

Exercise 11: Difference-in-Difference Design, New Orleans Population

ECON 256

Data Analysis and Visualization

Objective

Run a difference-in-difference regression and interpret the results.

1 Set up Your R Workspace

Set up a working directory and initialize TidyVerse.

2 Download/Load the Data

Download the “populations.csv” data from Laulima.

Use the `read_csv()` command to load the .csv file into R.

The data set contains the populations of New Orleans and Shreveport, Louisiana for 1990, 2000 and 2010.

3 Setup and Run a “Difference-in-Difference” Regression

In 2005 Hurricane Katrina had a devastating effect on New Orleans. Many people who lost their homes simply left the city, causing the population of New Orleans to fall dramatically. Shreveport is an other city in Louisiana, but was essentially unaffected by Katrina. If we assume New Orleans and Shreveport experienced similar population growth pressures, and the only significantly different factor was Hurricane Katrina, we could use a difference-in-difference design to estimate the population loss in New Orleans by using Shreveport as a comparison.

Add a few new variables to the data:

1. Add a variable that equals 1 for New Orleans observations and 0 for Shreveport observations. Something like this: `pop2<-mutate(pop,neworleans = ifelse(city=="New Orleans",1,0))`
2. Add another variable that equals 1 if the observation was made after 2005 and 0 if it was before 2005.
3. Add a third variable that is the *interaction* of the two variables you just made. In other words, it equals one if the observation is in New Orleans AND was made after 2005. You can create this variable by multiplying the two variables you just made.

Now run a difference-in-difference regression. We need to control for the general population difference between New Orleans and Shreveport, and the general change in population growth across years, and then use our interaction term to estimate the unique impact of being IN New Orleans AFTER Katrina on population. This will equal the population loss we can attribute to the hurricane.

The regression will look something like this, but will depend on the names you chose for your objects and variables:

```
summary(lm(formula= population ~ poststorm + neworleans + interaction,data=popsdata))
```

What is your estimate of how many people left New Orleans because of Katrina? (answer in a comment)

4 Graph the Effect

Using ggplot, create a graph that shows population over time, with one line for each city. Is the parallel trend assumption reasonable here? (answer in a comment)

5 Send me Your Code

Save your R code.

Name the R code with your last name, followed by the exercise number.

Submit it on Laulima.