A3

Yue Han

4/24/2021

```
library(tidyverse)
```

```
## -- Attaching packages ---------------------------------------------------------------- tidyverse 1.3
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.3     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.3.1     v forcats 0.5.0
```

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

```
library(xts)
```

```
## Loading required package: zoo
```

```
##
## Attaching package: 'zoo'
```

```
## The following objects are masked from 'package:base':
##
##     as.Date, as.Date.numeric
```

```
##
## Attaching package: 'xts'
```

```
## The following objects are masked from 'package:dplyr':
##
##     first, last
```

```
library(plm)
```

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

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

Exercise 1 Links to the datasets

```
population <- read.csv("https://www.dropbox.com/s/s38cde88670y5mw/population.csv?dl=1")
crime_long <- read.csv("https://www.dropbox.com/s/t3vushurhm3s5my/crime_long.csv?dl=1")
officers <- read.csv("https://www.dropbox.com/s/8q2fpdb7phy86m8/officers.csv?dl=1")
```

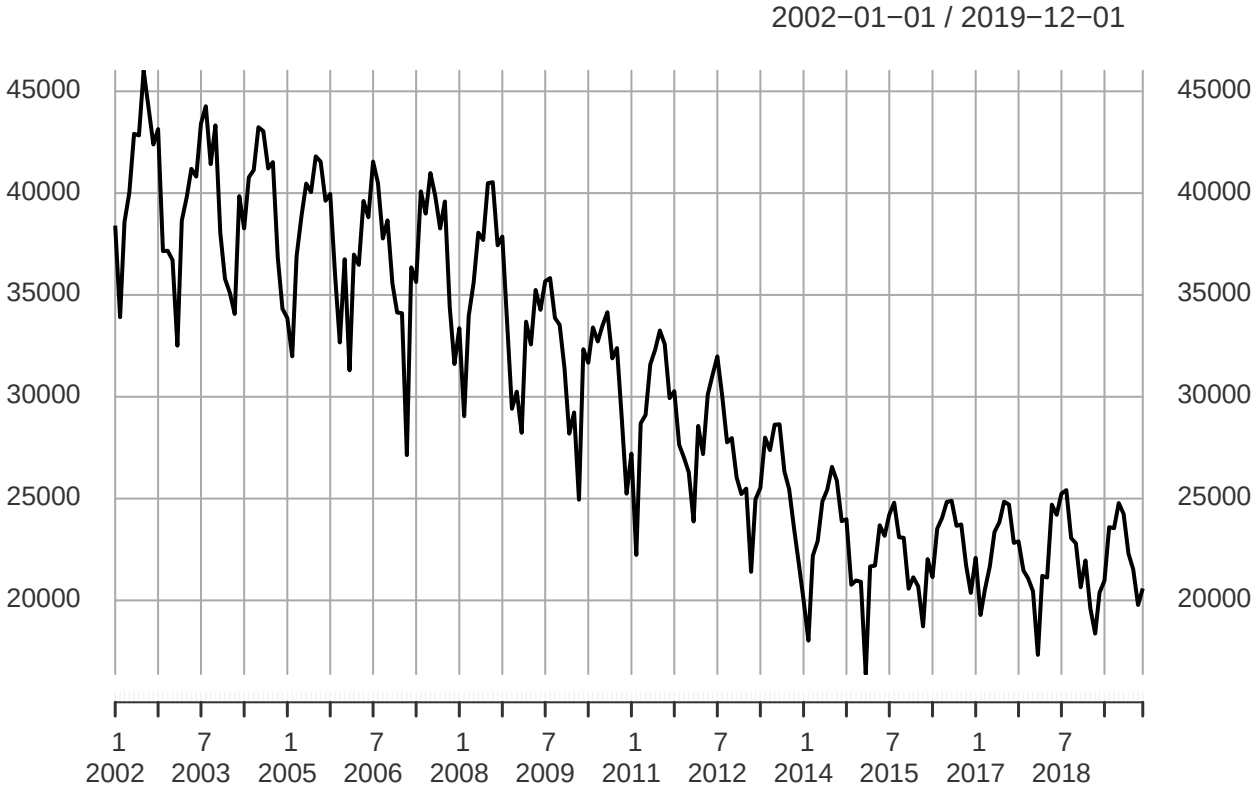
Exercise 2 Data Manipulation

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

```
# total crime per month
crime_monthly <- crime_long %>%
  group_by(crime_month) %>%
  summarize(crime_by_month = sum(crimes))
```

```
## `summarise()` ungrouping output (override with `.groups` argument)
```

```
# plot time series of crime
crime_monthly_xts <- xts(crime_monthly$crime_by_month, as.Date(crime_monthly$crime_month, format='%Y-%m
plot(crime_monthly_xts,type = 'l',main=' ')
```



Merge the two datasets by districts-units and period

```
crime_population <- merge(crime_long, population,by.x = c("district","crime_month"), by.y = c("district
```

Construct a panel data of unit over time with the following variables

– Total crimes per resident – Violent crimes per resident – Property crimes per resident – Median income – Share of black, Hispanic, and white residents

```
panel_data <- crime_population %>%
  mutate(
    violent_crimes = case_when(
      crime_type == "violent" ~ crimes,
      TRUE ~ 0L
    ),
```

```
    property_crimes = case_when(
      crime_type == "property" ~ crimes,
      TRUE ~ 0L
    )
  ) %>%
  group_by(district,crime_month) %>%
  summarize(
    total_crimes = sum(crimes),
    violent_crimes = sum(violent_crimes),
    property_crimes = sum(property_crimes),
    median_income = p50_inc,
    share_of_black = tot_black/tot_pop,
    share_of_hisp = tot_hisp/tot_pop,
    share_of_white = tot_white/tot_pop
  ) %>%
  distinct()
```

```
## `summarise()` regrouping output by 'district', 'crime_month' (override with `.groups` argument)
```

Exercise 3 Panel Data: Introduction

```
df <- merge(officers, panel_data, by.x = c("month", "unit"), by.y=c("crime_month", "district"), all.x=TR
panel_df <- pdata.frame(df,index=c("NUID","month"))
```

```
# use lm
lm_pooled <- lm(
  formula = arrest ~ tenure + total_crimes + median_income + share_of_black + share_of_hisp + share_of_w
  data = df
)
# estimators
# beta
lm_pooled$coefficients[1]
```

```
##       tenure
## 2.878869e-05
```

```
# gamma
lm_pooled$coefficients[2:6]
```

```
##   total_crimes  median_income share_of_black  share_of_hisp share_of_white
##  -1.364283e-05   7.210106e-07   5.028186e-01   5.172976e-01   5.151714e-01
```

```
# check
pooled <- plm(
  formula = arrest ~ tenure + total_crimes + median_income + share_of_black + share_of_hisp + share_of_w
  data = panel_df,
  model = "pooling"
)
# estimator
# beta
pooled$coefficients[1]
```

```
##       tenure
## 2.878869e-05
```

```r
# gamma
pooled$coefficients[2:6]
```

```
##   total_crimes  median_income share_of_black  share_of_hisp share_of_white
##  -1.364283e-05   7.210106e-07   5.028186e-01   5.172976e-01   5.151714e-01
```

Exercise 4 Panel Data: More controls

```r
# use lm
fe1_lm <- lm(
    formula = arrest ~ tenure + total_crimes + median_income + share_of_black + share_of_hisp + share_o
    data = df
)
# estimators
# beta
fe1_lm$coefficients[1]
```

```
##        tenure
## -3.809782e-06
```

```r
# gamma
fe1_lm$coefficients[2:6]
```

```
##   total_crimes  median_income share_of_black  share_of_hisp share_of_white
##  -6.320360e-06  -4.910055e-07  -9.200883e-02  -1.398480e-01  -1.012976e-01
```

```r
# psi
fe1_lm$coefficients[7:31]
```

```
##  factor(unit)1  factor(unit)2  factor(unit)3  factor(unit)4  factor(unit)5
##      0.6361725      0.6117946      0.6161604      0.6361362      0.6203707
##  factor(unit)6  factor(unit)7  factor(unit)8  factor(unit)9 factor(unit)10
##      0.6223150      0.6150872      0.6595989      0.6392821      0.6438540
## factor(unit)11 factor(unit)12 factor(unit)13 factor(unit)14 factor(unit)15
##      0.6248365      0.6366138      0.6331086      0.6605815      0.6158985
## factor(unit)16 factor(unit)17 factor(unit)18 factor(unit)19 factor(unit)20
##      0.6434606      0.6372529      0.6401890      0.6451079      0.6207241
## factor(unit)21 factor(unit)22 factor(unit)23 factor(unit)24 factor(unit)25
##      0.5994560      0.6355380      0.6252008      0.6207761      0.6581988
```

```r
# kappa
fe1_lm$coefficients[32:length(fe1_lm$coefficients)]
```

```
## factor(month)2007-02-01 factor(month)2007-03-01 factor(month)2007-04-01
##            2.746517e-03            5.425086e-03           -4.433449e-03
## factor(month)2007-05-01 factor(month)2007-06-01 factor(month)2007-07-01
##            9.642377e-03           -1.518439e-02           -2.476077e-03
## factor(month)2007-08-01 factor(month)2007-09-01 factor(month)2007-10-01
##           -9.283480e-03            2.831905e-03            5.531776e-03
## factor(month)2007-11-01 factor(month)2007-12-01 factor(month)2008-01-01
##            3.270094e-03           -8.522393e-03           -1.580319e-02
## factor(month)2008-02-01 factor(month)2008-03-01 factor(month)2008-04-01
##           -7.935551e-03           -7.042028e-03            7.880029e-03
## factor(month)2008-05-01 factor(month)2008-06-01 factor(month)2008-07-01
##            1.889847e-02            1.042877e-03           -1.992463e-05
## factor(month)2008-08-01 factor(month)2008-09-01 factor(month)2008-10-01
```

```
##            1.040177e-03            9.368277e-03           -3.356443e-03
## factor(month)2008-11-01 factor(month)2008-12-01 factor(month)2009-01-01
##           -1.541507e-03            1.093272e-02           -6.007364e-03
## factor(month)2009-02-01 factor(month)2009-03-01 factor(month)2009-04-01
##           -5.783149e-03           -8.894307e-03           -4.911125e-03
## factor(month)2009-05-01 factor(month)2009-06-01 factor(month)2009-07-01
##           -1.544614e-03            4.205706e-03            5.471119e-03
## factor(month)2009-08-01 factor(month)2009-09-01 factor(month)2009-10-01
##           -4.666546e-03           -4.795034e-03           -1.586495e-03
## factor(month)2009-11-01 factor(month)2009-12-01 factor(month)2010-01-01
##           -5.210196e-03           -9.885677e-03            1.814697e-03
## factor(month)2010-02-01 factor(month)2010-03-01 factor(month)2010-04-01
##            3.619692e-03           -8.878372e-03            7.082318e-03
## factor(month)2010-05-01 factor(month)2010-06-01 factor(month)2010-07-01
##           -3.414388e-03           -5.868300e-03           -1.353438e-02
## factor(month)2010-08-01 factor(month)2010-09-01 factor(month)2010-10-01
##            9.875346e-03            2.215222e-03           -2.548071e-03
## factor(month)2010-11-01 factor(month)2010-12-01 factor(month)2011-01-01
##           -9.174658e-03           -4.759482e-03           -8.207417e-03
## factor(month)2011-02-01 factor(month)2011-03-01 factor(month)2011-04-01
##           -5.389484e-03           -4.548749e-03           -1.189441e-03
## factor(month)2011-05-01 factor(month)2011-06-01 factor(month)2011-07-01
##           -5.307481e-04            1.206606e-04            4.004173e-03
## factor(month)2011-08-01 factor(month)2011-09-01 factor(month)2011-10-01
##           -1.042689e-04           -1.134552e-02            1.049381e-02
## factor(month)2011-11-01 factor(month)2011-12-01 factor(month)2012-01-01
##           -6.974074e-03            2.879175e-04           -1.044974e-02
## factor(month)2012-02-01 factor(month)2012-03-01 factor(month)2012-04-01
##            2.973167e-03            1.819509e-03            4.576250e-03
## factor(month)2012-05-01 factor(month)2012-06-01 factor(month)2012-07-01
##            4.361146e-03           -2.065034e-03            6.780647e-04
## factor(month)2012-08-01 factor(month)2012-09-01 factor(month)2012-10-01
##           -5.126210e-03            2.322100e-03           -1.457656e-02
## factor(month)2012-11-01 factor(month)2012-12-01 factor(month)2013-01-01
##            3.462557e-03           -1.687488e-02            3.795055e-03
## factor(month)2013-02-01 factor(month)2013-03-01 factor(month)2013-04-01
##           -8.077931e-03           -1.585271e-02            3.512273e-03
## factor(month)2013-05-01 factor(month)2013-06-01 factor(month)2013-07-01
##            1.410776e-02            7.059393e-03            1.753106e-02
## factor(month)2013-08-01 factor(month)2013-09-01 factor(month)2013-10-01
##           -2.731977e-03            3.036804e-03           -8.699381e-03
## factor(month)2013-11-01 factor(month)2013-12-01 factor(month)2014-01-01
##           -5.810358e-03            3.319051e-03           -1.496478e-02
## factor(month)2014-02-01 factor(month)2014-03-01 factor(month)2014-04-01
##            2.499075e-03           -4.441004e-03           -1.862457e-02
## factor(month)2014-05-01 factor(month)2014-06-01 factor(month)2014-07-01
##           -5.901253e-03           -1.760679e-03            5.574852e-03
## factor(month)2014-08-01 factor(month)2014-09-01 factor(month)2014-10-01
##           -1.112962e-03            1.054709e-03           -1.729819e-03
## factor(month)2014-11-01 factor(month)2014-12-01 factor(month)2015-01-01
##           -8.539207e-03           -7.021181e-03           -4.922362e-03
## factor(month)2015-02-01 factor(month)2015-03-01 factor(month)2015-04-01
##           -9.307792e-03            4.154443e-03           -3.524368e-03
## factor(month)2015-05-01 factor(month)2015-06-01 factor(month)2015-07-01
```

```
##           -6.722439e-03           -7.111330e-03           -8.631208e-03
## factor(month)2015-08-01 factor(month)2015-09-01 factor(month)2015-10-01
##           -5.340895e-03           -1.012199e-03           -8.334859e-03
## factor(month)2015-11-01 factor(month)2015-12-01 factor(month)2016-01-01
##            2.320336e-03           -7.372761e-03           -5.343684e-03
## factor(month)2016-02-01 factor(month)2016-03-01 factor(month)2016-04-01
##           -7.098580e-03           -1.453481e-02            9.817929e-03
## factor(month)2016-05-01 factor(month)2016-06-01 factor(month)2016-07-01
##            9.858467e-03           -5.146654e-03           -1.082352e-02
## factor(month)2016-08-01 factor(month)2016-09-01 factor(month)2016-10-01
##           -1.921383e-02           -1.541889e-03            3.117914e-03
## factor(month)2016-11-01 factor(month)2016-12-01 factor(month)2017-01-01
##           -1.397966e-02           -1.289825e-02            8.783617e-05
## factor(month)2017-02-01 factor(month)2017-03-01 factor(month)2017-04-01
##           -5.838321e-03            6.054365e-03           -5.737714e-03
## factor(month)2017-05-01 factor(month)2017-06-01 factor(month)2017-07-01
##            8.409860e-03           -1.219203e-02           -6.123435e-03
## factor(month)2017-08-01 factor(month)2017-09-01 factor(month)2017-10-01
##           -2.664221e-03           -8.493195e-03           -9.683101e-03
## factor(month)2017-11-01 factor(month)2017-12-01
##           -1.684569e-02           -8.924383e-03
```

```
# check
fe1 <- plm(
    formula = arrest ~ tenure + total_crimes + median_income + share_of_black + share_of_hisp + share_o
    effect = "time",
    data = panel_df,
    model = "within"
)
# estimators
# beta
fe1$coefficients[1]
```

```
##        tenure
## -3.809782e-06
```

```
# gamma
fe1$coefficients[2:6]
```

```
##   total_crimes  median_income share_of_black  share_of_hisp share_of_white
##  -6.320360e-06  -4.910055e-07  -9.200883e-02  -1.398480e-01  -1.012976e-01
```

```
# psi
fe1$coefficients[7:30]
```

```
##  factor(unit)1  factor(unit)2  factor(unit)3  factor(unit)4  factor(unit)5
##   -0.022026285   -0.046404197   -0.042038392   -0.022062613   -0.037828069
##  factor(unit)6  factor(unit)7  factor(unit)8  factor(unit)9 factor(unit)10
##   -0.035883828   -0.043111605    0.001400111   -0.018916707   -0.014344782
## factor(unit)11 factor(unit)12 factor(unit)13 factor(unit)14 factor(unit)15
##   -0.033362271   -0.021585041   -0.025090190    0.002382670   -0.042300270
## factor(unit)16 factor(unit)17 factor(unit)18 factor(unit)19 factor(unit)20
##   -0.014738192   -0.020945932   -0.018009774   -0.013090894   -0.037474745
## factor(unit)21 factor(unit)22 factor(unit)23 factor(unit)24
##   -0.058742814   -0.022660816   -0.032998025   -0.037422669
```

```
# kappa
fixef(fe1)
```

```
## 2007-03-01 2007-04-01 2007-05-01 2007-06-01 2007-07-01 2007-08-01 2007-09-01
##    0.66362    0.65377    0.66784    0.64301    0.65572    0.64892    0.66103
## 2007-10-01 2007-11-01 2007-12-01 2008-01-01 2008-02-01 2008-03-01 2008-04-01
##    0.66373    0.66147    0.64968    0.64240    0.65026    0.65116    0.66608
## 2008-05-01 2008-06-01 2008-07-01 2008-08-01 2008-09-01 2008-10-01 2008-11-01
##    0.67710    0.65924    0.65818    0.65924    0.66757    0.65484    0.65666
## 2008-12-01 2009-01-01 2009-02-01 2009-03-01 2009-04-01 2009-05-01 2009-06-01
##    0.66913    0.65219    0.65242    0.64930    0.65329    0.65665    0.66240
## 2009-07-01 2009-08-01 2009-09-01 2009-10-01 2009-11-01 2009-12-01 2010-01-01
##    0.66367    0.65353    0.65340    0.65661    0.65299    0.64831    0.66001
## 2010-02-01 2010-03-01 2010-04-01 2010-05-01 2010-06-01 2010-07-01 2010-08-01
##    0.66182    0.64932    0.66528    0.65478    0.65233    0.64466    0.66807
## 2010-09-01 2010-10-01 2010-11-01 2010-12-01 2011-01-01 2011-02-01 2011-03-01
##    0.66041    0.65565    0.64902    0.65344    0.64999    0.65281    0.65365
## 2011-04-01 2011-05-01 2011-06-01 2011-07-01 2011-08-01 2011-09-01 2011-10-01
##    0.65701    0.65767    0.65832    0.66220    0.65809    0.64685    0.66869
## 2011-11-01 2011-12-01 2012-01-01 2012-02-01 2012-03-01 2012-04-01 2012-05-01
##    0.65122    0.65849    0.64775    0.66117    0.66002    0.66278    0.66256
## 2012-06-01 2012-07-01 2012-08-01 2012-09-01 2012-10-01 2012-11-01 2012-12-01
##    0.65613    0.65888    0.65307    0.66052    0.64362    0.66166    0.64132
## 2013-01-01 2013-02-01 2013-03-01 2013-04-01 2013-05-01 2013-06-01 2013-07-01
##    0.66199    0.65012    0.64235    0.66171    0.67231    0.66526    0.67573
## 2013-08-01 2013-09-01 2013-10-01 2013-11-01 2013-12-01 2014-01-01 2014-02-01
##    0.65547    0.66124    0.64950    0.65239    0.66152    0.64323    0.66070
## 2014-03-01 2014-04-01 2014-05-01 2014-06-01 2014-07-01 2014-08-01 2014-09-01
##    0.65376    0.63957    0.65230    0.65644    0.66377    0.65709    0.65925
## 2014-10-01 2014-11-01 2014-12-01 2015-01-01 2015-02-01 2015-03-01 2015-04-01
##    0.65647    0.64966    0.65118    0.65328    0.64889    0.66235    0.65467
## 2015-05-01 2015-06-01 2015-07-01 2015-08-01 2015-09-01 2015-10-01 2015-11-01
##    0.65148    0.65109    0.64957    0.65286    0.65719    0.64986    0.66052
## 2015-12-01 2016-01-01 2016-02-01 2016-03-01 2016-04-01 2016-05-01 2016-06-01
##    0.65083    0.65286    0.65110    0.64366    0.66802    0.66806    0.65305
## 2016-07-01 2016-08-01 2016-09-01 2016-10-01 2016-11-01 2016-12-01 2017-01-01
##    0.64738    0.63898    0.65666    0.66132    0.64422    0.64530    0.65829
## 2017-02-01 2017-03-01 2017-04-01 2017-05-01 2017-06-01 2017-07-01 2017-08-01
##    0.65236    0.66425    0.65246    0.66661    0.64601    0.65208    0.65553
## 2017-09-01 2017-10-01 2017-11-01 2017-12-01 2007-01-01 2007-02-01
##    0.64971    0.64852    0.64135    0.64927    0.65820    0.66095
```

Exercise 5 Panel Data: Individual fixed effects

Implement a within, between, and first difference estimator for the parameter beta. Then, compare the estimated values.

```
# within
fe2 <- plm(
    formula = arrest ~ tenure + total_crimes + median_income + share_of_black + share_of_hisp + share_o
    effect = "twoway",
    data = panel_df,
    model = "within"
)
```

```r
# between
fe3 <- plm(
    formula = arrest ~ tenure + total_crimes + median_income + share_of_black + share_of_hisp + share_o
    effect = "individual",
    data = panel_df,
    model = "between"
)
```

```r
# fd
fe4 <- plm(
    formula = arrest ~ tenure + total_crimes + median_income + share_of_black + share_of_hisp + share_o
    effect = "individual",
    data = panel_df,
    model = "fd"
)
```

```r
# compare beta
est_betas <- c(fe2$coefficients[1], fe3$coefficients[1], fe4$coefficients[1])
names(est_betas) <- c("within", "between", "fd")
est_betas
```

```
##        within       between            fd
## -2.767569e-04 -1.754975e-05  5.430383e-03
```

within and between estimated beta estimator are both negative, first difference estimated beta is positive.

Use a GMM approach to estimate all parameters (including fixed effects) in one step.

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.