Supporting Information (SI)

Title: Air quality and urban form in US urban areas

Authors: Lara P. Clark[1], Dylan B. Millet[1,2], Julian D. Marshall[1]*

*Corresponding author email: julian@umn.edu.

[1]Department of Civil Engineering, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, MN, 55455.

[2]Department of Soil, Water, and Climate, University of Minnesota, 1991 Upper Buford Circle, St. Paul, MN, 55108.

Number of pages: 32

Number of tables: 17

Number of figures: 9

1. Methods Supporting Information

In addition to the core models for ozone and $PM_{2.5}$ in the main paper, here we present details for the stepwise linear regression models for the remaining criteria pollutants (and alternate measures of ozone and of the long-term air quality index).

1.1. Dependent Variables. As described in **Table S1**, EPA monitor data are for year-1990 for the six criteria pollutants reported in that year. For the two criteria pollutants (PM_{10} and $PM_{2.5}$) not reported in 1990, EPA data are for year-1995 (PM_{10}) and year-2000 ($PM_{2.5}$). For the six criteria pollutants monitored year-round, we calculate annual average concentrations. For the two criteria pollutants (ozone and carbon monoxide) monitored seasonally, we calculate the 5-month seasonal average concentrations (5-month summer average for ozone; 5-month winter average for carbon monoxide). Models for lead and for PM_{10} predict the natural log of population-weighted concentrations because lead and PM_{10} concentration data are log-normally distributed. The remaining models predict population-weighted concentrations. Additionally, we present stepwise linear regression models for 4 alternate population-weighted summer ozone concentration metrics ([1] 24-hour average; [2] 8-hour nighttime average (22:00-06:00); [3] 8-hour maximum average, which is the EPA's regulatory metric; [4] year-2000 8-hour daytime average) and 2 alternate population-weighted long-term air quality indices ([1] LAQI of two priority pollutants: ozone and $PM_{2.5}$; and [2] LAQI of 8 pollutants: carbon monoxide, lead, nitrogen dioxide, ozone, PM_{10}, $PM_{2.5}$, sulfur dioxide, total suspended particulates).

Equation S1 presents the calculation of population-weighted air pollutant concentration (C) for each UA, where c_i is the interpolated concentration (using inverse distance-weighting of the long-term average concentrations from the 3 nearest monitors within 50 km) for each 1-km gridcell center, i, within the UA; p_i is the estimated population in each 1-km gridcell, i, within the UA; and n is the number of 1-km gridcell centers within the UA. **Figure S2** presents boxplots of the population-weighted pollutant concentrations.

(Equation S1)

$$C = \frac{\sum_{i=1}^{n} c_i p_i}{\sum_{i=1}^{n} p_i}$$

1.2. Independent Variables. Figure S3 presents boxplots of the independent variables, including measures of urban form, transportation infrastructure, climate, region, income and land area.

1.2.1. Urban form datasets. **Table S2** summarizes ten published datasets of empirical measures of urban form for US cities.

1.2.2. Temperature metrics. We tested the following temperature metrics in preliminary regressions for time periods matching the air pollution data: Heating Degree Days (HDD; base 18.3°C [65 °F]), Cooling Degree Days (CDD; base 18.3°C [65 °F]), average daily temperature, average maximum daily temperature, average minimum daily temperature. In the final regression models, we employed the temperature metric with highest predictive power for each pollutant: 5-month summer average daily maximum temperature (ozone) and annual HDD (NO_2; SO_2). (CDD, HDD and average daily temperature have similar predictive power for NO_2 and SO_2. Models for NO_2 and SO_2 employing CDD, HDD, or average daily temperature yield consistent results. None of the temperature metrics tested were statistically significant predictors of CO, lead, $PM_{2.5}$, PM_{10}, or TSP.)

1.3. Stepwise Linear Regression Models. This analysis focused on daytime only concentrations of ozone to control for the effect of NO_x titration at night. **Table S3** illustrates the effect of including nighttime ozone concentrations by comparing models for alternate population-weighted ozone concentration metrics: daytime only, nighttime only, and 24-hour concentrations. The two ozone metrics that include night hours (24-hour average and nighttime only average) are negatively associated ($p<0.05$) with annual VKT (i.e., UAs with higher VKT have lower population-weighted ozone concentrations), whereas the metrics that do not include night hours show no association with VKT. This apparently reflects NO_x titration of ozone at night.

For comparison, we generated a year-2000 5-month summer daytime only ozone model (**Table S3**). The year-2000 ozone model results are consistent with the year-1990 ozone model results (positive association with temperature [$p<0.01$]; negative association with population centrality [$p<0.01$] and dilution rate [$p<0.01$]), and with the year-2000 $PM_{2.5}$ results (positive association with population density [$p<0.05$]; negative association with transit supply [$p<0.05$]).

As part of a multicollinearity analysis, **Table S4** presents a correlation matrix of independent variables. The highest correlation between independent variables included in models is for transit supply and population density in the $PM_{2.5}$ model. As shown in **Table S5**, multicollinearity is avoided (variance inflation factor <5) for the $PM_{2.5}$ model including both transit supply and population density, with consistent results (and variance inflation factor <2) for alternate $PM_{2.5}$ models including either transit supply or population density (but not both metrics).

2. Results and Discussion Supporting Information

As discussed for ozone and $PM_{2.5}$ in the main paper, our results for the additional six criteria pollutants (**Figure 4; Tables S6-S16)** support the findings that: (1) urban form is associated with air quality, even after accounting for other common explanatory variables, and (2) the magnitude of impact is significant compared to climatic factors widely considered to be important for air pollution. Although the range of model adjusted R^2 (0.06 to 0.51) suggests limited model predictive power across criteria pollutants, the model *p*-values ($p<0.001$ for all pollutants except lead (ln) [$p<0.05$] and the LAQI of the 8 pollutants [$p<0.05$]) indicate that the models illustrate underlying trends in the datasets with statistical significance. **Figures S5-S8** present regression model residual plots, which illustrate that the model residuals are approximately normally distributed.

Considering all pollutants (**Figure S4; Table S6**) the most robust urban form findings are for population density, road density, and population centrality. For those three metrics, results are statistically significant for two or more pollutants, with all regression coefficients in the same direction. Greater density of people, and of roads, is associated with higher levels of population-weighted air pollutant concentrations, whereas greater population centrality is associated with lower levels. Population density is associated with increased levels of population-weighted carbon monoxide ($p<0.1$), nitrogen dioxide ($p<0.01$), $PM_{2.5}$ ($p<0.01$) and PM_{10} ($p<0.05$), and road density is associated with increased levels of population-weighted PM_{10} ($p<0.05$) and total suspended particulates ($p<0.01$). Greater population centrality (i.e., greater share of population living close to the urban core) is associated with lower levels of population-weighted ozone ($p<0.01$), $PM_{2.5}$ ($p<0.01$), and PM_{10} ($p<0.01$). Additionally, transit supply and city shape (i.e., circularity of urban form) are associated with lower levels of population-weighted air pollutants, but results are statistically significant for only one pollutant. Transit supply is associated with decreased levels of population-weighted $PM_{2.5}$ ($p<0.1$) and city shape is associated with decreased levels of population-weighted carbon monoxide ($p<0.01$).

Climatic factors are statistically significant predictors of air pollution, and in the expected direction. Dilution rate is associated with decreased levels of population-weighted carbon monoxide ($p<0.01$), lead ($p<0.05$), nitrogen dioxide ($p<0.01$), ozone ($p<0.01$), $PM_{2.5}$ ($p<0.01$), PM_{10} ($p<0.01$), and sulfur dioxide ($p<0.05$). Average daily maximum temperature is associated with increased levels of population-weighted ozone ($p<0.01$) (i.e., higher temperatures yield higher daytime ozone concentrations), and annual heating degree days are associated with increased levels of population-weighted nitrogen dioxide ($p<0.05$) and sulfur dioxide ($p<0.05$) (i.e., increased need for heating of buildings [i.e., lower temperature] is associated with increased population-weighted air pollutant concentrations; perhaps a reflection of greater fuel-use for winter heating or more frequent inversions in colder climates).

The predicted magnitude of impact of urban form on air pollution is significant compared to the predicted magnitude of impact of climatic factors. **Figure S4** shows the percent change in population-weighted air pollutant levels associated with increasing the independent variables across the interquartile range (IQR), holding all other variables constant at arithmetic mean value. Increasing individual urban form factors by 1-IQR is associated with 4% to 27% changes in population-weighted air pollutant levels, and increasing individual climate factors by 1-IQR is associated with 7% to 30% changes in population-weighted air pollutant levels.

Here, we provide details for the Web of Knowledge search results presented in the main text. Web of Knowledge identified 24 articles on ambient air pollution and air quality [topic search terms: ("air quality" OR "air pollution") AND ("ambient" OR "outdoor") AND ("urban form" OR "urban design" OR "urban planning" OR "city form" OR "city design" OR "city planning")] compared to 671 articles on ambient air pollution and meteorology [topic search terms: ("air quality" OR "air pollution") AND ("ambient" OR "outdoor") AND ("meteorology" OR "climate")] on June 11, 2011. The 24 articles identified with the urban form search terms are listed below.

1. Bambrick, H.J.; Capon, A.G.; Barniett, G.B.; Beaty, R.M.; Burton, A.J. Climate Change and Health in the Urban Environment: Adaptation Opportunities in Australian Cities. *Asia-Pac. J. Public He.* **2011**, *23,* 67S-79S.

2. Binns, H.J.; Forman, J.A.; Karr, C.J.; Osterhoudt, K.; Paulson, J.A.; Roberts, J.R.; Sandel, M.T.; Seltzer, J.M.; Wright, R.O.; Comm Environm Hlth The Built Environment: Designing Communities to Promote Physical Activity in Children. *Pediatrics* **2009**, *123,* 1591-1598.

3. Campbell-Lendrum, D; Corvalan, C. Climate change and developing-country cities: implications for environmental health and equity. *J. Urban Health* **2007**, *84,* i109-17.

4. Dannenberg, A.L.; Jackson, R.J.; Frumkin, H.; Schieber, R.A.; Pratt, M.; Kochtitzky, C.; Tilson, H.H. The impact of community design and land-use choices on public health: A scientific research agenda. *Am. J. Public Health* **2003**, *93,* 1500-1508.

5. Deguen, S.; Zmirou-Navier, D. Social inequalities resulting from health risks related to ambient air quality-A European review. *Eur. J. Public Health* **2010**, *20,* 27-35.

6. Edwards, R.D.; Schweizer, C.; Llacqu, V.; Lai, H.K.; Jantunen, M.; Bayer-Oglesby, L.; Kunzli, N. Time-activity relationships to VOC personal exposure factors. *Atmos. Environ.* **2006**, *40,* 5685-5700.

7. Junk, J.; Helbig, A.; Krein, A. Screening and scenarios of traffic emissions at Trier, Germany. *Environ. Sci. Pollut. R.* **2004**, *11,* 297-301.

8. Keles, N.; Ilicali, O.C.; Deger, K. Impact of air pollution on prevalence of rhinitis in Istanbul. *Arch. Environ. Health* **1999**, *54,* 48-51.

9. Kousa, A.; Kukkonen, J.; Karppinen, A.; Aarnio, P.; Koskentalo, T. A model for evaluating the population exposure to ambient air pollution in an urban area. *Atmos. Environ.* **2002**, *36,* 2109-2119.

10. Kruger, E.L.; Minella, F.O.; Rasia, F. Impact of urban geometry on outdoor thermal comfort and air quality from field measurements in Curitiba, Brazil. *Build. Environ.* **2011**, *46,* 621-634.

11. Li, H.L.; Huang, G.H.; Zou, Y. An integrated fuzzy-stochastic modeling approach for assessing health-impact risk from air pollution. *Stoch. Env. Res. Risk A.* **2008**, *22,* 789-803.

12. Marshall, J.D.; Brauer, M.; Frank, L.D. Healthy Neighborhoods: Walkability and Air Pollution. *Environ. Health Perspect.* **2009**, *117,* 1752-1759.

13. Marshall, J.D.; Granvold, P.W.; Hoats, A.S.; McKone, T.E.; Deakin, E.; Nazaroff, W.W. Inhalation intake of ambient air pollution in California's South Coast Air Basin. *Atmos. Environ.* **2006**, *40,* 4381-4392.

14. Mayer, H.; Holst, J.; Dostal, P.; Imbery, F.; Schindler, D. Human thermal comfort in summer within an urban street canyon in Central Europe. *Meteorologische Zeitschrift* **2008**, *17,* 241-250.

15. Mi, Y.; Norback, D.; Tao, J.; Mi, Y.; Ferm, M. Current asthma and respiratory symptoms among pupils in Shanghai, China: influence of building ventilation, nitrogen dioxide, ozone, and formaldehyde in classrooms. *Indoor Air* **2006**, *16,* 454-64.

16. Miles, G.H.; Jakeman, A.J.; Bai, J. A Method for Predicting the Frequency-Distribution of Air-Pollution from Vehicle Traffic, Basic Meteorology, and Historical Concentrations to Assist Urban-Planning. *Environ. Int.* **1991**, *17,* 575-580.

17. Pikhart, H.; Prikazsky, V.; Bobak, M.; Kriz, B.; Celko, M.; Danova, J.; Pyrl, K.; Pretel, J. Association between ambient air concentrations of nitrogen dioxide and respiratory symptoms in children in Prague, Czech Republic. Preliminary results from the Czech part of the SAVIAH Study. Small Area Variation in Air Pollution and Health. *Cent. Eur. J. Public Health* **1997**, *5,* 82-5.

18. Schweitzer, L.; Zhou, J.P. Neighborhood Air Quality, Respiratory Health, and Vulnerable Populations in Compact and Sprawled Regions. *J. Am. Plann. Assoc.* **2010**, *76,* 363-371.

19. Smargiassi, A.; Fournier, M.; Griot, C.; Baudouin, Y.; Kosatsky, T. Prediction of the indoor temperatures of an urban area with an in-time regression mapping approach. *J. Expo. Sci. Env. Epid.* **2008**, *18,* 282-288.

20. Stone, B. Urban sprawl and air quality in large US cities. *J. Environ. Manage.* **2008**, *86,* 688-698.

21. Synnefa, A.; Dandou, A.; Santamouris, M.; Tombrou, M.; Soulakellis, N. On the Use of Cool Materials as a Heat Island Mitigation Strategy. *J. Appl. Meteorol. Clim.* **2008**, *47,* 2846-2856.

22. Taha, H. Modeling the impacts of large-scale albedo changes on ozone air quality in the South Coast Air Basin. *Atmos. Environ.* **1997**, *31,* 1667-1676.

23. Vedal, S. Update on the health effects of outdoor air pollution. *Clin. Chest Med.* **2002**, *23,* 763-+.

24. Zou, B.; Wilson, J.G.; Zhan, F.B.; Zeng, Y.N. Spatially differentiated and source-specific population exposure to ambient urban air pollution. *Atmos. Environ.* **2009**, *43,* 3981-3988.

Table S1. Air pollution data inclusion criteria and descriptive statistics

Pollutant	Measure (Typical sampling frequency)	Study period	Monitors included[a] (of total monitors located in 111 UAs evaluated)	Number of UAs with at least 1 monitor (of 111 UAs evaluated)	Median (AM[b]) number of monitors per UA (of 111 UAs evaluated)	Median (AM[b]) number of monitors per UA (of UAs with 1 or more monitor)	Population-weighted concentration AM[b] (ASD[b])	Population-weighted concentration GM[c] (GSD[c])
Carbon Monoxide (Winter average)	24-hour concentration (sampled daily during 5-month winter)	1990 5-month winter (November through March)	193 (of 228)	90	2 (1.7)	2 (2.1)	1.3 (0.43) ppm	1.2 (1.4) ppm
Lead	24-hour concentration (sampled every sixth day)	1990 (annual)	137 (of 207)	52	0 (1.2)	2 (2.6)	0.051 (0.053) µg m^{-3}	0.035 (2.4) µg m^{-3}
Nitrogen Dioxide	24-hour concentration (sampled daily)	1990 (annual)	91 (of 145)	55	0 (0.8)	1 (1.7)	0.021 (0.0056) ppm	0.020 (1.3) ppm
Ozone (Summer daytime average; Core ozone metric)	8-hour daytime (10:00-18:00) concentration (sampled daily during 5-month summer)	1990 5-month summer (May through September)	267 (of 281)	100	2 (2.4)	2 (2.7)	45 (8.5) ppb	44 (1.2) ppb

Ozone (Summer daytime average; year-2000)	8-hour daytime (10:00-18:00) concentration (sampled daily during 5-month summer)	2000 5-month summer (May through September)	294 (of 318)	100	2 (2.6)	2 (2.9)	46 (8.3) ppb	45 (1.2) ppb
Ozone (Summer nighttime average; Alternate ozone metric)	8-hour nighttime (22:00 - 06:00) concentration (sampled daily during 5-month summer)	1990 5-month summer (May through September)	265 (of 281)	100	2 (2.4)	2 (2.7)	17 (4.1) ppb	18 (1.3) ppb
Ozone (Summer average; Alternate ozone metric)	24-hour concentration (sampled daily during 5-month summer)	1990 5-month summer (May through September)	257 (of 272)	99	1 (2.3)	2 (2.6)	30 (4.7) ppb	30 (1.2) ppb
Ozone (Summer average daily 8-hour maximum; Alternate ozone metric)	Daily 8-hour maximum concentration (sampled daily during 5-month summer)	1990 5-month summer (May through September)	270 (of 281)	100	2 (2.4)	2 (2.7)	47 (7.8) ppb	46 (1.2) ppb

Particulate Matter ($PM_{2.5}$)	24-hour concentration (sampled every third day)	2000 (annual)	344 (of 485)	107	2 (3.1)	3 (3.2)	14 (3.2) μg m^{-3}	14 (1.3) μg m^{-3}
Particulate Matter (PM_{10})	24-hour concentration (sampled every sixth day)	1995 (annual)	476 (of 519)	104	3 (4.3)	3 (4.6)	26 (6.5) μg m^{-3}	25 (1.3) μg m^{-3}
Sulfur Dioxide	24-hour concentration (sampled daily)	1990 (annual)	173 (260)	72	1 (1.6)	2 (2.4)	0.0072 (0.0037) ppm	0.0062 (1.9) ppm
Total Suspended Particulates (TSP)	24-hour concentration (sampled every sixth day)	1990 (annual)	183 (of 427)	57	1 (1.6)	3 (3.2)	55 (21) μg m^{-3}	53 (1.3) μg m^{-3}
Long-term Air Quality Index (LAQI)[d] of 8 pollutants	Long-term concentrations for each of the 8 pollutants evaluated divided by their respective long-term NAAQS	Composite (5-month summer 1990, 5-month winter 1990, year-1990, year-1995, year-2000)	498 (of 2,552 criteria pollutant monitors in the 111 UAs evaluated)	12 (UAs with at least 1 monitor for each of the 8 pollutants)	0 (4.5)	36 (42)	3.8 (0.53) [unitless]	3.7 (1.2) [unitless]

[a]Monitor inclusion criteria: (1) observations reported for at least 75% of expected sampling days in study period, where the number of expected sampling days is based on typical sampling frequency over the study period, and (2) not source-oriented (i.e., "ambient") monitor. Of the 2,552 criteria pollutant monitors located in the 111 UAs evaluated, 734 monitors (29%) were eliminated under the inclusion criteria. Of the 734 monitors

eliminated, 456 monitors were eliminated under the minimum observations reported criterion, 363 monitors were eliminated under the ambient criterion, and 85 monitors were eliminated under both inclusion criteria.

[b]Arithmetic Mean (Arithmetic Standard Deviation)

[c]Geometric Mean (Geometric Standard Deviation)

[d]LAQI is the population-weighted sum of the long-term concentrations for each of the 8 pollutants evaluated divided by their respective long-term NAAQS, as shown in **Table S17.**

Table S2. Summary of urban form datasets for United States cities

Publication	Sample[a]	Urban Form Metrics	Primary Data Sources
Bento et al. (2005)[S1]	114 UA	Urban form is measured in a set of 4 metrics: (1) population centrality, (2) road network density, (3) jobs-housing imbalance, and (4) city shape	US Census (1990) Nationwide Personal Transportation Survey (1990), Zip Codes Business Patterns
Burchfield et al. (2006)[S2]	275 MSA	The percentage of undeveloped land in the square kilometer surrounding an average residential development	National Land Cover Data Set (1992), Land Use and Land Cover aerial photographs (1976)
Cutsinger et al. (2005)[S3]	50 EUA	Urban form is measured in a set of 7 metrics: (1) density/continuity, (2) proximity, (3) job distribution, (4) mixed use, (5) housing centrality, (6) nuclearity, (7) housing unit concentration	National Land Use Cover Data Base (1992, 1993), Census Transportation Planning Package (1990)
Eid et al. (2008)[S4]	6,111 neighbor-hoods	Neighborhood-level urban form is measured in 2 metrics: (1) residential sprawl, as defined by Burchfield et al. (2006), (2) land use mix, measured as the count of retail shops and churches	Landsat Data (1992), Zip Code Business Patterns (1994)
Ewing et al. (2003)[S5]	83 MSA	Composite sprawl index, based on linear combination of 22 variables describing residential density, land use mix, degree of centering, and street accessibility	Census Transportation Planning Package, American Housing Survey, Zip Code Business Patterns, National Resources Inventory (1990, 2000)
Fulton et al. (2001)[S6]	281 MSA	Change in population and urbanized land (1982 to 1997), measured as population divided by urbanized land area (as defined by National Resource Inventory)	US Census (1980, 1990), National Resource Inventory (1982, 1997)
Glaeser et al. (2001)[S7]	106 MSA	Job decentralization, measured as (1) share of jobs within 3, 10, and 35 miles of Central Business District (CBD), (2) median employee distance from CBD, (3) job density gradient	Zip Code Business Patterns (1996)
Huang and Sellers (2007)[S8]	12 MSA	Metropolitan urban form is measured in a set of 7 metrics: 6 metrics describing physical characteristics of landscape (derived from satellite images) and 1 metric describing population density	Global Land Cover Facility satellite photographs (1999, 2000, or 2001), United Nations Revision Population Database (2000)

Lopez and Hynes (2003)[S9]	330 MSA	Relative share of metropolitan population living in US Census tracts with low population density (defined as 200 to 3,500 persons per square mile)	US Census (1990, 2000)
Tsai (2005)[S10]	291 MSA	Metropolitan urban form is measured in 4 metrics of population and employment: counts, density, degree of equal distribution, clustering	1995 Census Transportation Planning Package (CTPP)

[a]MSA = Metropolitan Statistical Areas, as defined by US Census; UA = Urban Area, as defined by US Census, EUA = Extended Urban Area, as defined by Cutsinger et al.[S3]

Jaret et al.[S11] provides a recent review of empirical measures of urban sprawl.

Table S3. Analysis of alternate summer ozone metrics. Standardized coefficients[a] for stepwise linear regression models predicting population-weighted ozone concentrations (5-month summer 1990; 5-month summer 2000).

Independent Variable	8-hour daytime average (10:00-18:00) (1990)	8-hour maximum average (EPA regulatory metric) (1990)	24-hour average (1990)	8-hour nighttime average (22:00-06:00) (1990)	8-hour daytime average (10:00-18:00) (2000)
Intercept	0.68	0.49	4.1***	4.0***	0.21
Urban Form					
City Shape	-	-	-	0.29***	-
Jobs-Housing Imbalance	-	-	-	-	-
Population Centrality	-0.29***	-0.34***	-0.29**	-	-0.30***
Population Density	-	0.22*	-	-0.31***	0.23**
Road Density	-	-	-	-	-
Climate					
Dilution Rate	-0.32***	-0.44***	-0.25***	0.18**	-0.30***
Temperature	0.62***	0.70***	0.45**	-	0.60***
Transportation					
Transit Supply	-	-	-	-	-0.12**
VKT	-	-	-0.16**	-0.16**	-
Other Urban					
Income	-	-	-	-	-
Land Area	-	-	-	-	-
Region	0.34**	0.56***	-	-	-
Model adjusted R^2	0.34	0.35	0.15	0.13	0.43
Model p-value	0.0000***	0.0000***	0.0004***	0.0015***	0.0000***
Sample (n)[b]	100	100	99	100	100

[a]Coefficient standardized to interquartile ranges (IQR) of dependent and independent variables. For example, a standardized coefficient of 0.2 would mean that a 1-IQR increase in the independent variable is associated with a 0.2-IQR increase in the dependent variable (population-weighted ozone concentration).

[b]The sample size (number of UAs with at least 1 monitor) differs across ozone metrics because there are 9 more EPA monitors with complete data (at least 75% of expected observations) for the 8-hour averages than for the 24-hour averages.

*Statistical significance: * p<0.1; ** p<0.05; *** p<0.01.

Table S4. Correlation matrix (Pearson's correlation coefficients) for independent variables (n = 111)

	City Shape	Population Centrality	Population Density	Jobs-Housing Imbalance	Road Density	Dilution Rate	Heating Degree Days	Ozone Season Temperature	Transit Supply	VKT	Income	Land Area	Region
City Shape	1.00												
Population Centrality	-0.33	1.00											
Population Density	-0.12	0.09	1.00										
Jobs-Housing Imbalance	0.07	-0.37	0.01	1.00									
Road Density	0.06	-0.17	0.38	0.09	1.00								
Dilution Rate	-0.20	-0.07	0.42	0.03	0.14	1.00							
Heating Degree Days	0.22	0.28	0.09	-0.39	-0.02	-0.23	1.00						
Ozone Season Temperature	-0.06	-0.39	-0.18	0.35	0.14	0.15	**-0.85**[b]	1.00					
Transit Supply	-0.08	0.08	**0.69**[a]	0.14	0.29	0.25	0.10	-0.18	1.00				
VKT	0.09	-0.12	-0.09	0.26	-0.01	-0.10	-0.18	0.16	-0.10	1.00			
Income	0.01	0.27	0.28	0.07	0.00	0.00	0.08	-0.26	0.38	0.00	1.00		
Land Area	-0.06	-0.08	0.27	0.37	0.04	0.06	-0.05	0.03	0.40	0.03	0.34	1.00	
Region	-0.03	0.21	-0.21	-0.38	**-0.57**[c]	0.02	0.15	-0.27	-0.12	-0.03	0.12	-0.09	1.00

[a]Transit supply and population density have moderate correlation (r=0.69) and were both selected in the $PM_{2.5}$ stepwise regression model. **Table S5** presents multicollinearity analysis for transit supply and population density in the $PM_{2.5}$ stepwise regression models. The condition VIF<5 is maintained in reported regression models.

[b]The two temperature variables have moderate correlation (r=-0.85) but were not offered together in stepwise regression models.

[c]Region (binary variable) and road density have moderate correlation (r=-0.57) but were not selected together in regression models.

Table S5. Multicollinearity analysis for $PM_{2.5}$[a] models. Standardized coefficients[b] and variance inflation factors (VIF) with and without transit supply and population density (r=0.69) included as independent variables.

	Standardized coefficient[b] (VIF[c])		
Independent Variable	Model 1: stepwise regression with 11 independent variables (core $PM_{2.5}$ model)	Model 2: stepwise regression without population density (10 independent variables)	Model 3: stepwise regression without transit supply (10 independent variables)
Intercept	2.7***	3.3***	2.9***
Urban Form			
City Shape	-	-	-
Jobs-Housing Imbalance	-	-	-
Population Centrality	-0.31*** (1.1)	-0.27*** (1.1)	-0.31*** (1.1)
Population Density	0.36*** (2.4)	-	0.21** (1.3)
Road Density	-	-	-
Climate			
Dilution Rate	-0.28*** (1.2)	-0.21*** (1.0)	-0.26*** (1.2)
Temperature	-	-	-
Transportation			
Transit Supply	-0.14* (2.0)	-	-
VKT	-	-	-
Other Urban Characteristics			
Income	-	-	-
Land Area	-	-	-
Region	0.67*** (1.2)	0.55*** (1.1)	0.65*** (1.2)
Model adjusted R^2	0.29	0.22	0.25
Model p-value	0.0000***	0.0000***	0.0000***

[a]Models predict population-weighted year-2000 annual average $PM_{2.5}$ concentrations ($\mu g\ m^{-3}$).

[b]Coefficient standardized to interquartile range (IQR) of dependent and independent variables. For example, a standardized coefficient of 0.2 would mean that a 1-IQR increase in the independent variable is associated with a 0.2-IQR increase in the dependent variable (population-weighted year-2000 annual average $PM_{2.5}$ concentration).

[c]Variance inflation factor

*Statistical significance: * p<0.1; ** p<0.05; *** p<0.01

Table S6. Standardized coefficients[a] for stepwise linear regression models predicting population-weighted air pollutant levels

Variable	Carbon Monoxide	Lead [ln][b]	Nitrogen Dioxide	Ozone (Daytime)	$PM_{2.5}$	PM_{10} [ln][b]	Sulfur Dioxide	TSP	Long-term Air Quality Index[c]
Intercept	3.5***	-3.0***	1.3***	0.68	2.7***	11***	0.76***	0.85*	4.1***
Urban Form									
City Shape	-0.25***	-	-	-	-	-	-	-	-
Jobs-Housing Imbalance	-	-	-	-	-	-	-	-	-
Population Centrality	-	-	-	-0.29***	-0.31***	-0.23***	-	-	-
Population Density	0.17*	-	0.56***	-	0.36***	0.24**	-	-	-
Road Density	-	-	-	-	-	0.17**	-	0.95***	-
Climate									
Dilution Rate	-0.29***	-0.19**	-0.21***	-0.32***	-0.27***	-0.22***	-0.18**	-	-
Temperature[d]	-	-	0.25**	0.62***	-	-	0.34**	-	-
Transportation									
Transit Supply	-	-	-	-	-0.14*	-	-	-	-
VKT	-	-	-	-	-	-	-	-	-
Other Urban Characteristics									
Income[e]	-	-	-	-	-	-0.32***	-	-	-
Land Area	-0.14**	-	-	-	-	0.20***	0.15**	-	0.52**
Region	-0.46***	-	-	0.34**	0.67***	-	0.42**	-	-
Model adjusted R^2	0.32	0.06	0.42	0.34	0.27	0.36	0.26	0.51	0.33
Model *p*-value	0.0000***	0.0485**	0.0000***	0.0000***	0.0000***	0.0000***	0.0001***	0.0000***	0.0285**
Sample (*n*)	90	52	55	100	107	104	72	57	12

[a]Coefficient standardized to interquartile range (IQR) of dependent and independent variables. For example, a standardized coefficient of 0.2 would mean that a 1-IQR increase in the independent variable is associated with a 0.2-IQR increase in the dependent variable (population-weighted air pollutant concentration).

[b]Models for lead and PM_{10} predict the natural log of population-weighted concentrations.

[c]Long-term air quality index is a population-weighted aggregate measure of the 8 pollutants as described in **Table S17.**

[d]Temperature metric is 1990 5-month summer average daily maximum temperature (ozone) or 1990 Heating Degree Days (NO_2; SO_2).

[e]Income variable is the square of average income per person.

*Statistical significance: *$p<0.1$; **$p<0.05$; ***$p<0.01$

Table S7. Stepwise linear regression model results for carbon monoxide[a] (n = 90)

Variable	Coefficient	Standardized Coefficient[1]	Δ (ppm) per IQR↑[§]	Δ (%) per IQR↑[§]	p-value	VIF
(Intercept)	2.16	3.52	-	-	0.0000***	-
City Shape (unitless)	-0.609	-0.248	-0.15	-11%	0.0065***	1.13
Dilution Rate (m^2 s^{-1})	-5.84e-4	-0.287	-0.18	-12%	0.0001***	1.25
Land Area (km^2)	-1.11e-4	-0.144	-0.088	-6.4%	0.0185**	1.05
Population Density (persons km^{-2})	2.84e-4	0.167	0.10	8.2%	0.0568*	1.32
Region (binary)	-0.283	-0.462	-0.28	-19%	0.0016***	1.12
Model adjusted R^2 = 0.36	Model p-value = 0.0000***					

[a]Population-weighted 1990 5-month winter (November through March) average carbon monoxide concentration (ppm). Dilution Rate is the 1990 5-month winter harmonic mean.

Table S8. Stepwise linear regression model results for lead [ln][b] (n = 52)

Variable	Coefficient	Standardized Coefficient[1]	Δ (µg m^{-3}) per IQR↑[§]	Δ (%) per IQR↑[§]	p-value	VIF
(Intercept)	-2.95	-2.99			0.0000***	-
Dilution Rate (m^2 s^{-1})	-5.43e-4	-0.187	-0.0067	-17%	0.0485**	-
Model adjusted R^2 = 0.06	Model p-value = 0.0485**					

[b]Natural log of population-weighted 1990 annual average lead concentration (µg m^{-3}). Dilution rate is the year-1990 harmonic mean.

Table S9. Stepwise linear regression model results for nitrogen dioxide[c] (n = 55)

Variable	Coefficient	Standardized Coefficient[1]	Δ (ppm) per IQR↑[§]	Δ (%) per IQR↑[§]	p-value	VIF
(Intercept)	0.0115	1.31	-	-	0.0000***	-
Dilution Rate (m^2 s^{-1})	-7.39e-6	-0.206	-0.0018	-8.0%	0.0002***	1.48
Population Density (persons km^{-2})	1.13e-5	0.562	0.0049	27%	0.0000***	1.36
Temperature (°C days)	1.54e-6	0.251	0.0022	11%	0.0202**	1.11
Model adjusted R^2 = 0.42	Model p-value = 0.0000***					

[c]Population-weighted 1990 annual average nitrogen dioxide concentration (ppm). Dilution Rate is 1990 annual harmonic mean. Temperature is 1990 annual heating degree days (base 18.3°C [65 °F]).

[1]Coefficient standardized to the interquartile range (IQR) of dependent and independent variables. For example, a standardized coefficient of 0.2 would mean that a 1-IQR increase in the independent variable is associated with a 0.2-IQR increase in the dependent variable (population-weighted air pollutant levels).

[§]Predicted change (or predicted percent change) in population-weighted air pollutant levels associated with increasing the independent variable across the interquartile range, holding all other variables constant at arithmetic mean value.

*Statistical significance: * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table S10. Stepwise linear regression model results for daytime ozone [d] (n = 100)

Variable	Coefficient	Standardized Coefficient[1]	Δ (ppm) per IQR↑[§]	Δ (%) per IQR↑[§]	p-value	VIF
(Intercept)	6.79e-3	0.683	-	-	0.5061	-
Dilution Rate ($m^2\ s^{-1}$)	-8.12e-6	-0.316	-0.0031	-6.6%	0.0000***	1.06
Population Centrality (unitless)	-2.37e-3	-0.290	-0.0029	-6.2%	0.0035***	1.19
Region (binary)	3.41e-3	0.343	0.0034	8.0%	0.0482**	1.18
Temperature (°C)	8.98e-4	0.618	0.0061	15%	0.0001***	1.40
Model adjusted R^2 = 0.35	Model p-value = 0.0000***					

[d]Population-weighted 1990 5-month summer (May through September) average daytime ozone concentration (10:00-18:00) (ppm). Temperature is the 1990 5-month summer average daily maximum. Dilution Rate is the 1990 5-month summer harmonic mean.

Table S11. Stepwise linear regression model results for fine particulate matter ($PM_{2.5}$)[e] (n = 107)

Variable	Coefficient	Standardized Coefficient[1]	Δ (μg m^{-3}) per IQR↑[§]	Δ (%) per IQR↑[§]	p-value	VIF
(Intercept)	11.6	2.68	-	-	0.0000***	-
Dilution Rate ($m^2\ s^{-1}$)	-3.73e-3	-0.274	-1.2	-7.9%	0.0000***	1.25
Population Centrality (unitless)	-1.10	-0.308	-1.3	-8.9%	0.0001***	1.10
Population Density (persons km^{-2})	4.08e-3	0.360	1.6	12%	0.0030***	2.37
Region (binary)	2.88	0.669	2.9	24%	0.0000***	1.17
Transit Supply (route-km km^{-2})	-4.36e-5	-0.138	-0.60	-4.1%	0.0580*	1.96
Model adjusted R^2 = 0.27	Model p-value = 0.0000***					

[e]Population-weighted 2000 annual average $PM_{2.5}$ concentration (μg m^{-3}). Dilution Rate is 2000 annual harmonic mean.

[1]Coefficient standardized to the interquartile range (IQR) of dependent and independent variables. For example, a standardized coefficient of 0.2 would mean that a 1-IQR increase in the independent variable is associated with a 0.2-IQR increase in the dependent variable (population-weighted air pollutant levels).

[§]Predicted change (or predicted percent change) in population-weighted air pollutant levels associated with increasing the independent variable across the interquartile range, holding all other variables constant at arithmetic mean value.

*Statistical significance: * $p<0.1$; ** $p<0.05$; *** $p<0.01$

Table S12. Stepwise linear regression model results for coarse particulate matter (PM_{10}) [ln][f] (n = 104)

Variable	Coefficient	Standardized Coefficient[1]	Δ (µg m^{-3}) per IQR↑[§]	Δ (%) per IQR↑[§]	p-value	VIF
(Intercept)	3.25	10.9	-	-	0.0000***	-
Dilution Rate (m^2 s^{-1})	-2.49e-4	-0.221	-1.7	-6.4%	0.0003***	1.31
[Income]^2 ($)	-1.309e-9	-0.321	-2.4	-9.1%	0.0001***	1.27
Land Area (km^2)	9.25e-5	0.201	1.5	6.2%	0.0006***	1.21
Population Centrality (unitless)	-0.0549	-0.233	-1.7	-6.7%	0.0089***	1.16
Population Density (persons km^{-2})	1.97e-4	0.244	1.8	7.5%	0.0174**	1.68
Road Density (%)	0.0339	0.173	1.3	5.3%	0.0191**	1.23
Model adjusted R^2 = 0.36	Model p-value = 0.0000***					

[f]Natural log of population-weighted 1995 annual average PM_{10} concentration (µg m^{-3}). Dilution Rate is the 1995 harmonic mean. Income is mean annual household income squared.

Table S13. Stepwise linear regression model results for sulfur dioxide[g] (n = 72)

Variable	Coefficient	Standardized Coefficient[1]	Δ (ppm) per IQR↑[§]	Δ (%) per IQR↑[§]	p-value	VIF
(Intercept)	4.43e-3	0.759	-	-	0.0060***	-
Dilution Rate (m^2 s^{-1})	-3.53e-6	-0.179	-0.0010	-13%	0.0315**	1.02
Land Area (km^2)	9.66e-7	0.148	0.00086	12%	0.0295**	1.03
Region (binary)	2.44e-3	0.418	0.0024	42%	0.0120**	1.16
Temperature (°C days)	1.06e-6	0.341	0.0020	30%	0.0111**	1.17
Model adjusted R^2 = 0.26	Model p-value = 0.0001***					

[g]Population-weighted 1990 annual average sulfur dioxide concentration (ppm). Dilution Rate is the 1990 annual harmonic mean. Temperature is the 1990 annual heating degree days (base 18.3°C [65 °F]).

[1]Coefficient standardized to the interquartile range (IQR) of dependent and independent variables. For example, a standardized coefficient of 0.2 would mean that a 1-IQR increase in the independent variable is associated with a 0.2-IQR increase in the dependent variable (population-weighted air pollutant levels).

[§]Predicted change (or predicted percent change) in population-weighted air pollutant levels associated with increasing the independent variable across the interquartile range, holding all other variables constant at arithmetic mean value.

*Statistical significance: * p<0.1; ** p<0.05; *** p<0.01

Table S14. Stepwise linear regression model results for total suspended particulates (TSP)[h] (n = 57)

Variable	Coefficient	Standardized Coefficient[1]	Δ (µg m^{-3}) per IQR↑[§]	Δ (%) per IQR↑[§]	*p*-value	VIF
(Intercept)	11.4	0.849	-	-	0.0647*	-
Road Density (%)	10.6	0.946	12.7	27%	0.0000***	-
Model adjusted R^2 = 0.51	Model *p*-value = 0.0000***					

[h]Population-weighted 1990 annual average TSP concentration (µg m^{-3}).

Table S15. Stepwise linear regression model results for long-term air quality index of 8 pollutants[i] (n = 12)

Variable	Coefficient	Standardized Coefficient[1]	Δ (unitless) per IQR↑[§]	Δ (%) per IQR↑[§]	*p*-value	VIF
(Intercept)	3.25	4.13	-	-	0.0000***	-
Land Area (km^2)	3.08e-4	0.515	0.41	12%	0.0285**	-
Model adjusted R^2 = 0.33	Model *p*-value = 0.0285**					

[i]Population-weighted index (unitless) of 8 pollutants (CO, lead, NO_2, ozone, $PM_{2.5}$, PM_{10}, SO_2, TSP) as described in **Table S17**.

Table S16. Stepwise linear regression model results for long-term air quality index of 2 pollutants[j] (n = 97)

Variable	Coefficient	Standardized Coefficient[1]	Δ (unitless) per IQR↑[§]	Δ (%) per IQR↑[§]	*p*-value	VIF
(Intercept)	0.797	2.11	-	-	0.0548*	-
Dilution Rate (m^2 s^{-1})	-4.00e-4	-0.332	-0.13	-7.5%	0.0000***	1.41
Population Centrality (unitless)	-0.0906	-0.295	-0.11	-6.8%	0.0015***	1.23
Population Density (persons km^{-2})	2.13e-4	0.201	0.076	4.9%	0.0327**	1.50
Region (binary)	0.247	0.654	0.25	18%	0.0004***	1.45
Temperature (°C)	1.46e-2	0.262	0.10	6.5%	0.0687*	1.58
Model adjusted R^2 = 0.29	Model *p*-value = 0.0000***					

[j]Population-weighted index (unitless) of 2 pollutants (ozone and $PM_{2.5}$) as described in **Table S17**. Dilution Rate is 2000 annual harmonic mean. Temperature is 1990 5-month summer average daily maximum.

[1]Coefficient standardized to interquartile ranges (IQR) of dependent and independent variables. For example, a standardized coefficient of 0.2 would mean that a 1-IQR increase in the independent variable is associated with a 0.2-IQR increase in the dependent variable (population-weighted air pollutant levels).

[§]Predicted change (or predicted percent change) in population-weighted air pollutant levels associated with increasing the independent variable across the interquartile range, holding all other variables constant at arithmetic mean value.

*Statistical significance: * p<0.1; ** p<0.05; *** p<0.01

Table S17. Calculation of long-term air quality index (LAQI)[a]

Pollutant	Calculation
Carbon Monoxide	Divide 1990 5-month winter (November through March) average concentration by current NAAQS[b] 8-hour standard (9 ppm)
Lead	Divide 1990 annual average concentration by current NAAQS rolling 3-month standard (0.15 $\mu g\ m^{-3}$)
Nitrogen Dioxide	Divide 1990 annual average concentration by current NAAQS annual standard (0.053 ppm)
Ozone	Divide 1990 5-month summer (May through September) average of daily 8-hour maximum concentrations by NAAQS 8-hour maximum standard (0.075 ppm)
Particulate Matter ($PM_{2.5}$)	Divide 2000 annual average by NAAQS annual standard (15 $\mu g\ m^{-3}$)
Particulate Matter (PM_{10})	Divide 1995 annual average concentration by current NAAQS 24-hour standard (150 $\mu g\ m^{-3}$)
Sulfur Dioxide	Divide 1990 annual average concentration by current NAAQS annual standard (0.03 ppm)
Total Suspended Particulates (TSP)	Divide 1990 annual average concentration by previous NAAQS annual standard (75 $\mu g\ m^{-3}$)

[a]Long-term air quality index (LAQI) is the population-weighted sum of long-term concentrations divided by long-term NAAQS (as in **Table S17** above). We calculate 2 alternate long-term air quality indices: (1) index of 2 pollutants: ozone and $PM_{2.5}$, (2) index of 8 pollutants: CO, lead, NO_2, ozone, $PM_{2.5}$, PM_{10}, SO_2, and TSP.

[b]National Ambient Air Quality Standards (United States Environmental Protection Agency)

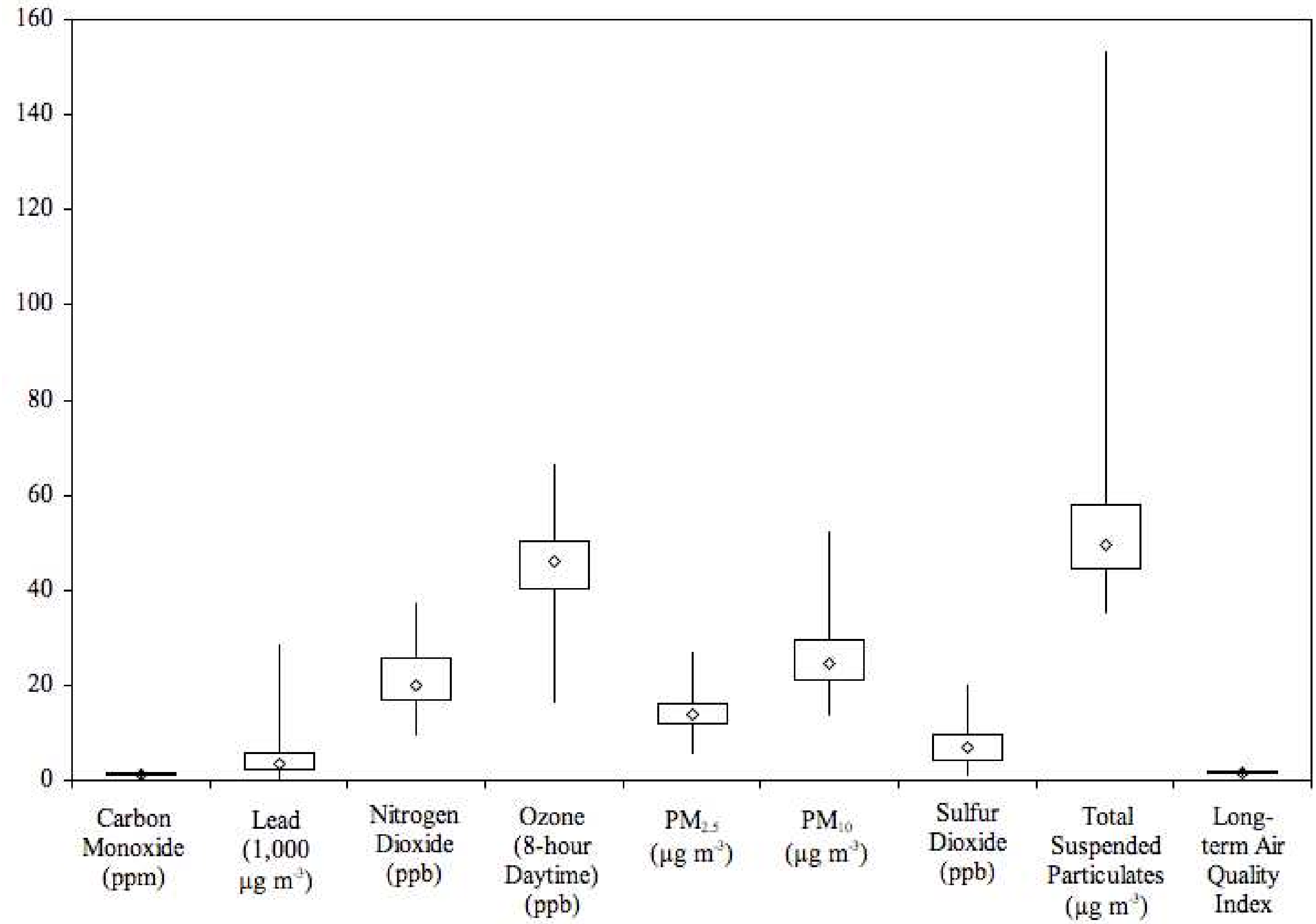

Figure S1. Dependent variable boxplots (minimum, 25th percentile, median, 75th percentile, maximum).

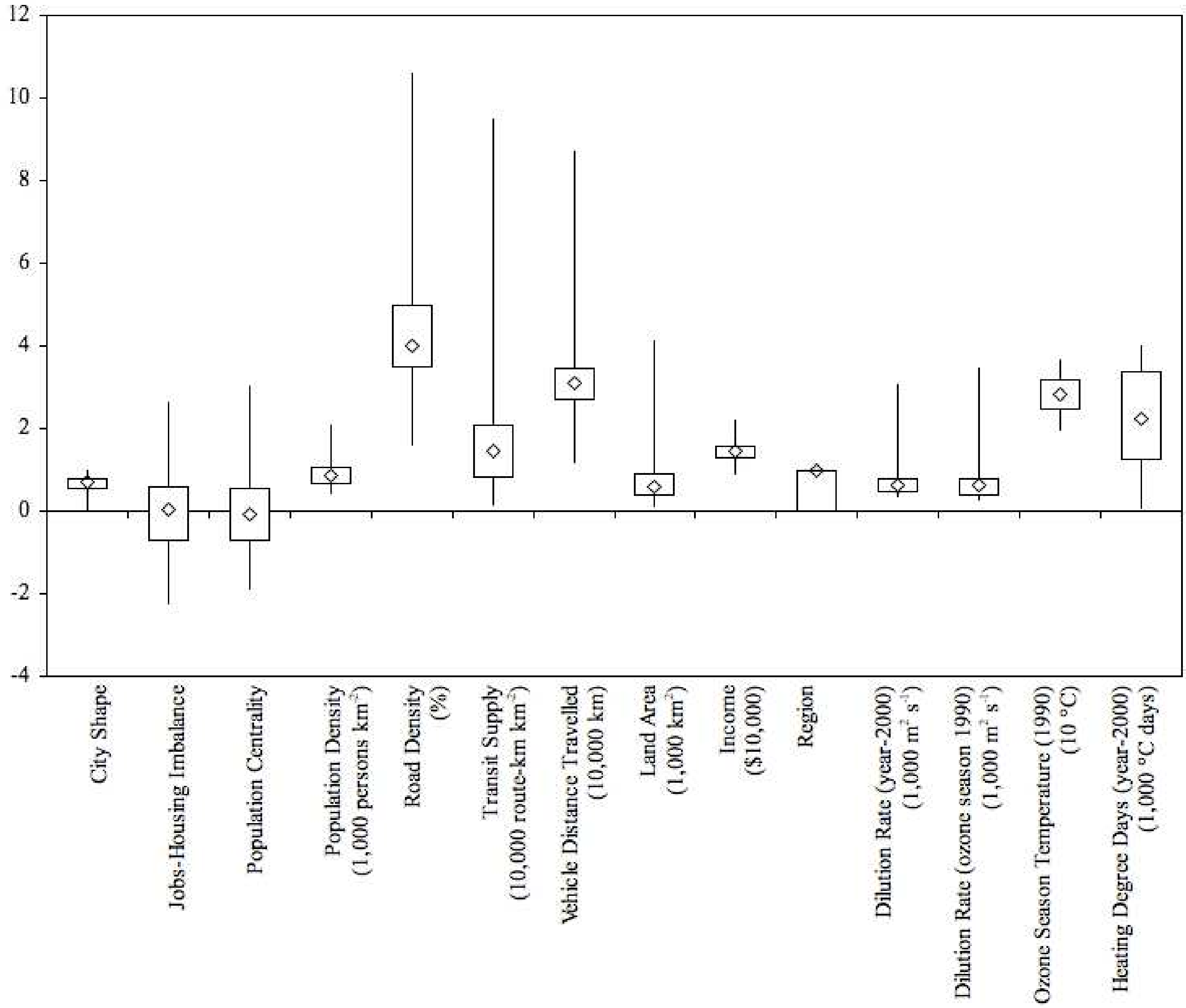

Figure S2. Independent variable boxplots (minimum, 25th percentile, median, 75th percentile, maximum).

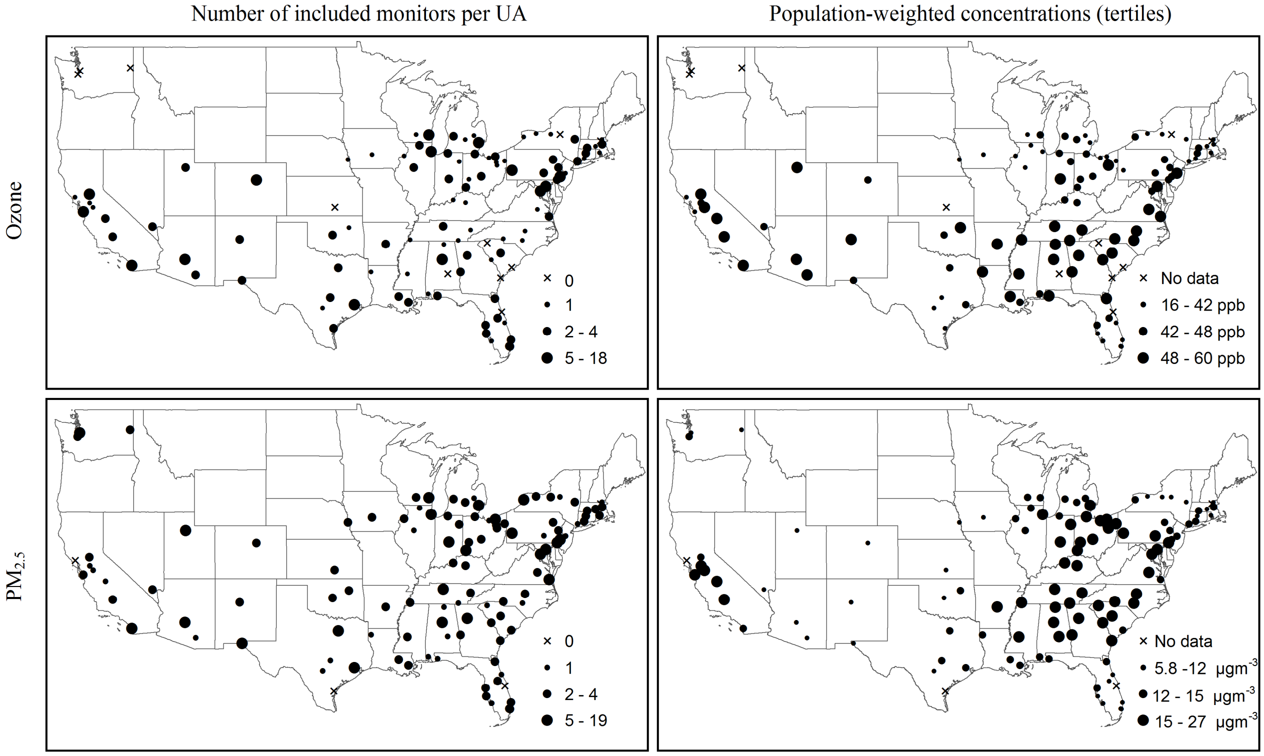

Figure S3. Number of included monitors and population-weighted concentrations (tertiles) for ozone (*n* = 100) and $PM_{2.5}$ (*n* = 107) for each of the 111 UAs evaluated.

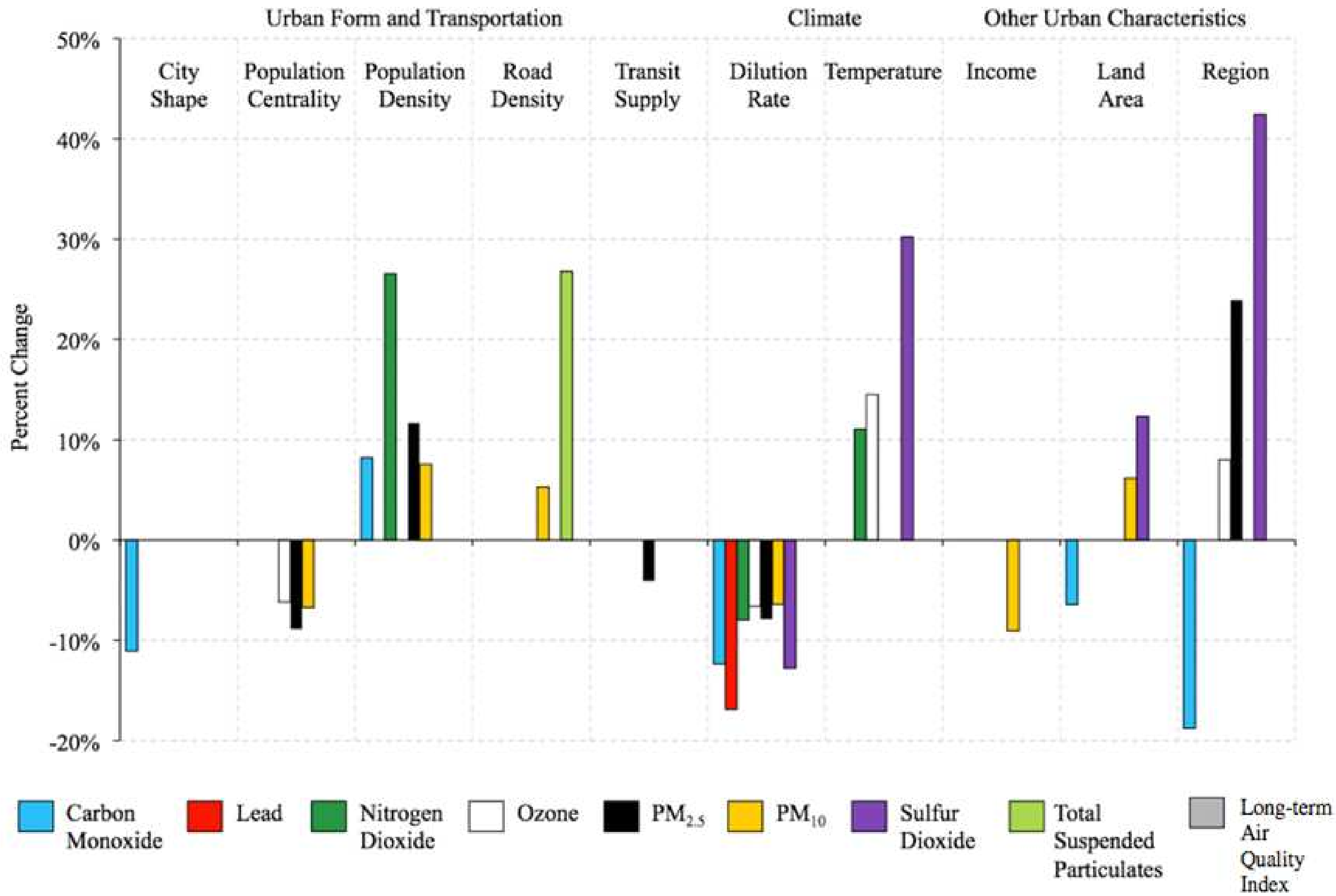

Figure S4. Percent change in population-weighted air pollution concentrations associated with increasing the independent variable across the interquartile range, holding all other variables constant at arithmetic mean value. Here, the long-term air quality index is a population-weighted aggregate measure of the 8 pollutants (see **Table S17**). Temperature is 5-month summer average maximum daily temperature (ozone) or annual heating degree days (nitrogen dioxide; sulfur dioxide). For region (binary variable), reported percent change in air pollution concentration is for a change from western to eastern region.

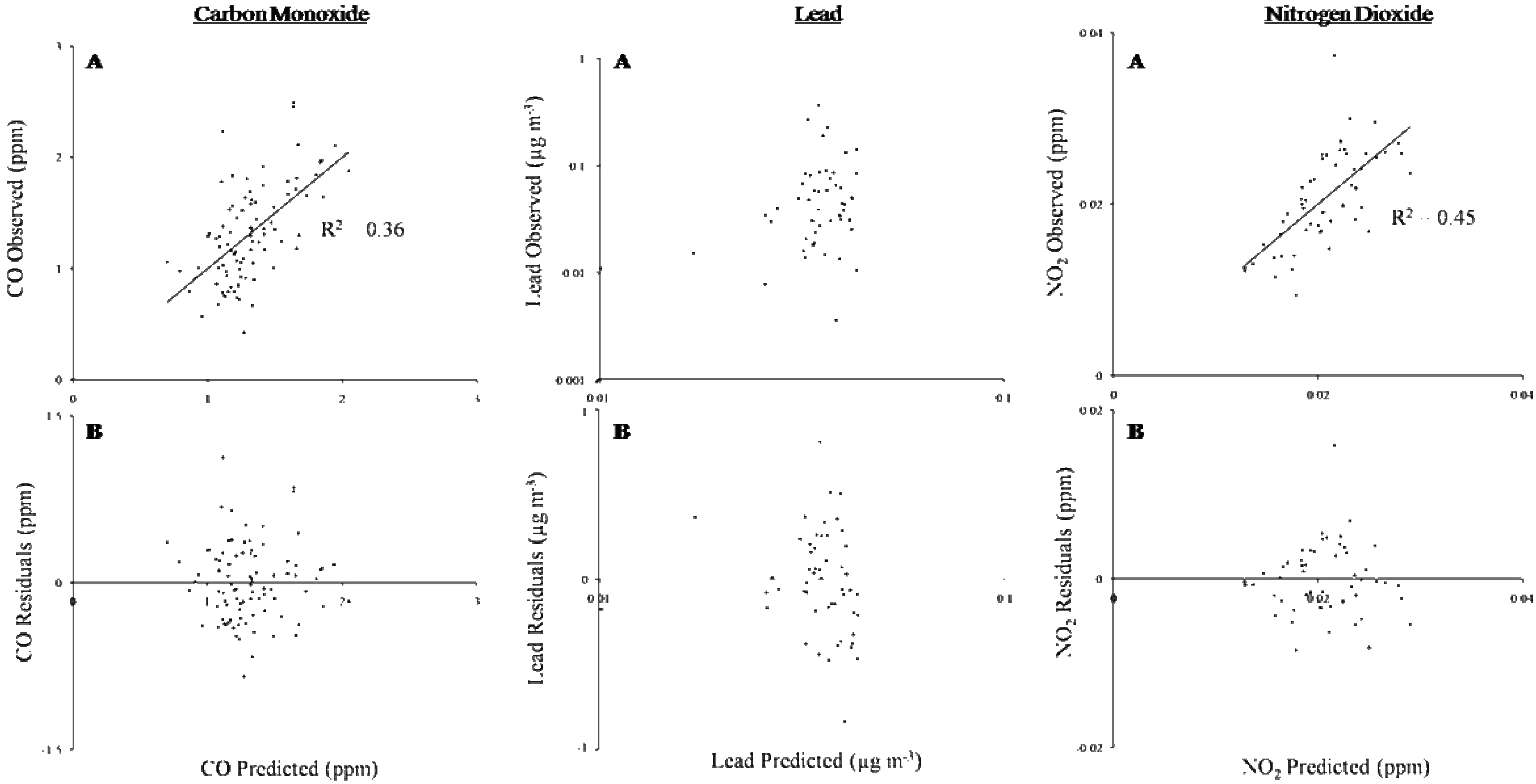

Figure S5. (**A**) Observed population-weighted concentrations versus predicted population-weighted concentrations, and (**B**) model residuals versus predicted population-weighted concentrations, for carbon monoxide, lead [ln], and nitrogen dioxide stepwise linear regression models.

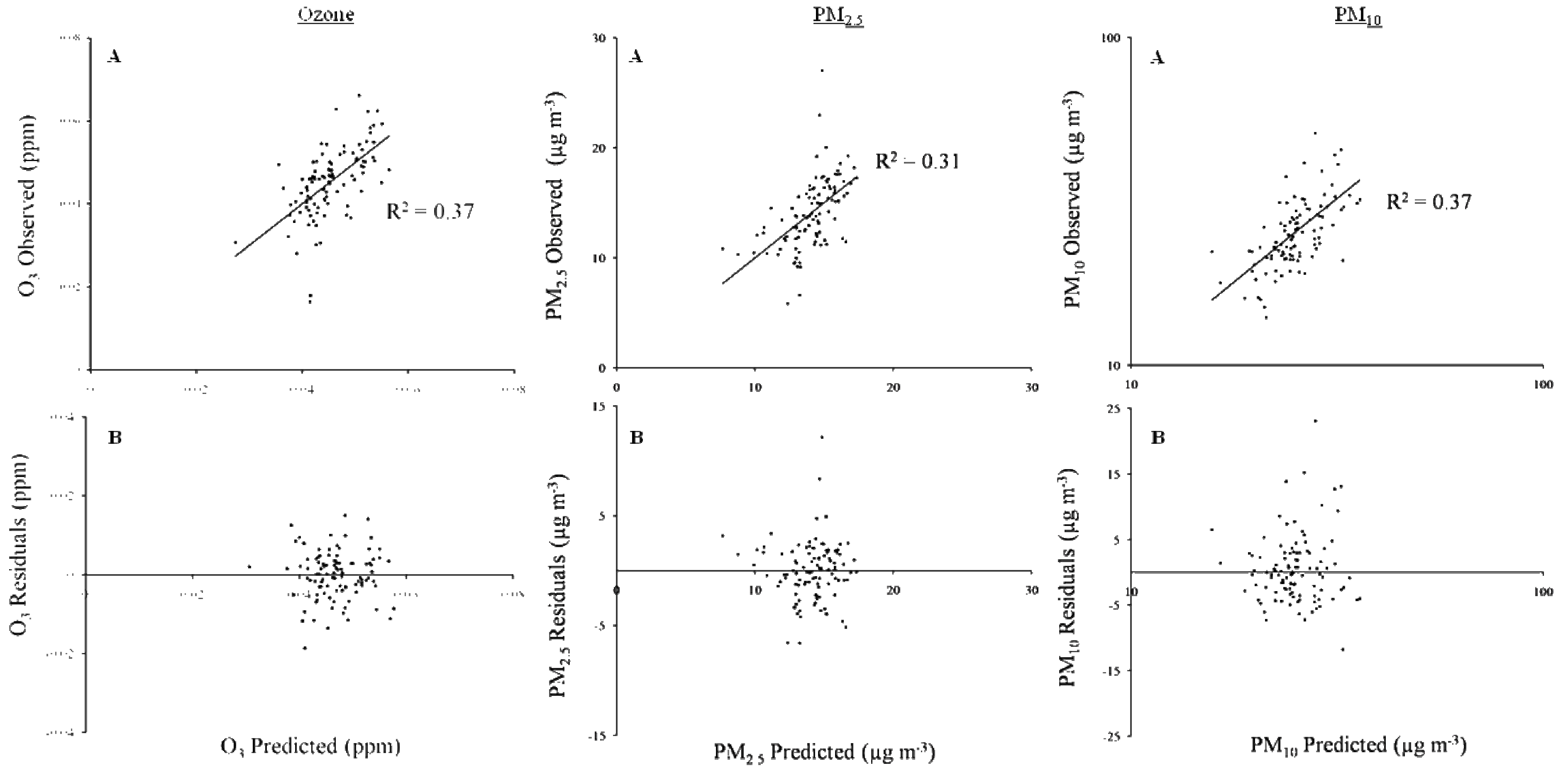

Figure S6. (A) Observed population-weighted concentrations versus predicted population-weighted concentrations, and (**B)** model residuals versus predicted population-weighted concentrations, for ozone (daytime), $PM_{2.5}$, and PM_{10} [ln] stepwise linear regression models.

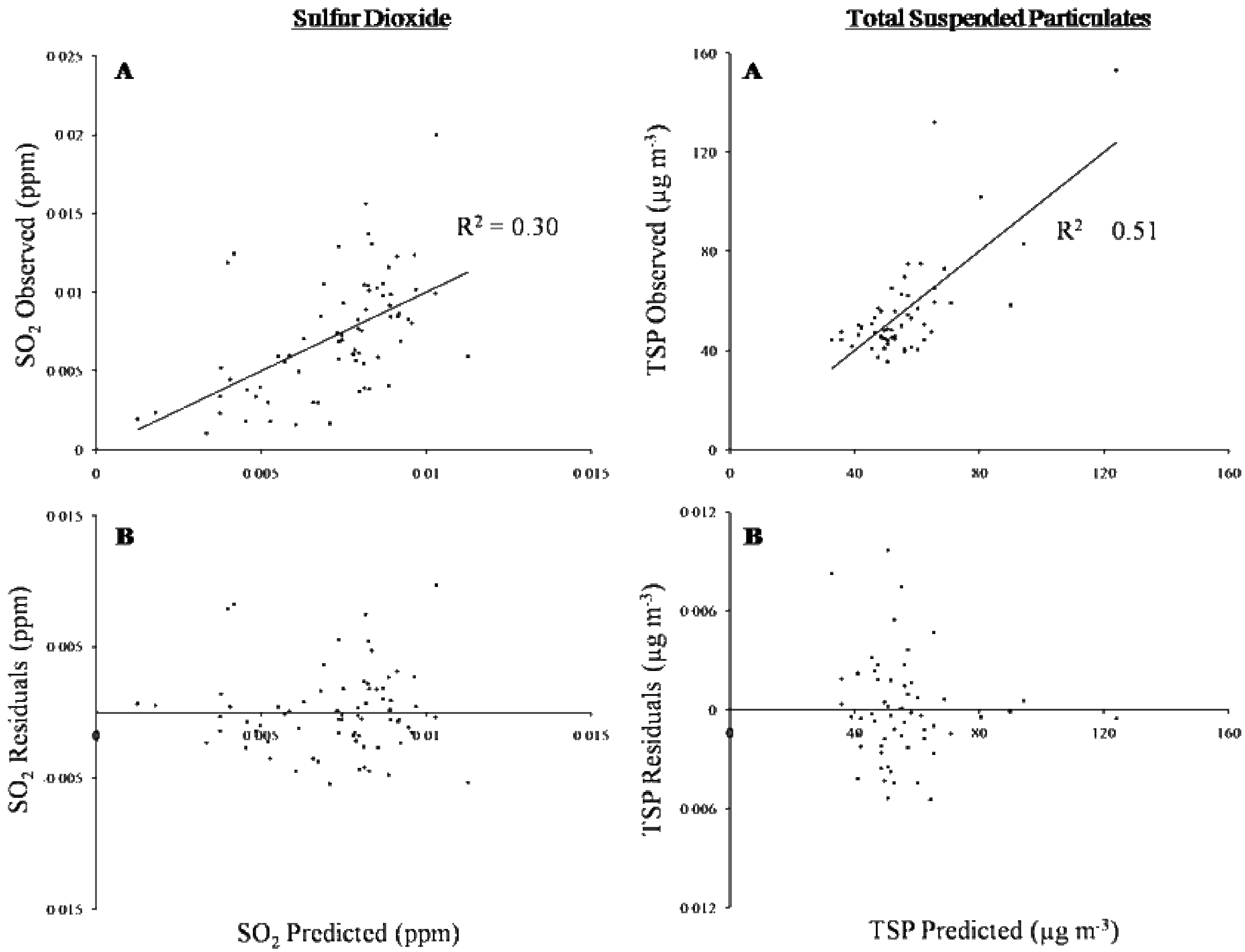

Figure S7. (A) Observed population-weighted concentrations versus predicted population-weighted concentrations, and (**B)** model residuals versus predicted population-weighted concentrations, for sulfur dioxide and total suspended particulates stepwise linear regression models.

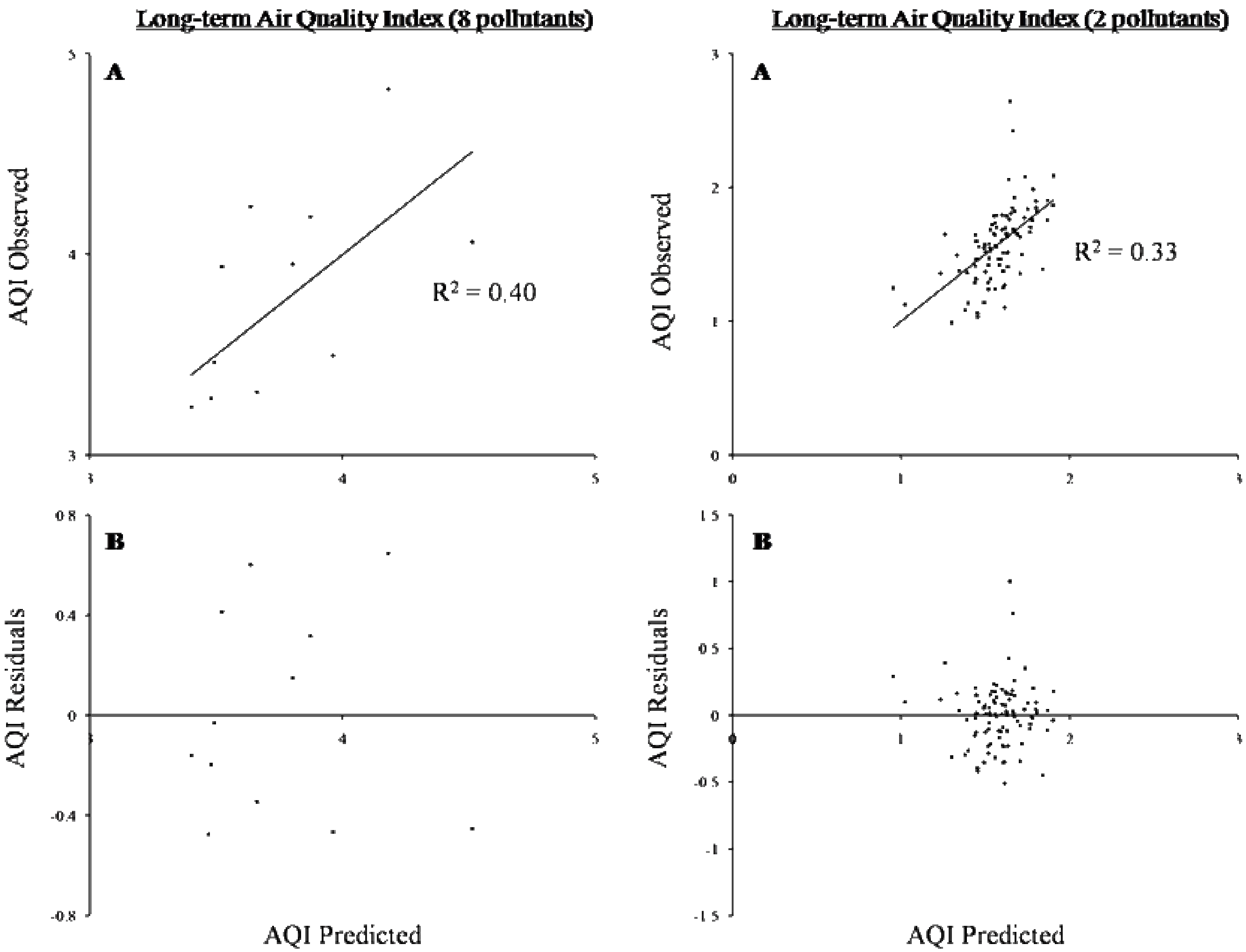

Figure S8. **(A)** Observed population-weighted concentrations versus predicted population-weighted concentrations, and (**B)** model residuals versus predicted population-weighted concentrations, for stepwise linear regression models of long-term air quality index of [1] 8 pollutants: CO, lead, NO_2, ozone, $PM_{2.5}$, PM_{10}, SO_2, TSP; and [2] 2 pollutants: ozone and $PM_{2.5}$.

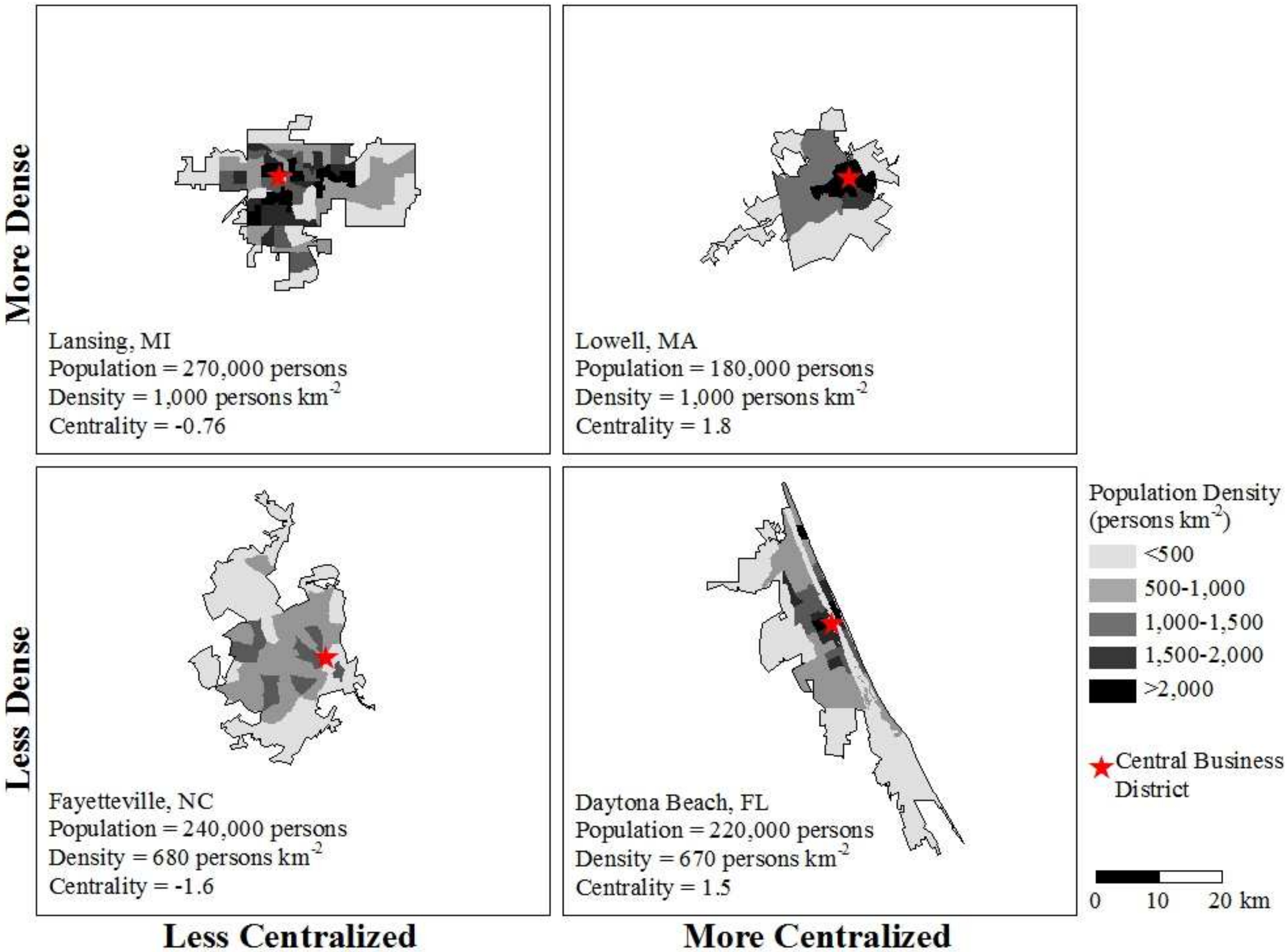

Figure S9. Four urban areas (each with population 180,000 to 270,000 persons in year-1990) illustrating high- and low- population density and centrality. Maps show the 1990 Census Tract population density. Higher values of centrality (Bento et al.[S1]) indicate that a greater fraction of population lives near the Central Business District.

3. References for Supporting Information

S1 Bento, A. M.; Cropper, M. L.; Mobarak, A. M.; Vinha, K. The effects of urban spatial structure on travel demand in the United States. *Rev. Econ. Stat.* **2005**, *87* (3), 466-478.

S2 Burchfield, M.; Overman, H. G.; Puga, D.; Turner, M. A. Causes of sprawl: a portrait from space. *Q. J. Econ.* **2006**, *121* (2), 587-633.

S3 Cutsinger, J.; Galster, G.; Wolman, H.; Hanson, R.; Towns, D. Verifying the multi-dimensional nature of metropolitan land use: advancing the understanding and measurement of sprawl. *J. Urban Aff.* **2005**, *27* (3), 235-259.

S4 Eid, J.; Overman, H. G.; Puga, D.; Turner, M. A. Fat city: questioning the relationship between urban sprawl and obesity. *J. Urban Econ*. **2008**, *63* (2), 385-404.

S5 Ewing, R.; Pendall, R.; Chen, D. Measuring sprawl and its transportation impacts. *Transport. Res. Rec.* **2003**, *1831*, 175-183.

S6 Fulton, W.; Pendall, R.; Nguyen, M.; Harrison, A. Who sprawls most? How growth patterns differ across the US. *Brookings Institute.* **2001.**

S7 Glaeser, E. L.; Kahn, M. E.; Arnott, R.; Mayer, C. Decentralized employment and the transformation of the American city. *Brookings-Wharton Urban Aff*. **2001**,1-63.

S8 Huang, J.; Lu, X. X.; Sellers, J. M. A global comparative analysis of urban form: applying spatial metrics and remote sensing. *Landscape Urban Plan.* **2007**, *82* (4), 184-197.

S9 Lopez, R.; Hynes, H. P. Sprawl in the 1990s: measurement, distribution, and trends. *Urban Aff. Rev.* **2003**, *38* (3), 325-355.

S10 Tsai, Y. H. Quantifying urban form: compactness versus 'sprawl'. *Urban Stud*. **2005**, *42* (1), 141-161.

S11 Jaret, C.; Ghadge, R.; Reid, L.W.; Adelman, R.M. The measurement of suburban sprawl: an evaluation. *City Community.* **2009**, *8* (1), 65-84.