Lab #20 - Panel Data (MLDA Revisited)

Econ 224

November 15th, 2018

Introduction

This lab revisits the MLDA example using panel data methods (state and year effects) rather than regression discontinuity. Before beginning, please download the file `deaths.dta` from the Mastering 'Metrics website under "Killer Apps > Chapter 5." Here is a description of the variables that you will need for this exercise:

Name	Description
`state`	Indicator for US States and DC
`year`	Year
`pop`	Population in state s in year t
`legal`	Prop. of 18-20 year olds who can legally drink in state s in year t
`agegr`	Indicator for age ranges (2 = 18-20 year olds)
`mrate`	Mortality rate in state s in year t
`dtype`	Indicator for *which* mortality category `mrate` contains (1 = all deaths)
`beertaxa`	Measure of per-unit beer taxes in state s in year t

Exercises

1. Preliminaries:

   (a) Use an appropriate package to open `deaths.dta` in R.
   (b) Convert `year` to factor using `as.factor`
   (c) Use `as.factor` to create a new variable called `year_factor` containing the same information as `year` but stored as a factor.

   (d) Restrict the sample to years before 1984, 18-20 year olds, and "all deaths" mortality rates.

2. (a) Use `lm_robust` to estimate the effect of `legal` on `mrate` including state and year effects. Use cluster robust standard errors by setting `clusters = state` and `se_type = 'stata'`.
   (b) Repeat (a), but run a *weighted* regression by setting `weights = pop`.
   (c) Repeat (b) but allow for *state-specific* effects by including an interaction between `state` and `year`. Why is this different from including an interaction between `state` and `year_factor`?
   (d) Come up with an appropriate way to display *only* the coefficient estimates and standard errors for `legal`, and not all the estimates of state and year effects. Discuss your findings.

3. Repeat 2, but control for beer taxes. Discuss your findings.

Solutions

```r
# 1- Preliminaries
library(tidyverse)
library(haven)
library(estimatr)

mlda <- read_dta('~/econ224/labs/deaths.dta')
mlda <- mlda %>%
  filter(year <= 1983, agegr == 2, dtype == 1) %>%
  mutate(year_factor = factor(year), state = factor(state))

# 2
reg1 <- lm_robust(mrate ~ legal + state + year_factor - 1,
                  data = mlda, clusters = state, se_type = 'stata')
reg2 <- lm_robust(mrate ~ legal + state + year_factor + state:year - 1,
                  data = mlda, clusters = state, se_type = 'stata')
reg3 <- lm_robust(mrate ~ legal + state + year - 1,
                  data = mlda, weights =  pop, clusters = state, se_type = 'stata')

# 3
reg4 <- lm_robust(mrate ~ legal + beertaxa + state + year_factor - 1,
                  data = mlda, clusters = state, se_type = 'stata')
reg5 <- lm_robust(mrate ~ legal + beertaxa + state + year_factor + state:year - 1,
                  data = mlda, clusters = state, se_type = 'stata')
reg6 <- lm_robust(mrate ~ legal + beertaxa + state + year - 1,
                  data = mlda, weights =  pop, clusters = state, se_type = 'stata')

# Results
estimates <- c(coef(reg1)[1], coef(reg2)[1], coef(reg3)[1],
               coef(reg4)[1], coef(reg5)[1], coef(reg6)[1])
std_errors <- c(reg1$std.error[1], reg2$std.error[1], reg3$std.error[1],
                reg4$std.error[1], reg5$std.error[1], reg6$std.error[1])

results <- cbind(estimates, std_errors)
row.names(results) <- paste0('reg', 1:6)
results
```

```
     estimates std_errors
reg1 10.804141   4.592205
reg2  8.466624   5.097812
reg3 12.000347   3.346856
reg4 10.982723   4.691735
reg5 10.029325   4.915832
reg6 12.292449   3.283094
```