

Panel Data 2: Implementation in R

Instructor: Yuta Toyama

Last updated: 2020-06-29

Section 1

Panel

Preliminary:

- ▶ I use the following package
 - ▶ lfe package.

Panel Data Regression

- ▶ I use the dataset Fatalities in AER package.
 - ▶ See <https://www.rdocumentation.org/packages/AER/versions/1.2-6/topics/Fatalities> for details.

```
library(AER)
```

```
## Warning: package 'AER' was built under R version 3.6.3
```

```
## Loading required package: car
```

```
## Warning: package 'car' was built under R version 3.6.3
```

```
## Loading required package: carData
```

```
## Loading required package: lmtest
```

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

```
## Warning: package 'zoo' was built under R version 3.6.3
```

```
##
```

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

```
## The following objects are masked from 'package:base':
```

```
##
```

- ▶ As a preliminary analysis, let's plot the relationship between fatality rate and beer tax in 1998.

```
library("dplyr")
```

```
## Warning: package 'dplyr' was built under R version 3.6.3
```

```
##
```

```
## Attaching package: 'dplyr'
```

```
## The following object is masked from 'package:car':
```

```
##
```

```
##      recode
```

```
## The following objects are masked from 'package:stats':
```

```
##
```

```
##      filter, lag
```

```
## The following objects are masked from 'package:base':
```

```
##
```

```
##      intersect, setdiff, setequal, union
```

```
Fatalities %>%
```

```
  mutate(fatal_rate = fatal / pop * 10000) %>%
```

```
  filter(year == "1988") -> data
```

- ▶ Run fixed effect regression using `felm` command in `lfe` package.
 - ▶ <https://www.rdocumentation.org/packages/lfe/versions/2.8-3/topics/felm>

```
library("lfe")
```

```
## Warning: package 'lfe' was built under R version 3.6.2
```

```
## Loading required package: Matrix
```

```
##
```

```
## Attaching package: 'lfe'
```

```
## The following object is masked from 'package:lmtest':
```

```
##
```

```
##      waldtest
```

```
Fatalities %>%
```

```
  mutate(fatal_rate = fatal / pop * 10000) -> data
```

```
# OLS
```

```
result_ols <- felm( fatal_rate ~ beertax | 0 | 0 | 0, data = data )
```

```
summary(result_ols, robust = TRUE)
```

```
##
```

▶ What if we do not use the cluster-robust standard error?

```
# State FE w.o. CRS
```

```
result_wo_CRS <- felm( fatal_rate ~ beertax | state | 0 | 0, data = data )
```

```
# State FE w. CRS
```

```
result_w_CRS <- felm( fatal_rate ~ beertax | state | 0 | state, data = data )
```

```
# Report heteroskedasticity robust standard error and cluster-robust standard error
```

```
stargazer::stargazer(result_wo_CRS, result_w_CRS, type = "text", se = list("SE", "Heteroskedasticity-Robust", "Cluster-Robust"), add.lines = list(c("SE", "Heteroskedasticity-Robust", "Cluster-Robust")))
```

```
##
```

```
## =====
```

```
##                               Dependent variable:
```

```
##                               -----
```

```
##                               fatal_rate
```

```
##                               (1)                (2)
```

```
## -----
```

```
## beertax                       -0.656***          -0.656**
```

```
##                               (0.203)            (0.292)
```

```
##
```

```
## -----
```

```
## SE                               Heteroskedasticity-Robust Cluster-Robust
```

Section 2

Panel + IV

Panel Data with Instrumental Variables

- ▶ Revisit the demand for Cigaretts
- ▶ Consider the following model

$$\log(Q_{it}) = \beta_0 + \beta_1 \log(P_{it}) + \beta_2 \log(\text{income}_{it}) + u_i + e_{it}$$

where

- ▶ Q_{it} is the number of packs per capita in state i in year t ,
- ▶ P_{it} is the after-tax average real price per pack of cigarettes, and
- ▶ income_{it} is the real income per capita. This is demand shifter.
- ▶ As an IV for the price, we use the followings:
 - ▶ SalesTax_{it} : the proportion of taxes on cigarettes arising from the general sales tax.
 - ▶ Relevant as it is included in the after-tax price
 - ▶ Exogenous(independent) since the sales tax does not influence demand directly, but indirectly through the price.
 - ▶ CigTax_{it} : the cigarett-specific taxes

```
# load the data set and get an overview  
library(AER)  
data("CigarettesSW")  
CigarettesSW %>%  
  mutate( rincome = (income / population) / cpi) %>%  
  mutate( rprice = price / cpi ) %>%  
  mutate( salestax = (taxes - tax) / cpi ) %>%  
  mutate( cigtax = tax/cpi ) -> Cigdata
```

▶ Run IV regression with panel data.

```

# OLS
result_1 <- felm( log(packs) ~ log(rprice) + log(rincome) | 0 | 0 | state,

# State FE
result_2 <- felm( log(packs) ~ log(rprice) + log(rincome) | state | 0 | sta

# IV without FE
result_3 <- felm( log(packs) ~ log(rincome) | 0 | (log(rprice) ~ salestax

# IV with FE
result_4 <- felm( log(packs) ~ log(rincome) | state | (log(rprice) ~ sales

stargazer::stargazer(result_1, result_2, result_3, result_4, type = "text")

```

```

##
## =====
##                               Dependent variable:
##                               -----
##                               log(packs)
##                               (1)           (2)           (3)
## -----
## log(rprice)           -1.334***           -1.210***

```

Section 3

felm command

How to report heteroskedasticity robust standard error in stargazer

```
# Run felm command without specifying cluster.
```

```
result_1 <- felm( log(packs) ~ log(rprice) + log(rincome) | 0 | 0 | state,
```

```
# `result_1$rse` contains heteroskedasticity robust standard error. Put the
```

```
stargazer::stargazer(result_1, type = "text",  
                      se = list(result_1$rse) ) )
```

```
##  
## =====  
##                               Dependent variable:  
##                               -----  
##                               log(packs)  
##                               -----  
## log(rprice)                   -1.334***  
##                               (0.154)  
##  
## log(rincome)                   0.318**  
##                               (0.154)  
##  
## Constant                       10.067***  
##                               (0.502)  
##
```

How to conduct F test after felm

Run felm command without specifying cluster.

```
result_1 <- felm( packs ~ rprice + rincome | 0 | 0 | 0, data = Cigdata )
```

The following tests $H_0: _b[rincome] = 0 \ \& \ _b[rprice] = 0$

```
fctest1 = waldtest(result_1, ~ rincome | rprice )
```

```
fctest1
```

```
##           p           chi2           df1           p.F           F
## 4.180596e-22 9.845284e+01 2.000000e+00 2.621701e-15 4.922642e+01 9.300000
## attr(,"formula")
## ~rincome | rprice
## <environment: 0x0000000017ced520>
```

fctest[5] corresponds to F-value

```
fval1 = fctest[5]
```

The following tests $H_0: _b[rincome] - 1 = 0 \ \& \ _b[rprice] = 0$

```
fctest2 = waldtest(result_1, ~ rincome - 1 | rprice )
```

```
fctest2
```

```
##           p           chi2           df1           p.F           F
```