CA5aVdVeen

Ties van der Veen

27-9-2019

Rebecca Costello - Ties van der Veen - Niels van Opstal

I. Introduction to the assignment

Packages

```
library(foreign)
library(tidyverse)
library(ggdag)
library(dplyr)
library(tinytex)
library(jtools)
library(huxtable)
library(summarytools)
library(ggstance)
library(pwr)
library(knitr)
library(lemon)
library(AER)
library(lubridate)
```

```
knit_print.data.frame <- lemon_print
```

```
st_options(plain.ascii = FALSE, style = "rmarkdown")
st_css()
```

```
## <style type="text/css">
##   img {   background-color: transparent;   border: 0; }  .st-table td, .st-table th {   padding: 8px;
```

Data

```
theUrl_ca5a <- "https://surfdrive.surf.nl/files/index.php/s/7Mn58Z3iMOJWwTV/download"
minwage <- read.dta (file = theUrl_ca5a)
```

II. Potential outcomes

(a)

Y(0,i) = no increase in minimum wage, no increase in employees.

Y(1,i) = increase in minimum wage, increase in employees.

(b)

Employment is affected by a change in minimum wage because of the costs for the employer. Higher costs per employee per hour could lead to a necessity of firing an employee to keep costs similar to what they were before the change. On the other hand, a higher minimum wage means that all firms that paid a wage inbetween the previous minimum and the new one (so above 4,25 but below 5,05) will all be at a level playing field. This removes competition on employee cost level, making this a less profitable area to cut costs in. Thus, employment would not decrease.

III. Descriptive statistics

(a)

```
minwage %>%
  group_by(time, sample, state) %>%
  summarise(number = n())
```

```
## # A tibble: 8 x 4
## # Groups:   time, sample [4]
##    time sample state number
##   <int>  <int> <int>  <int>
## 1     0      0     0     13
## 2     0      0     1     46
## 3     0      1     0     66
## 4     0      1     1    285
## 5     1      0     0     13
## 6     1      0     1     46
## 7     1      1     0     66
## 8     1      1     1    285
```

As we can see in this table, the overall number of restaurants interviewed is 410 (add up all results at time 0 or 1). Number of restaurants within the relevant sample for wave 1 is 285 + 66 = 351. 285 of the relevant restaurants are located in NJ, rest (66) is in PA.

(b)

```
minwage_sample <- minwage[minwage$sample==1, ]
min(minwage_sample$fte, na.rm=TRUE)
```

```
## [1] 3
```

```
max(minwage_sample$fte, na.rm=TRUE)
```

```
## [1] 80
```

So the minimum is 3, and the maximum is 80 fte.

(c)

```r
min(minwage_sample$wage_st, na.rm=TRUE)
```

```
## [1] 4.25
```

```r
max(minwage_sample$wage_st, na.rm=TRUE)
```

```
## [1] 6.25
```

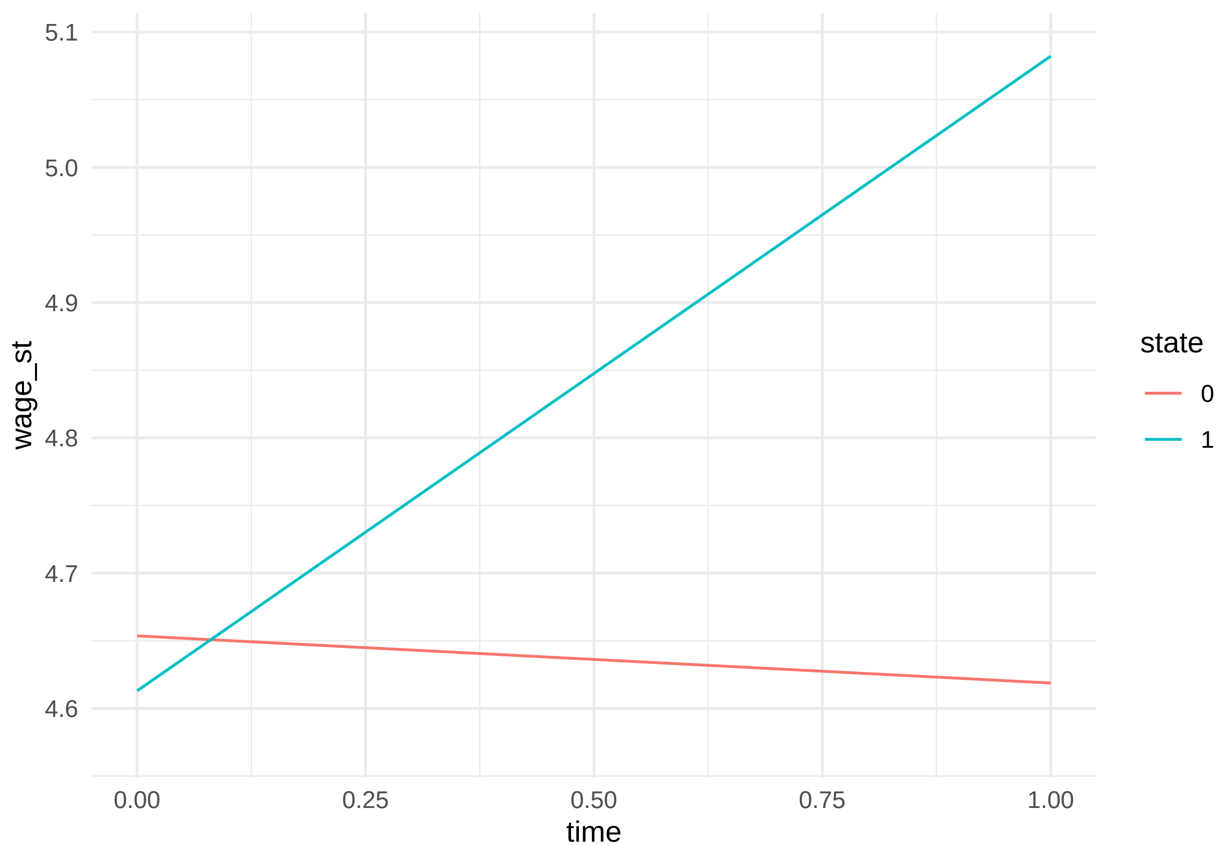
So the minimum is 4.25, and the maximum is 6.25.

IV. D-I-D plot

```r
minwage_sample$state <- as.factor(minwage_sample$state)

ggplot(minwage_sample, aes(time, wage_st, col = state)) +
  stat_summary(geom = 'line') +
  theme_minimal()
```

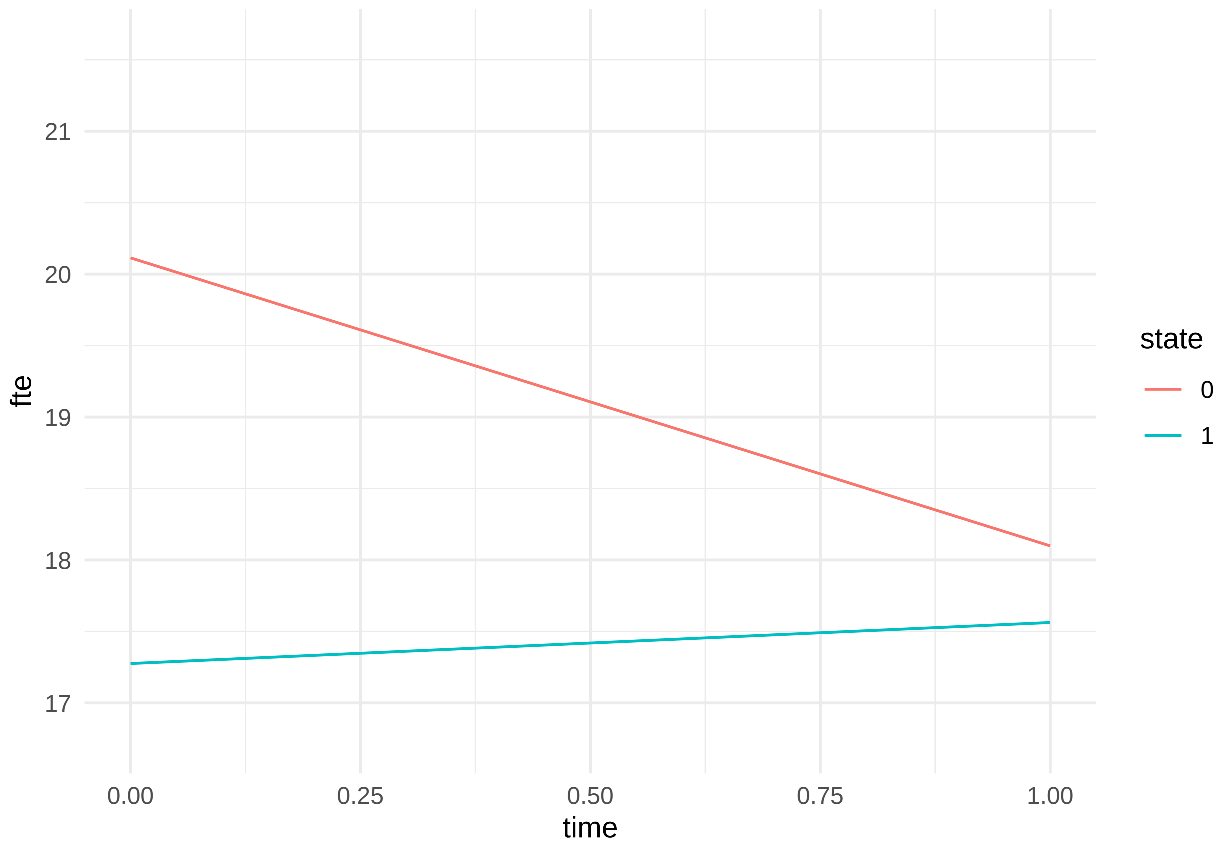
```
## No summary function supplied, defaulting to `mean_se()
```



```r
minwage_sample$state <- as.factor(minwage_sample$state)

ggplot(minwage_sample, aes(time, fte, col = state)) +
  stat_summary(geom = 'line') +
  theme_minimal()
```

```
## No summary function supplied, defaulting to `mean_se()
```



V. D-I-D by hand

```r
NJ0=sapply(subset(minwage_sample, time==0 & state==1, select=wage_st), mean)
NJ1=sapply(subset(minwage_sample, time==1 & state==1, select=wage_st), mean)
PE0=sapply(subset(minwage_sample, time==0 & state==0, select=wage_st), mean)
PE1=sapply(subset(minwage_sample, time==1 & state==0, select=wage_st), mean)
```

```r
(NJ1-NJ0)-(PE1-PE0)
```

```
##   wage_st
## 0.5040066
```

There is a positive difference in the differences of starting wage between NJ and PE. This tells us that either NJ's wages increased, PE's wages decreased, or both. Either way, NJ's new wages are higher than PE's.

This can be considered a causal estimate if all other factors are kept consistent.

```
PE0=sapply(subset(minwage_sample, time==0 & state==0, select=fte), mean)
PE1=sapply(subset(minwage_sample, time==1 & state==0, select=fte), mean)
NJ0=sapply(subset(minwage_sample, time==0 & state==1, select=fte), mean)
NJ1=sapply(subset(minwage_sample, time==1 & state==1, select=fte), mean)
(NJ1-NJ0)-(PE1-PE0)
```

```
##      fte
## 2.301994
```

Here we see that the fte is relatively higher in NJ after the minimum wage change than in PE.

VI. D-I-D regression

(a)

```
didwage <- lm(wage_st ~ time*state, data=minwage_sample)
summ(didwage)
```

```
## MODEL INFO:
## Observations: 702
## Dependent Variable: wage_st
## Type: OLS linear regression
##
## MODEL FIT:
## F(3,698) = 156.91, p = 0.00
## R² = 0.40
## Adj. R² = 0.40
##
## Standard errors: OLS
## ------------------------------------------------
##                      Est.   S.E.   t val.      p
## ----------------- -------- ------ -------- ------
## (Intercept)          4.65   0.03   136.34   0.00
## time                -0.03   0.05    -0.72   0.47
## state1              -0.04   0.04    -1.07   0.28
## time:state1          0.50   0.05     9.41   0.00
## ------------------------------------------------
```

```
didfte <- lm(fte ~ time*state, data=minwage_sample)
summ(didfte)
```

```
## MODEL INFO:
## Observations: 702
## Dependent Variable: fte
## Type: OLS linear regression
##
## MODEL FIT:
## F(3,698) = 1.87, p = 0.13
```

```
## R² = 0.01
## Adj. R² = 0.00
##
## Standard errors: OLS
## ------------------------------------------------
##                       Est.   S.E.   t val.      p
## ----------------- -------- ------ -------- ------
## (Intercept)          20.11   1.10    18.23   0.00
## time                 -2.02   1.56    -1.29   0.20
## state1               -2.84   1.22    -2.32   0.02
## time:state1           2.30   1.73     1.33   0.18
## ------------------------------------------------
```

As can be seen from these regression estimates (time:state1), the earlier predictions were accurate. However, the one for fte is not statistically significant.

(b)

```
didwage <- lm(wage_st ~ time*state + co_owned + factor(chain), data=minwage_sample)
summ(didwage)
```

```
## MODEL INFO:
## Observations: 702
## Dependent Variable: wage_st
## Type: OLS linear regression
##
## MODEL FIT:
## F(7,694) = 73.61, p = 0.00
## R² = 0.43
## Adj. R² = 0.42
##
## Standard errors: OLS
## ---------------------------------------------------
##                          Est.   S.E.   t val.      p
## -------------------- -------- ------ -------- ------
## (Intercept)              4.59   0.04   128.51   0.00
## time                    -0.03   0.05    -0.73   0.46
## state1                  -0.04   0.04    -0.98   0.33
## co_owned                 0.07   0.02     3.04   0.00
## factor(chain)2           0.02   0.03     0.58   0.56
## factor(chain)3           0.05   0.03     1.60   0.11
## factor(chain)4           0.12   0.03     3.50   0.00
## time:state1              0.50   0.05     9.57   0.00
## ---------------------------------------------------
```

```
didfte <- lm(fte ~ time*state + co_owned + factor(chain), data=minwage_sample)
summ(didfte)
```

```
## MODEL INFO:
## Observations: 702
## Dependent Variable: fte
## Type: OLS linear regression
```

```
##
## MODEL FIT:
## F(7,694) = 26.22, p = 0.00
## R² = 0.21
## Adj. R² = 0.20
##
## Standard errors: OLS
## ---------------------------------------------------
##                          Est.   S.E.   t val.      p
## -------------------- -------- ------ -------- ------
## (Intercept)             22.56   1.05    21.42   0.00
## time                    -2.02   1.40    -1.44   0.15
## state1                  -2.14   1.10    -1.95   0.05
## co_owned                -1.01   0.72    -1.41   0.16
## factor(chain)2         -10.16   0.84   -12.04   0.00
## factor(chain)3          -1.35   0.86    -1.57   0.12
## factor(chain)4          -1.37   0.97    -1.41   0.16
## time:state1              2.30   1.55     1.48   0.14
## ---------------------------------------------------
```

(c)

For starting wage, there is no change other than a minor one in the t-value. For fte, there is a change in the standard error, t-value and significance, but these changes do not change the interpretation. We did not expect a change in results, as that would signal significant differences between these stores and their hiring policies, which would be odd seeing as they are in the same branch of business.