EDS241: Assignment 02 - Air Quality Regulations

Julia Parish

2022/02/06

1 EDS241 Environmental Policy Evaluation Assignment 02

This statistical test was completed as an assignment for the course, Environmental Data Science 241: Environmental Policy Evaluation. The goal of this assignment was to provide a simple test of whether the effects of air quality regulations are the same across various regions with diverse demographics. the test will analyze 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.

Source:

The data was sourced from researched conducted by Dr. Deschenes et al..

Glossary

FIPS - Federal Information Processing System codes 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
Dnox_masstons - change in annual NOx emissions from all power plants in a county between 2000 and 2008 (in tons).

```
# read in the NOx Budget data

nox_data <- read_excel("data/NBP.xls") %>%
  clean_names()

# check for NA values
#map(nox_data, ~sum(is.na(.)))
```

1.1 Question A. Make a histogram depicting the distribution of Dnox_masstons.

```
dnox_histogram <- ggplot(data = nox_data, aes(x = dnox_masstons)) +
  geom_histogram(color = "#1C366B",
                 fill = "#1DACE8",
                 binwidth = 2) +
  labs(title = "Distribution of Change in NOx emissions (t)",
       subtitle = "2000 - 2008",
       y = "Frequency",
       x = "NOx Emissions (tons)") +
  theme_minimal()
```

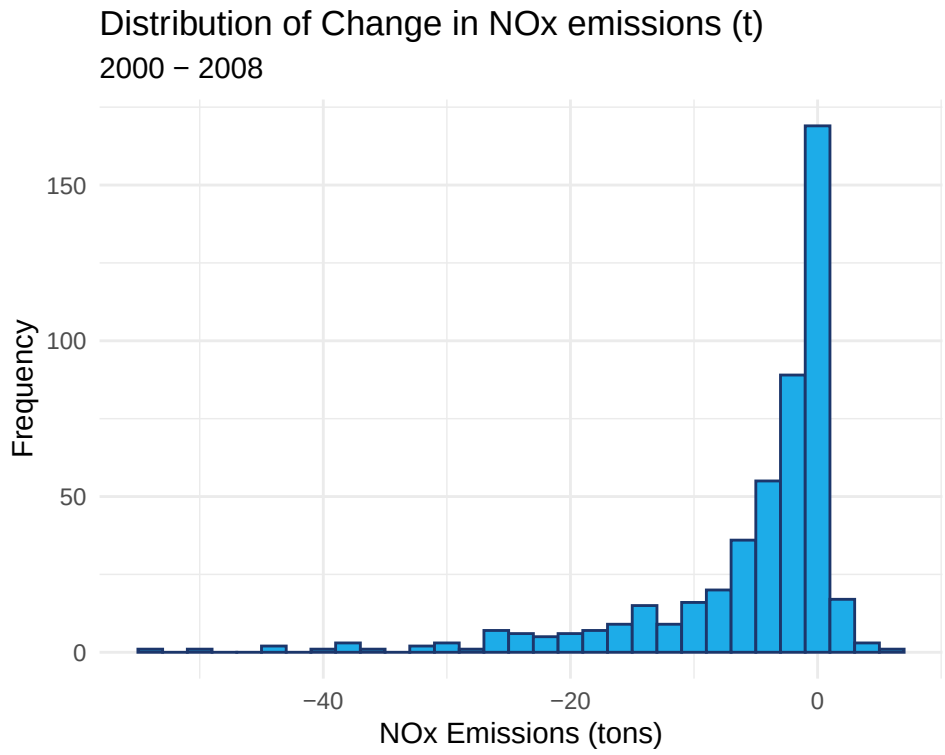

Figure 1: This histogram displays the distribution of change in Nitrogen Oxides emission between the years 2000 and 2008 for 485 counties in the United States. It shows a reduction in NOx emmissions between 2000 & 2008. The NOx Budget Trading Program (NBP) was implemented in 2003 as a cap and trade program to reduce NOx emissions from power plants and other large sources on the east coast. The NBP regulated emissions in 212 of the 485 counties in this dataset. This histogram shows that the NBP may have reduced NOx emissions in New England.

1.2 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?

```
# create median percent black values for counties in dataset
medpctblk <- median(nox_data$pct_black)
medpctblk
```

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

```
# create d variable for +/- percent black population in counties
nox_data <- nox_data %>%
  mutate(d = case_when(
    pct_black > medpctblk ~ 1,
    pct_black <= medpctblk ~ 0
  ))

# filter d variable for + percent black population to create mean
avgpctblk <- nox_data %>%
  filter(d == 1) %>%
  summarize(mean = mean(pct_black))
```

```
avgpctblk
```

mean
19.3

1.2.0.1 Answer

The average percent of black residents in counties above the median is 19.3%.

1.3 Question C. Estimate a regression of Dnox_masstons on NBP. Interpret the estimated intercept and the coefficient on NBP.

$$Y_i = \beta_0 + \beta_1 D_{1i} + u_i \tag{1}$$

The regression being evaluated is for Y_i as the average rate of change in annual NO_x emissions in tons in counties, i, contained in the dataset. β_0 is the intercept and refers to a value of Y_i when $X = 0$. $\beta_1 D_{1i}$ is the binary variable indicating whether or not a county was regulated under NBP. u_i is the regression error term.

```
# robust regression
dnox_mod_robust <- lm_robust(dnox_masstons ~ nbp, data = nox_data)

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

term	estimate	std.error	p.value	conf.low	conf.high
(Intercept)	-3.622031	0.4203230	0.0000000	-4.447918	-2.796144
nbp	-3.920467	0.7959108	0.0000012	-5.484342	-2.356591

```
# linear regression for stargazer
dnox_lm_mod <- lm(data = nox_data, formula = dnox_masstons ~ nbp)
se_dnoxmod <- starprep(dnox_lm_mod)
```

```
# stargazer table
stargazer(dnox_lm_mod, se = se_dnoxmod,
          type = "latex",
          ci = FALSE,
          no.space = TRUE,
          header = FALSE,
          omit.stat = c("adj.rsq", "ser", "f"),
          covariate.labels = c("Regulated", "Unregulated (Intercept)"),
          dep.var.labels = c("Change in NOx emissions (t)"),
          dep.var.caption = c(""),
          initial.zero = FALSE,
          title = "Change in Annual NOx Influenced by NBP Regulations",
          table.placement = "H")
```

Table 1: Change in Annual NOx Influenced by NBP Regulations

	Change in NOx emissions (t)
Regulated	−3.920***
	(.796)
Unregulated (Intercept)	−3.622***
	(.420)
Observations	485
R^2	.052
Note:	*p<0.1; **p<0.05; ***p<0.01

1.3.0.1 Answer For the regression of Dnox_masstons on NBP, β_0, the estimated intercept, is -3.622. This translates to an annual *reduction* in NO_x emissions of 3.662 tons for the average unregulated county in the Eastern US between 2000 and 2008 when all else is held constant.

For the regression of Dnox_masstons on NBP, β_1 , the estimated coefficent, is -3.920 and represents NBP regulated counties. This translates to for every additional year, the average NO_x emissions is *reduced* by an **additional** 3.920 tons more than unregulated counties when all else is held constant.

1.4 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.

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_{3i} D_{3i} + u_i \tag{2}$$

The regression being evaluated is for Y_i is the average rate of change in annual NO_x emissions, in tons, in a county (i). D_{1i} is a binary variable of whether a county was regulated or unregulated by NBP. D_{2i} is a binary variable of whether a county's residences are predominantly white or black. D_{3i} is the interaction between the two binary variables of D_{1i} and D_{2i}, which evaluates the interaction of whether a county was NBP regulated and if the residential population is predominantly white or black. u_i is the regression error term.

```
# robust interaction regression
nox_d_robust <- lm_robust(dnox_masstons ~ nbp + d + nbp:d, data = nox_data)

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

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
d	-2.588031	0.8533574	0.0025542	-4.264800	-0.9112619
nbp:d	6.371798	1.6144274	0.0000910	3.199597	9.5439997

```
# for stargazer table
nox_d_mod <- lm(data = nox_data,
             formula = dnox_masstons ~ nbp + d + nbp:d)

se_nox_d_mod <- starprep(nox_d_mod)
```

Table 2: Influence of NBP on Annual NOx Pollution

	Change in NOx emissions (t)
NBP Regulated, predominantly white	-7.141^{***}
	(1.257)
Unregulated, predominantly African American	-2.588^{***}
	(.853)
NBP Regulated, predominantly African American	6.372^{***}
	(1.614)
Unregulated, predominantly white	-2.418^{***}
	(.442)
Observations	485
R^2	.086
Note:	*p<0.1; **p<0.05; ***p<0.01

1.4.0.1 Answer The β_0 intercept, -2.418, represents that for NBP unregulated counties that had predominately white residents had an average annual *decrease* in NO_x emissions of 2.4 tons between 2000 and 2008.

The β_1 coefficient, -7.141, represents that for NBP regulated counties (NBP = 1) and the population is predominately white (D = 0) had an average annual *decrease* in NO_x emissions of 7.14 tons more than unregulated, predominately white counties (β_0) between 2000 and 2008.

The β_2, or D, coefficient, -2.588, represents that for unregulated counties (NBP = 0) and the population is predominately black (D = 1) had an average annual *decrease* in NO_x emissions of 2.588 tons more than unregulated, predominately white counties (β_0) between 2000 and 2008.

The β_3 coefficient, or the NBP:D interaction, is 6.372. This represents that for regulated counties (NBP = 1) with a population that is predominately black (D = 1), the average annual *decrease* in NO_x emissions is 6.4 tons less than for counties that are predominately white.

1.5 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.

```
# predict change in annual NOx emissions from all power plants in a county between 2000 & 2008
# when NBP = 0 and D = 1.

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

ci <- predict(nox_d_robust, newdata = dnox_predict,
                      se.fit = T, interval = 'confidence')

ci
```

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

```
## 0.7297841
```

1.5.0.1 Answer The predicted change in annual NO_x emissions (dnox_masstons) from all power plants in a county that was not regulated under NBP and where the population is predominately black (D = 1) is a decrease of 5.01 tons. It is with 95% confidence that the predicted change in annual NO_x emissions (dnox_masstons) is a decrease between 3.57 and 6.44 tons.