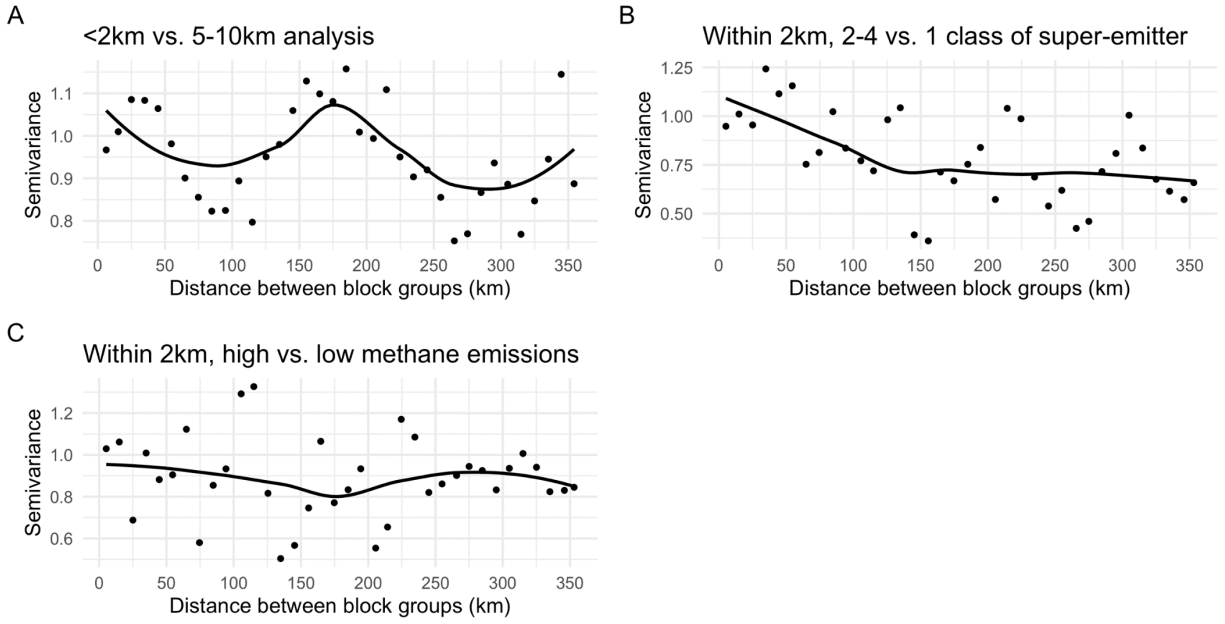
Supplemental Figure 2: Example of block groups exposed to multiple classes of super-emitter. For example, the block group outlined in a dashed white line contains populated areas located within 2km of at least two classes of super-emitter (wastewater treatment and oil and gas distribution). White hexagons represent wastewater treatment facilities, squares are oil and gas production sites, circles are oil and gas distribution sites, and triangles are landfills. Grey polygons are populated areas within block groups and the larger polygons bounded in black are block groups. Analyses were conducted at the block group-level, but only those block groups with a populated area located within 2km of a super-emitter were included. O&G, oil and gas.



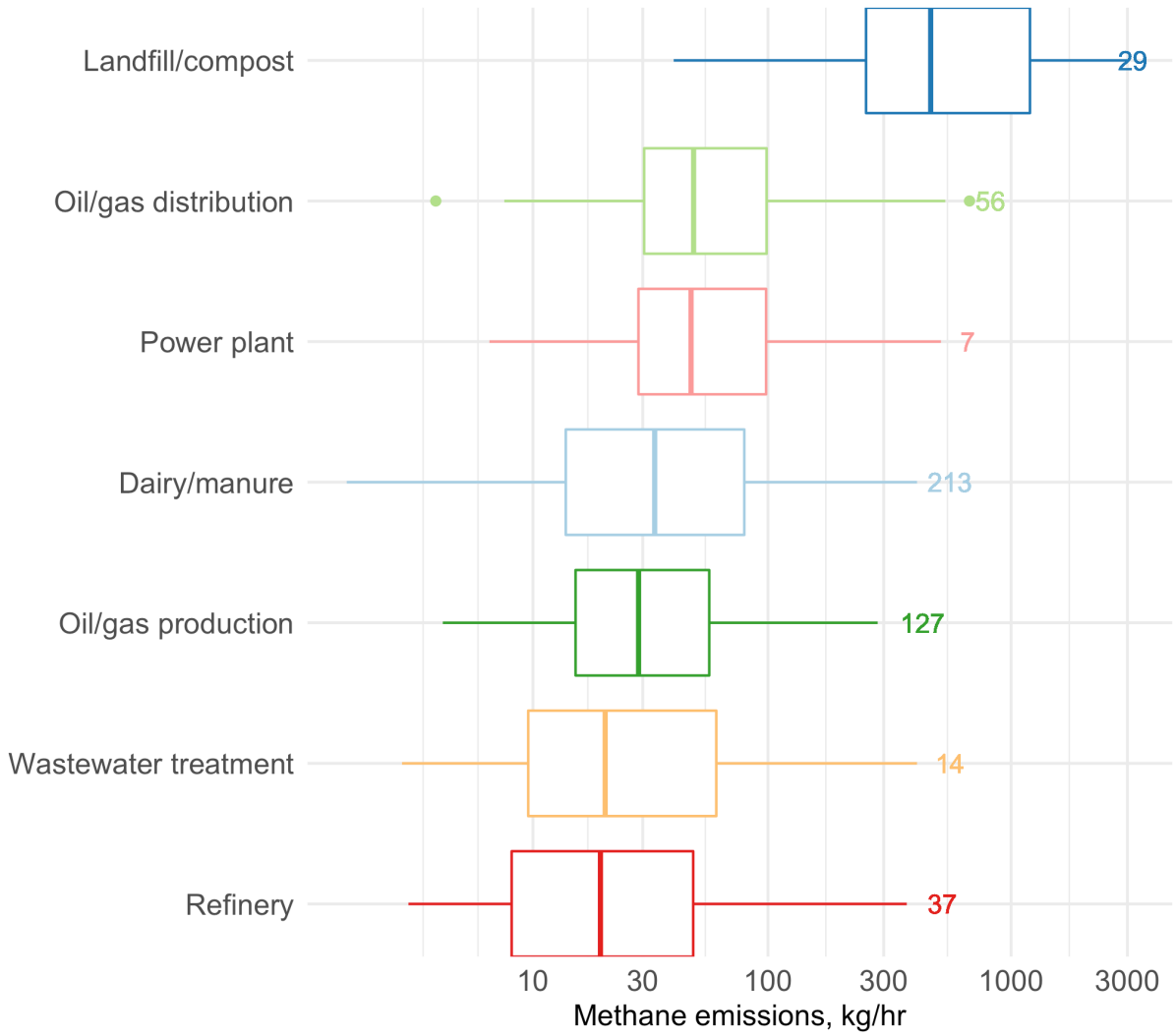
Supplemental Figure 3: Spearman correlation matrix for block group-level sociodemographic variables.



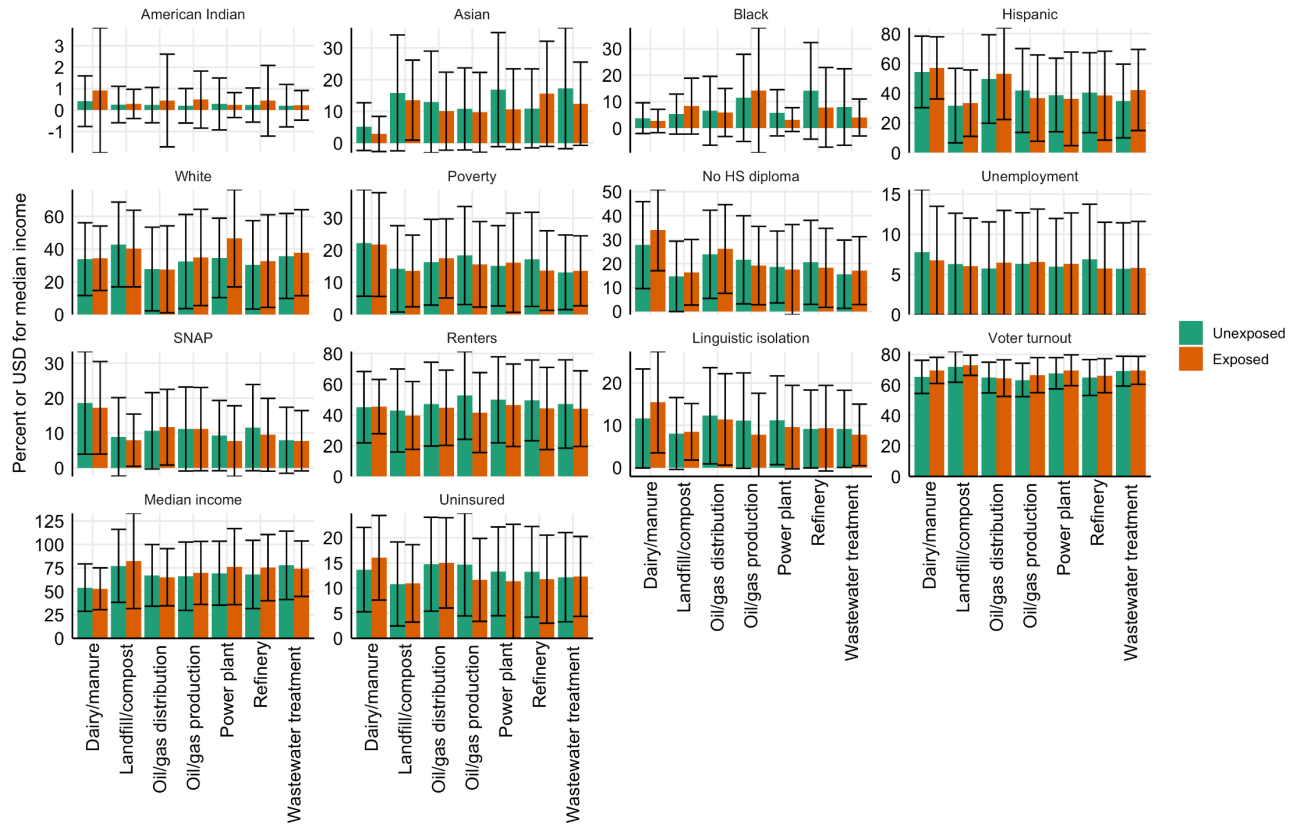
Supplemental Figure 4: Semivariograms for the three analyses: A. Main 2km vs. 5-10km; B. 2-4 vs. 1 class of super-emitter within 2km; C. High (>3rd quartile) vs. low (quartiles 1-3) CH₄ emissions within 2km. The shapes of the semivariograms are consistent with limited residual spatial autocorrelation. Based on residuals from logistic mixed models with a random intercept for county adjusted for block group-level for population density, percent individuals of non-Hispanic Asian, Black, and Native American race/ethnicity, and percent individuals of Hispanic race/ethnicity, percent individuals living below the federal poverty threshold, percent voters, percent renters, percent limited English speaking households, and percent uninsured individuals.



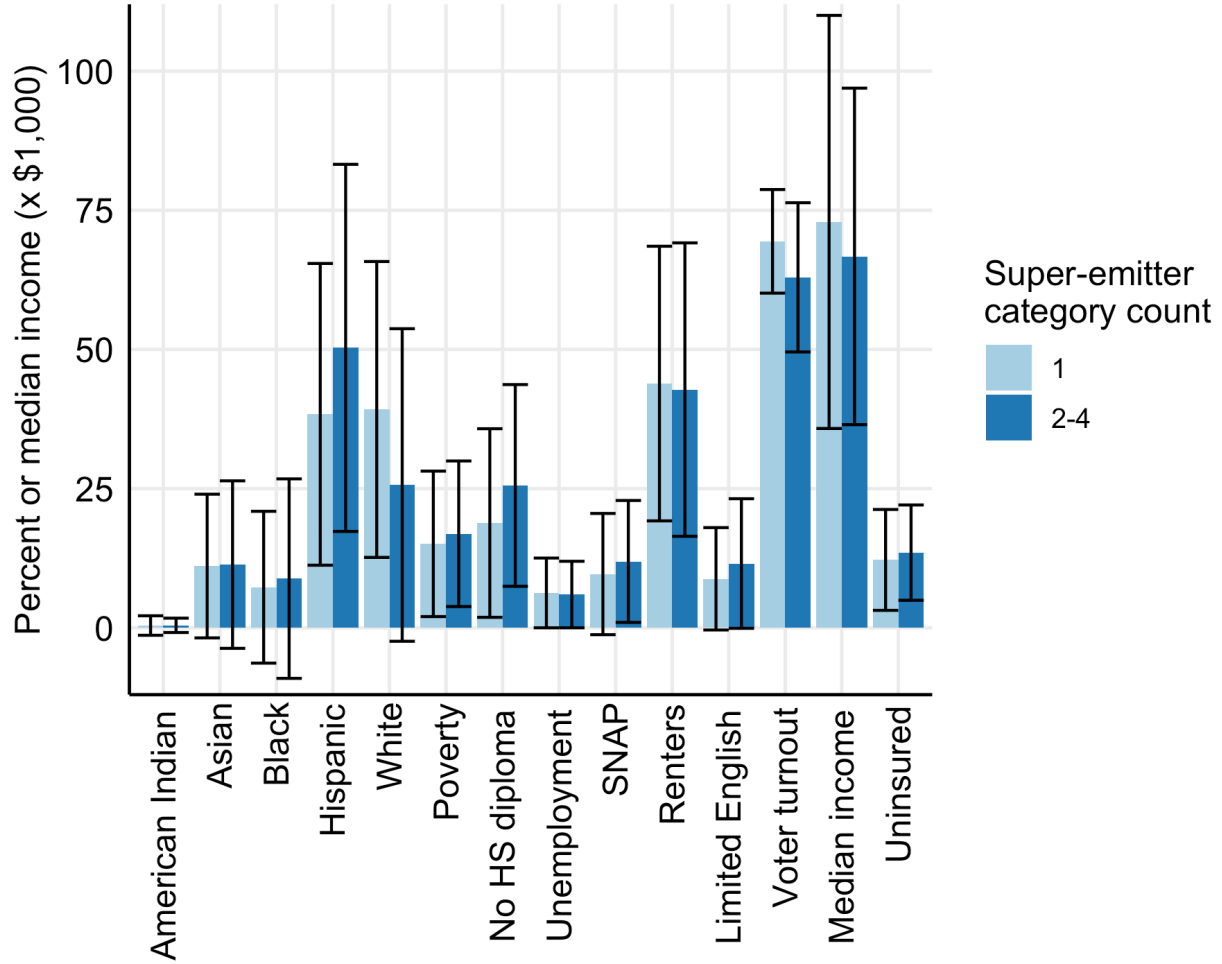
Supplemental Figure 5: Distribution of methane emissions (kg/hr) by super-emitter category. Numbers indicate the count of super-emitters in each category. The x-axis is log-scale.



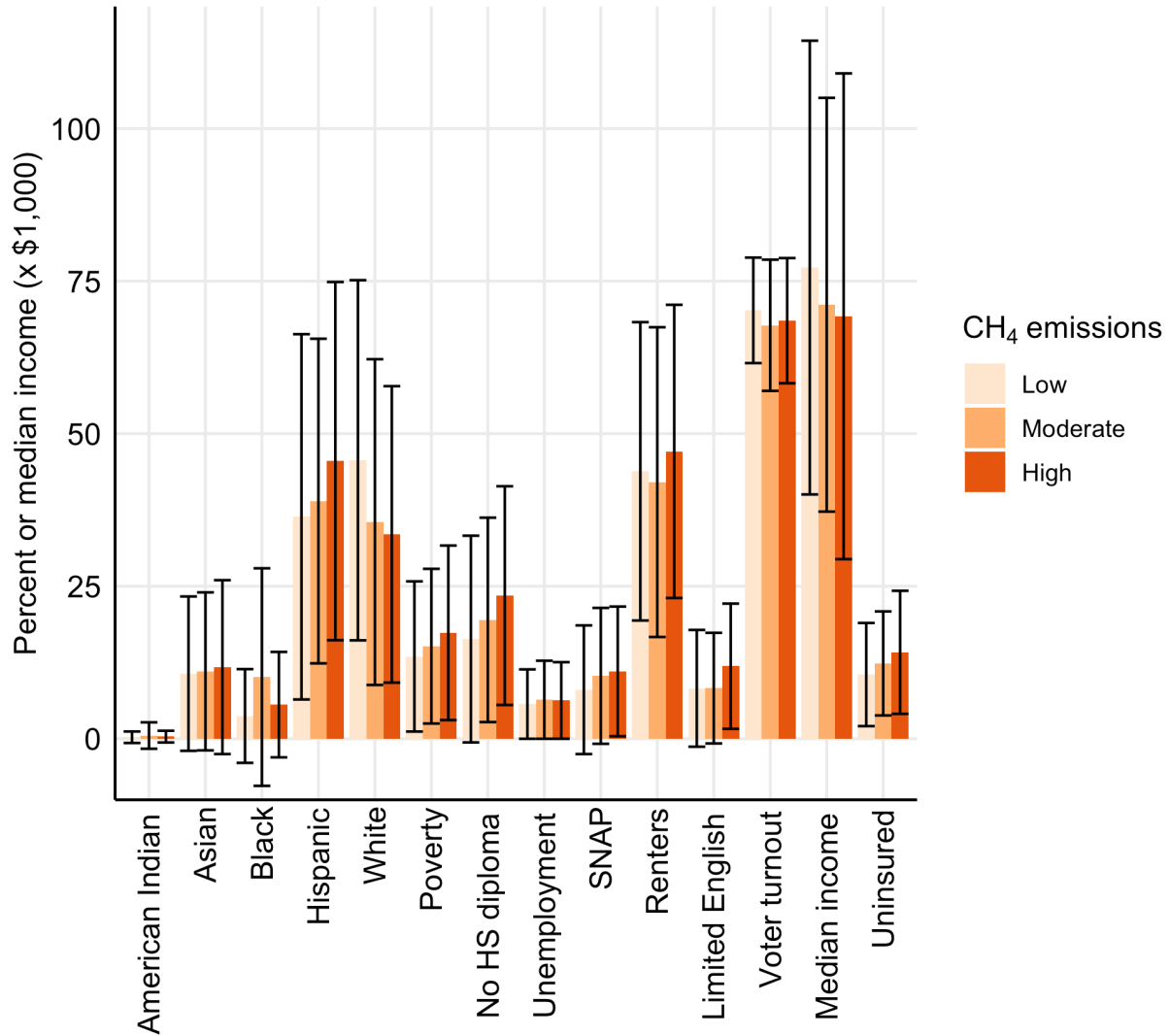
Supplemental Figure 6: Mean block group-level sociodemographic characteristics by super-emitter class. Exposed block groups were those with a populated area located within 2km of a super-emitter and unexposed those located 5-10km from a super-emitter. Bars represent 1-SD.



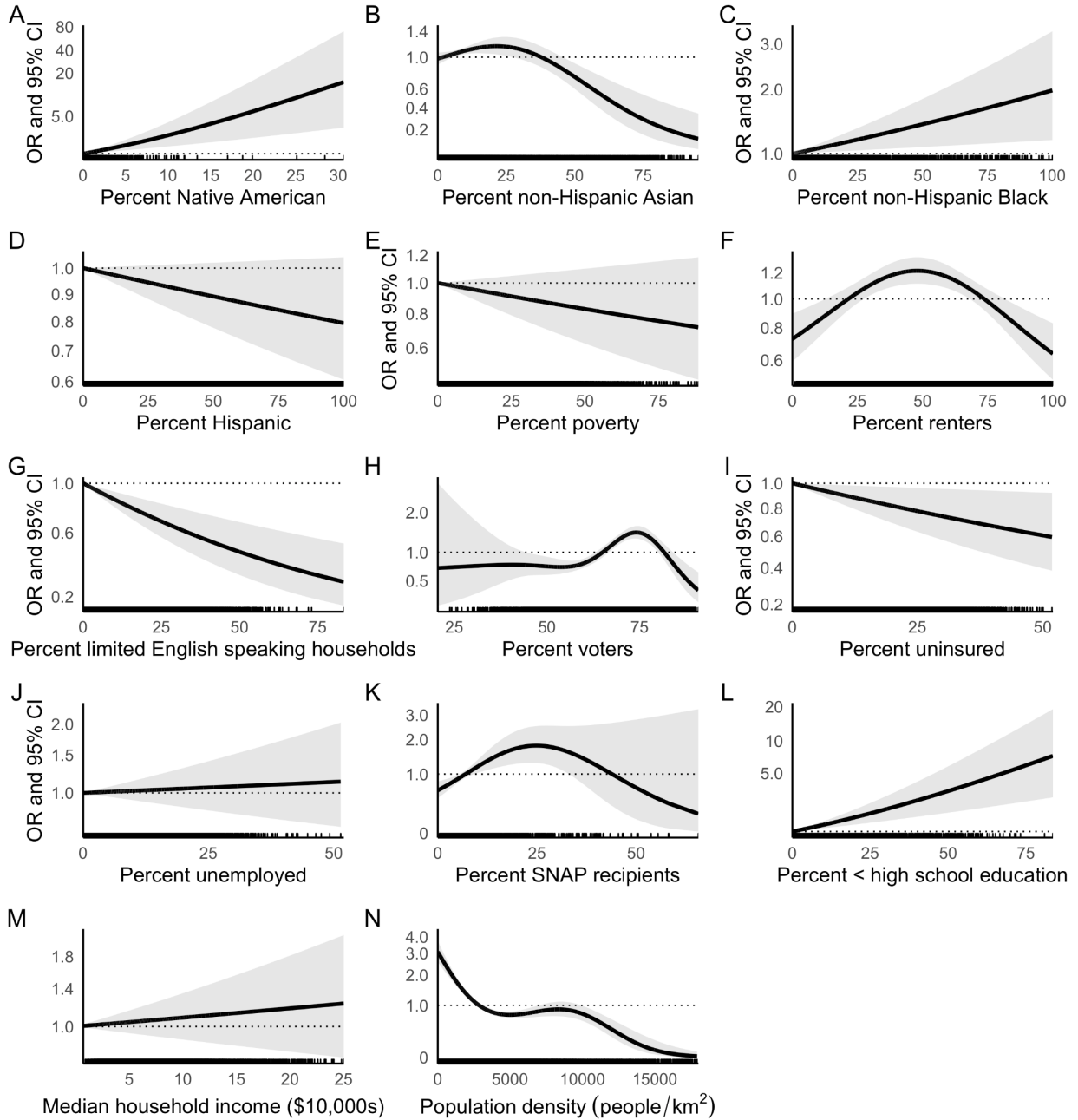
Supplemental Figure 7: Mean block group-level sociodemographic characteristics among block groups located within 2km of a super-emitter, stratified by the number of categories of super-emitter located within 2km. For example, block groups located within 2km of a refinery and a dairy would fall in the 2-4 category. Bars represent 1-SD.



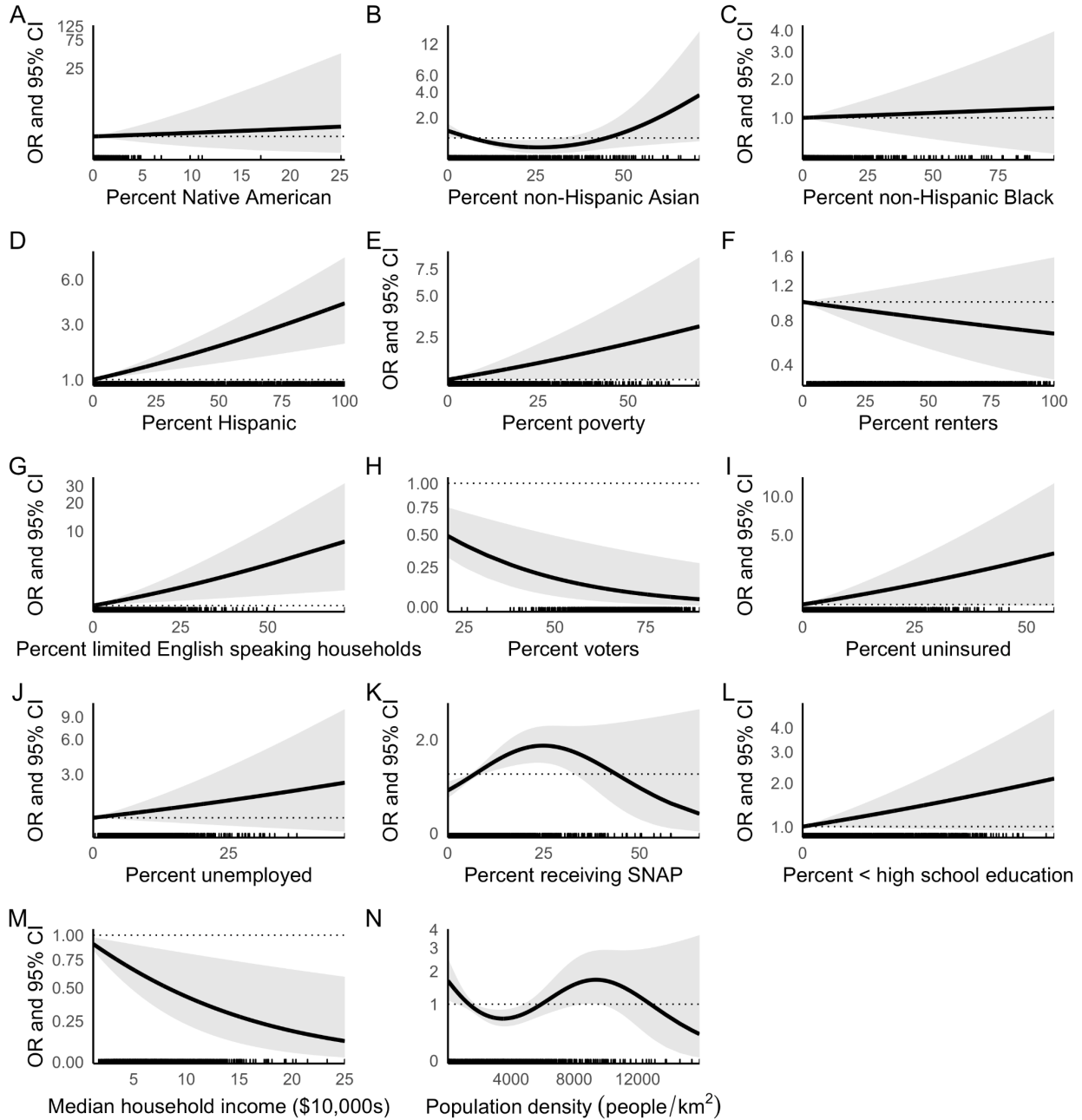
Supplemental Figure 8: Mean block group-level sociodemographic characteristics by the sum of super-emitter CH₄ emissions (kg/hr) within 2km of the block group. CH₄ emissions were categorized based on their distribution into low (<40 kg/hr, first tertile), moderate (40 to <185/hr, second tertile), and high (≥185 kg/hr, third tertile).



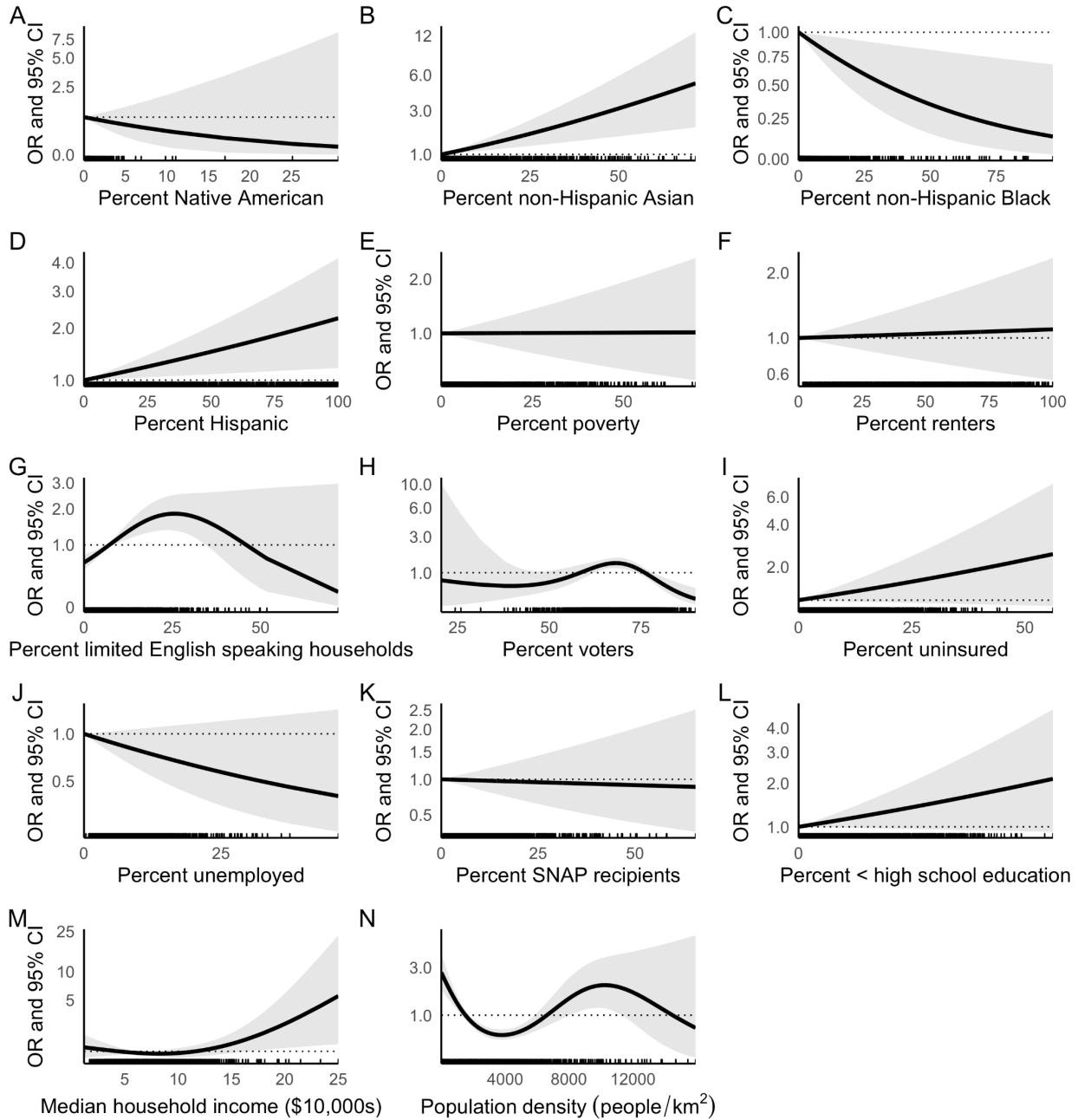
Supplemental Figure 9: Unadjusted association between sociodemographic variables and odds of being located within 2km versus 5-10km from a CH₄ super-emitter. Includes n = 951 exposed and n = 8722 unexposed block groups. Black lines are odds ratios and grey areas represent the 95% confidence intervals. Results from a generalized additive mixed model with a logit link and a random intercept for county adjusted only for block group-level population density. Rug plot displayed along the x-axis shows the number of observations at each level of the respective sociodemographic variable. CI, confidence interval; OR, odds ratio. Non-linear associations in panels B, F, H, K, and N were statistically significant at the $\alpha=0.05$ level.



Supplemental Figure 10: Unadjusted association between sociodemographic variables and odds of being located within 2km of 2-4 versus 1 category of CH₄ super-emitter, among block groups located within 2km of at least 1 super-emitter (n = 951). Black lines are odds ratios and grey areas represent the 95% confidence interval. Results from a generalized additive mixed model with a logit link and a random intercept for county adjusted only for population density. Rug plot displayed along the x-axis shows the number of observations at each level of the respective sociodemographic variable. CI, confidence interval; OR, odds ratio. Non-linear associations in panels B, K, and N were statistically significant at the $\alpha=0.05$ level.

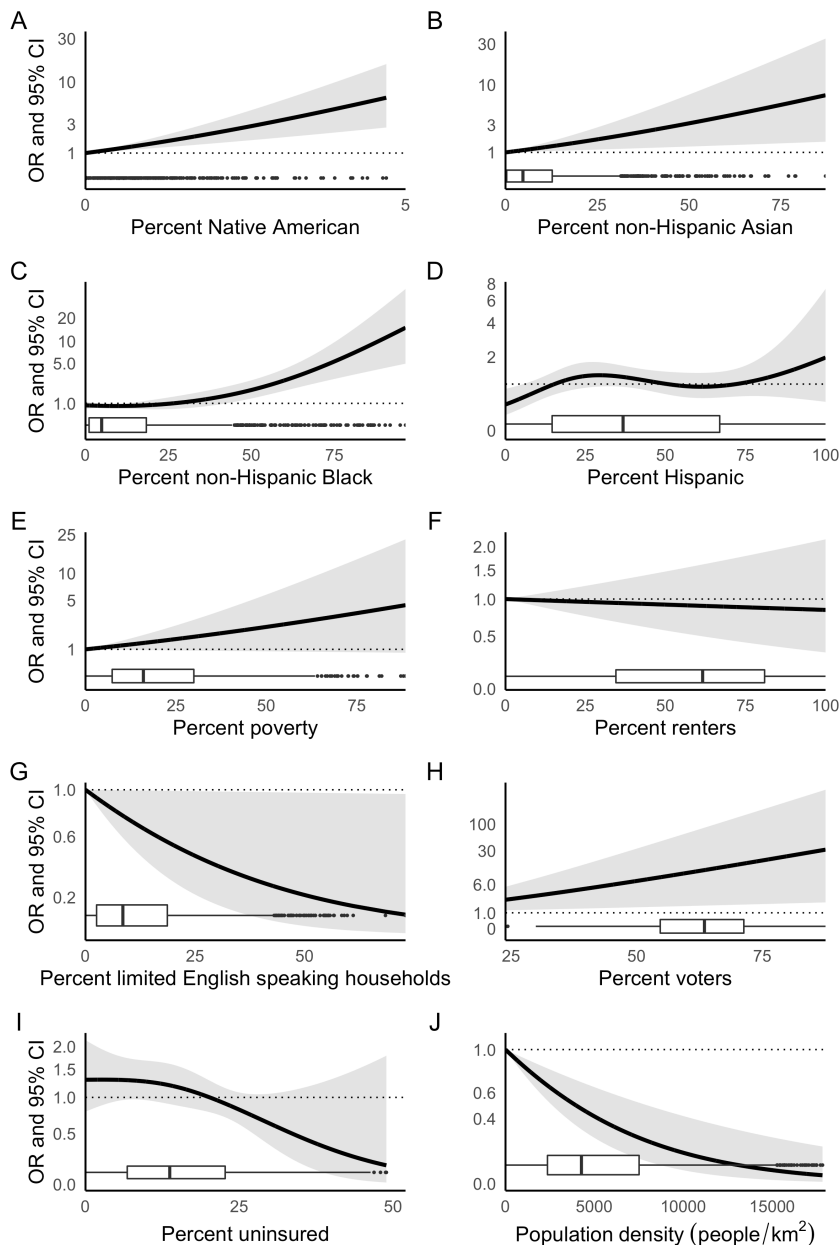


Supplemental Figure 11: Unadjusted association between sociodemographic variables and odds of being exposed to high (>quartile 3 [185 kg/hr]) versus low (quartile 1-3 [2.8-185 kg/hr]) CH₄ emissions, among block groups located within 2km of at least 1 super-emitter (n = 951). Black lines are odds ratios and grey areas represent the 95% confidence interval. Results from a generalized additive mixed model with a logit link and a random intercept for county adjusted only for population density. Rug plot displayed along the x-axis shows the number of observations at each level of the respective sociodemographic variable. CI, confidence interval; OR, odds ratio. Non-linear associations in panels G, H, M, and N were statistically significant at the $\alpha=0.05$ level.

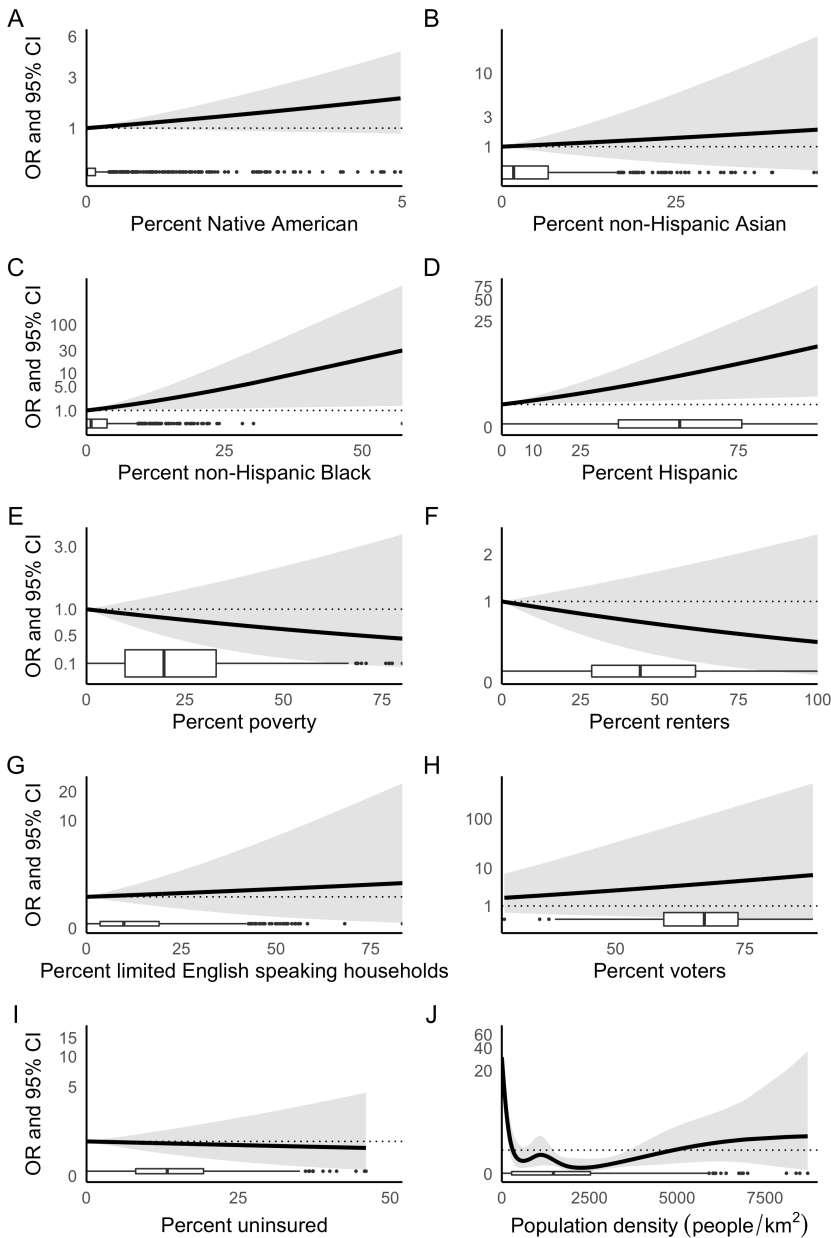


Supplemental Figure 12: Association between sociodemographic variables and odds of being located within 2km versus 5-10km from an oil and gas production CH₄ super-emitter. Includes n = 177 exposed and n = 1382 unexposed block groups. Black lines are odds ratios and grey areas represent the 95% confidence intervals. Results from a generalized additive mixed model with a logit link and a random intercept for county adjusted for block group-level percent individuals of non-Hispanic Native American, Asian, and Black race/ethnicity, and percent individuals of Hispanic race/ethnicity, percent individuals living below the federal poverty threshold, percent renters, percent limited English speaking households, percent voter turnout, percent uninsured individuals, and population density. Boxplot displayed along the x-axis shows the number of observations at each level of the respective sociodemographic variable.

CI, confidence interval; OR, odds ratio. Non-linear associations in panels C, D, I, and J were statistically significant at the $\alpha=0.05$ level.



Supplemental Figure 13: Association between sociodemographic variables and odds of being located within 2km versus 5-10km from dairy or manure CH₄ super-emitter. Includes n = 87 exposed and n = 697 unexposed block groups. Black lines are odds ratios and grey areas represent the 95% confidence intervals. Results from a generalized additive mixed model with a logit link and a random intercept for county adjusted for block group-level percent individuals of non-Hispanic Native American, Asian, and Black race/ethnicity, and percent individuals of Hispanic race/ethnicity, percent individuals living below the federal poverty threshold, percent renters, percent limited English speaking households, percent voter turnout, percent uninsured individuals, and population density. Boxplot displayed along the x-axis shows the number of observations at each level of the respective sociodemographic variable. CI, confidence interval; OR, odds ratio. The non-linear association in panel J was statistically significant at the $\alpha=0.05$ level.



Supplemental Figure 14: Spearman correlation between 2018 California Air Resources Board Pollution Mapping Tool annual reported CH₄ emissions in MT CO₂e and co-pollutant emissions.

	Refineries (n=21)	Utility power plants (n=145)
VOCs (ton)	0.78	0.47
NOx (ton)	0.80	0.60
SOx (ton)	0.76	0.56
PM _{2.5} (ton)	0.79	0.59

