Chapter 5 of MHE: Fixed Effects, Diff-in-Diff and Panel Data

05 April, 2018

### (MHE) Section 5.2.1 - Regression Differnce in Difference

We want to reproduce Figure 5.2.4 from MHE. This figure was originally produced by [Autor (2003)](http://economics.mit.edu/files/589) (Figure 3 in the original) to show the estimated impact of a law (implied-contract exceptions to the employment-at-will doctrine) over an outcome (the use of temporary workers).

## Download and load the data

First download the data from the author’s website (will migrate to external data repo in the future!), and load it in to R.

# Download the data and unzip it  
download.file('http://economics.mit.edu/~dautor/outsourcingatwill\_table7.zip',   
 'outsourcingatwill\_table7.zip')  
unzip('outsourcingatwill\_table7.zip')

# Load the data  
autor <- read.dta('table7/autor-jole-2003.dta')

The data downloaded data set contains 24 years (from 74 to 97) and 50 states, for a total of 1200 observations.

## “Clean” the data

Restrict sample to years between 1979 and 1995 (inclusive), and deleting Guam from the sample (state = 98).

# Restrict sample  
autor <- autor[which(autor$year >= 79 & autor$year <= 95), ]  
autor <- autor[which(autor$state != 98), ]

The clean data now contains 17 years (from 79 to 95) and 50 states, for a total of 850 observations.

## “Build” analytic file

In addition to the variables already defined in the data, we need to construct the following variables:  
- Log of total employment  
- Normalize the year variable to 1978

# Log total employment: from BLS employment & earnings  
autor$lnemp <- log(autor$annemp)  
  
# Normalize year to 1978  
autor$t <- autor$year - 78  
  
# Create state and year factors (required format for fe reg package)  
autor$state <- factor(autor$state)  
autor$year <- factor(autor$year)

## Define model to estimate

We want to estimate a fixed effect model with lead and lag treatment variables:

Where:  
- is the (log) number of jobs under the catergory of Temporary Help Services for each state.  
- are the state fixed effect, are the time fixed effec, and are state specific time trends.  
- contains: log state nonfarm employment, and leads and lags of adoption of the public policy and good faith exceptions.  
- Treatment variables in contemporary () leads and lags

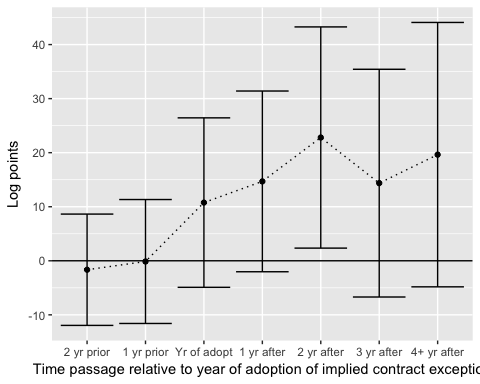
The estimation of this model is presented in column 2 of table 7 of the original paper.

# Diff-in-diff regression  
did <- felm(lnths ~ lnemp + admico\_2 + admico\_1 + admico0 + admico1 + admico2 +   
 admico3 + mico4 + admppa\_2 + admppa\_1 + admppa0 + admppa1 +  
 admppa2 + admppa3 + mppa4 + admgfa\_2 + admgfa\_1 + admgfa0 +  
 admgfa1 + admgfa2 + admgfa3 + mgfa4  
 | state + year + state:t | 0 | state, data = autor)

## Vizualize the results

And this estimates are then used to create figure 3 of the original paper, which is figure 5.2.4 in MHE.

# Plot results  
lags\_leads <- c("admico\_2", "admico\_1", "admico0",  
 "admico1" , "admico2" , "admico3",  
 "mico4")  
labels <- c("2 yr prior", "1 yr prior", "Yr of adopt",  
 "1 yr after", "2 yr after", "3 yr after",  
 "4+ yr after")  
results.did <- data.frame(label = factor(labels, levels = labels),  
 coef = summary(did)$coef[lags\_leads, "Estimate"] \* 100,  
 se = summary(did)$coef[lags\_leads, "Cluster s.e."] \* 100)  
g <- ggplot(results.did, aes(label, coef, group = 1))  
p <- g + geom\_point() +  
 geom\_line(linetype = "dotted") +   
 geom\_errorbar(aes(ymax = coef + 1.96 \* se,  
 ymin = coef - 1.96 \* se)) +  
 geom\_hline(yintercept = 0) +  
 ylab("Log points") +  
 xlab(paste("Time passage relative to year of",  
 "adoption of implied contract exception"))  
print(p)



ggsave(p, file = "Figure 5-2-4-R.png", height = 6, width = 8, dpi = 300)  
  
# End of script