

Homework 4

Introduction to Econometrics, Fall 2018

1/13/2018

1 Due Date:

- *Due to Jan.30 2019, 12:00 pm*

2 Learning Objectives

- Use *Stata* or *R* to understand and practice following Econometric Methods
 - **Regression Discontinuity Design**
 - **Fixed Effects Model**
 - **Difference-in-Differences**

3 Empirical Exercise

3.1 RDD: Minimum Legal Drinking Age and Death Rates

- Christopher Carpenter and Carlos Dobkin(2009), “The Effect of Alcohol Consumption on Mortality: Regression Discontinuity Evidence from the Minimum Drinking Age”, *Am Econ J Appl Econ.* 2009 January 1; 1(1): 164–182.
- Data and Variable Descriptions: The file `mlda.dta` (in Stata format) contains data the National Center for Health Statistics (NCHS) confidential mortality detail files for 1997-2004. These data are derived from death certificates and cover all deaths in the United States in the study period.
- The sample is restricted to fatalities of young adults aged 19-22. The data used here consist of averages in 48 cells defined by age in 30-day intervals.
- Causes of death are divided into internal and external, with the latter split into mutually exclusive subcategories: homicide, suicide, motor vehicle accidents, and other external causes. A separate category for alcohol-related causes covers all deaths for which alcohol was mentioned on the death certificate.
- Outcomes are mortality rates per 100,000, where the denominator comes from census population estimates.

Variables	Description
all	Death rate from all causes(per 100,000)
internal	Death rate from internal causes (per 100,000)
external	Death rate from external causes (per 100,000)
alcohol	Death rate from alcohol-related causes (per 100,000)
homicide	Death rate from homicide causes (per 100,000)
suicide	Death rate from suicide causes (per 100,000)
mva	Death rate from motor vehicle accidents (per 100,000)
drugs	Death rate from drugs-related causes (per 100,000)
externalother	Death rate from other external causes (per 100,000)
agecell	age in month

- 1) Describe the role of the running variable in the case and generate the treatment variable and call it *over21*.
- 2) Create a scatter plot of death rate from all causes against age. Describe the relationship between these two variables based on the scatter plot.(Figure_1)
- 3) Use death rate from all causes as an outcome variable and estimate the following RD regression:

$$deathrate_i = \alpha + \rho Age21_i + \gamma AgeCell_i + u_i$$

Interpret the estimated intercept, coefficient on the treatment variable, and coefficient on the running variable.

- 4) To make the estimated intercept easier to interpret, center the running variable agecell by subtracting the cutoff, 21. In other words, generate a new variable age, which is equivalent to $agecell - 21$. Run the “simple” RD regression

$$deathrate_i = \alpha + \rho Age21_i + \gamma Age_i + u_i$$

report the estimate of α and explain its meaning.

- 5) Superimpose fitted values from the regression you just run on the scatter plot you created.
- 6) Run a RD regression of death rates from all causes on the treatment variable, controlling for *age* (the centered running variable), age^2 , age interacted with *over21*, and age^2 interacted with *over21*. Note that in order to run the regression, you need to generate three more variables: *age2*, *age_{over21}*, and *age2_{over21}*.
- 7) Use the fitted values from the *quadratic* regression you run.
- 8) Rerun the equation(1) and equation(2) RD regression, but includes only ages *20-21* instead of *19-22*. Explain the coefficients on the treatment variable. What happens to the robust standard errors compared to the regressions with a bigger sample?

3.2 Fixed Effects Model

- S.W. Textbook Chapter 10. Empirical Exercises E10.2

3.3 DID: Minimum Wages and Employment

- Card, D. and A.B. Krueger (1994). Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania. American Economic Review, Vol. 84, No. 4, pp. 772-793.
- On April 1, 1992, the minimum wage in New Jersey was raised from 4.25 dollar per hour to 5.05 dollar per hour. In the neighboring state of Pennsylvania, however, the minimum wage remained constant at \$4.25. David Card and Alan Krueger (1994) analyze the impact of the minimum wage increase on employment in the fast-food industry, since this is a sector which employs many low-wage workers.

- 1) Show the distribution of wages in both states, New Jersey and Pennsylvania, before and after the reform. Is the minimum wage increase visible in the data?(Hint: use kernel density function to estimate the distribution of wage and draw a vertical line at the minimum wage)
- 2) Write down the regression equation and calculate the difference-in-differences (DID) estimator. Under which identification assumption does the DID estimator give you the causal effect of interest?
- 3) Now, instead of a dummy treatment variable, generate a variable of **treatment intensity, Gap**. This variable measures how much a store is affected by the minimum wage increase. This variable is defined as:

$$Gap = \begin{cases} 0 & \text{if } treated = 0 \\ 0 & \text{if } treated = 0, wage_0 \geq 5.5 \\ \frac{(5.5 - wage_0)}{wage_0} & \text{if } treated = 1, wage_0 < 5.5 \end{cases}$$

- 4) What is the effect of the minimum wage in this case? Should you also control for the state in the regression? Why? Why not?
- 5) Do restaurants that were imposed a wage increase raise their prices for fast-food?(Hint: You could use the same methods as in questions 2)