

# R Notebook

```
setwd("D:/ECON613/A3")
require(tidyverse)
```

```
## Loading required package: tidyverse
```

```
## Warning: package 'tidyverse' was built under R version 4.0.5
```

```
## -- Attaching packages ----- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.3      v purrr   0.3.4
## v tibble  3.1.1      v dplyr   1.0.5
## v tidyr   1.1.3      v stringr 1.4.0
## v readr   1.4.0      v forcats 0.5.1
```

```
## Warning: package 'tibble' was built under R version 4.0.5
```

```
## Warning: package 'tidyr' was built under R version 4.0.4
```

```
## Warning: package 'dplyr' was built under R version 4.0.4
```

```
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
```

```
require(ggplot2)
require(dplyr)
require(plm)
```

```
## Loading required package: plm
```

```
## Warning: package 'plm' was built under R version 4.0.5
```

```
##
## Attaching package: 'plm'
```

```
## The following objects are masked from 'package:dplyr':
##
##   between, lag, lead
```

```
require(fixest)
```

```
## Loading required package: fixest
```

```
## Warning: package 'fixest' was built under R version 4.0.5
```

```
population <- read.csv("population.csv")
crime_long <- read.csv("crime_long.csv")
officers <- read.csv("officers.csv")

crime_long <- crime_long %>%
  group_by(crime_month, district, crime_type) %>%
  summarise(crimes=sum(crimes)) %>% tibble()
```

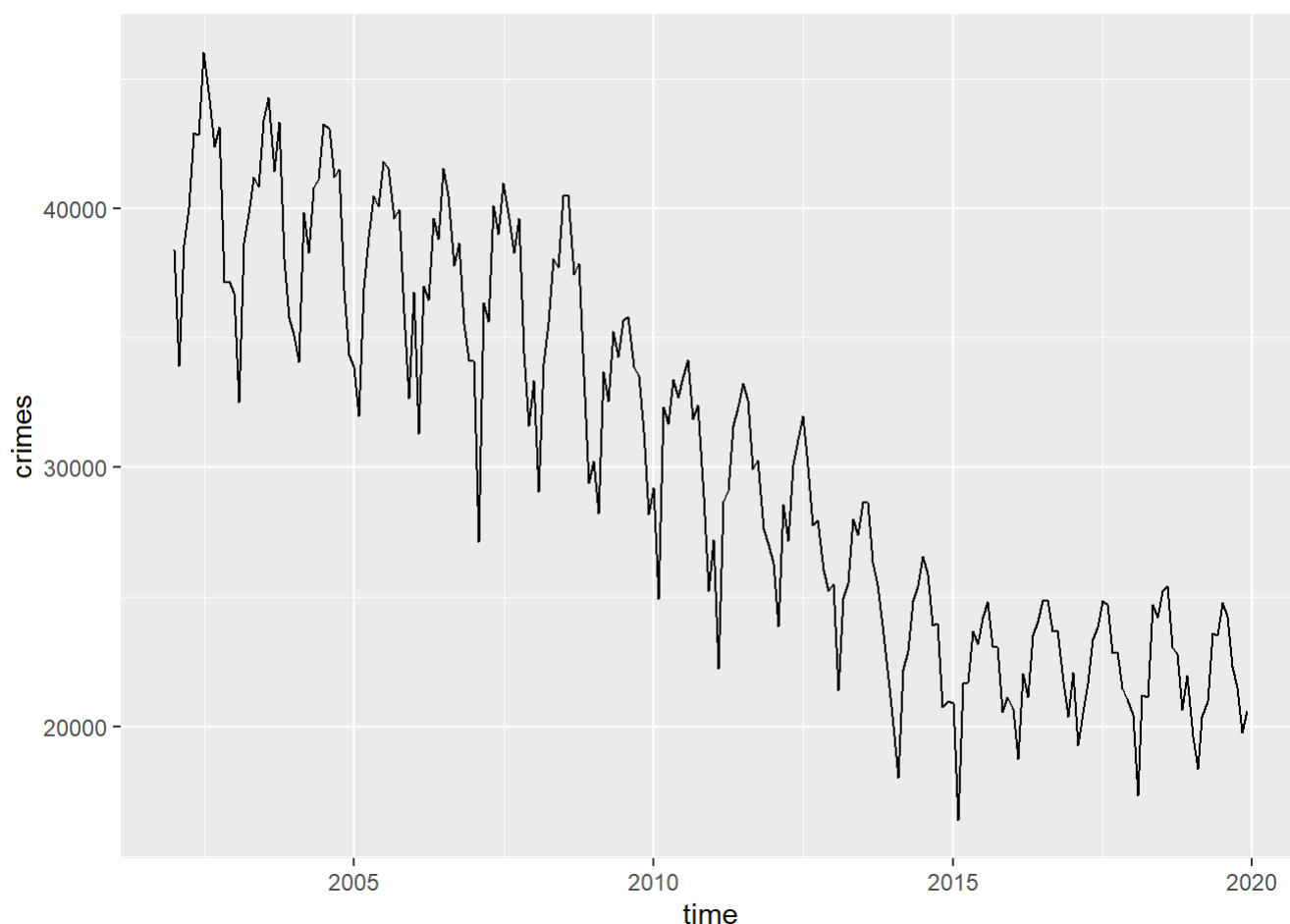
```
## `summarise()` has grouped output by 'crime_month', 'district'. You can override using the `.groups` argument.
```

## exercise 2

Calculate total crime per month and plot the time series of crime.

```
crime_long$crime_month <- as.Date(crime_long$crime_month)
```

```
crime_long %>%
  select(crime_month, crimes) %>%
  group_by(by=crime_month) %>% summarise(crimes.sum=sum(crimes)) %>%
  mutate(month=by) %>%
  ggplot(aes(month, crimes.sum)) +
  geom_line() +
  xlab("time") +
  ylab("crimes")
```



## Merge the two datasets by districts-units and period.

```
population$month <- as.Date(population$month)
```

```
population.crime <- population %>% left_join(crime_long, by=c("month"="crime_month",
  "district"="district"))
```

```
population.crime$white.share = round(population.crime$tot_white/population.crime$tot_pop, 3)
population.crime$hisp.share = round(population.crime$tot_hisp / population.crime$tot_pop, 3)
population.crime$black.share = round(population.crime$tot_black / population.crime$tot_pop, 3)
```

```
population.crime <- population.crime %>% drop_na() %>%
  select(-tot_pop, -tot_white, -tot_black, -tot_hisp) %>%
  spread(crime_type, crimes) %>%
  mutate(tot.crimes=drug + other + property + violent) %>%
  select(-c("other", "drug", "period")) %>%
  select(district, everything()) %>% arrange(district, month)
```

```
head(population.crime)
```

district	month	p50_inc	white.share	hisp.share	black.share	property	violent	tot
<int>	<date>	<dbl>	<dbl>	<dbl>	<dbl>	<int>	<int>	

	<b>district</b> <int>	<b>month</b> <date>	<b>p50_inc</b> <dbl>	<b>white.share</b> <dbl>	<b>hisp.share</b> <dbl>	<b>black.share</b> <dbl>	<b>property</b> <int>	<b>violent</b> <int>	<b>tot</b>
1	1	2005-01-01	91084.91	0.588	0.066	0.129	784	212	
2	1	2005-02-01	91084.91	0.588	0.066	0.129	557	175	
3	1	2005-03-01	91084.91	0.588	0.066	0.129	655	228	
4	1	2005-04-01	91084.91	0.588	0.066	0.129	734	243	
5	1	2005-05-01	91084.91	0.588	0.066	0.129	781	250	
6	1	2005-06-01	91084.91	0.588	0.066	0.129	774	283	

6 rows

## Exercise 3 Panel Data: Introduction

```

officers <- read.csv("officers.csv")
officers$month <- as.Date(officers$month, format = "%m/%d/%Y")

officers.crime <- officers %>%
  left_join(population.crime, by=c("month"="month", "unit"="district")) %>%
  drop_na()
head(officers.crime)

```

	<b>N...</b> <int>	<b>month</b> <date>	<b>unit</b> <int>	<b>tenure</b> <int>	<b>arrest</b> <int>	<b>p50_inc</b> <dbl>	<b>white.share</b> <dbl>	<b>hisp.share</b> <dbl>	<b>black.share</b> <dbl>	
1	1	2007-03-01	14	18	2	61849.46	0.417	0.461	0.072	
2	1	2007-04-01	14	19	0	61849.46	0.417	0.461	0.072	
3	1	2007-05-01	14	20	1	61849.46	0.417	0.461	0.072	
4	1	2007-06-01	14	21	1	61849.46	0.417	0.461	0.072	
5	1	2007-07-01	14	22	0	61849.46	0.417	0.461	0.072	
6	1	2007-08-01	14	23	0	61849.46	0.417	0.461	0.072	

6 rows | 1-10 of 13 columns

```

# ols regression
officers.crime.3.lm <- lm(arrest~tenure+p50_inc+white.share+hisp.share+black.share+tot.crimes+prop
erty+violent - 1, officers.crime) #
## coefficients of ols
### tenture
officers.crime.3.lm$coefficients[1]

```

```
##      tenure
## 3.047387e-05
```

```
#### others
officers.crime.3.lm$coefficients[2:length(officers.crime.3.lm$coefficients)]
```

```
##      p50_inc  white.share  hisp.share  black.share  tot.crimes
## 9.723698e-07 5.012935e-01 5.087375e-01 4.876706e-01 -7.778593e-06
##      property      violent
## -3.061504e-05 3.078785e-05
```

```
# panel regression, pooling, used to verify the result above.
officers.crime.3.plm <- plm(arrest~tenure+p50_inc+white.share+hisp.share+black.share+tot.crimes+pr
operty+violent - 1, officers.crime,model = "pooling", index = c("NUID", "month")) #
```

```
## coefficients of plm, pooling
```

```
#### tenure
officers.crime.3.plm$coefficients[1]
```

```
##      tenure
## 3.047387e-05
```

```
#### others
officers.crime.3.plm$coefficients[2:length(officers.crime.3.lm$coefficients)]
```

```
##      p50_inc  white.share  hisp.share  black.share  tot.crimes
## 9.723698e-07 5.012935e-01 5.087375e-01 4.876706e-01 -7.778593e-06
##      property      violent
## -3.061504e-05 3.078785e-05
```

## Exercise 4 Panel Data: More controls

```
# ols
officers.crime.4.lm <- lm(arrest~tenure+p50_inc+white.share+hisp.share+black.share+tot.crimes+prop
erty+violent +factor(month) + factor(unit) - 1, officers.crime) #
```

```
oc.4.lm.f <- grep("factor", as.vector(names(officers.crime.4.lm$coefficients)))
## coefficients of ols
#### tenure
officers.crime.4.lm$coefficients[1]
```

```
##      tenure
## -1.839865e-06
```

```
### Z
officers.crime.4.lm$coefficients[2:oc.4.lm.f[1]-1]
```

```
##      tenure      p50_inc  white.share  hisp.share  black.share
## -1.839865e-06 -6.074772e-07 -7.596157e-02 -1.117428e-01 -8.009815e-02
##      tot.crimes      property      violent
## -2.908547e-05  4.497553e-05  2.580770e-06
```

```
### district fixed effects
district.f <- grep("factor.unit.", as.vector(names(officers.crime.4.lm$coefficients)))
officers.crime.4.lm$coefficients[district.f]
```

```
## factor(unit)2 factor(unit)3 factor(unit)4 factor(unit)5 factor(unit)6
## -0.0165852054 -0.0080025995  0.0062590472 -0.0050577975 -0.0009885826
## factor(unit)7 factor(unit)8 factor(unit)9 factor(unit)10 factor(unit)11
## -0.0069657462  0.0192506643  0.0025134913  0.0097577816  0.0090497520
## factor(unit)12 factor(unit)13 factor(unit)14 factor(unit)15 factor(unit)16
## -0.0035241168 -0.0030673850  0.0191375700 -0.0038139670  0.0043663972
## factor(unit)17 factor(unit)18 factor(unit)19 factor(unit)20 factor(unit)21
## -0.0013455432 -0.0015958960  0.0040683061 -0.0152137409 -0.0281271721
## factor(unit)22 factor(unit)23 factor(unit)24 factor(unit)25
##  0.0053627750 -0.0132694124 -0.0140083265  0.0206905945
```

```
### year and month fixed effects
time.f <- grep("factor.month.", as.vector(names(officers.crime.4.lm$coefficients)))
officers.crime.4.lm$coefficients[time.f]
```

```
## factor(month)2007-01-01 factor(month)2007-02-01 factor(month)2007-03-01
##          0.6214141          0.6294927          0.6287721
## factor(month)2007-04-01 factor(month)2007-05-01 factor(month)2007-06-01
##          0.6161938          0.6348352          0.6079623
## factor(month)2007-07-01 factor(month)2007-08-01 factor(month)2007-09-01
##          0.6190805          0.6095102          0.6292321
## factor(month)2007-10-01 factor(month)2007-11-01 factor(month)2007-12-01
##          0.6263041          0.6236486          0.6108848
## factor(month)2008-01-01 factor(month)2008-02-01 factor(month)2008-03-01
##          0.6060934          0.6148399          0.6151064
## factor(month)2008-04-01 factor(month)2008-05-01 factor(month)2008-06-01
##          0.6321893          0.6416472          0.6244949
## factor(month)2008-07-01 factor(month)2008-08-01 factor(month)2008-09-01
##          0.6222314          0.6218115          0.6291827
## factor(month)2008-10-01 factor(month)2008-11-01 factor(month)2008-12-01
##          0.6169322          0.6185283          0.6317326
## factor(month)2009-01-01 factor(month)2009-02-01 factor(month)2009-03-01
##          0.6144517          0.6162632          0.6136925
## factor(month)2009-04-01 factor(month)2009-05-01 factor(month)2009-06-01
##          0.6162772          0.6230426          0.6249418
## factor(month)2009-07-01 factor(month)2009-08-01 factor(month)2009-09-01
##          0.6247718          0.6166997          0.6166521
## factor(month)2009-10-01 factor(month)2009-11-01 factor(month)2009-12-01
##          0.6150505          0.6152180          0.6093529
## factor(month)2010-01-01 factor(month)2010-02-01 factor(month)2010-03-01
##          0.6233165          0.6261844          0.6125543
## factor(month)2010-04-01 factor(month)2010-05-01 factor(month)2010-06-01
##          0.6270820          0.6177752          0.6121133
## factor(month)2010-07-01 factor(month)2010-08-01 factor(month)2010-09-01
##          0.6070168          0.6292832          0.6235422
## factor(month)2010-10-01 factor(month)2010-11-01 factor(month)2010-12-01
##          0.6182416          0.6103124          0.6156992
## factor(month)2011-01-01 factor(month)2011-02-01 factor(month)2011-03-01
##          0.6105496          0.6115033          0.6193840
## factor(month)2011-04-01 factor(month)2011-05-01 factor(month)2011-06-01
##          0.6178900          0.6204345          0.6207161
## factor(month)2011-07-01 factor(month)2011-08-01 factor(month)2011-09-01
##          0.6249620          0.6167380          0.6074181
## factor(month)2011-10-01 factor(month)2011-11-01 factor(month)2011-12-01
##          0.6268214          0.6116403          0.6187861
## factor(month)2012-01-01 factor(month)2012-02-01 factor(month)2012-03-01
##          0.6122728          0.6253858          0.6223464
## factor(month)2012-04-01 factor(month)2012-05-01 factor(month)2012-06-01
##          0.6239884          0.6246277          0.6192927
## factor(month)2012-07-01 factor(month)2012-08-01 factor(month)2012-09-01
##          0.6209638          0.6136969          0.6194293
## factor(month)2012-10-01 factor(month)2012-11-01 factor(month)2012-12-01
##          0.6017567          0.6233991          0.6022260
## factor(month)2013-01-01 factor(month)2013-02-01 factor(month)2013-03-01
##          0.6212240          0.6124246          0.6054525
## factor(month)2013-04-01 factor(month)2013-05-01 factor(month)2013-06-01
##          0.6253034          0.6344246          0.6274609
## factor(month)2013-07-01 factor(month)2013-08-01 factor(month)2013-09-01
```

```
##           0.6381276           0.6164643           0.6231493
## factor(month)2013-10-01 factor(month)2013-11-01 factor(month)2013-12-01
##           0.6088992           0.6118126           0.6224321
## factor(month)2014-01-01 factor(month)2014-02-01 factor(month)2014-03-01
##           0.6043203           0.6234869           0.6133104
## factor(month)2014-04-01 factor(month)2014-05-01 factor(month)2014-06-01
##           0.6002403           0.6160394           0.6208670
## factor(month)2014-07-01 factor(month)2014-08-01 factor(month)2014-09-01
##           0.6251092           0.6158498           0.6197939
## factor(month)2014-10-01 factor(month)2014-11-01 factor(month)2014-12-01
##           0.6178644           0.6118953           0.6111764
## factor(month)2015-01-01 factor(month)2015-02-01 factor(month)2015-03-01
##           0.6166966           0.6125955           0.6262616
## factor(month)2015-04-01 factor(month)2015-05-01 factor(month)2015-06-01
##           0.6168094           0.6134227           0.6117509
## factor(month)2015-07-01 factor(month)2015-08-01 factor(month)2015-09-01
##           0.6105852           0.6165148           0.6189486
## factor(month)2015-10-01 factor(month)2015-11-01 factor(month)2015-12-01
##           0.6114049           0.6227367           0.6108374
## factor(month)2016-01-01 factor(month)2016-02-01 factor(month)2016-03-01
##           0.6131590           0.6119276           0.6042766
## factor(month)2016-04-01 factor(month)2016-05-01 factor(month)2016-06-01
##           0.6298911           0.6296974           0.6142544
## factor(month)2016-07-01 factor(month)2016-08-01 factor(month)2016-09-01
##           0.6055479           0.5984030           0.6166462
## factor(month)2016-10-01 factor(month)2016-11-01 factor(month)2016-12-01
##           0.6187095           0.6024098           0.6024920
## factor(month)2017-01-01 factor(month)2017-02-01 factor(month)2017-03-01
##           0.6189461           0.6110851           0.6257720
## factor(month)2017-04-01 factor(month)2017-05-01 factor(month)2017-06-01
##           0.6136808           0.6261398           0.6056502
## factor(month)2017-07-01 factor(month)2017-08-01 factor(month)2017-09-01
##           0.6115586           0.6147641           0.6085598
## factor(month)2017-10-01 factor(month)2017-11-01 factor(month)2017-12-01
##           0.6077035           0.6006984           0.6077938
```

## Exercise 5 Panel Data: Individual fixed effects

```
# within
officers.crime.5.plm <- plm(arrest~tenure+p50_inc+white.share+hisp.share+black.share+tot.crimes+
property+violent+factor(unit),
                           data= officers.crime, model = "within", index = c("NUID", "month"))
```

```
# between
officers.crime.5.plm.b <- plm(arrest~tenure+p50_inc+white.share+hisp.share+black.share+tot.crimes+
property+violent+factor(unit),
                              data= officers.crime, model = "between", index = c("NUID", "month"
))
```



```
# first difference
officers.crime.5.plm.d <- plm(arrest~tenure+p50_inc+white.share+hisp.share+black.share+tot.crimes+
property+violent+factor(unit),
                             data= officers.crime, model = "fd", index = c("NUID", "month"))
```

```
coef.within.tenure <- as.vector(officers.crime.5.plm$coefficients[1])
coef.between.tenure <- as.vector(officers.crime.5.plm.b$coefficients[2])
coef.fd.tenure <- as.vector(officers.crime.5.plm.d$coefficients[2])
pander::pandoc.table(data.frame(coef.within.tenure ,coef.between.tenure ,coef.fd.tenure))
```

```
##
## -----
##  coef.within.tenure   coef.between.tenure   coef.fd.tenure
## -----
##      -3.258e-05           9.225e-06           0.0005986
## -----
```

As we known,  $\beta_{within,tenure}$  is  $-3.2581665 \times 10^{-5}$ ,  $\beta_{between,tenure}$  is  $9.2254365 \times 10^{-6}$ ,  $\beta_{fd,tenure}$  is  $5.9861314 \times 10^{-4}$ . so, estimated effect by between and fd both are positive effects on dependent variable, estimated effect by within is negative. And estimated level by first difference is the largest.

## b

```
officers.crime.5.p <- feols(arrest~tenure+p50_inc+white.share+hisp.share+black.share+tot.crimes |
month + unit + NUID, officers.crime, panel.id = c("NUID", "month"))
etable(officers.crime.5.p)
```

officers.crime.5.p <chr>	
Dependent Var.:	arrest
tenure	0.0007 (0.0023)
p50_inc	-6.39e-7 (7.8e-7)
white.share	-0.1069 (0.1769)
hisp.share	-0.1290 (0.1919)
black.share	-0.1241 (0.1051)
tot.crimes	-6.48e-6 (5.25e-6)
Fixed-Effects:	-----
month	Yes
1-10 of 17 rows	Previous 1 2 Next

```
officers.crime.5.p.coef <- fixef(officers.crime.5.p)
summary(officers.crime.5.p.coef)
```

```
## Fixed_effects coefficients
##               month    unit    NUID
## Number of fixed-effects    132     25   12621
## Number of references        0      1      1
## Mean                      0.59 -0.016 -0.0275
## Standard-deviation         0.0279 0.0138   0.165
##
## COEFFICIENTS:
##   month: 2007-01-01 2007-02-01 2007-03-01 2007-04-01 2007-05-01
##           0.6327      0.6397      0.6391      0.6259      0.6443
##
##   ... 127 remaining
## -----
##   unit:      1      2      3      4      5
##           0.001294 -0.03529 -0.01585 0.0002887 -0.02093 ... 20 remaining
## -----
##   NUID: 1      2      6      7      16
##          0 0.002663 -0.241 -0.07698 0.0002425 ... 12,616 remaining
```

one-step GMM:

$$\hat{\beta}_{2SLS} = [X'Z(Z'Z)^{-1}Z'X]^{-1}X'Z(Z'Z)^{-1}Z'y$$

if  $X = Z$ , it is just estimator of OLS.