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)

```
## 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':

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

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 %>%
```

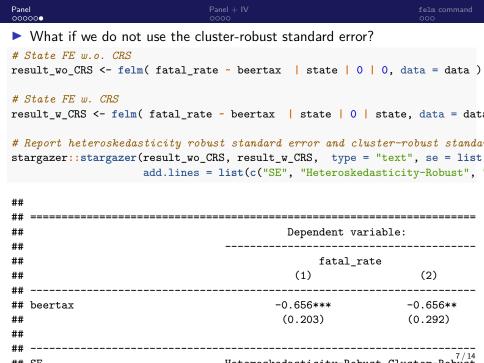
mutate(fatal_rate = fatal / pop * 10000) %>%

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)
```

Panel



Section 2

Panel + IV

Panel Data with Instrumental Variables

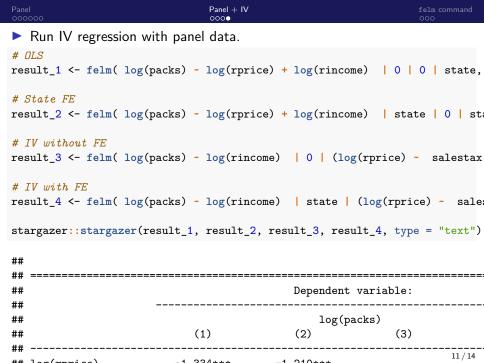
- Revisit the demand for Cigaretts
- Consider the following model

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

where

- \triangleright 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(indepndent) 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 = (taxs - tax) / cpi ) %>%
  mutate( cigtax = tax/cpi ) -> Cigdata
```



Section 3

felm command

-1.334***

(0.154)

0.318**

(0.154)

10.067***

(0.502)

13 / 14

log(rprice)

log(rincome)

Constant

##

##

##

##

##

##

##

```
# Run felm command without specifying cluster.
result_1 <- felm( packs ~ rprice + rincome | 0 | 0 | 0, data = Cigdata )
# The following tests HO: _b[rincome] = 0 & _b[rprice] = 0
ftest1 = waldtest(result_1, ~ rincome | rprice )
ftest1</pre>
```

```
## p chi2 df1 p.F F
## 4.180596e-22 9.845284e+01 2.000000e+00 2.621701e-15 4.922642e+01 9.30000
## attr(,"formula")
## ~rincome | rprice
## <environment: 0x000000017ced520>
# ftest[5] corresponds to F-value
fval1 = ftest1[5]
```

The following tests HO: _b[rincome] - 1 = 0 & _b[rprice] = 0
ftest2 = waldtest(result_1, ~ rincome - 1 | rprice)
ftest2

14 / 14