Lecture 19 EEP118

General Multi Year Panel Data

Assumptions for Fixed Effect (FE) Model

Policy analysis with Panel Data

End chapter 13. Read 14.1

See R code on bcourses and also notebook in datahub for this lecture

Daily Assignment 19 posted as well as solutions

Pset 4 posted see due date in bcourses Pset 5 posted soon

Last class, with a panel of 46 cities and two years of data, 82 and 87, we estimated the following model $crimerate\_it = \alpha\_i + \beta\_1\ unem\_it + \beta\_2\ d87 + v\_it$

β_1 = controlling for all the characteristics of the cities that do not vary over time (α_i fixed effects) and for the general effect of time common to all cities (d87) a one percent point increase in unemployment induces a β_1 increase in crime rate (number crimes per 100 people)

we used Lecture18_CRIME2 .dta

and estimated a model

reg6n <- lm(crmrte ~ unem + d87+ +offarea+lawexpc+pcinc+factor(city) , data = mydata) summary(reg6n)

How did we interpret beta hat for unemployment?

Controlling for all constant characteristics of cities (that do not vary over time) =city fixed effects, and for effect of time (dummy for year 87) that captures changes in 1987 relative to 1982 factors that affect crime rate that are common to all cities, and controlling for officer per area, law espenditures, per capita income, a 1 percentage point increase in unemployment rate induces a significant 2.932 per thousand increase in crime rate

Last lecture we had two years of data

Generalizing to many years is mechanically very easy and will open the door to many new analyses

Illustration of multi-year panel data analysis:

We will measure the impact of Enterprise Zones (EZ) on employment

Data: 22 cities in Indiana, from 1980 to 1988 22 cities and many years of data

Multi-year panel data: Impact of enterprise zones on employment Source: data file EZUNEM (Wooldridge). 22 cities in Indiana, from 1980 to 1988

Six city enterprise zones created in 1984, and 4 more city EZ created in 1985.

In [1]:

```
# Load the 'pacman' package
library(pacman)
#packages to use load them now using the pacman "manager"
p_load(dplyr, haven, readr)
#Another great feature of p_load(): if you try to load a package that is not
p_load(ggplot2)


pacman::p_load(lfe, lmtest, haven, sandwich, tidyverse)
# lfe for running fixed effects regression
# lmtest for displaying robust SE in output table
# haven for loading in dta files
# sandwich for producing robust Var-Cov matrix
# tidyverse for manipulating data and producing plots


#change into Lecture 19 directory
#setwd("/Users/sofiavillas-boas/Dropbox/EEP118_Spring2023/Lectures/Lecture19


pacman::p_load(lfe, haven, tidyverse)
```

In [2]:

```
#read in a Stata dataset DATA LECTURE 19
mydata <- read_dta("Lecture19_ezunem.dta")
head(mydata)

#when did city 1 get an EZ?
```

A tibb

year	uclms	ez	d81	d82	d83	d84	d85	d86	d87	⋯
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	⋯
1980	166746	0	0	0	0	0	0	0	0	⋯
1981	83561	0	1	0	0	0	0	0	0	⋯
1982	158146	0	0	1	0	0	0	0	0	⋯
1983	83572	0	0	0	1	0	0	0	0	⋯
1984	45949	1	0	0	0	1	0	0	0	⋯
1985	48848	1	0	0	0	0	1	0	0	⋯

City 1 got EZ=1 in 1984 onwards, so in 1984

In [3]:

```
# summarize data
summary(mydata)
```

```
      year          uclms              ez              d81        
 Min.   :1980   Min.   : 12360   Min.   :0.0000   Min.   :0.0000  
 1st Qu.:1982   1st Qu.: 43922   1st Qu.:0.0000   1st Qu.:0.0000  
 Median :1984   Median : 69170   Median :0.0000   Median :0.0000  
 Mean   :1984   Mean   : 95383   Mean   :0.2323   Mean   :0.1111  
 3rd Qu.:1986   3rd Qu.:114443   3rd Qu.:0.0000   3rd Qu.:0.0000  
 Max.   :1988   Max.   :667208   Max.   :1.0000   Max.   :1.0000  

      d82              d83              d84              d85        
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.0000  
 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000  
 Median :0.0000   Median :0.0000   Median :0.0000   Median :0.0000  
 Mean   :0.1111   Mean   :0.1111   Mean   :0.1111   Mean   :0.1111  
 3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:0.0000  
 Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.0000  

      d86              d87              d88               c1         
 Min.   :0.0000   Min.   :0.0000   Min.   :0.0000   Min.   :0.00000  
 1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.0000   1st Qu.:0.00000  
 Median :0.0000   Median :0.0000   Median :0.0000   Median :0.00000  
 Mean   :0.1111   Mean   :0.1111   Mean   :0.1111   Mean   :0.04545  
 3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:0.0000   3rd Qu.:0.00000  
 Max.   :1.0000   Max.   :1.0000   Max.   :1.0000   Max.   :1.00000  

       c2                c3                c4                c5         
 Min.   :0.00000   Min.   :0.00000   Min.   :0.00000   Min.   :0.00000  
 1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.00000  
 Median :0.00000   Median :0.00000   Median :0.00000   Median :0.00000  
 Mean   :0.04545   Mean   :0.04545   Mean   :0.04545   Mean   :0.04545  
 3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.00000  
 Max.   :1.00000   Max.   :1.00000   Max.   :1.00000   Max.   :1.00000  

       c6                c7                c8                c9         
 Min.   :0.00000   Min.   :0.00000   Min.   :0.00000   Min.   :0.00000  
 1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.00000  
 Median :0.00000   Median :0.00000   Median :0.00000   Median :0.00000  
 Mean   :0.04545   Mean   :0.04545   Mean   :0.04545   Mean   :0.04545  
 3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.00000  
 Max.   :1.00000   Max.   :1.00000   Max.   :1.00000   Max.   :1.00000  

      c10               c11               c12               c13         
 Min.   :0.00000   Min.   :0.00000   Min.   :0.00000   Min.   :0.00000  
 1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.00000  
 Median :0.00000   Median :0.00000   Median :0.00000   Median :0.00000  
 Mean   :0.04545   Mean   :0.04545   Mean   :0.04545   Mean   :0.04545  
 3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.00000  
 Max.   :1.00000   Max.   :1.00000   Max.   :1.00000   Max.   :1.00000  

      c14               c15               c16               c17         
 Min.   :0.00000   Min.   :0.00000   Min.   :0.00000   Min.   :0.00000  
 1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.00000  
 Median :0.00000   Median :0.00000   Median :0.00000   Median :0.00000  
 Mean   :0.04545   Mean   :0.04545   Mean   :0.04545   Mean   :0.04545  
 3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.00000  
 Max.   :1.00000   Max.   :1.00000   Max.   :1.00000   Max.   :1.00000  
```

```
     c18                c19                c20                c21
Min.   :0.00000   Min.   :0.00000   Min.   :0.00000   Min.   :0.00000
1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.00000   1st Qu.:0.00000
Median :0.00000   Median :0.00000   Median :0.00000   Median :0.00000
Mean   :0.04545   Mean   :0.04545   Mean   :0.04545   Mean   :0.04545
3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.00000   3rd Qu.:0.00000
Max.   :1.00000   Max.   :1.00000   Max.   :1.00000   Max.   :1.00000

     c22              luclms            guclms               cez
Min.   :0.00000   Min.   : 9.422   Min.   :-0.84730   Min.   :0.00000
1st Qu.:0.00000   1st Qu.:10.690   1st Qu.:-0.38671   1st Qu.:0.00000
Median :0.00000   Median :11.144   Median :-0.21562   Median :0.00000
Mean   :0.04545   Mean   :11.191   Mean   :-0.15939   Mean   :0.05682
3rd Qu.:0.00000   3rd Qu.:11.648   3rd Qu.: 0.02655   3rd Qu.:0.00000
Max.   :1.00000   Max.   :13.411   Max.   : 0.79429   Max.   :1.00000
                                   NA's   :22         NA's   :22
     city
Min.   : 1.0
1st Qu.: 6.0
Median :11.5
Mean   :11.5
3rd Qu.:17.0
Max.   :22.0
```

We want to estimate the Impact of establishing Enterprise Zones (EZ) on Unemployment

The variable $uclms\_it$ = Unemployment claims at city i and year t

The variable EZ_it=0 if no EZ and=1 if there is an EZ in city i at time t

uclms is number of unemployment claims file during the year in a city

year 1980 to 1988

ez =1 if have enterprise zone, =0 o.w.

city a city identifier from 1 to 22

What is the number of observations? 9 years times 22 cities = 198

What is the unit of observation? -> a city in a year

What percent of observations have EZ=1? ---> .2323232

the average of EZ is .2323232

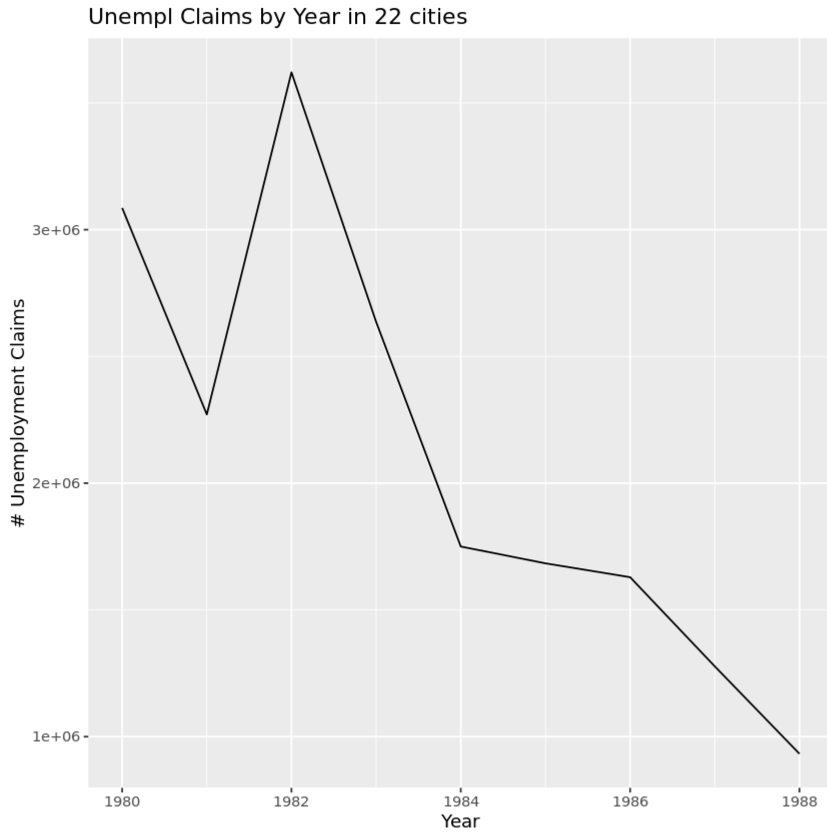
Create a graph to let us look at evolution of Total unemployment over time

In [4]:
```
# first generate sum of claims by year
#The command below creates a column in the original data with sum (over all

lec19df <- mydata %>%
  group_by(year) %>%
  mutate(annual_claims = sum(uclms)) %>%
  ungroup
```

make graph of total (sum over all cities) umemployment claims by year

In [5]:
```
# plot
ggplot(lec19df, aes(x = year, y = annual_claims)) +
  geom_line() +
  labs(title = "Unempl Claims by Year in 22 cities",
       x = "Year",
       y = "# Unemployment Claims")
```

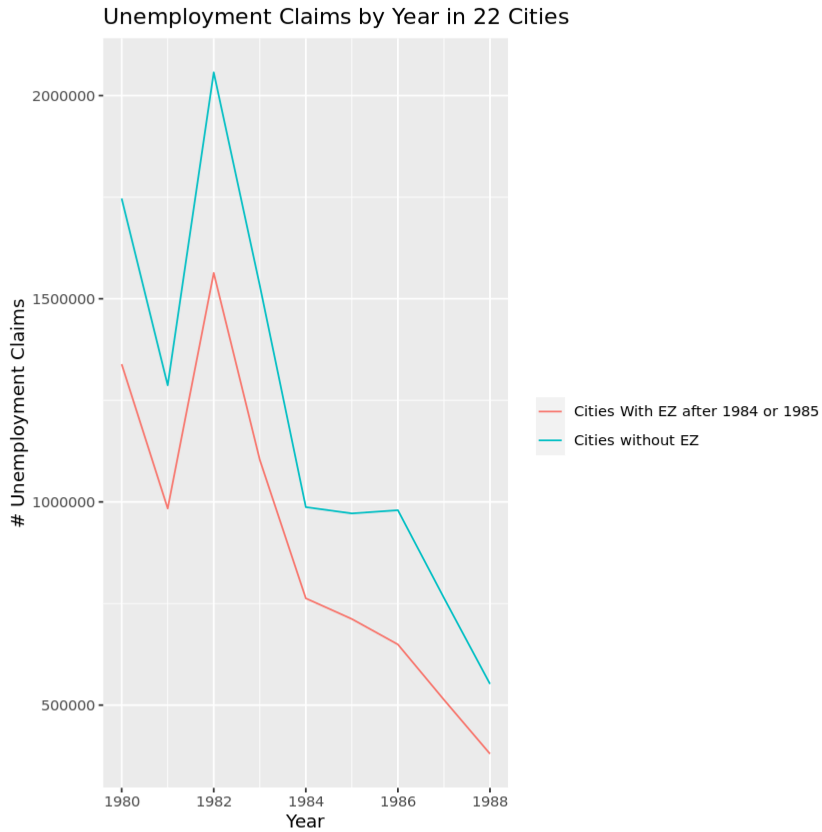


Graph Total unemployment claims (Sum of for EZ cities) and then (Sum for non EZ cities) over time

In [6]:
```
# variables for graph for EZ and non EZ cities
lec19df <- mydata %>%
  group_by(city) %>%
  mutate(cityez = ifelse(max(ez)==1, "EZ", "Non-EZ")) %>%
  group_by(year, cityez) %>%
```

```
  mutate(annual_claims_ez = sum(uclms)) %>%
  ungroup()
```

In [7]:
```
# plot for EZ and non EZ cities
ggplot(lec19df, aes(x = year, y = annual_claims_ez, group = cityez, color =
  geom_line() +
  scale_color_discrete(name="",
                       breaks=c("EZ", "Non-EZ"),
                       labels=c("Cities With EZ after 1984 or 1985", "Cities
  labs(title = "Unemployment Claims by Year in 22 Cities",
       x = "Year",
       y = "# Unemployment Claims")
```



What do you see?

1. A drop in 1982

2. B. EZ cities always had and have lower unemployment
3. The difference between red line (cities without EZ) and blue line before 1984 had nothing to do with EZ. Since EZ only happens after 1984 for some cities

4. The switches to EZ are in 1984 and in 1985

5. The path of the blue line cannot be fully attributed to EZ

The empirical question is :

Does the difference in the two lines increase or decrease when cities adopt EZs?

Regression

Lets estimate a regression model to investigate whether EZ caused an impact on unemployment

Lets consider the relationship between unemployment and EZ over time and across the 22 cities in the data using T years (T=9 here)

Model- city j year t

ln$(uclms)_jt$= β_1 EZ_jt + {aj } + {dt } + v_jt

Unemployment marginal effect: (∂ $\boldsymbol{ln(uclms)}$)/($\boldsymbol{EZ\ 0\ to\ 1}$) = β_1 percent change in unemployment claims, controlling for constant factors by city and factors common to all cities by year, with city fixed effects, and also year fixed effects, respectively.

aj captures all the characteristics of city j that are constant over time and affect unemployment -> that is, it controls for the overall level of unemployment in a city j, a city fixed effect.

dt captures overall time patterns, year by year, that is the factors changing over time, that change year by year that affect unemployment and that are common to all cities, a year fixed effect.

Short writing for one dummy for ALL BUT ONE city: a2, a3, …,a22.

And one dummy for all but ONE year (d81, d82, …dT) exclude 1980

For city j in 1980 to 1983 no EZ (EZ=0 for all j and t<1984)

ln$(uclms)_j$80=β_0+aj +v_j80 in 1980

---> constant β_0 is 1980 (when all other dyears=0

ln$(uclms)_j$81= aj +d81 +v_j81 in 1981

For city j at time t, notation,

ln$(uclms)jt$= β_0+ β_1 EZ_jt +aj +dt +v_jt

In [8]:
```
#regression in levels
reg1 <- lm(uclms ~ ez, data = lec19df)
summary(reg1)
```

```
Call:
lm(formula = uclms ~ ez, data = lec19df)

Residuals:
   Min     1Q Median     3Q    Max 
-94003 -48961 -20850  24958 560845 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   106364       7066  15.052  < 2e-16 ***
ez            -47262      14661  -3.224  0.00148 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 87120 on 196 degrees of freedom
Multiple R-squared:  0.05035,	Adjusted R-squared:  0.04551 
F-statistic: 10.39 on 1 and 196 DF,  p-value: 0.001482
```

In [9]:
```
#regression in log unemployment
reg0 <- lm(luclms ~ ez, data = lec19df)
summary(reg0)
```

```
Call:
lm(formula = luclms ~ ez, data = lec19df)

Residuals:
     Min       1Q   Median       3Q      Max 
-1.87835 -0.49217 -0.01882  0.48605  2.11029 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 11.30057    0.05575 202.685  < 2e-16 ***
ez          -0.47257    0.11567  -4.085 6.41e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6874 on 196 degrees of freedom
Multiple R-squared:  0.07847,	Adjusted R-squared:  0.07377 
F-statistic: 16.69 on 1 and 196 DF,  p-value: 6.411e-05
```

We see that EZ from 0 to 1 reduced ln of log unemployment claims by -0.47

Significantly. Or reduced unclms by 47%. Significant at 1% level. p value is p-value: 6.411e-05

Fixed effects regression

And one dummy for all but ONE year (d81, d82, ...dT) exclude 1980

One dummy for ALL BUT ONE city: a2, a3, ...,a22. exclude city1

In [10]:
```
#reg with fixed effects and log of unemployment claims as Y variable
reg <- lm(luclms ~ ez + factor(year) + factor(city), data = lec19df)
```

```
summary(reg)
```

```
Call:
lm(formula = luclms ~ ez + factor(year) + factor(city), data = lec19df)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.57618 -0.10837 -0.00977  0.11364  0.49623 

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)    
(Intercept)      11.67615    0.08008 145.807  < 2e-16 ***
ez               -0.10441    0.05542  -1.884 0.061291 .  
factor(year)1981 -0.32163    0.06046  -5.320 3.30e-07 ***
factor(year)1982  0.13550    0.06046   2.241 0.026332 *  
factor(year)1983 -0.21926    0.06046  -3.627 0.000381 ***
factor(year)1984 -0.57915    0.06232  -9.293  < 2e-16 ***
factor(year)1985 -0.59179    0.06550  -9.036 3.92e-16 ***
factor(year)1986 -0.62126    0.06550  -9.486  < 2e-16 ***
factor(year)1987 -0.88895    0.06550 -13.573  < 2e-16 ***
factor(year)1988 -1.22763    0.06550 -18.744  < 2e-16 ***
factor(city)2    -0.19349    0.09941  -1.946 0.053292 .  
factor(city)3    -0.37894    0.09941  -3.812 0.000194 ***
factor(city)4    -0.54118    0.09941  -5.444 1.83e-07 ***
factor(city)5     0.01103    0.09472   0.116 0.907423    
factor(city)6     0.55458    0.09452   5.867 2.32e-08 ***
factor(city)7     0.75006    0.09452   7.935 2.90e-13 ***
factor(city)8    -0.05876    0.09472  -0.620 0.535888    
factor(city)9     0.35342    0.09472   3.731 0.000261 ***
factor(city)10    1.64501    0.09941  16.548  < 2e-16 ***
factor(city)11   -0.13032    0.09941  -1.311 0.191694    
factor(city)12   -0.03498    0.09941  -0.352 0.725377    
factor(city)13   -0.83258    0.09941  -8.375 2.15e-14 ***
factor(city)14   -0.87363    0.09472  -9.223  < 2e-16 ***
factor(city)15   -0.23542    0.09941  -2.368 0.019020 *  
factor(city)16    0.43574    0.09941   4.383 2.06e-05 ***
factor(city)17   -0.44523    0.09452  -4.710 5.18e-06 ***
factor(city)18   -0.04289    0.09941  -0.431 0.666677    
factor(city)19    0.09341    0.09941   0.940 0.348780    
factor(city)20   -0.35098    0.09452  -3.713 0.000279 ***
factor(city)21    0.45779    0.09452   4.843 2.90e-06 ***
factor(city)22    0.21864    0.09941   2.199 0.029229 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2005 on 167 degrees of freedom
Multiple R-squared:  0.9332,	Adjusted R-squared:  0.9212 
F-statistic: 77.75 on 30 and 167 DF,  p-value: < 2.2e-16
```

EZ from 0 to 1 reduced log of unemployment claims by -0.10

Significantly. Or reduced unclms by 10%. Significant at 6% level.

what does the intercept estimate mean?

Ln uclms hat for city=1 and 1980

dummy for all but ONE year to take into account annual differences- exclude 1980

One dummy for ALL BUT ONE city exclude city1, to take into account city constant differences

If we want to control for FE but not necessarily have all the estimates printed out as output, an equivalent command is in R:

In [11]:
```
#xtreg equivalent using felm
reg2 <- felm(luclms ~ ez | year + city, data = lec19df)
summary(reg2)
```

```
Call:
   felm(formula = luclms ~ ez | year + city, data = lec19df)

Residuals:
     Min       1Q   Median       3Q      Max 
-0.57618 -0.10837 -0.00977  0.11364  0.49623 

Coefficients:
   Estimate Std. Error t value Pr(>|t|)  
ez -0.10441    0.05542  -1.884   0.0613 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2005 on 167 degrees of freedom
Multiple R-squared(full model): 0.9332   Adjusted R-squared: 0.9212 
Multiple R-squared(proj model): 0.02081   Adjusted R-squared: -0.1551 
F-statistic(full model):77.75 on 30 and 167 DF, p-value: < 2.2e-16 
F-statistic(proj model):  3.55 on 1 and 167 DF, p-value: 0.06129 
```

Make a Table

In [13]:
```
#make table
#library(stargazer)
#stargazer(list(reg2), type="text",keep.stat=c("n","rsq"))

#the above work in R Studio
```

Panel Data Estimates of the Effects of Death Penalty Laws, and Executions on Murder Rates 1960-2000

Another data set now:

TODAY, as an example, “Panel Data Estimates of the Effects of Death Penalty Laws on Murder Rates 1960-2000”

Using data from FBI that I collected in March 22, 2017

To try to perform a Replication of Donohue and Wolfers (2006) analysis

In [14]:
```
#open data
mydata2 <- read_dta("dataLecture19murder.dta")
head(mydata2)
# summarize data
summary(mydata2)

#with variables to use
#by number of executions

#ex10=1 if ex10>0

#active=legal*ex10

#passive=legal*(ex10==0)
```

year	state	popul	pc_mur	pc_rob	pc_assa	pc_burg	pc_larc	pc_auto	r_in
<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<dbl>	<db
1930	NA	NA	NA	NA	NA	NA	NA	NA	N
1931	NA	NA	NA	NA	NA	NA	NA	NA	N
1932	NA	NA	NA	NA	NA	NA	NA	NA	N
1933	2	NA	NA	NA	NA	NA	NA	NA	N
1934	2	NA	NA	NA	NA	NA	NA	NA	N
1935	2	NA	NA	NA	NA	NA	NA	NA	N

```
      year          state          popul             pc_mur      
 Min.   :1930   Min.   : 1    Min.   :   96.0   Min.   : 0.000  
 1st Qu.:1948   1st Qu.:13    1st Qu.:  983.5   1st Qu.: 2.900  
 Median :1967   Median :26    Median : 2657.0   Median : 5.400  
 Mean   :1967   Mean   :26    Mean   : 3940.5   Mean   : 6.958  
 3rd Qu.:1986   3rd Qu.:39    3rd Qu.: 4757.5   3rd Qu.: 9.400  
 Max.   :2004   Max.   :51    Max.   :35484.4   Max.   :80.600  
                NA's   :153   NA's   :250       NA's   :322     
     pc_rob          pc_assa          pc_burg          pc_larc      
 Min.   :   0.4   Min.   :   0.5   Min.   :  67.8   Min.   :  59.9  
 1st Qu.:  30.3   1st Qu.:  44.8   1st Qu.: 387.3   1st Qu.: 553.9  
 Median :  67.5   Median : 119.8   Median : 642.5   Median :1151.2  
 Mean   : 101.9   Mean   : 170.3   Mean   : 764.5   Mean   :1499.7  
 3rd Qu.: 126.2   3rd Qu.: 246.9   3rd Qu.:1059.9   3rd Qu.:2394.3  
 Max.   :1635.1   Max.   :1557.6   Max.   :2906.7   Max.   :5372.7  
 NA's   :284      NA's   :296      NA's   :308      NA's   :994     
    pc_auto           r_infd          prisoner         r_execut      
 Min.   :  32.2   Min.   : 382.1   Min.   :   129   Min.   : 0.0000  
 1st Qu.: 163.6   1st Qu.: 940.3   1st Qu.:  1262   1st Qu.: 0.0000  
 Median : 251.4   Median :1510.0   Median :  2952   Median : 0.0000  
 Mean   : 311.8   Mean   :1688.6   Mean   :  7858   Mean   : 0.3258  
 3rd Qu.: 403.5   3rd Qu.:2360.0   3rd Qu.:  7922   3rd Qu.: 0.0742  
 Max.   :1839.9   Max.   :5658.7   Max.   :164933   Max.   :12.4046  
 NA's   :289      NA's   :1136     NA's   :265      NA's   :365      
    r_death          cri_prop          cri_viol           pc_pris       
 Min.   : 0.000   Min.   : 124.2   Min.   :   3.21   Min.   :  20.32  
 1st Qu.: 2.031   1st Qu.: 565.5   1st Qu.:  93.80   1st Qu.:  78.34  
 Median : 3.078   Median : 925.3   Median : 203.80   Median : 112.59  
 Mean   : 3.749   Mean   :1078.1   Mean   : 281.68   Mean   : 161.00  
 3rd Qu.: 4.870   3rd Qu.:1472.0   3rd Qu.: 388.10   3rd Qu.: 184.22  
 Max.   :32.527   Max.   :4342.3   Max.   :2865.70   Max.   :1785.01  
 NA's   :606      NA's   :311      NA's   :328       NA's   :327      
    pcpris_1          c_pris_1          region         r_black      
 Min.   :  20.32   Min.   :0.0104   Min.   :1.000   Min.   :0.0000  
 1st Qu.:  78.34   1st Qu.:0.0740   1st Qu.:2.000   1st Qu.:0.0112  
 Median : 112.59   Median :0.1185   Median :3.000   Median :0.0506  
 Mean   : 161.00   Mean   :0.1357   Mean   :2.667   Mean   :0.0965  
 3rd Qu.: 184.22   3rd Qu.:0.1799   3rd Qu.:4.000   3rd Qu.:0.1429  
 Max.   :1785.01   Max.   :0.8702   Max.   :4.000   Max.   :0.7107  
 NA's   :327       NA's   :415      NA's   :153     NA's   :386     
    r_urban          r_15_24          r_25_44          r_45_64      
 Min.   :0.1433   Min.   :0.0177   Min.   :0.2042   Min.   :0.1223  
 1st Qu.:0.4942   1st Qu.:0.1510   1st Qu.:0.2589   1st Qu.:0.1812  
 Median :0.6293   Median :0.1679   Median :0.2812   Median :0.1948  
 Mean   :0.6180   Mean   :0.1662   Mean   :0.2813   Mean   :0.1916  
 3rd Qu.:0.7464   3rd Qu.:0.1815   3rd Qu.:0.3032   3rd Qu.:0.2053  
 Max.   :1.0000   Max.   :0.2340   Max.   :0.3898   Max.   :0.2374  
 NA's   :357      NA's   :896      NA's   :386      NA's   :896     
     r_0_14           r_0_24            r_ue           stname         
 Min.   :0.1622   Min.   :0.2716   Min.   : 0.100   Length:3825       
 1st Qu.:0.2393   1st Qu.:0.3889   1st Qu.: 1.900   Class :character  
 Median :0.2725   Median :0.4272   Median : 2.800   Mode  :character  
 Mean   :0.2739   Mean   :0.4280   Mean   : 3.192                     
 3rd Qu.:0.3063   3rd Qu.:0.4635   3rd Qu.: 4.100                     
 Max.   :0.4454   Max.   :0.6441   Max.   :13.900                     
 NA's   :896      NA's   :386      NA's   :1053                       
```

```
    execs           pc_exec           execrate           execrt      
Min.   : 0.000   Min.   :   0.00   Min.   :  0.000   Min.   :0.0000  
1st Qu.: 0.000   1st Qu.:   0.00   1st Qu.:  0.000   1st Qu.:0.0000  
Median : 0.000   Median :   0.00   Median :  0.000   Median :0.0000  
Mean   : 1.151   Mean   :  40.09   Mean   :  6.047   Mean   :0.0605  
3rd Qu.: 1.000   3rd Qu.:  22.38   3rd Qu.:  2.954   3rd Qu.:0.0295  
Max.   :40.000   Max.   :1935.48   Max.   :359.755   Max.   :3.5976  
NA's   :255      NA's   :301       NA's   :377       NA's   :377     
     st              klssample         fips         _merge_age   
Length:3825        Min.   :1      Min.   : 1.00   Min.   :1.000  
Class :character   1st Qu.:1      1st Qu.:16.00   1st Qu.:1.000  
Mode  :character   Median :1      Median :29.00   Median :1.000  
                   Mean   :1      Mean   :28.96   Mean   :1.164  
                   3rd Qu.:1      3rd Qu.:42.00   3rd Qu.:1.000  
                   Max.   :1      Max.   :56.00   Max.   :3.000  
                   NA's   :1981   NA's   :3111    NA's   :408    
  _merge_race    _merge_census2000  _merge_urban     age15to19     
Min.   :1.000   Min.   :1.000     Min.   :1.000   Min.   :0.0565  
1st Qu.:1.000   1st Qu.:1.000     1st Qu.:1.000   1st Qu.:0.0741  
Median :1.000   Median :1.000     Median :1.000   Median :0.0841  
Mean   :1.298   Mean   :1.015     Mean   :1.029   Mean   :0.0836  
3rd Qu.:1.000   3rd Qu.:1.000     3rd Qu.:1.000   3rd Qu.:0.0923  
Max.   :3.000   Max.   :2.000     Max.   :3.000   Max.   :0.1262  
NA's   :408     NA's   :357       NA's   :357     NA's   :411     
   age20to24         nonwhite        _merge_DScensus _merge_CPSunemp
Min.   :0.0513   Min.   :0.00000   Min.   :1.000   Min.   :1.00   
1st Qu.:0.0701   1st Qu.:0.03620   1st Qu.:1.000   1st Qu.:1.00   
Median :0.0783   Median :0.08196   Median :1.000   Median :1.00   
Mean   :0.0786   Mean   :0.13046   Mean   :1.204   Mean   :1.75   
3rd Qu.:0.0863   3rd Qu.:0.17886   3rd Qu.:1.000   3rd Qu.:3.00   
Max.   :0.1241   Max.   :0.75895   Max.   :3.000   Max.   :3.00   
NA's   :411      NA's   :264       NA's   :357     NA's   :153    
_merge_ManpowerUnemp       ur           _merge_ur      population      
Min.   :1.000        Min.   : 1.800   Min.   :1.00   Min.   :  226167  
1st Qu.:1.000        1st Qu.: 4.300   1st Qu.:3.00   1st Qu.: 1099000  
Median :1.000        Median : 5.400   Median :3.00   Median : 3096612  
Mean   :1.556        Mean   : 5.668   Mean   :2.75   Mean   : 4531744  
3rd Qu.:3.000        3rd Qu.: 6.700   3rd Qu.:3.00   3rd Qu.: 5387250  
Max.   :3.000        Max.   :17.400   Max.   :3.00   Max.   :35484453  
NA's   :153          NA's   :1530     NA's   :153    NA's   :1581      
murderandnonnegligentmanslaughte violentcrimerate   murderrate    
Min.   :   1                     Min.   :   9.5   Min.   : 0.200  
1st Qu.:  46                     1st Qu.: 177.2   1st Qu.: 3.300  
Median : 180                     Median : 328.1   Median : 5.850  
Mean   : 348                     Mean   : 396.2   Mean   : 7.016  
3rd Qu.: 443                     3rd Qu.: 539.4   3rd Qu.: 9.300  
Max.   :4096                     Max.   :2921.8   Max.   :80.600  
NA's   :1581                     NA's   :1581     NA's   :1581    
forcibleraperate  robberyrate     aggravatedassaultrate propertycrimerate
Min.   :  0.80   Min.   :   1.9   Min.   :   3.6        Min.   : 573.1   
1st Qu.: 14.60   1st Qu.:  38.8   1st Qu.: 103.0        1st Qu.:2660.5   
Median : 25.75   Median :  96.4   Median : 193.7        Median :3834.2   
Mean   : 27.30   Mean   : 131.5   Mean   : 230.4        Mean   :3791.3   
3rd Qu.: 37.40   3rd Qu.: 164.8   3rd Qu.: 317.9        3rd Qu.:4755.5   
Max.   :102.20   Max.   :1635.1   Max.   :1557.6        Max.   :9512.1   
NA's   :1581     NA's   :1581     NA's   :1581          NA's   :1581     
```

```
  burglaryrate    larcenytheftrate motorvehicletheftrate   _merge_ucr
 Min.   : 182.6   Min.   : 293.3   Min.   :  48.3        Min.   :1.000
 1st Qu.: 625.1   1st Qu.:1715.2   1st Qu.: 198.2        1st Qu.:1.000
 Median : 914.6   Median :2508.9   Median : 328.4        Median :3.000
 Mean   : 978.9   Mean   :2440.7   Mean   : 371.6        Mean   :2.222
 3rd Qu.:1246.8   3rd Qu.:3052.0   3rd Qu.: 481.4        3rd Qu.:3.000
 Max.   :2906.7   Max.   :5833.8   Max.   :1839.9        Max.   :3.000
 NA's   :1581     NA's   :1581     NA's   :1581          NA's   :153
      pop              murders           rpc_inc        _merge_realy
 Min.   :  226167   Min.   :   0.0   Min.   :  9.769   Min.   :3
 1st Qu.: 1099000   1st Qu.:  38.0   1st Qu.: 62.327   1st Qu.:3
 Median : 3096612   Median : 158.0   Median : 95.563   Median :3
 Mean   : 4531744   Mean   : 297.5   Mean   : 98.216   Mean   :3
 3rd Qu.: 5387250   3rd Qu.: 400.0   3rd Qu.:131.880   3rd Qu.:3
 Max.   :35484452   Max.   :4096.0   Max.   :274.235   Max.   :3
 NA's   :1581       NA's   :322      NA's   :187       NA's   :153
  statename          _merge_infantmortality _merge_prisoners     c_pris
 Length:3825        Min.   :1.000          Min.   :1.000    Min.   :0.0105
 Class :character   1st Qu.:1.000          1st Qu.:1.000    1st Qu.:0.0727
 Mode  :character   Median :1.000          Median :1.000    Median :0.1179
                    Mean   :1.667          Mean   :1.778    Mean   :0.1351
                    3rd Qu.:3.000          3rd Qu.:3.000    3rd Qu.:0.1788
                    Max.   :3.000          Max.   :3.000    Max.   :0.8702
                    NA's   :153            NA's   :153      NA's   :415
 _merge_executions  prisondeaths    _merge_prisondeaths   nfpayrolls
 Min.   :1.000     Min.   :  0.00   Min.   :1.000       Min.   :   34.8
 1st Qu.:3.000     1st Qu.:  5.00   1st Qu.:1.000       1st Qu.:  368.0
 Median :3.000     Median : 15.00   Median :1.000       Median :  833.5
 Mean   :2.639     Mean   : 35.94   Mean   :1.694       Mean   : 1537.8
 3rd Qu.:3.000     3rd Qu.: 45.00   3rd Qu.:3.000       3rd Qu.: 1905.3
 Max.   :3.000     Max.   :454.00   Max.   :3.000       Max.   :14601.9
 NA's   :153       NA's   :2554     NA's   :153         NA's   :535
  lnnfpayrolls   trendpayrollslambda10     empgap        _merge_nfpayrolls
 Min.   :3.550   Min.   :3.591         Min.   :-0.2028   Min.   :1.000
 1st Qu.:5.908   1st Qu.:5.914         1st Qu.:-0.0117   1st Qu.:3.000
 Median :6.726   Median :6.721         Median :-0.0003   Median :3.000
 Mean   :6.731   Mean   :6.731         Mean   : 0.0000   Mean   :2.792
 3rd Qu.:7.552   3rd Qu.:7.552         3rd Qu.: 0.0118   3rd Qu.:3.000
 Max.   :9.589   Max.   :9.593         Max.   : 0.0805   Max.   :3.000
 NA's   :535     NA's   :535           NA's   :535       NA's   :153
      demp            age1519           age2024        _merge_dsdemographics
 Min.   :-0.1237   Min.   :0.05651   Min.   :0.05128   Min.   :1.000
 1st Qu.: 0.0068   1st Qu.:0.07447   1st Qu.:0.07037   1st Qu.:3.000
 Median : 0.0240   Median :0.08451   Median :0.07879   Median :3.000
 Mean   : 0.0254   Mean   :0.08401   Mean   :0.07903   Mean   :2.853
 3rd Qu.: 0.0413   3rd Qu.:0.09263   3rd Qu.:0.08675   3rd Qu.:3.000
 Max.   : 0.3995   Max.   :0.13111   Max.   :0.12411   Max.   :3.000
 NA's   :586       NA's   :264       NA's   :264
     police      _merge_statepolice    ipolice          legal
 Min.   :   71   Min.   :1.000      Min.   :   71   Min.   :0.000
 1st Qu.:  474   1st Qu.:1.000      1st Qu.:  446   1st Qu.:1.000
 Median :  946   Median :3.000      Median :  884   Median :1.000
 Mean   : 1442   Mean   :2.093      Mean   : 1378   Mean   :0.755
 3rd Qu.: 1723   3rd Qu.:3.000      3rd Qu.: 1659   3rd Qu.:1.000
 Max.   :10626   Max.   :3.000      Max.   :10626   Max.   :1.000
 NA's   :2016                       NA's   :1816
```

```
     decade         statdec          regyear            stid       sample_ds    
 Min.   :193.0   Min.   :  1.0   Min.   :  1.0   Min.   : 1   Min.   :1      
 1st Qu.:194.0   1st Qu.:102.0   1st Qu.: 95.0   1st Qu.:13   1st Qu.:1      
 Median :196.0   Median :204.0   Median :164.0   Median :26   Median :1      
 Mean   :196.3   Mean   :204.3   Mean   :156.5   Mean   :26   Mean   :1      
 3rd Qu.:198.0   3rd Qu.:306.0   3rd Qu.:218.0   3rd Qu.:39   3rd Qu.:1      
 Max.   :200.0   Max.   :408.0   Max.   :288.0   Max.   :51   Max.   :1      
                                 NA's   :153                  NA's   :1816   
   sample_kls     decade_ds         ex10            active      
 Min.   :1      Min.   :1930   Min.   :0.0000   Min.   :0.0000  
 1st Qu.:1      1st Qu.:1940   1st Qu.:0.0000   1st Qu.:0.0000  
 Median :1      Median :1960   Median :1.0000   Median :1.0000  
 Mean   :1      Mean   :1963   Mean   :0.5948   Mean   :0.5464  
 3rd Qu.:1      3rd Qu.:1980   3rd Qu.:1.0000   3rd Qu.:1.0000  
 Max.   :1      Max.   :2000   Max.   :1.0000   Max.   :1.0000  
 NA's   :1917                                                   
    passive      
 Min.   :0.0000  
 1st Qu.:0.0000  
 Median :0.0000  
 Mean   :0.2086  
 3rd Qu.:0.0000  
 Max.   :1.0000  
```

Run regressions and Make table of results

In [15]:

```
# death penalty and murders
REG1<-lm(pc_mur~legal,mydata2)
REG2<-felm(pc_mur~legal| state+year+decade_ds, data=mydata2)
REG3<-felm(pc_mur~legal+rpc_inc+ ur+ ipolice+ nonwhite +age15to19 +age20to24
REG4<-felm(pc_mur~active+ passive| state+year+decade_ds, data=mydata2)
REG5<-felm(pc_mur~active+ passive+rpc_inc+ ur+ ipolice+ nonwhite +age15to19
```

In [19]:

```
#make table
#stargazer(list(REG1,REG2,REG3, REG4, REG5), type="text",keep.stat=c("n","rs
```

```
==============================================================
                            Dependent variable:
             -------------------------------------------------
                                   pc_mur
               OLS                      felm
               (1)       (2)       (3)       (4)       (5)
--------------------------------------------------------------
legal        0.903*** -1.945***  -0.100
             (0.249)   (0.258)   (0.143)

rpc_inc                          0.018***            0.016***
                                  (0.005)             (0.005)

ur                               -0.099***           -0.106***
                                  (0.032)             (0.032)

ipolice                          -0.00001            -0.00000
                                 (0.0001)            (0.0001)

nonwhite                         9.989***            9.410***
                                  (1.835)             (1.827)

age15to19                        62.965***           66.959***
                                 (13.681)            (13.613)

age20to24                        41.779***           44.250***
                                  (8.088)             (8.049)

active                                     -2.186*** -0.527***
                                            (0.289)   (0.165)

passive                                    -1.815***   0.028
                                            (0.268)   (0.144)

Constant     6.286***
             (0.214)

--------------------------------------------------------------
Observations  3,503     3,503     2,009     3,503     2,009
R2            0.004     0.738     0.866     0.738     0.868
==============================================================
Note:                              *p<0.1; **p<0.05; ***p<0.01
```

Active=1 if legal and there are executions, passive =1 if legal but no executions

What do you conclude?

Placebos

In [20]:

```
#placebo on other murders
PREG1<-felm(pc_mur~active+ passive+rpc_inc+ ur+ ipolice+ nonwhite +age15to19
PREG2<-felm(violentcrimerate~active+ passive+rpc_inc+ ur+ ipolice+ nonwhite
PREG3<-felm(pc_rob~active+ passive+rpc_inc+ ur+ ipolice+ nonwhite +age15to19
PREG4<-felm(pc_larc~active+ passive+rpc_inc+ ur+ ipolice+ nonwhite +age15to1
```

```
PREG5<-felm(pc_burg~active+ passive+rpc_inc+ ur+ ipolice+ nonwhite +age15to1
PREG6<-felm(pc_auto~active+ passive+rpc_inc+ ur+ ipolice+ nonwhite +age15to1

#make table
stargazer(list(PREG1,PREG2,PREG3, PREG4, PREG5, PREG6), type="text",keep.sta
```

	Dependent variable:					
	pc_mur	violentcrimerate	pc_rob	pc_larc	pc_burg	pc_auto
	(1)	(2)	(3)	(4)	(5)	(6)
active	-0.527***	18.579*	-1.501	205.816***	-17.976	47.203***
	(0.165)	(10.098)	(4.570)	(39.999)	(18.658)	(10.940)
passive	0.028	11.167	-3.272	99.714***	-2.860	11.022
	(0.144)	(8.854)	(4.007)	(34.761)	(16.358)	(9.592)
rpc_inc	0.016***	0.874***	0.703***	6.990***	0.063	0.907***
	(0.005)	(0.313)	(0.142)	(1.474)	(0.578)	(0.339)
ur	-0.106***	2.944	2.306***	71.184***	39.580***	5.974***
	(0.032)	(1.952)	(0.883)	(7.905)	(3.606)	(2.114)
ipolice	-0.00000	0.059***	0.029***	-0.008	0.031***	0.050***
	(0.0001)	(0.005)	(0.002)	(0.028)	(0.009)	(0.005)
nonwhite	9.410***	-3.595	-93.149*	4,276.654***	-3,558.661***	-133.465
	(1.827)	(111.964)	(50.674)	(784.116)	(206.867)	(121.299)
age15to19	66.959***	2,460.854***	2,116.539***	9,794.005**	3,940.529**	3,422.675***
	(13.613)	(834.453)	(377.662)	(3,945.919)	(1,541.748)	(904.023)
age20to24	44.250***	-1,911.505***	-157.160	-10,137.260***	-552.886	3,065.014***
	(8.049)	(493.377)	(223.295)	(2,339.760)	(911.569)	(534.510)
Observations	2,009	2,009	2,009	1,519	2,009	2,009
R2	0.868	0.882	0.857	0.944	0.87	

```
9        0.808
============================================================================
==================
Note:                                                               *p<0.1; *
*p<0.05; ***p<0.01
```

Even more convincing, show that active legal death penalty has deterrence on murder but not on other crimes (e.g., auto theft)

Upcoming Lectures

Learn methods to approach a policy impact analysis like the soda tax

https://news.berkeley.edu/2019/04/02/berkeleys-soda-tax-election-changed-drinking-habits-months-before-prices-went-up/?fbclid=IwAR0NeswpSOdW82qoLZas348ZJpxyQLJ5oF3ZBl0j2pyrG3Olk3rvOKK1v5g

Correlation is not causation , check these crazy spurious correlations: http://www.tylervigen.com/spurious-correlations