EDS241: Final Assingment

Jake Eisaguirre

03/15/2022

```
data <- read_csv(here("KM_EDS241.csv"))
```

0.0.0.0.1 Clean Data

```
## Rows: 321 Columns: 7
```

```
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## dbl (7): year, age, rooms, area, land, nearinc, rprice
```

```
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

0.0.0.0.2 (a) Using the data for 1981, estimate a simple OLS regression of real house values on the indicator for being located near the incinerator in 1981. What is the house value "penalty" for houses located near the incinerator? Does this estimated coefficient correspond to the 'causal' effect of the incinerator (and the negative amenities that come with it) on housing values? Explain why or why not.

```
data_81 <- data %>%
  filter(year == 1981)

mod1 <- lm_robust(rprice ~ nearinc, data = data_81)
```

```
tidy(mod1) %>%
  kable()
```

term	estimate	std.error	statistic	p.value	conf.low	conf.high	df	outcome
(Intercept)	101307.51	2944.810	34.402059	0.0e+00	95485.47	107129.6	140	rprice
nearinc	-30688.27	6243.167	-4.915498	2.4e-06	-43031.35	-18345.2	140	rprice

Answer: The model shows that the penatly for being near the incinerator for the year 1981 is about $30688, or in other words the average housing prices for houses near the incineratror is $30688 less than the average housing prices far away from the incinerator. No I do not believe this is representative of the causal effect of adding an incinerator because to begin with the incinerator would not have been added near places with expensive houses due to push back from home owners.

0.0.0.0.3 (b) Using the data for 1978, provide some evidence the location choice of the incinerator was not "random", but rather selected on the basis of house values and characteristics. [Hint: in the 1978 sample, are house values and characteristics balanced by nearinc status?]

```
data_78 <- data %>%
  filter(year == 1978)

mod2 <- lm_robust(rprice ~ nearinc, data = data_78)
```

```
tidy(mod2) %>%
  kable()
```

term	estimate	std.error	statistic	p.value	conf.low	conf.high	df	outcome
(Intercept)	82517.23	1878.277	43.932406	0.0000000	78810.53	86223.927	177	rprice
nearinc	-18824.37	6010.014	-3.132167	0.0020309	-30684.88	-6963.864	177	rprice

Answer: The model shows that housing prices in 1978 from the surrounding area prior to the incinerator location were already on average worth $18,824 less that houses that are considered farther away. This points to the incinerator location not being random and was selected to be placed in a less wealthy area.

0.0.0.0.4 (c) Based on the observed differences in (b), explain why the estimate in (a) is likely to be biased downward (i.e., overstate the negative effect of the incinerator on housing values).

Answer: Since there was a housing value difference in 1978 and one in 1981, Our first model estimate is biased downward. We can not assume that all the ~$30,000 difference can be due to the presence of the incinerator because we already saw a difference of ~$18000 prior to the incinerator being installed. Our models can not dis-associate the effect of the incinerator and two different locations, so the incinerator effect gets added on to make it ~$30000

0.0.0.0.5 (d) Use a difference-in-differences (DD) estimator to estimate the causal effect of the incinerator on housing values without controlling for house and lot characteristics. Interpret the magnitude and sign of the estimated DD coefficient.

0.0.0.0.6 (e) Report the 95% confidence interval for the estimate of the causal effect on the incinerator in (d).

0.0.0.0.7 (f) How does your answer in (d) changes when you control for house and lot characteristics? Test the hypothesis that the coefficients on the house and lot characteristics are all jointly equal to 0.

0.0.0.0.8 (g) Using the results from the DD regression in (f), calculate by how much did real housing values change on average between 1978 and 1981.

0.0.0.0.9 (h) Explain (in words) what is the key assumption underlying the causal interpretation of the DD estimator in the context of the incinerator construction in North Andover.