PC project 4: Differences-in-Differences

You are expected to solve this PC-Project until and including Exercise 2b prior to 21.03.2022. Submit the PDF with your answers as well as your Python files with a reproducible code in pc4.py format. All functions should be submitted as pc4_functions.py. Upload your solution in a zip file named pc4_yournames.zip to the designated module in Canvas. Please, provide all 3 files, namely the pc4.py, pc4_functions.py and pc4.pdf. Both the codes and the written answers are graded.

General information

In this PC project we use a random sample of data coming from the case study of David Card and Alan B. Krueger (1994). The study analyses the effect of higher minimum wages on employment in the US fast-food sector based on panel data. It uses a rise in *New Jersey*'s minimum wage in 1992 from $4.25 to $5.05 per hour to evaluate employment changes in restaurants induced by the policy change compared to employment in *Pennsylvania*, where the minimum wage remained constant. The data set includes the following variables:

Variable	Description
STOREID	Unique restaurant identifier
YEAR	1992 or 1993
STATE	Dummy: 1 if New Jersey (NJ), 0 if Pennsylvania (PA)
SOUTHJ	Dummy: 1 if southern NJ
CENTRALJ	Dummy: 1 if central NJ
NORTHJ	Dummy: 1 if northern NJ
PA1	1 if in PA, northern suburbs of PA.
PA2	1 if in PA, eastern PA
CHAIN	Fast food chain: 1 if Burgerking, 2 if kfc, 3 if Royrogers, 4 if Wendys
CO_OWNED	1 if company-owned, 0 if franchisee-owned
HRSOPEN	Number of hours open per day
PRICE	Price of a full (standard) meal
FTE	Size of restaurant in terms of full-time equivalent employment
WAGE_ST	Starting wage per hour in US$

Exercise 1: Descriptive Statistics

Download the dataset *data_pc4.csv* and save it in a folder on your computer. Download also the *pc4.py* and *pc4_functions.py* and save them in the same folder on your computer. You may re-use the functions from previous PC projects and include them in *pc4_functions.py.*

  a. Open the *pc4.py* in Spyder. Specify the path variable accordingly and make sure that *pc4_functions* is imported. Load the *data_pc4.csv* into your environment as Pandas DataFrame object.

  b. Code new, adjust or re-use the summary statistics function developed in previous PC projects to inspect the dataset. Can you detect any anomalies? Does the dataset contain any missing or implausible values?

  c. Plot a histogram of the dependent variable *fte*, i.e., number of full-time equivalent employment separately for each year and state and comment on its distribution. Additionally, plot a histogram of the change in *fte* across time for both states. Can you detect any outliers?

  d. Check if the regional dummy variables *southj, centralj, northj, pa1*, and *pa2* are well defined, i.e., if they sum up to 1. Code a small function for this and present your results in a table.

  e. Code a function which creates a table with the mean values of *fte* and the following covariates: *wage_st*, *hrsopen*, and *price*, together with the number of observations separately for each year and state. What information can you infer from this table?

  f. Recode the variable *chain* into dummies. Also recode the variable *year* into a dummy which equals 1 for the year 1993. Use the Pandas module for this.

Exercise 2: Difference-in-Differences

  a. Estimate the effect of higher minimum wages on full time equivalent employment in fast food restaurants by mean difference between New Jersey and Pennsylvania after the policy change. You can use the *ate_md()* function

provided in *pc4_functions.py*. Interpret the estimated effect and discuss the potential problems in this approach.

b. Estimate the effect of higher minimum wages on full time equivalent employment in fast food restaurants by mean difference in New Jersey before and after the policy change. You can use the *ate_md()* function provided in *pc4_functions.py*. Interpret the estimated effect and discuss the potential problems in this approach.

c. Estimate the ATET of higher minimum wages on full time equivalent employment in fast food restaurants by simple difference of mean differences between treated and controls across time and provide the standard error. Code your own function to estimate these parameters. Are the results in line with the economic theory that minimum wages reduce employment? (Discuss e.g. the study by Dube, A., Lester, T. W., & Reich, M. (2010) on the effect of minimum wages on employment.) Furthermore, explain why we are interested in the ATET parameter, instead of the ATE. How would you identify the ATE?

d. Consider a linear regression version of a DiD model *without* covariates that yields the ATET of higher minimum wages on full time equivalent employment in fast food restaurants. Estimate this model by the pooled OLS method and present a table stating the estimated regression coefficients, their standard errors, and the corresponding t-statistics and p-values. Code your own function to estimate these parameters. Compare the results of exercise 2c to the results obtained now. Do you notice any differences, and if so, can you explain them?

e. Now, consider a linear version of the DiD model with covariates and repeat the estimation of the ATET by pooled OLS. Include the following set of covariates: *wage_st, hrsopen, price, centralj, northj, pa2, co_owned, chain_2, chain_3,* and *chain_4*. Discuss the role of covariates in the DiD model. How does the estimated ATET change after including the covariates? Comment on the obtained results.

References

Card, David, and Alan B. Krueger (1994). Minimum Wages and Employment: A Case Study of the New Jersey and Pennsylvania Fast Food Industries. *American Economic Review*, 84:4, 772–793.

Dube, A., Lester, T. W., & Reich, M. (2010). Minimum wage effects across state borders: Estimates using contiguous counties. *The review of economics and statistics*, 92(4), 945-964.