

**Midterm Examination**  
**Katherine Wolf**  
PH250C Spring 2020  
March 31, 2020

1. **Background.** Both exposure to fine particulate matter  $\leq 2.5$  microns in diameter ( $PM_{2.5}$ ) and racial residential segregation, the separation of two or more groups into different neighborhoods by race, of African Americans in the United States (US) have been associated with common negative health outcomes including cardiovascular disease, respiratory disease, lung cancer, low birthweight, preterm birth, and death. Evidence is mounting that the toxicity of  $PM_{2.5}$  varies according to its chemical composition. Moreover, component concentrations can be used to identify  $PM_{2.5}$  sources.
2. **Specific aim.** We will estimate the association between racial residential segregation and  $PM_{2.5}$  component concentrations in census tracts in urban areas of the United States.

3. **Approach.**

- a. **Study population.** The population for this study will include the 886 census tracts both (1) classified as urban by Department of Agriculture Rural-Urban Commuting Area codes and (2) that contain Environmental Protection Agency (EPA) Chemical Speciation Network (CSN) monitors with at least 3 years and 180 days of observations from 2005 to 2015.
- b. **Outcome.** The outcomes for this study will be the 2005-2015 ten-year mean concentration in micrograms ( $\mu g$ ) per cubic meter ( $m^3$ ) of  $PM_{2.5}$  components<sup>1</sup> generated by calculating the mean of monthly means of daily means measured by the Environmental Protection Agency Chemical Speciation Network of  $PM_{2.5}$  component monitors within a census tract.

**Exposure.** The exposure for this study will be the Black racial isolation in a census tract as measured by the spatial isolation index developed by Anthopolos et al.<sup>2</sup>. The index ranges from 0 (no spatial isolation, in which a Black census tract resident only encounters non-Black neighbors) to 1 (complete spatial isolation, in which a Black census tract resident only encounters Black neighbors).

**Confounders.** Models will consider covariate adjustment for

- Total census tract population (persons)
- Population density (persons per square kilometer)
- Sex (percent female)
- Race/ethnicity, including percent white (not Hispanic or Latinx), Percent Hispanic or Latinx (any race), percent African American (not Hispanic or Latinx), percent Asian (not Hispanic or Latinx), and percent other
- Median home value of owner-occupied housing units (dollars)
- Median household income (dollars)
- Age, including percent 0-19 years old and percent 65 years old and older
- Housing tenure (percent of homes renter-occupied)
- Linguistic isolation (percent of households in which no one 14 and over either (1) speaks English only or (2) speaks a language other than English at home and speaks English "very well")
- Education (percent of adults  $\geq 25$  years of age with less than a high school education)
- Poverty (percent of the population in poverty)
- Unemployment (percent of households with one family member unemployed)
- Region (Midwest, Northeast, South, West)

- c. **Target parameter.** We seek to estimate the change in the concentration (in  $\mu g/m^3$ ) of each  $PM_{2.5}$  component per 0.2-unit change in African American spatial isolation, holding the values of the other covariates listed above constant. We will use a restricted polynomial spline model to avoid the assumption of a linear dose-response relationship.

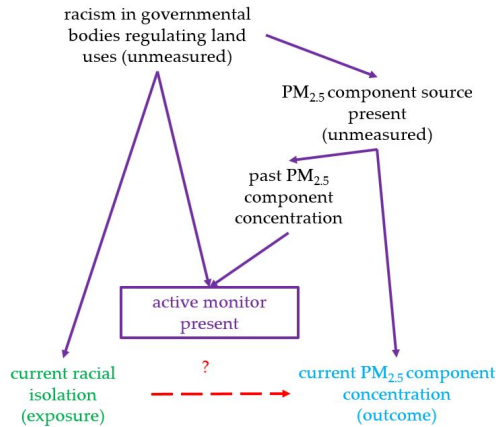
---

<sup>1</sup>Components include (where  $n$  indicates the number of urban census tracts with at least 180 daily observations of that component over at least 3 years) total  $PM_{2.5}$  ( $n = 886$ ), aluminum ( $n = 276$ ), arsenic (As) ( $n = 276$ ), bromine (Br) ( $n = 274$ ), cadmium ( $n = 276$ ), calcium ( $n = 276$ ), chlorine ( $n = 276$ ), copper (Cu) ( $n = 275$ ), iron (Fe) ( $n = 276$ ), lead (Pb) ( $n = 276$ ), mercury ( $n = 162$ ), nickel (Ni) ( $n = 276$ ), silicon ( $n = 276$ ), sodium ( $n = 264$ ), titanium ( $n = 276$ ), vanadium (V) ( $n = 276$ ), zinc (Zn) ( $n = 276$ ), ammonium ion ( $NH_4^+$ ) ( $n = 213$ ), sodium ion ( $n = 213$ ), nitrate ion ( $NO_3^-$ ) ( $n = 267$ ), sulfate ion ( $SO_4^{2-}$ ) ( $n = 277$ ), and elemental carbon (EC) ( $n = 201$ ).

<sup>2</sup>Anthopolos R, James SA, Gelfand AE, Miranda ML. A spatial measure of neighborhood level racial isolation applied to low birthweight, preterm birth, and birthweight in North Carolina. *Spat Spatiotemporal Epidemiol* 2011;2(4):235–46.

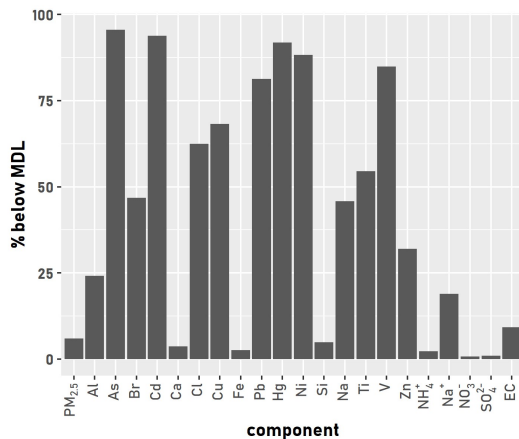
**4. Methodological issue.** One important limitation of the proposed analysis is selection bias induced by the limitation of the study to census tracts containing EPA CSN PM<sub>2.5</sub> component monitors. For example, racism in governmental bodies regulating land uses might affect Black racial isolation in neighborhoods (via decisions on acceptable land uses, mortgages, etc.), monitor siting and activity, and the presence of PM<sub>2.5</sub> sources. Meanwhile, the EPA explicitly prioritizes monitoring in areas where they expect to find exceedances of federal PM<sub>2.5</sub> standards, particularly in areas with known PM<sub>2.5</sub> sources that have produced PM<sub>2.5</sub> exceedances in the past. The selection bias results from conditioning (selecting census tracts based) on an effect (presence of an active monitor) of both an unmeasured cause of the exposure (racism in governmental bodies) and an unmeasured cause of the outcome (presence of a PM<sub>2.5</sub> source). (Racism in governmental bodies, by affecting siting decisions for PM<sub>2.5</sub> component sources, also induces ordinary unmeasured confounding by opening a backdoor path between the exposure and the outcome that runs exclusively through unmeasured variables in the current study). A simplified directed acyclic graph (DAG) illustrating the problem is shown below in Figure 1.<sup>3</sup>

**Figure 1. DAG showing selection bias from including only census tracts with PM<sub>2.5</sub> component monitors.**



Another limitation of the proposed analysis is the missing (but bounded) data in the outcome variables, ten-year mean PM<sub>2.5</sub> component concentrations, due to the relationship between the ambient concentrations of the PM<sub>2.5</sub> components and the monitor method detection limit (MDL) for each component. Figure 1 shows the percentage of observations (daily means) below the MDL for each PM<sub>2.5</sub> component.<sup>4</sup>

**Figure 2. Percentage of observations (daily means) below the MDL for each PM<sub>2.5</sub> component.**



<sup>3</sup>Hernán MA, Hernández-Díaz S, Robins JM. A Structural Approach to Selection Bias: Epidemiology 2004;15(5):615–25. doi:10.1097/01.ede.0000135174.63482.43. This particular form of selection bias takes the form of the bias illustrated in Figure 6c on page 617, but with additional ordinary unmeasured confounding introduced by the arrow from “racism in governmental bodies regulating land uses (unmeasured)”, which causes the exposure, to “PM<sub>2.5</sub> component source present (unmeasured)”, which causes the outcome.

<sup>4</sup>I realize that I am only supposed to identify one, but I add this second one because I would genuinely love advice on how to deal with it.