Lab 12 - Treatment effects (experiments)

Julieth Santamaria

April 27, 2018

Panel data

PSID example taken from the Econometrics Academy webpage.

1. Load the data and set the dataset as panel

```
library(plm)
```

```
## Warning: package 'plm' was built under R version 3.5.3
```

```
## Loading required package: Formula
```

```
setwd("G:/My Drive/U of M/TA/TA APEC3003/APEC 3003 - 2019/APEC 3003 R work/labs/")
load("../data/panelwage.Rdata")
panel_wage <- pdata.frame(panel_wage, index=c("id","t"))
```

2. Estimate the effect of experience using a pooled model

```
# Pooled OLS estimator
pooled <- plm(lwage ~ exp+exp2+wks+ed, data=panel_wage, model= "pooling")
summary(pooled)
```

```
## Pooling Model
##
## Call:
## plm(formula = lwage ~ exp + exp2 + wks + ed, data = panel_wage,
##     model = "pooling")
##
## Balanced Panel: n = 595, T = 7, N = 4165
##
## Residuals:
##        Min.     1st Qu.      Median     3rd Qu.        Max.
## -2.16057670 -0.25034526  0.00027256  0.26792139  2.12969386
##
## Coefficients:
##                Estimate  Std. Error  t-value  Pr(>|t|)
## (Intercept)  4.9080e+00  6.7330e-02  72.8945 < 2.2e-16 ***
## exp          4.4675e-02  2.3929e-03  18.6701 < 2.2e-16 ***
## exp2        -7.1563e-04  5.2794e-05 -13.5552 < 2.2e-16 ***
## wks          5.8270e-03  1.1826e-03   4.9271 8.673e-07 ***
## ed           7.6041e-02  2.2266e-03  34.1511 < 2.2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    886.9
## Residual Sum of Squares: 635.41
## R-Squared:      0.28356
## Adj. R-Squared: 0.28287
## F-statistic: 411.624 on 4 and 4160 DF, p-value: < 2.22e-16
```

3. How do your results change using fixed effects at the individual level?

```
fixed1 <- plm(lwage ~ exp+exp2+wks+ed, data=panel_wage, model= "within")
summary(fixed1)
```

```
## Oneway (individual) effect Within Model
##
## Call:
## plm(formula = lwage ~ exp + exp2 + wks + ed, data = panel_wage,
##     model = "within")
##
## Balanced Panel: n = 595, T = 7, N = 4165
##
## Residuals:
##       Min.    1st Qu.     Median    3rd Qu.       Max.
## -1.8120879 -0.0511128  0.0037112  0.0614250  1.9434065
##
## Coefficients:
##         Estimate  Std. Error t-value  Pr(>|t|)
## exp   1.1379e-01  2.4689e-03 46.0888 < 2.2e-16 ***
## exp2 -4.2437e-04  5.4632e-05 -7.7678 1.036e-14 ***
## wks   8.3588e-04  5.9967e-04  1.3939    0.1634
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    240.65
## Residual Sum of Squares: 82.632
## R-Squared:      0.65663
## Adj. R-Squared: 0.59916
## F-statistic: 2273.74 on 3 and 3567 DF, p-value: < 2.22e-16
```

4. Add both, individual and time fixed effects to the last model.

```
fixed2 <- plm(lwage ~ exp+exp2+wks+ed, data=panel_wage, model= "within", effect = "twoways")
summary(fixed2)
```

```
## Twoways effects Within Model
##
## Call:
## plm(formula = lwage ~ exp + exp2 + wks + ed, data = panel_wage,
##     effect = "twoways", model = "within")
##
## Balanced Panel: n = 595, T = 7, N = 4165
##
## Residuals:
##      Min.   1st Qu.    Median   3rd Qu.      Max.
## -1.827312 -0.053372  0.003455  0.063260  1.925563
##
## Coefficients:
##         Estimate  Std. Error t-value  Pr(>|t|)
## exp2 -4.0505e-04  5.4568e-05 -7.4230 1.425e-13 ***
## wks   6.7996e-04  5.9893e-04  1.1353    0.2563
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    83.178
```

```
## Residual Sum of Squares: 81.854
## R-Squared:      0.015918
## Adj. R-Squared: -0.1504
## F-statistic: 28.8088 on 2 and 3562 DF, p-value: 3.8779e-13
```

5. Estimate now the first differences estimator

```
firstdiff <- plm(lwage ~ exp+exp2+wks+ed, data=panel_wage, model= "fd")
summary(firstdiff)
```

```
## Oneway (individual) effect First-Difference Model
##
## Call:
## plm(formula = lwage ~ exp + exp2 + wks + ed, data = panel_wage,
##     model = "fd")
##
## Balanced Panel: n = 595, T = 7, N = 4165
## Observations used in estimation: 3570
##
## Residuals:
##       Min.    1st Qu.     Median    3rd Qu.       Max.
## -2.1131555 -0.0654718 -0.0095751  0.0483881  2.3295637
##
## Coefficients:
##                Estimate  Std. Error t-value  Pr(>|t|)
## (Intercept)  0.11706540  0.00631057 18.5507 < 2.2e-16 ***
## exp2        -0.00053212  0.00013927 -3.8207 0.0001354 ***
## wks         -0.00026826  0.00056483 -0.4749 0.6348525
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Total Sum of Squares:    118.06
## Residual Sum of Squares: 117.58
## R-Squared:      0.004108
## Adj. R-Squared: 0.0035496
## F-statistic: 7.35691 on 2 and 3567 DF, p-value: 0.0006479
```

Treatment Effects

Policy makers are usually interested in causal effects. In experiments, assignment to treatment is usually done at random.. We will calculate three treatment effects that differ on the groups we compare:

- ATE (**A**verage **T**reatment **E**ffect): $E(y_{1i}) - E(y_{0i})$
- ITT (**I**ntention **T**o **T**reat): $E(y|g = 1) - E(y|g = 0)$
- ATT (**A**verage **T**reatment effect on the **T**reated): $E(y_{1i}|w = 1) - E(y_{0i}|w = 1)$

To calculate the causal effect in a randomized experiment we just need to calculate the difference in means using the following model:

$Y_i = \beta_0 + \beta_1 Treatment_i + \varepsilon_i$

ABC Issues

- Attition
- Balance
- Compliance

Application

This lab uses a dataset used in the paper by Woodbury and Spiegelman (AER, 1987).

Motivation

UI may prolong jobless spell beyond what it would be in the absence of unemployment benefit:

1. UI benefits may act as subsidy to additional job search
2. UI benefits are also a subsidy to leisure

Experiments

Between mid 1984 and mid 1985, New claimants for Unemployment Insurance were randomly assigned to one of two experiments that were designed to speed up the return to work.

Claimant Experiment: "A random sample were instructed that they would qualify for a cash bonus of $500 if they found a job (of 30 hours or more per week) within 11 weeks of filing the UI claim, and if they held that job for 4 months"

Employer Experiment: "A second random sample were told that their next employer would qualify for a cash bonus of $500 if they, the claimants, found a job within 11 weeks of filing the UI claim, and if they retained that job for four months.

Eligibility Criteria

An individual had to:

1. File a claim for UI between July 29, 1984 and November 17,1984
2. Be eligible for 26 weeks of UI benefits
3. Register with one of the 22 Job service offices in northern and central Illinois
4. Be at least 20 years old, but less than 55

Each claimant was assigned to one of three groups: control group, Claimant Experiment and Employer Experiment based on the last two digits of his/her SSN. Claimants were asked to sign an "agreement to participate".

Treatment effects

1. Upload the dataset. Notice that the original dataset is in Stata format (.dta). Stata is a statistical program similar to R.

```
setwd("G:/My Drive/U of M/TA/TA APEC3003/APEC 3003 - 2019/APEC 3003 R work/labs/")
load("../data/illinois.Rdata")
```

2. Identify the number of observations in each experimental arm and the proportion of treated that signed the agreement.

control: Control group

jsie: Job Search Incentive experiment group (Claimant Experiment - CE)

hie: Hiring Incentive Experiment Group (Employer Experiment - EE)

lagree: Agree to participate

```
table(illinois$control)
```

```
##
##    0    1
## 8149 3952
```

```r
table(illinois$jsie)
```

```
##
##    0    1
## 7915 4186
```

```r
table(illinois$hie)
```

```
##
##    0    1
## 8138 3963
```

```r
table(illinois$lagree,illinois$jsie) # Check who agreed, given that you are in the CE
```

```
##
##        0    1
##   0 1211  443
##   1 2586 3527
```

```r
table(illinois$lagree,illinois$hie)  # Check who agreed, given that you are in the EE
```

```
##
##        0    1
##   0  443 1211
##   1 3527 2586
```

```r
table(illinois$jsie,illinois$hie)   # You're either in one experiment or the other
```

```
##
##        0    1
##   0 3952 3963
##   1 4186    0
```

3. Calculate the 'ITT' on the following outcomes:

- Benefits paid ($)
  - *srben1*: State regular benefits, first spell
  - *benepaid*: Total benefits paid, first spell. It refers to the sum of state regular benefits and Federal Supplemental Compensation
  - *srbenyr*: State regular benefits paid during the full year for each claimant
  - *benpdbye*: Total benefits paid during the full year
- Weeks of insured unemployment
  - *wkpaid*: Weeks of benefits, first spell
  - *wkspdbye*: Weeks of benefits, benefit year

```r
# Check the slides that come with this lab


illinois <- within(illinois,{
  # Create group treatments

  group.jsie<-NA
  group.jsie[jsie==1]<-1    #1 if in the claimant experiment
  group.jsie[control==1]<-0 #0 if in the control group

  group.hie<-NA
  group.hie[hie==1]<-1      #1 if in the employer experiment
  group.hie[control==1]<-0  #0 if in the control group
```

```
})
```

State regular benefits, first spell

```
ITT.CE<-lm(srben1~group.jsie,data=illinois)  #ITT for CE
summary(ITT.CE)
```

```
##
## Call:
## lm(formula = srben1 ~ group.jsie, data = illinois)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2266.7 -1576.7  -341.5  1425.3  3359.8
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2266.74      27.42  82.671  < 2e-16 ***
## group.jsie   -192.53      38.23  -5.036 4.85e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1724 on 8136 degrees of freedom
##   (3963 observations deleted due to missingness)
## Multiple R-squared:  0.003108,   Adjusted R-squared:  0.002985
## F-statistic: 25.36 on 1 and 8136 DF,  p-value: 4.855e-07
```

```
ITT.EE<-lm(srben1~group.hie,data=illinois)   # ITT for EE
summary(ITT.EE)
```

```
##
## Call:
## lm(formula = srben1 ~ group.hie, data = illinois)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2266.7 -1603.7  -312.7  1436.3  3275.3
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2266.74      27.63  82.028  < 2e-16 ***
## group.hie    -108.01      39.05  -2.766  0.00569 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1737 on 7913 degrees of freedom
##   (4186 observations deleted due to missingness)
## Multiple R-squared:  0.0009657,  Adjusted R-squared:  0.0008394
## F-statistic: 7.649 on 1 and 7913 DF,  p-value: 0.005695
```

4. Now calculate the ATT for the same outcomes

```
# Create "treated" dummies
illinois<-within(illinois,{
  treated.jsie<-NA
```

```
  treated.jsie[group.jsie==1 & lagree==1]<-1   # 1 only if actually participated in CE
  treated.jsie[control==1]<-0

  treated.hie<-NA
  treated.hie[group.hie==1 & lagree==1]<-1 # 1 only if actually participated in EE
  treated.hie[control==1]<-0
})

ATT.CE<-lm(srben1~treated.jsie,data=illinois)  #ATT for CE
summary(ATT.CE)
```

```
##
## Call:
## lm(formula = srben1 ~ treated.jsie, data = illinois)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2266.7 -1567.7  -348.7  1424.3  3383.3
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2266.74      27.46  82.556  < 2e-16 ***
## treated.jsie  -216.06      39.98  -5.404 6.72e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1726 on 7477 degrees of freedom
##   (4622 observations deleted due to missingness)
## Multiple R-squared:  0.00389,    Adjusted R-squared:  0.003757
## F-statistic:  29.2 on 1 and 7477 DF,  p-value: 6.72e-08
```

```
ATT.EE<-lm(srben1~treated.hie,data=illinois)   # ATT for EE
summary(ATT.EE)
```

```
##
## Call:
## lm(formula = srben1 ~ treated.hie, data = illinois)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -2266.7 -1603.7  -324.3  1421.9  3341.7
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  2266.74      27.52  82.364  < 2e-16 ***
## treated.hie  -174.47      43.76  -3.987 6.76e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 1730 on 6536 degrees of freedom
##   (5563 observations deleted due to missingness)
## Multiple R-squared:  0.002426,   Adjusted R-squared:  0.002274
## F-statistic:  15.9 on 1 and 6536 DF,  p-value: 6.765e-05
```