Assignment 2 - EDS 241: Environmental Policy Evaluation

Juliet Cohen

02/03/2022

The goal of this assignment is to provide a simple test of whether the effects of air quality regulations are the same across locations with different racial mix. To this end you will test if the NOx Budget Program, a cap-and-trade market for nitrogen oxides (NOx) emissions from power plants lead to similar effects in counties that are predominantly white versus counties that are predominantly African American. The data are a subset of data sample I used in the following paper: https://olivierdeschenes.weebly.com/uploads/1/3/5/0/135068654/defensive-investmentsand-the-demans-for-air-quality.pdf. You can also get more information on the NOx Budget Program, here: https://www.epa.gov/airmarkets/nox-budget-trading-program

The data included in the file NBP.xls, which is available on Gauchospace, are: fips (fips code identifying each county), NBP (indicator =1 if the county was regulated under the NOx Budget Program), `PctBlack` (fraction of the county population that is African American), and Dnox_masstons (change in annual NOx emissions from all power plants in a county between 2000 and 2008 (in tons)). Note that the NBP market was in effect in 212 of the 485 counties in the sample from 2003 to 2008, so the 2008-2000 change give us a sense of the program's effect on emissions. If emissions of NOx from power plants declined in a county, then Dnox_masstons should be negative.

0.0.1 Load Packages

```
library(here)
library(dplyr)
library(tidyverse)
library(estimatr)
library(modelsummary)
library(knitr)
library(readxl)
```

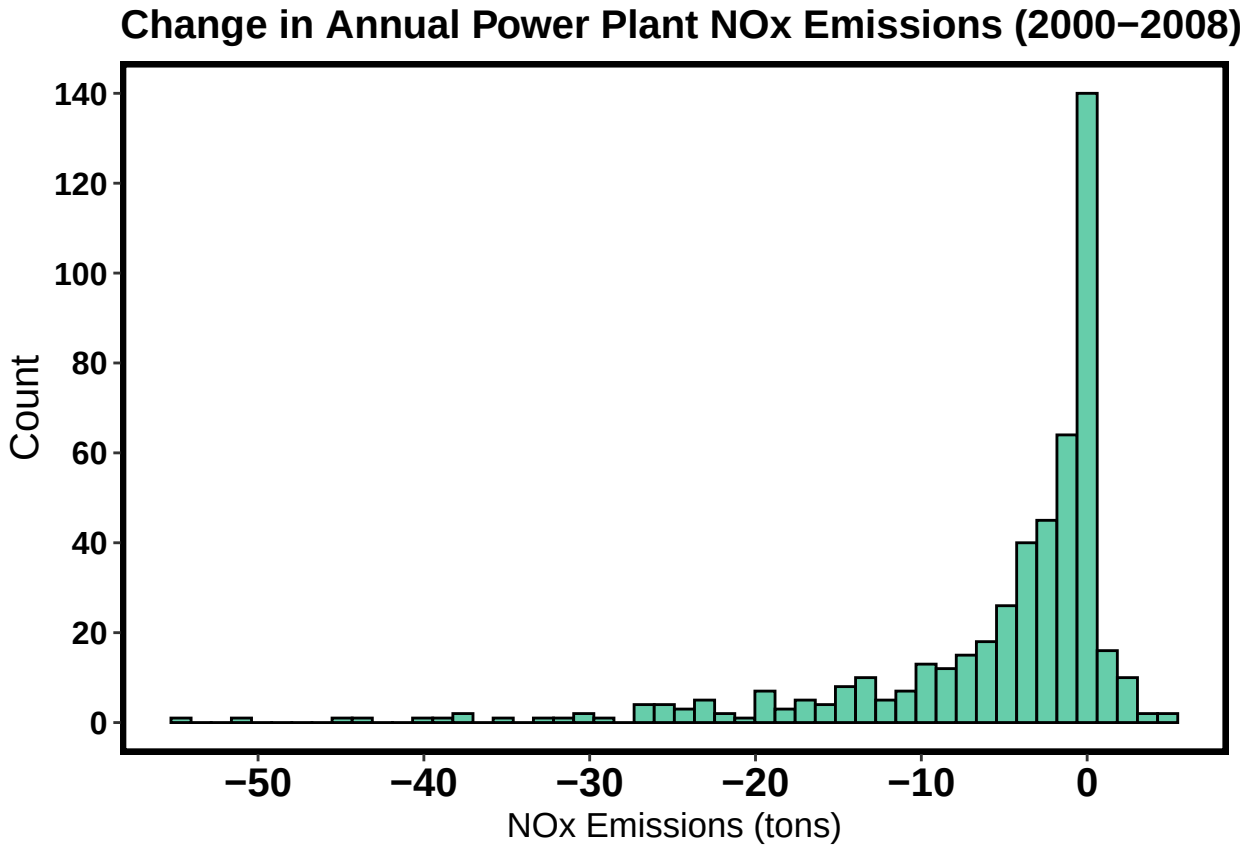
0.0.2 Read in Data

```
data <- readxl::read_excel(here("NBP.xls"))

# check for NA values
#map(data, ~sum(is.na(.)))
# there are none
```

0.1 (a) Make a histogram depicting the distribution of Dnox_masstons.

```
emissions_histogram <- ggplot(data = data, aes(x = Dnox_masstons)) +
  geom_histogram(bins = 50, color = "black", fill = "aquamarine3") +
  ggtitle("Change in Annual Power Plant NOx Emissions (2000-2008)") +
   xlab("NOx Emissions (tons)") +
   ylab("Count") +
   theme(panel.background = element_blank(),
```

```
        axis.title.x = element_text(color = "black", size = 13),
        axis.text.x = element_text(face = "bold", color = "black", size = 15),
        axis.title.y = element_text(color = "black", size = 15),
        axis.text.y = element_text(face = "bold", color = "black", size = 12),
        plot.title = element_text(color="black", size = 15, face = "bold"),
        panel.border = element_rect(colour = "black", fill = NA, size = 2)) +
  scale_y_continuous(breaks = seq(0, 160, by = 20)) +
  scale_x_continuous(breaks = seq(-60, 6, by = 10))

emissions_histogram
```



0.2 (b) Create an indicator =1 if the county has `PctBlack` above the sample median, and =0 otherwise (in the rest of the assignment, I refer to this variable as 'D'). What is the average of `PctBlack` for counties above the median, i.e. counties for which D=1?

```
# find median of the PctBlack values
median_pct_blk <- median(data$PctBlack)
median_pct_blk
```

```
## [1] 4.8
```

```
data_blk_med <- data %>%
  mutate(abv_bel_med_blk = case_when(
    PctBlack > median_pct_blk ~ 1,
    PctBlack <= median_pct_blk ~ 0))
```

```
data_blk_med_summary <- data_blk_med %>%
  group_by(abv_bel_med_blk) %>%
  summarise(mean = mean(PctBlack))

data_blk_med_summary$mean[1]
```

```
## [1] 1.556735
```

```
data_blk_med_summary$mean[2]
```

```
## [1] 19.31375
```

The average of `PctBlack` for counties above the median (D=1) is .

0.3 (c) Estimate a regression of Dnox_masstons on NBP. Interpret the estimated intercept and the coefficient on NBP.

```
nox_nbp_model <- lm_robust(formula = Dnox_masstons ~ NBP, data = data)

#nox_nbp_model$coefficients[1]
#nox_nbp_model$coefficients[2]

nox_nbp_model_table <- tidy(nox_nbp_model)
nox_nbp_model_table %>%
  select(term, estimate, std.error, p.value, conf.low, conf.high) %>%
  kable()
```

term	estimate	std.error	p.value	conf.low	conf.high
(Intercept)	-3.622031	0.4203230	0.0e+00	-4.447918	-2.796144
NBP	-3.920467	0.7959108	1.2e-06	-5.484342	-2.356591

```
nox_nbp_model_table
```

```
##          term  estimate std.error statistic      p.value  conf.low conf.high
## 1 (Intercept) -3.622031 0.4203230 -8.617257 9.830089e-17 -4.447918 -2.796144
## 2         NBP -3.920467 0.7959108 -4.925761 1.156072e-06 -5.484342 -2.356591
##    df       outcome
## 1 483 Dnox_masstons
## 2 483 Dnox_masstons
```

- For the regression of `Dnox_masstons` on `NBP`, the estimated intercept is -3.622031. This means that for the average county without these NBP Budget Program regulations (NBP is held constant at 0), the amount of change in annual NOx emissions from all power plants between 2000 and 2008 was -3.622031 tons. Since this number is negative, we can conclude that annual NOx emissions decreased by that amount when all else is held constant.

- The estimated coefficient on NBP is -3.9204667. This means that for the average county with these regulations under the NOx Budget Program, the amount of change in annual NOx emissions from all power plants between 2000 and 2008 was -3.9204667 tons. Since this number is negative, we can conclude that the NOx emissions decreased by that amount when counties were regulated but all else is held constant.

0.4 (d) Create an interaction between the variables NBP and D. Estimate a regression of Dnox_masstons on NBP, D, and this interaction. Interpret each estimated regression coefficient, including the intercept.

```
nox_nbp_d_model <- lm_robust(formula = Dnox_masstons ~ NBP + abv_bel_med_blk + NBP:abv_bel_med_blk, data

nox_nbp_d_model_table <- tidy(nox_nbp_d_model)
nox_nbp_d_model_table %>%
  select(term, estimate, std.error, p.value, conf.low, conf.high) %>%
  kable()
```

term	estimate	std.error	p.value	conf.low	conf.high
(Intercept)	-2.418075	0.4423052	0.0000001	-3.287164	-1.5489862
NBP	-7.141242	1.2572938	0.0000000	-9.611709	-4.6707748
abv_bel_med_blk	-2.588031	0.8533574	0.0025542	-4.264800	-0.9112619
NBP:abv_bel_med_blk	6.371798	1.6144274	0.0000910	3.199597	9.5439997

- The intercept is -2.4180753, which represents the average county's change in annual NOx emissions from all power plants between 2000 and 2008 (in tons) when the county was not regulated under the NOx Budget Program and the county's `PctBlack` is below the sample median, meaning the fraction of the county population that is African American is below the sample median.

- The estimated regression coefficient for NBP is -7.1412417, which represents the average county's change in annual NOx emissions from all power plants between 2000 and 2008 (in tons) when the county is regulated under the NOx Budget Program and the county's `PctBlack` is below the sample median, meaning the fraction of the county population that is African American is below the sample median.

- The estimated regression coefficient for D (which is represented by the variable `abv_bel_med_blk` in my code) is -2.5880307, which represents the average county's change in annual NOx emissions from all power plants between 2000 and 2008 (in tons) when the county is not regulated under the NOx Budget Program and the county's `PctBlack` is above the sample median, meaning the fraction of the county population that is African American is above the sample median.

- The estimated regression coefficient for the interaction between NBP and D is 6.3717981, which represents the difference in the average county's change in annual NOx emissions from all power plants between 2000 and 2008 (in tons) when counties are regulated under the NOx Budget Program and the `PctBlack` is above the sample median, meaning the fraction of the county population that is African American is above the sample median, compared to the average county's change in annual NOx emissions from all power plants between 2000 and 2008 (in tons) when there are regulations under the NOx Budget Program and the county's `PctBlack` is below the sample median, meaning the fraction of the county population that is African American is below the sample median. This interaction term is the difference in the effect of regulation between predominantly black and predominantly white counties. We are investigating if the regulations bring the same benefits the these two categories of counties from an environmental justice perspective. Because this interaction term is not 0, we know there is indeed a difference for predominantly black and white communities, and because this interaction term is positive, we know that predominantly black communities are not getting as much benefit from these regulations compared to white communities. Emissions are improving for both predominantly white and predominantly black counties, but not as much for predominantly black counties.

0.5 (e) What is the predicted Dnox_masstons in a county that was not regulated under NBP and where PctBlack is above the sample median (i.e., where D=1)? Report the 95% confidence interval for this prediction. Make sure to use "heteroskedasticity-robust" standard errors.

```
not_reg_abv_blk <- data.frame(NBP = c(0), abv_bel_med_blk = c(1))

CI <- predict(object = nox_nbp_d_model, newdata = not_reg_abv_blk, se.fit = TRUE, interval = "confidenc
CI
```

```
## $fit
##            fit       lwr       upr
## [1,] -5.006106 -6.440065 -3.572147
##
## $se.fit
##         1
## 0.7297841
```

The predicted Dnox_masstons in a county that is not regulated under NBP and where `PctBlack` is above the sample median is -5.006106 tons, which represents the change in annual NOx emissions from all power plants in a county between 2000 and 2008 (in tons). The 95% confidence interval for this prediction is [-6.440065, -3.572147].