EDS241: Assignment 2

Wylie Hampson

02/04/2022

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 I 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

Import and clean the data

```
nox_data <- read_excel(here("data", "NBP.xls")) %>%
  clean_names()
```

Question a: Make a histogram depicting the distribution of Dnox_masstons.

```
dnox_hist <- ggplot(nox_data, aes(x = dnox_masstons)) +
  geom_histogram() +
  labs(x = "Change in Annual NOx Emissions (tons)",
       y = "Count",
       title = "Change in Annual NOx Emissions From All Power Plants in a County Between
2000 and 2008") +
  theme(title = element_text(size = 6))

dnox_hist
```

Change in Annual NOx Emissions From All Power Plants in a County Between 2000 and 2008
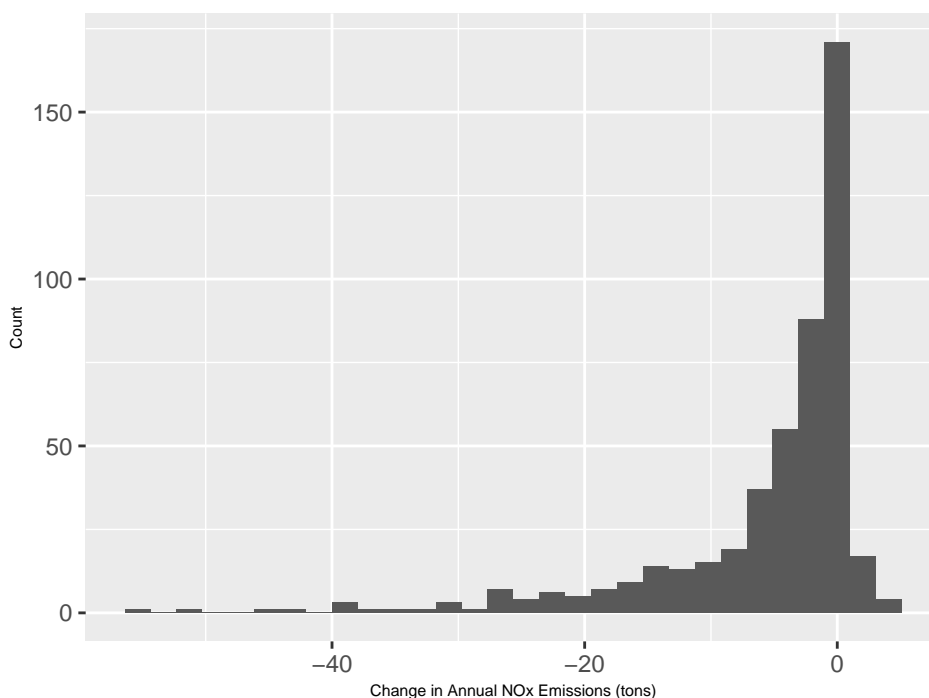
Question b: Create an indicator =1 if the county has PctBlack above the sample median, and =0 otherwise (in the rest of the assignment, I refer to this variable as 'D'). What is the average of PctBlack for counties above the median, i.e. counties for which D=1?

```
pct_black_median <- median(nox_data$pct_black)

nox_data <- nox_data %>%
  mutate(d = ifelse((pct_black > pct_black_median), 1, 0))

counties_above_median <- nox_data %>%
  filter(d == 1)

avg_above_median <- mean(counties_above_median$pct_black)
avg_above_median
```

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

For counties that have a percentage of African Americans in its population greater than the median for all counties, the average percentage of African Americans is 19.31375%.

Question c: Estimate a regression of Dnox_masstons on NBP. Interpret the estimated intercept and the coefficient on NBP.

```
model_1 <- lm_robust(formula = dnox_masstons ~ nbp, data = nox_data)

model_1_table <- broom::tidy(model_1) %>%
  dplyr::select(term, estimate, std.error, p.value) %>%
  knitr::kable()

model_1_table
```

term	estimate	std.error	p.value
(Intercept)	-3.622031	0.4203230	0.0000000
nbp	-3.920467	0.7959108	0.0000012

*The regression of **dnox_masstons** on nbp gives us an intercept coefficient of -3.62, which means that for a county that was not regulated under the NOx budget program, we would expect their change in annual NOx emissions from all power plants between 2000 and 2008 to be -3.62 tons on average. The nbp coefficient in this regression is -3.92, which means that we would expect counties that were regulated under the NOx budget program to have their change in annual NOx emissions from all power plants between 2000 and 2008 be -3.92 tons more than counties that were not regulated under the NOx budget program. The nbp coefficient is the difference in **dnox_masstons** between counties that we under NBP regulations and those that were not.*

Question 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.

```
model_2 <- lm_robust(formula = dnox_masstons ~ nbp + d + (nbp * d), data = nox_data)

model_2_table <- broom::tidy(model_2) %>%
  dplyr::select(term, estimate, std.error, p.value) %>%
  knitr::kable()

model_2_table
```

term	estimate	std.error	p.value
(Intercept)	-2.418075	0.4423052	0.0000001
nbp	-7.141242	1.2572938	0.0000000
d	-2.588031	0.8533574	0.0025542
nbp:d	6.371798	1.6144274	0.0000910

*In this regression the intercept coefficient is -2.42. This says that for counties that are not under NBP regulations and whose percentage of black population is below the median, on average they would expect to have a **dnox_masstons** of -2.42. The nbp coefficient is -7.14, which is the difference of **dnox_masstons** between counties that are under NBP regulations with a black population percentage below the median, and counties that are not under NBP regulations, with a black population percentage below the median. The d coefficient is -2.59, which is the difference of **dnox_masstons** between counties that have a black population percentage above the median, and not under NBP regulatiosn, and counties that have a black population percentage below the median, and not under NBP regulations. Finally, the nbp:d coefficient is 6.37, which is the change in **dnox_masstons** in counties that are under NBP regulations with with a black population percentage above the median, in addition to the change associated with the other coefficients.*

Question 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.

```
nox_pred <- data.frame(nbp = 0, d = 1)

predict(model_2, newdata = nox_pred, se.fit=TRUE, interval='confidence')
```

```
## $fit
```

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

In a county that was not regulated under NBP, and where **pct_black** *is above the sample median, the predicted* **dnox_masstons** *(the change in annual NOx emissions from all power plants in a county between 2000 and 2008 (in tons)) is -5.01 tons. We are 95% confident that the predicted* **dnox_masstons** *is between -6.44 tons and -3.57 tons.*