

EDS241: Assignment 2

Allie Cole

02/06/2022

1 Clean and plot data

The following code loads and cleans the data.

```
# Load data

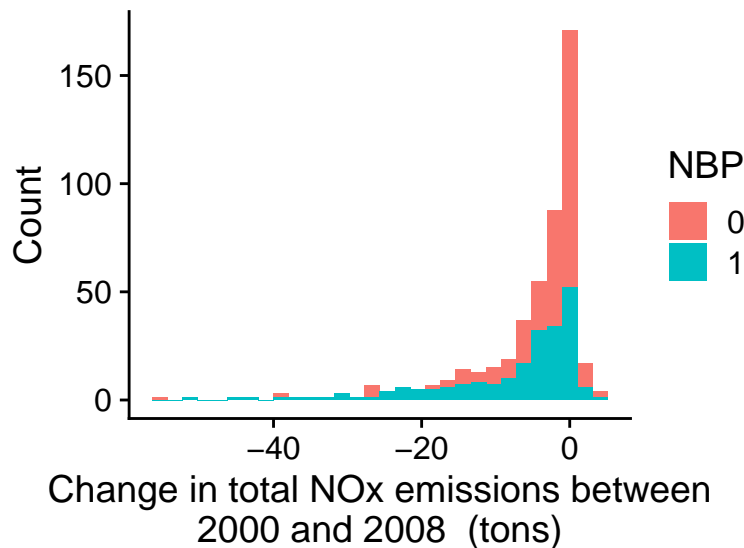
library(readxl)
data_nox <- read_excel("NBP.xls", sheet = 1, na = "NA") %>%
  clean_names()
```

The data included in the file NBP.xls 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 total 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.

2 Homework Questions

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

```
ggplot(data = data_nox, aes(x = dnox_masstons)) +
  geom_histogram(aes(fill = as.factor(nbp), position = 'dodge')) +
  theme_cowplot(14) +
  labs(x = "Change in total NOx emissions between
2000 and 2008 (tons)",
       y = "Count",
       fill = "NBP")
```



2.0.2 (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?

```
median_pctblack <- median(data_nox$pct_black)

data_nox <- data_nox %>%
  mutate(D = if_else(pct_black > median_pctblack, 1, 0)) #creating a new column that will have a 1 if t

average <- data_nox %>%
  select(pct_black, D) %>%
  filter(D == 1)

mean <- mean(average$pct_black)
```

The average of PctBlack for counties above the median is 19.31375

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

coefficient on NBP.

```
nbp_model <- lm(data = data_nox, dnox_masstons ~ nbp)

table1 <- broom::tidy(nbp_model) %>%
  knitr::kable() %>%
  kableExtra::kable_styling(full_width = FALSE,
                             latex_options = "HOLD_position")
table1
```

term	estimate	std.error	statistic	p.value
(Intercept)	-3.622031	0.5044385	-7.180322	0.0000000
nbp	-3.920467	0.7629766	-5.138384	0.0000004

The intercept here is the average rate of change from 2000 to 2008 in all nox emissions within counties

that are unregulated, so the predicted change in for those unregulated counties is -3.62 tons (when nbp is zero). The coefficient predicts that the change in annual NOx emissions was on average -3.92 ton higher for counties that were regulated, when compared to those that were unregulated. We know that a negative dnox_masstons gives us a decrease in NOx, so from this we can say that the NOx Budget Program was working and annual Nox emissions were decreasing by 3.92 tons when the county is regulated.

2.0.4 (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.

```
nbp_d_model <- lm_robust(data = data_nox, dnox_masstons ~ nbp + D + nbp:D)
```

```
table2 <- broom::tidy(nbp_d_model) %>%
  knitr::kable() %>%
  kableExtra::kable_styling(full_width = FALSE,
                             latex_options = "HOLD_position")
```

table2

term	estimate	std.error	statistic	p.value	conf.low	conf.high	df	outcome
(Intercept)	-2.418075	0.4423052	-5.466984	0.0000001	-3.287164	-1.5489862	481	dnox_masstons
nbp	-7.141242	1.2572938	-5.679851	0.0000000	-9.611709	-4.6707748	481	dnox_masstons
D	-2.588031	0.8533574	-3.032763	0.0025542	-4.264800	-0.9112619	481	dnox_masstons
nbp:D	6.371798	1.6144274	3.946785	0.0000910	3.199597	9.5439997	481	dnox_masstons

The intercept here shows the average rate of change of all nox emissions within a given county between 2000 to 2008 for a county that is unregulated and with a percentage of the African American population below the sample median is -2.42 tons. Or in terms of the variables, when nbp is zero and the pct_black is below 19.31 and D=0, the change in NOx emissions in the counties between 2000-2008 will be -2.42 tons

The estimate for nbp is saying that when everything is held constant, counties that were regulated under the Nox budget plan (nbp = 1) will have an average change from all power plants in the counties between 2000- 2008 of -7.14 tons compared to the intercept (Nbp= 0, pct_black = below 19.31 and D=0).

The estimate for D is saying that on average counties that have an average of African American populations higher than the median (19.31) will have a change in NOx emissions from all power plants in the counties between 2000 to 2008 of -2.59 compared to the intercept (Nbp= 0, pct_black = below 19.31 and D=0).

The estimate of nbp:D is the intersection between the implementation of the policy and the percent black above sample mean. This is saying that the NOx budget program raised the NOx emissions from all power plants in the counties between 2000- 2008 by 6.37 tons in counties that have an average percent of the African American population above 19.31% compared to counties that have an average percent of the African American population below 19.31%.

2.0.5 (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.

```
model_e <- lm_robust(data = data_nox, dnox_masstons ~nbp + D)
```

```
predict_data <- data.frame(nbp = c(0), D = c(1))
```

```
predict(model_e, newdata = predict_data, se.fit = T, interval = 'confidence')
```

```
## $fit
##           fit           lwr           upr
## [1,] -3.516179 -4.818788 -2.213571
##
## $se.fit
##           1
## 0.6629394
```

The predicted Dnox_masstons in a county that was not regulated under NBP and where PctBlack is above the sample median is -3.52 tons, with a 95% confidence interval, using heteroskedasticity robust standard errors, of -4.82 and -2.21 tons.