Supplemental Material

Race/Ethnicity, Socioeconomic Status, Residential Segregation, and Spatial Variation in Noise Exposure in the Contiguous United States

Joan A. Casey, Rachel Morello-Frosch, Daniel J. Mennitt, Kurt Fristrup, Elizabeth L. Ogburn, and Peter James

Table of Contents

References

Table S1. Results of noise model cross-validation. The root mean squared error (RMSE), median absolute deviation (MAD), and percent variance explained calculated from the residuals of the null model and a pooled set of residuals from the 995 cross validated geospatial sound models (GSM). This table was adapted from Mennitt and Fristrup 2016.

Metric	GSM RMSE (dB)	Null RMSE (dB)	GSM MAD (dB)	Null MAD (dB)	Variance explained (%)
L_{50}, nighttime	4.91	11.38	2.40	7.6	81
L_{10}, daytime	4.51	10.04	2.36	7.9	80
L_{50}, daytime	4.40	11.12	2.29	8.7	84

Multiple models were constructed to evaluate predictive performance of the geospatial sound modeling method applied to the various A-weighted metrics (Mennitt and Fristrup 2016). For each metric of interest, a full model was constructed using the training set containing all available observations. These fits showed excellent correlation with the empirical data. While the ensemble nature of the random forest algorithm helped reduce overfitting, any statistical model is at risk of overfitting the training data. The geospatial sound model was cross-validated using a leave-one-out process that omitted each site (all of its seasonal observations), constructed a new model (995 in total), and tested the model output against the observations for the omitted site. Removing collocated observations from the training set eliminated bias from temporal correlation. Table S1 shows statistics comparing cross-validated residuals to those from the null model.

Table S2. Distribution of anthropogenic L_{50} nighttime, L_{50} daytime, and L_{10} daytime noise among suburban/rural residents by race/ethnicity and socioeconomic characteristics

Characteristic	Total, n (%)	Median (IQR) anthropogenic noise, dBA[a]		
		L_{50} nighttime	L_{50} daytime	L_{10} daytime
Total population	45,999,315 (100)	38.8 (33.7-42.2)	37.0 (32.8-43.2)	43.7 (40.2-49.6)
Population <5 years	2,863,364 (6.2)	39.3 (34.2-42.5)	37.7 (33.2-43.8)	44.3 (40.7-50.1)
Population ≥ 5 years	43,135,951 (93.8)	38.8 (33.6-42.2)	37.0 (32.8-43.2)	43.7 (40.2-49.6)
Race/ethnicity[b]				
Hispanic	3,264,608 (7.1)	40.0 (34.2-43.0)	40.4 (34.7, 45.2)	47.1 (42.2-51.1)
Non-Hispanic				
American Indian	731,771 (1.6)	32.9 (24.9-40.0)	33.2 (28.7-38.5)	41.4 (37.9-45.8)
Asian	305,915 (0.7)	41.0 (36.1-43.3)	42.0 (35.7-45.7)	48.5 (42.6-51.8)
Black	3,843,814 (8.4)	41.2 (34.9-43.8)	40.1 (33.1-46.3)	46.1 (40.6-51.4)
White	37,130,640 (80.7)	38.7 (33.6-42.0)	36.9 (32.8-42.8)	43.5 (40.2-49.3)
Income ≤ poverty threshold[c]	7,443,243 (16.8)	39.7 (34.3-42.9)	38.1 (32.9-44.6)	44.6 (40.4-50.6)
Income > poverty threshold[c]	36,936,461 (83.2)	38.6 (33.5-42.1)	36.8 (32.7-42.9)	43.5 (40.2-49.3)
Total population ≥ 25 years	6,584,831 (100)	39.1 (33.7-42.4)	37.3 (32.7-43.8)	43.9 (40.1-50.0)
< High school education	1,248,645 (19.0)	39.4 (34.2-42.7)	37.2 (32.5-44.1)	43.8 (40.0-50.2)
≥ High school education	5,336,186 (81.0)	39.0 (33.6-42.4)	37.3 (32.7-43.7)	43.9 (40.1-50.0)
Total households	17,944,466 (100)	38.9 (33.7-42.3)	37.2 (32.8-43.4)	43.8 (40.3-49.8)
Median household income				
Quartile 1 ($3,981-$32,199)	4,489,489 (25.0)	40.8 (35.6-43.4)	40.8 (33.5-45.7)	47.3 (40.9-51.6)
Quartiles 2-4 (32,200-$185,682)	13,454,977 (75.0)	38.3 (33.1-41.8)	36.5 (32.7-42.2)	43.3 (40.1-48.7)
Linguistically isolated households	285,223 (1.6)	40.4 (34.8-43.3)	40.5 (34.7-45.4)	47.2 (42.0-51.2)
Non-linguistically isolated households	17,659,243 (98.4)	38.5 (33.4-42.1)	36.7 (32.6-42.7)	43.3 (40.1-49.1)
Housing tenure				
Owner-occupied homes	13,112,987 (73.1)	38.2 (33.1-41.9)	36.4 (32.4-42.2)	43.1 (39.9-48.7)
Renter-occupied homes	4,831,479 (26.9)	40.6 (35.5-43.2)	40.6 (34.5-43.2)	47.1 (41.7-51.4)
Total families	12,167,035 (100)	38.6 (33.4-42.1)	36.7 (32.6-42.8)	43.4 (40.1-49.2)
Unemployed families	574,953 (4.7)	39.2 (34.2-42.5)	37.2 (32.9-43.6)	43.7 (40.2-49.9)
Employed families	11,592,082 (95.3)	38.5 (33.4-42.1)	36.7 (32.6-42.7)	43.4 (40.1-49.2)

[a] Population-weighted by block group population (population < 5 years, and race/ethnicity), by number of families (unemployment), by households (household income, linguistic isolation, renters/owners), by population for whom poverty status was determined (poverty), and by population $\geq$ 25 years (< high school education). Weighting methods are described in the methods section of the main text.

[b] Race/ethnicity does not sum to total; 1,619,611 individuals were of mixed or other race/ethnicity.

[c] 722,567 people did not have poverty status determined and thus are not included in the poverty summary.

Table S3. Anthropogenic L_{50} nighttime noise fitted values in urban (n = 175,373) and suburban/rural (n = 38,732) block groups by race/ethnicity and socioeconomic characteristics

Block group characteristic	Median (IQR) anthropogenic nighttime noise, dBA[a]	
	Urban	Suburban/rural
Population <5 years, %		
0	44.6 (42.3-46.7)	38.3 (33.0-42.3)
1	44.3 (42.0-46.2)	37.5 (32.8-41.3)
5	44.2 (41.8-46.1)	37.1 (32.8-40.9)
10	44.8 (42.6-46.8)	38.8 (34.7-42.5)
< High school education, %		
0	44.2 (42.6-45.8)	39.1 (33.0-42.8)
10	44.0 (41.5-45.8)	37.5 (32.9-40.9)
20	44.7 (42.2-46.8)	38.0 (34.0-41.8)
50	46.7 (44.6-48.4)	40.0 (32.3-44.0)
Median household income, $		
$25,000	45.9 (43.8-47.9)	39.4 (35.0-43.2)
$50,000	44.3 (41.6-46.3)	37.2 (33.0-40.9)
$75,000	43.9 (41.4-45.9)	37.3 (33.6-40.8)
$100,000	44.0 (42.2-45.6)	40.9 (36.0-42.5)
Household income below poverty threshold, %		
0	44.0 (42.2-45.5)	38.4 (33.9-41.7)
10	44.2 (41.5-46.1)	36.9 (32.7-40.3)
20	44.9 (42.6-47.0)	37.8 (33.5-41.6)
50	46.9 (44.7-48.5)	41.8 (36.2-44.0)
Unemployed families, %		
0	44.6 (42.5-46.5)	38.2 (33.5-42.2)
10	44.6 (42.2-46.7)	38.1 (34.2-42.2)
20	45.7 (43.7-49.1)	40.8 (35.6-43.5)
50	47.6 (44.3-49.1)	43.6 (42.4-44.2)
Race/ethnicity		
Hispanic, %		
0	43.7 (40.1-45.9)	37.6 (33.5-41.3)
25	45.1 (43.4-46.9)	39.3 (35.1-42.7)
50	45.7 (44.0-47.6)	39.7 (35.8-43.6)
75	46.5 (43.7-48.3)	-
Non-Hispanic		
American Indian, %		
0	44.5 (42.3-46.5)	38.1 (34.0-41.9)
10	45.1 (41.9-47.1)	39.1 (32.3-42.6)
20	43.6 (41.5-45.9)	39.0 (29.6-42.0)
50	-	27.6 (26.7-28.4)
Asian, %		
0	44.0 (41.2-46.3)	37.7 (33.4-41.6)

10	45.6 (44.4-47.2)	39.9 (34.1-43.8)
20	46.0 (44.8-47.1)	42.2 (37.0-45.1)
50	46.2 (45.8-47.0)	-
Black, %		
0	42.3 (40.4-45.5)	37.0 (32.6-40.6)
25	45.1 (43.4-47.0)	40.1 (33.7-43.1)
50	45.9 (44.2-47.3)	40.2 (34.3-43.2)
75	46.3 (44.3-47.8)	38.2 (31.0-43.2)
White, %		
0	47.0 (45.4-48.7)	43.5 (31.6-45.7)
25	46.5 (44.8-48.1)	38.8 (33.5-44.2)
50	46.1 (44.4-47.6)	37.8 (32.4-42.2)
75	44.8 (43.0-46.5)	39.5 (34.4-43.4)
Linguistically isolated households, %		
0	43.7 (41.1-45.8)	37.7 (33.4-41.5)
10	45.4 (43.6-47.1)	39.6 (33.5-43.5)
20	46.4 (44.2-47.9)	40.0 (37.4-44.5)
50	46.6 (44.3-49.1)	-
Renter-occupied homes, %		
0	43.8 (42.1-45.2)	36.7 (31.5-40.2)
25	43.7 (41.2-45.8)	37.8 (33.6-41.8)
50	45.3 (43.5-47.2)	42.3 (38.3-44.5)
75	46.2 (44.4-47.9)	43.8 (41.4-44.9)
Owner-occupied homes, %		
0	46.4 (44.9-48.7)	41.1 (37.6-44.8)
25	46.2 (44.4-47.9)	43.9 (42.4-44.9)
50	45.3 (43.5-47.2)	42.3 (38.3-44.5)
75	43.7 (41.2-45.8)	37.8 (33.6-41.8)

[a]Median and interquartile range of fitted values from spatial error models, which were run separately for each characteristic and adjusted for block group population and population density. Models used a queen neighbor definition and variance-stabilizing weights. Noise levels were not estimated when <100 data points were available (e.g., for block groups with 50% American Indian population in urban areas).

Table S4. Degree of polynomial used to model the independent variable in each spatial error model for L_{50} nighttime (L_{50} N), L_{50} daytime (L_{50} D), and L_{10} daytime (L_{10} D) noise in urban and suburban/rural block groups

Model	Polynomial Degree/*p* value					
	Urban			Suburban/Rural		
	L_{50} N	L_{50} D	L_{10} D	L_{50} N	L_{50} D	L_{10} D
Population <5 years	6***	9***	9***	1**	1***	1***
< High school education	2***	1**	2***	2***	6***	1
Median household income	6***	4***	4***	8***	9***	3***
Income < poverty threshold	2***	1**	1**	3***	3***	9***
Unemployed families	3***	2***	2***	1*	1	1
Race/ethnicity						
Hispanic	1**	5***	6***	1	2***	2***
Non-Hispanic						
American Indian	9***	8***	2***	7***	8***	2***
Asian	4***	9***	7***	2***	2***	2***
Black	8***	7***	8***	2***	4***	4***
White	5***	7***	6***	2***	8***	8***
Linguistically isolated households	8***	10***	10***	3***	5***	5***
Renter-occupied homes	5***	7***	7***	6***	3***	5***

D, daytime noise; N, nighttime noise. Each model also included block group level total population and population density as predictor variables as linear terms. All models used a queen neighbor definition and variance-stabilizing weights.

*P value <0.05; **p value <0.01; ***p value <0.001 from likelihood ratio test comparing full polynomial model to model without predictor or from a t test of the single coefficient in linear models.

Table S5. Degree of polynomial used to model the independent variable in each spatial error model in the racial segregation analysis for L_{50} nighttime (L_{50} N), L_{50} daytime (L_{50} D), and L_{10} daytime (L_{10} D) noise in urban block groups.

|	Polynomial Degree											
	D_m											
Model	<0.4			0.4 to <0.5			0.5 to <0.6			≥0.6		
	L_{50} N	L_{50} D	L_{10} D	L_{50} N	L_{50} D	L_{10} D	L_{50} N	L_{50} D	L_{10} D	L_{50} N	L_{50} D	L_{10} D
Race/ethnicity												
Hispanic	3***	5***	5***	4***	6***	6***	4***	7***	5***	1***	1***	1***
Non-Hispanic												
Asian	1***	8***	8***	4***	6***	6***	3***	7***	7***	4***	4***	4***
Black	4***	7***	7***	8***	7***	8***	8***	7***	7***	5***	7***	8***
White	4***	4***	4***	5***	5***	6***	5***	5***	5***	2***	4***	4***

D, daytime noise; N, nighttime noise. Each model also included block group level total population and population density as predictor variables as linear terms. All models used a queen neighbor definition and variance-stabilizing weights. American Indians were excluded from this analysis due to small numbers in urban areas.

*P value <0.05; **p value <0.01; ***p value <0.001 from likelihood ratio test comparing full polynomial model to model without the predictor or from a t test of the single coefficient in linear models.

Figure S1. Correlation matrix (Spearman's rank) with LOESS lines for independent and
dependent variables in urban block groups (n = 175,373). *P value <0.05; **p value <0.01; ***p
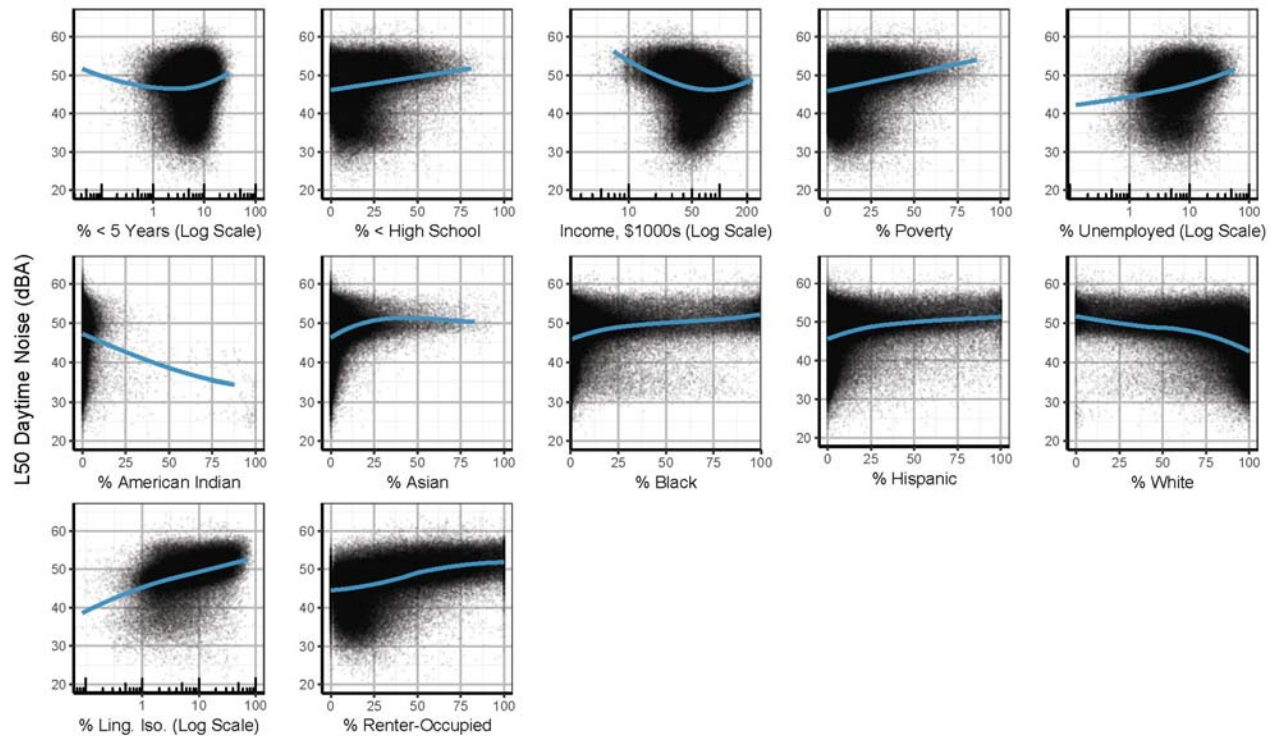value <0.001.
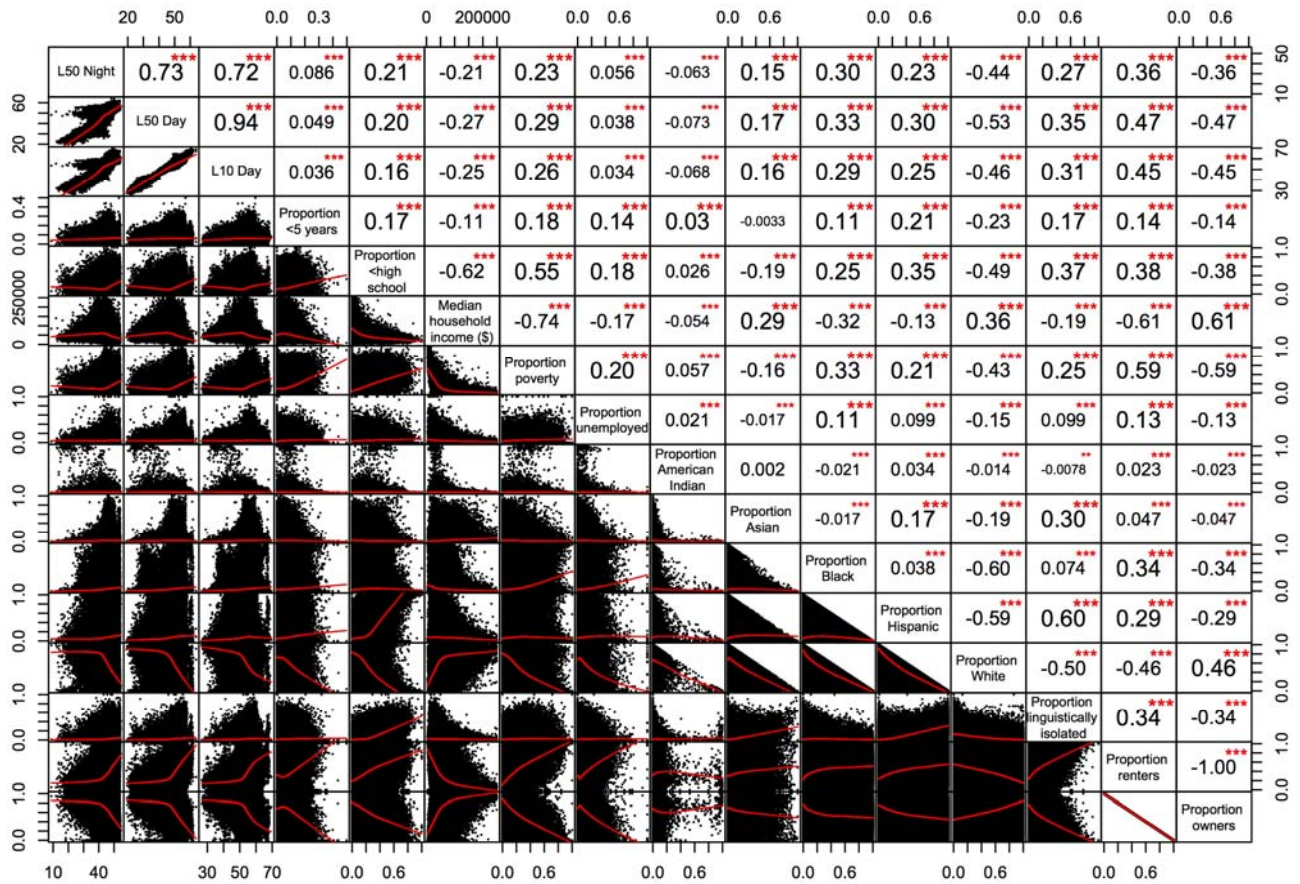
Figure S2. Spatial distribution of (A) anthropogenic L_{50} daytime noise; (B) anthropogenic L_{10} daytime noise at the block group level in the continental United States. Shapefiles used to generate these maps downloaded from the NHGIS site: http://www.nhgis.org.
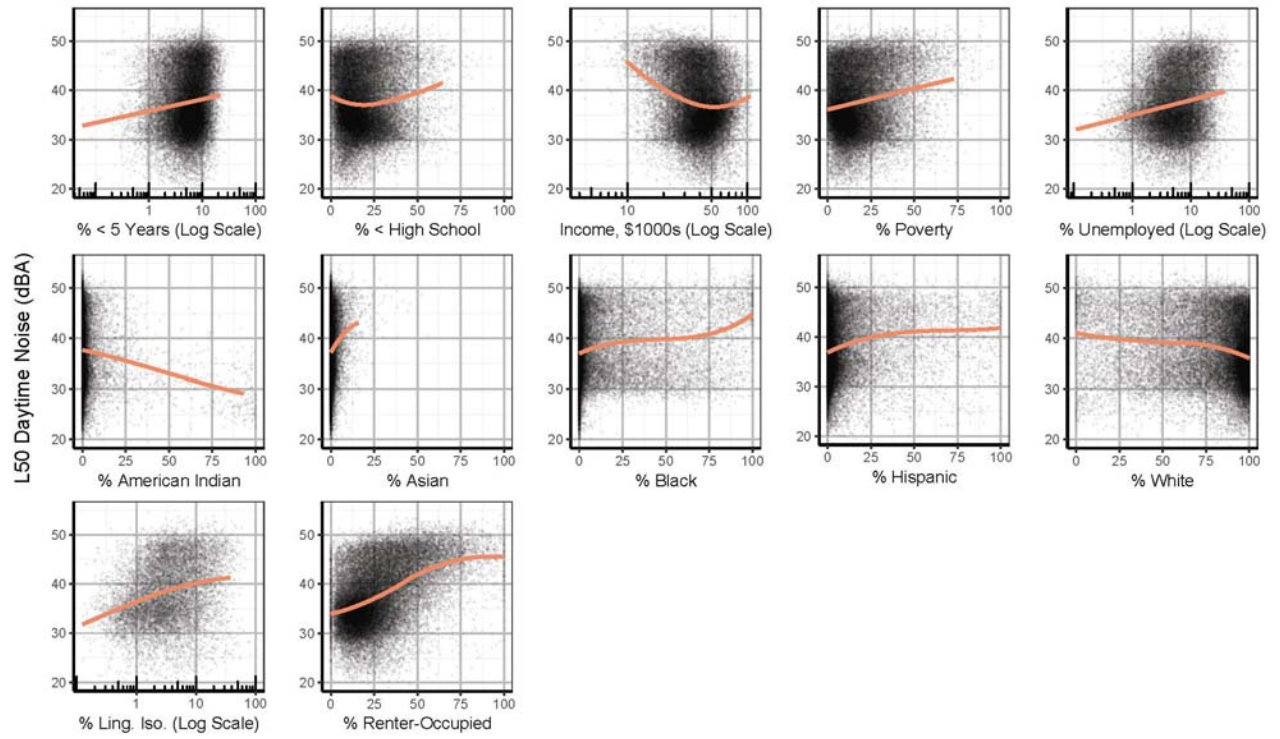
(A)

(B)
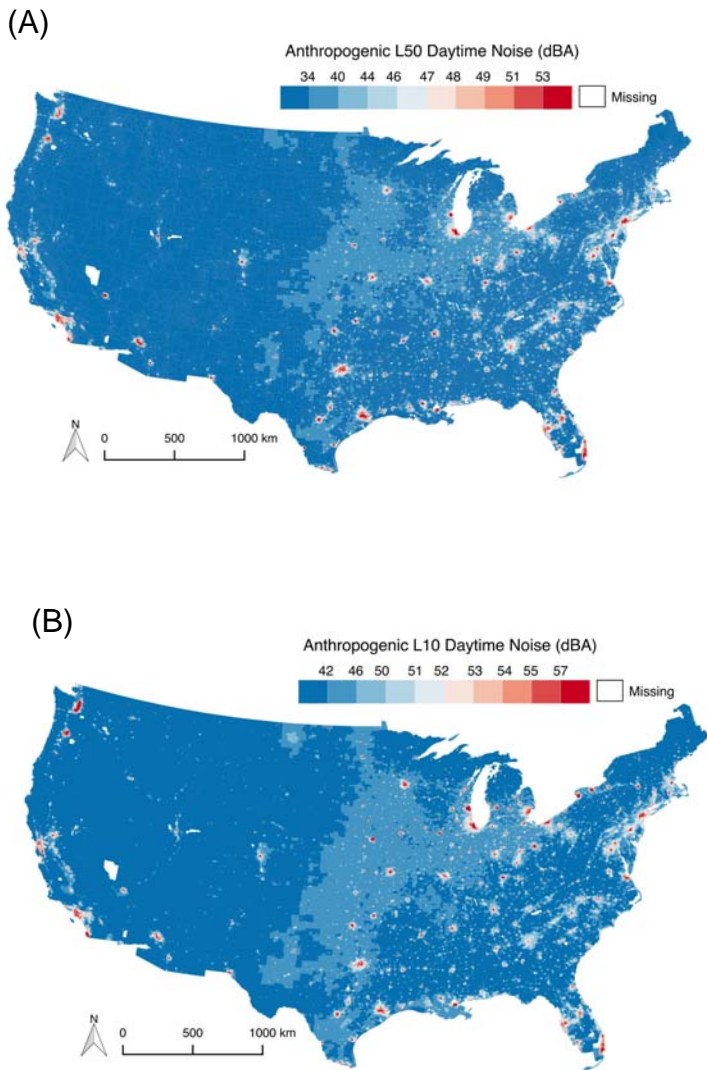
Figure S3. Race/ethnicity and socioeconomic characteristics and anthropogenic L_{50} daytime noise in (A) urban block groups (n = 175,373); and (B) suburban/rural block groups (n = 38,732). The Figure displays the fitted values (points) showing the relationship between noise and each of 12 demographic characteristics adjusted for block group population and population density and using a queen neighbor definition and variance-stabilizing weights. Four of the plots (% < 5 years, median household income ($1000s), % unemployed, and % linguistically isolated) use a log scale x-axis as noted on the figure. The LOESS line was only estimated when there were >100 observations.
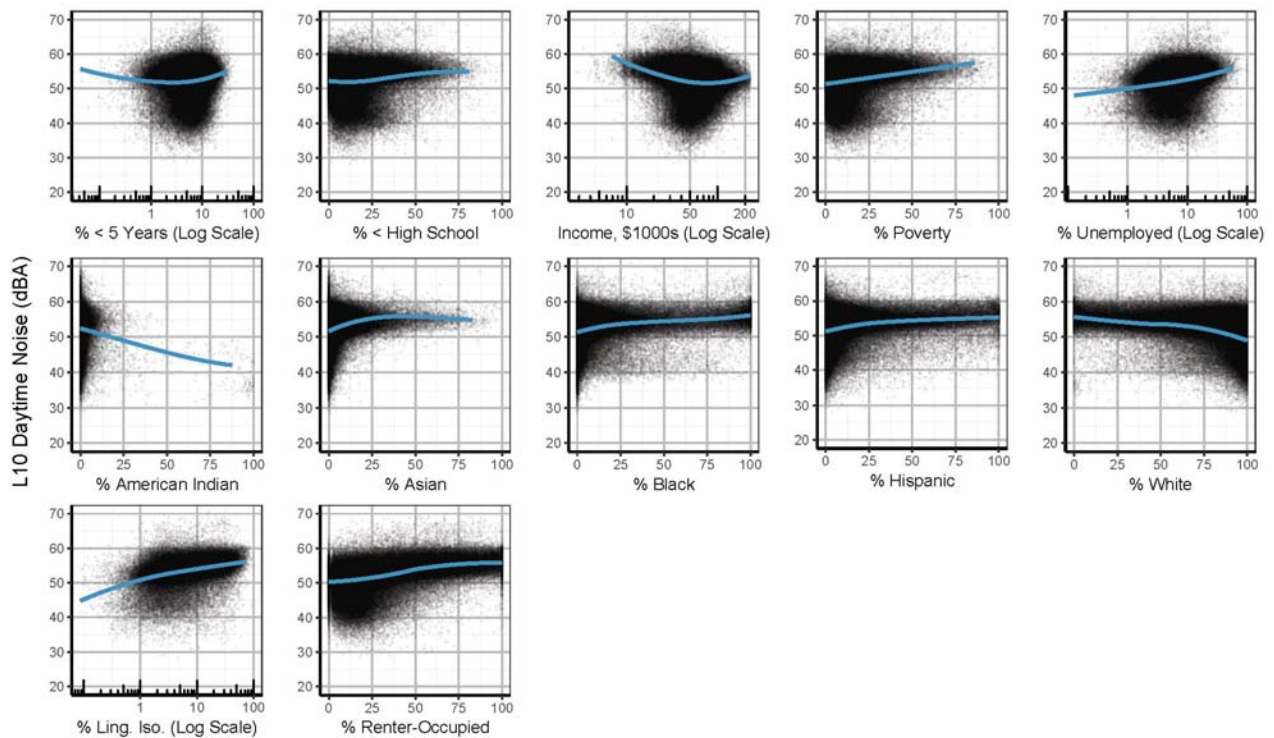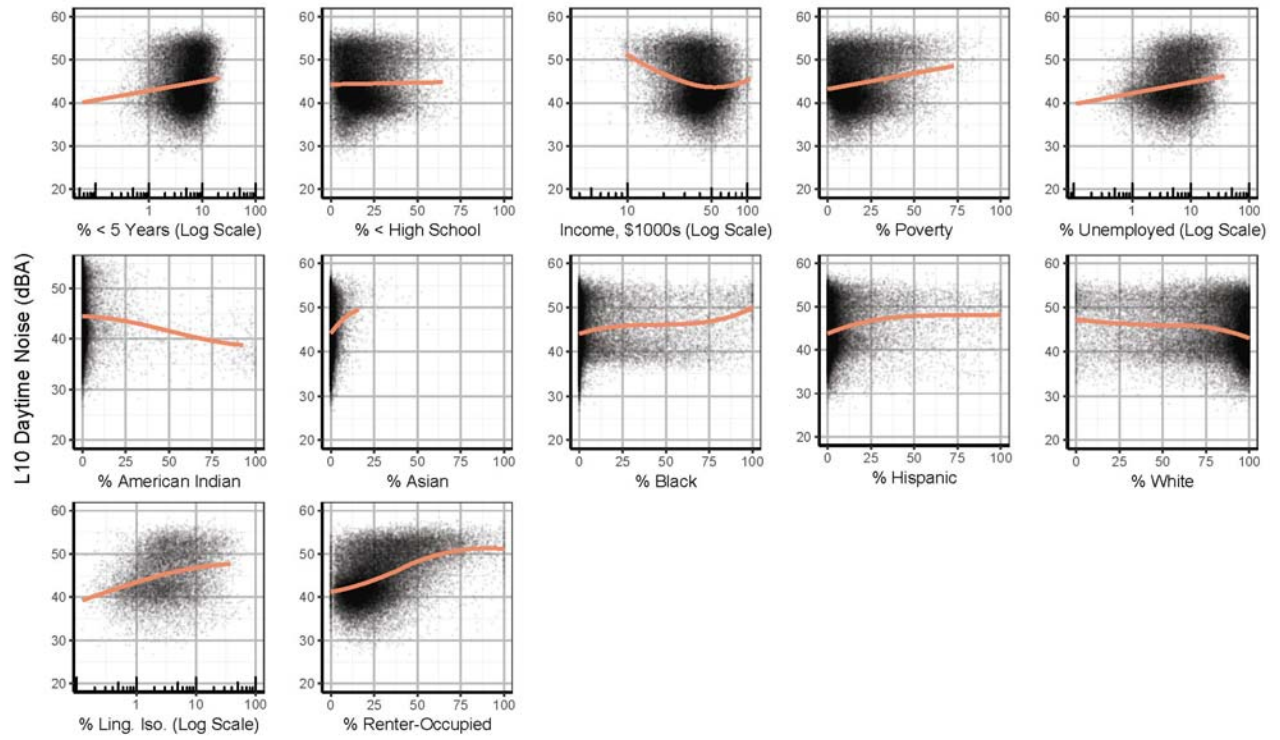
(A) Urban block groups

(B) Suburban/rural block groups

Figure S4. Race/ethnicity and socioeconomic characteristics and anthropogenic L_{10} daytime noise in (A) urban block groups (n = 175,373); and (B) suburban/rural block groups (n = 38,732). The Figure displays the fitted values (points) showing the relationship between noise and each of 12 demographic characteristics adjusted for block group population and population density and using a queen neighbor definition and variance-stabilizing weights. Four of the plots (% < 5 years, median household income ($1000s), % unemployed, and % linguistically isolated) use a log scale x-axis as noted on the figure. The LOESS line was only estimated when there were >100 observations.
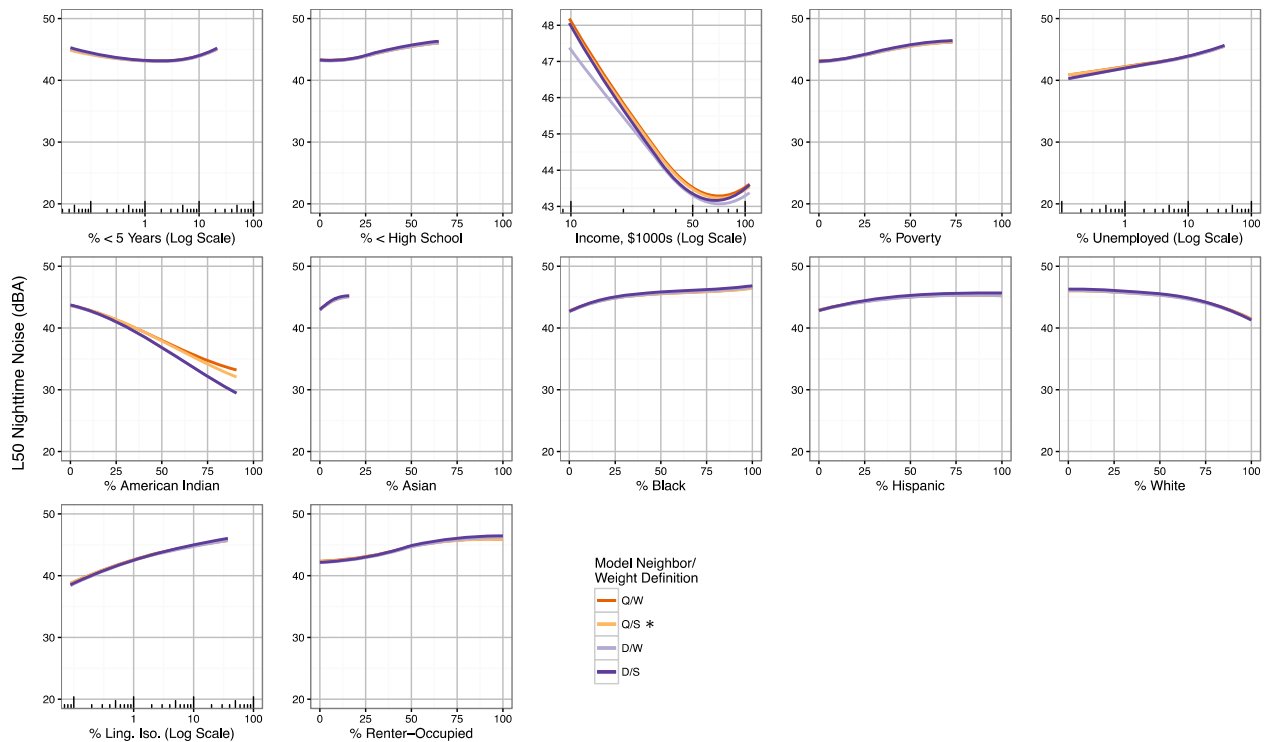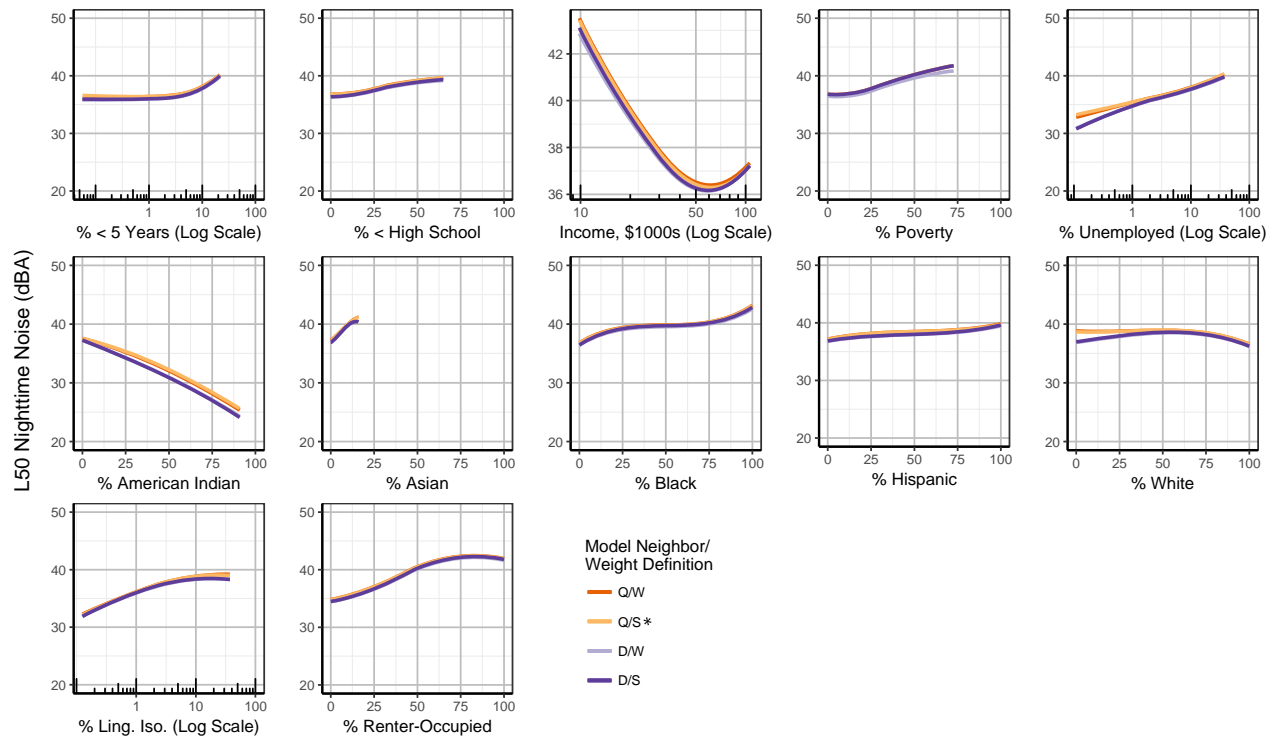
(A) Urban block groups

(B) Suburban/rural block groups

Figure S5. Weight/neighbor definition sensitivity analyses for race/ethnicity and socioeconomic characteristics and anthropogenic L_{50} nighttime noise in (A) urban block groups (n = 175,373); and (B) suburban/rural block groups (n = 38,732). The Figure displays the LOESS line of the fitted values from 48 (i.e., 4 different weight/neighbor definitions and 12 variables) separate spatial error models using 4 neighbor/weight definitions: queen/*W*-coding (orange); queen/*S*-coding (yellow; *we used this neighbor/weight definition in the main analysis); distance/*W*-coding (light purple); and distance/*S*-coding (dark purple). All models were adjusted for block group population and population density. The LOESS line was only estimated when there were >100 observations.
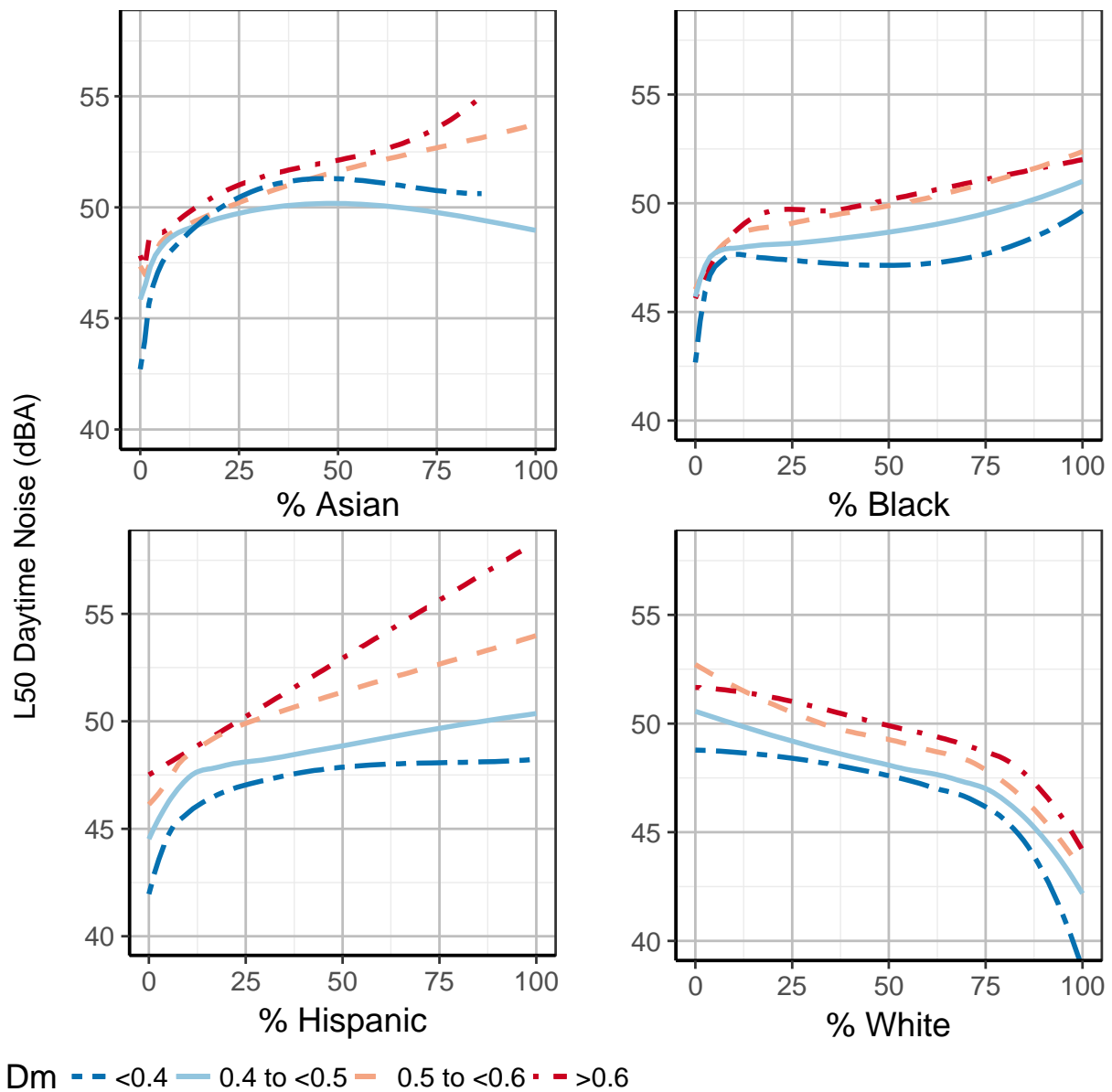
(A) Urban block groups

(B) Suburban/rural block groups

Figure S6. Race/ethnicity and anthropogenic L_{50} daytime noise in urban block groups (n = 175,373), stratified by multigroup racial/ethnic segregation (D_m) for (A) Asians; (B) Blacks; (C) Hispanics; and (D) Whites. American Indians were excluded due to small numbers in urban areas. The Figure displays the LOESS line of the fitted values from 16 (i.e., 4 categories of segregation and 4 race/ethnicities) separate spatial error models adjusted for block group population and population density and using a queen neighbor definition and variance-stabilizing weights. The LOESS line was only estimated when there were >100 observations.

Figure S7. Race/ethnicity and anthropogenic L_{10} daytime noise in urban block groups (n = 175,373), stratified by multigroup racial/ethnic segregation (D_m) for (A) Asians; (B) Blacks; (C) Hispanics; and (D) Whites. American Indians were excluded due to small numbers in urban areas. The Figure displays the LOESS line of the fitted values from 16 (i.e., 4 categories of segregation and 4 race/ethnicities) separate spatial error models adjusted for block group population and population density and using a queen neighbor definition and variance-stabilizing weights. The LOESS line was only estimated when there were >100 observations.
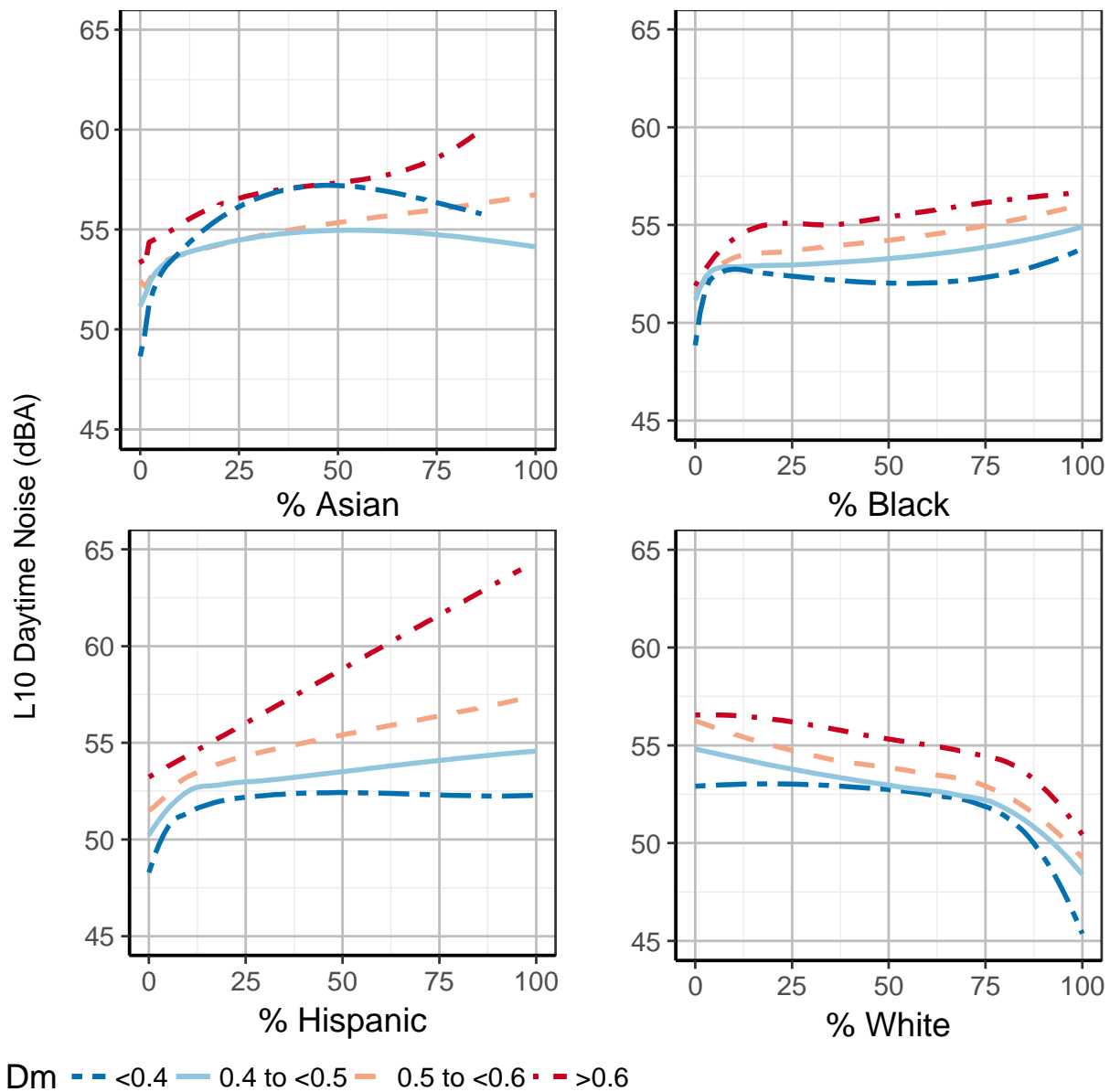
References

Mennitt DJ, Fristrup KM. 2016. Influence factors and spatiotemporal patterns of environmental sound levels in the contiguous United States. Noise Control Eng J 64(3):342-353.