Lab14

## Setup

setwd("C:/Users/22700/Desktop")  
library(data.table)  
library(ggplot2)  
library(stargazer)

##   
## Please cite as:

## Hlavac, Marek (2018). stargazer: Well-Formatted Regression and Summary Statistics Tables.

## R package version 5.2.2. https://CRAN.R-project.org/package=stargazer

load("fastfood.RData")  
load("fastfood3.RData")  
load("fastfood4.RData")

## Analysis and Results

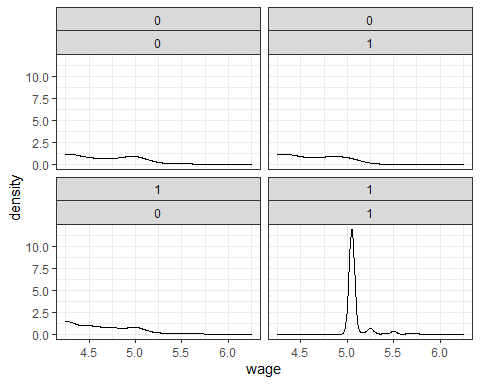
head(dt.fastfood)

## emptot gap demp state chain co\_owned atmin meals wage hrsopen pmeal  
## 1: 40.50 0 -16.50 0 1 0 NA 2 NA 16.5 2.58  
## 2: 13.75 0 -2.25 0 2 0 NA 2 NA 13.0 4.26  
## 3: 8.50 0 2.00 0 2 1 NA 2 NA 10.0 4.02  
## 4: 34.00 0 -14.00 0 4 1 0 2 5.0 12.0 3.48  
## 5: 24.00 0 11.50 0 4 1 0 3 5.5 12.0 3.29  
## 6: 20.50 0 NA 0 4 1 0 2 5.0 12.0 2.59  
## fracft time id  
## 1: 0.7407407 0 1  
## 2: 0.4727273 0 2  
## 3: 0.3529412 0 3  
## 4: 0.5882353 0 4  
## 5: 0.2500000 0 5  
## 6: 0.0000000 0 6

Plots # The following plot shows the change in the distribution of wages for the treatment and control group, before and after the change. We can see that both groups had a very similar wage distribution before the change in minimum wage was implemented in NJ. After the change, all restaurants in NJ that were paying less than $5.05/hour started paying at that rate, complying with the new legislation.

plot1 <- ggplot( data = dt.fastfood, aes(x = wage))  
plot1 + geom\_density() + facet\_wrap( ~ state + time) + theme\_bw()

## Warning: Removed 37 rows containing non-finite values (stat\_density).



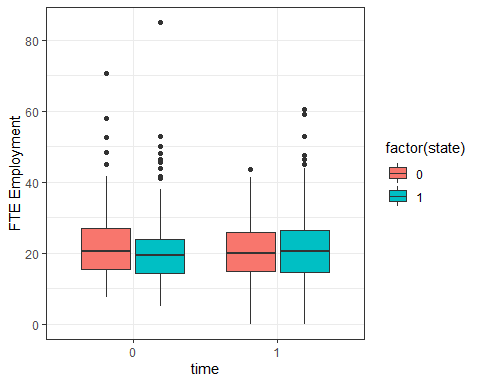
# Please note that the plot above is not testing the assumption of a common trend prior to treatment. In order to test the common trend assumption, one needs to have data for the control and treatment groups at more than one time point before the treatment takes place. This data set does not allow us to test that assumption as we only have two data points - before and after treatment.

# Change in employment:

# What kind of plot does economic theory predict? • How does this plot differ from our expectations?

qplot( data = dt.fastfood, x = factor(time), y = emptot  
, fill = factor(state)  
, geom = "boxplot") + theme\_bw() + xlab("time") + ylab("FTE Employment")

## Warning: Removed 21 rows containing non-finite values (stat\_boxplot).



# Economic theory predicts a reduction in employment resulting from the increase in minimum wage. The plot shows us that this did not happen, on the contrary, employment seems to have increased slightly for the treated group (NJ).

## Mean of the variables

# Use the full data set to build a table with the before and after means for treatment and control groups. For this purpose we convert our table to ‘data.table’ format. Data tables have all the features of data frames and more. As in data frames, you can write the table name followed by square brackets: ‘tablename[ , ]’, where the first space before the comma refers to the table’s rows, and the space after the comma refers to the table’s columns. You can also add a second comma ‘tablename[ , , ]’ where the space after the second comma refers to the grouping variables: ‘tablename[rows,columns, group by]’.

dt.bf.aft <- data.table(dt.fastfood) # Create a new table called dt.bf.aft  
dt.bf.aft <- dt.bf.aft[, list( # Create a list of the columns of your new table  
mean\_emptot = mean(emptot , na.rm=TRUE)  
, mean\_wage = mean(wage , na.rm=TRUE)  
, mean\_pmeal = mean(pmeal , na.rm=TRUE)  
, mean\_hrsopen = mean(hrsopen , na.rm=TRUE)  
), by=list(state, time)] # Specifiy the list of grouping variables  
dt.bf.aft

## state time mean\_emptot mean\_wage mean\_pmeal mean\_hrsopen  
## 1: 0 0 23.33117 4.630132 3.042368 14.52532  
## 2: 1 0 20.44557 4.610971 3.356471 14.42025  
## 3: 0 1 21.16558 4.617465 3.026620 14.65385  
## 4: 1 1 21.02743 5.080947 3.416809 14.41484

# At T1, average employment was 23.3 full-time equivalent (FTE) workers per store in Pennsylvania, compared with an average of 20.4 in New Jersey. Starting wages were very similar among stores in the two states, although the average price of a “full meal” was significantly higher in New Jersey. There were no significant cross-state differences in average hours of operation, or the fraction of full-time workers (Card and Krueguer, 1993). Despite the increase in wages, FTE employment increased in NJ relative to PA. Whereas NJ stores were initially smaller, employment gains in NJ coupled with losses in PA led to a small and statistically insignificant interstate difference at T2. Only two other variables show a relative change between T1 and T2: the fraction of full-time employees and the price of a meal. Both variables increased in NJ relative to PA.

## Create the same table, now using only the clean data:

dt.bf.aft.clean <- dt.fastfood[!is.na(wage),]  
dt.bf.aft.clean <- dt.bf.aft.clean[!is.na(pmeal),]  
dt.bf.aft.clean <- dt.bf.aft.clean[!is.na(emptot),]  
dt.bf.aft.clean <- dt.bf.aft.clean[!is.na(hrsopen),]  
dt.bf.aft.clean <- dt.bf.aft.clean[!is.na(emptot),]  
dt.bf.aft.clean <- data.table(dt.fastfood.clean)  
dt.bf.aft.clean <- dt.bf.aft.clean[, list(  
mean\_emptot = mean(emptot , na.rm=TRUE)  
, mean\_wage = mean(wage , na.rm=TRUE)  
, mean\_pmeal = mean(pmeal , na.rm=TRUE)  
, mean\_hrsopen = mean(hrsopen , na.rm=TRUE)  
), by=list(state, time)]  
dt.bf.aft.clean

## state time mean\_emptot mean\_wage mean\_pmeal mean\_hrsopen  
## 1: 0 0 23.62687 4.651343 3.054062 14.57463  
## 2: 1 0 20.51397 4.609655 3.377033 14.41207  
## 3: 0 1 21.50000 4.618788 3.006406 14.72727  
## 4: 1 1 20.71293 5.082141 3.451808 14.40053

# We can use t-tests to check if differences in means between NJ and PA are statistically significant: Difference in FTE employment between NJ and PA at T0.

t.test( dt.fastfood.clean[state==0 & time==0, emptot]  
, dt.fastfood.clean[state==1 & time==0, emptot])

##   
## Welch Two Sample t-test  
##   
## data: dt.fastfood.clean[state == 0 & time == 0, emptot] and dt.fastfood.clean[state == 1 & time == 0, emptot]  
## t = 1.9515, df = 84.174, p-value = 0.05432  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -0.05909098 6.28489129  
## sample estimates:  
## mean of x mean of y   
## 23.62687 20.51397

#Difference in FTE employment between NJ and PA at T1

t.test( dt.fastfood.clean[state==0 & time==1, emptot]  
, dt.fastfood.clean[state==1 & time==1, emptot])

##   
## Welch Two Sample t-test  
##   
## data: dt.fastfood.clean[state == 0 & time == 1, emptot] and dt.fastfood.clean[state == 1 & time == 1, emptot]  
## t = 0.66779, df = 103.74, p-value = 0.5058  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## -1.550250 3.124388  
## sample estimates:  
## mean of x mean of y   
## 21.50000 20.71293

## Difference in Differences

# The differences-in-differences strategy amounts to comparing the change in mean FTE in NJ to the change in mean FTE in PA.

(21.02743-20.44557) - (21.16558-23.33117)

## [1] 2.74745

## Using the clean clean data (balanced sub sample):

(20.71293-20.51397) - (21.50000-23.62687)

## [1] 2.32583

# Surprisingly, employment rose in NJ relative to PA after the minimum wage change. NJ stores were initially smaller than their PA counterparts but grew relative to PA stores after the rise in the minimum wage. The relative gain (the “difference in differences” of the changes in employment) is 2.75 FTE employees. The relative change between NJ and PA stores is virtually identical when the analysis is restricted to the balanced sub sample (2.32 FTE).

## Regression

# We can estimate the diff-in-diff estimator in a regression framework. The advantages are:

# It is easy to calculate standard errors.

# We can control for other variables which may reduce the residual variance (lead to smaller standarderrors).

# It is easy to include multiple periods.

# We can study treatments with different treatment intensity. (e.g. varying increases in the minimum

# wage for different states).

# Effect on employment

# How do we go from the table/plot to the regression?

lm1 <- lm( emptot ~ time + state + time\*state, data = dt.fastfood.clean)   
stargazer(lm1, type = "text")

##   
## ===============================================  
## Dependent variable:   
## ---------------------------  
## emptot   
## -----------------------------------------------  
## time -2.127   
## (1.639)   
##   
## state -3.113\*\*   
## (1.286)   
##   
## time:state 2.326   
## (1.818)   
##   
## Constant 23.627\*\*\*   
## (1.159)   
##   
## -----------------------------------------------  
## Observations 714   
## R2 0.009   
## Adjusted R2 0.005   
## Residual Std. Error 9.486 (df = 710)   
## F Statistic 2.116\* (df = 3; 710)   
## ===============================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## Are the coefficients statistically significant? How do we interpret the regression coefficients?

# emptot00: average FTE employment at T1 in PA (β0)

# emptot01: average FTE employment at T1 in NJ (β0 + β2)

# emptot10: average FTE employment at T2 in PA (β0 + β1)

# emptot11: average FTE employment at T2 in PA (β0 + β1 + β2 + β3)

# What is the correspondence between the betas and the values from the table?

dt.bf.aft.clean

## state time mean\_emptot mean\_wage mean\_pmeal mean\_hrsopen  
## 1: 0 0 23.62687 4.651343 3.054062 14.57463  
## 2: 1 0 20.51397 4.609655 3.377033 14.41207  
## 3: 0 1 21.50000 4.618788 3.006406 14.72727  
## 4: 1 1 20.71293 5.082141 3.451808 14.40053

# Beta 1 is equal

21.50000 - 23.62687

## [1] -2.12687

# Beta 2 is equal

21.50000 - 23.6268

## [1] -2.1268

# Beta 3 is equal

(20.71293 - 20.51397) - (21.50000 - 23.62687)

## [1] 2.32583

## Add controls for chain and ownership

lm <- lm( emptot ~ time + state + time\*state + factor(chain) + co\_owned , data = dt.fastfood.clean)  
stargazer(lm, type = "text")

##   
## ===============================================  
## Dependent variable:   
## ---------------------------  
## emptot   
## -----------------------------------------------  
## time -2.127   
## (1.479)   
##   
## state -2.400\*\*   
## (1.163)   
##   
## factor(chain)2 -10.440\*\*\*   
## (0.895)   
##   
## factor(chain)3 -1.768\*   
## (0.903)   
##   
## factor(chain)4 -1.235   
## (1.033)   
##   
## co\_owned -1.192   
## (0.754)   
##   
## time:state 2.326   
## (1.641)   
##   
## Constant 26.237\*\*\*   
## (1.115)   
##   
## -----------------------------------------------  
## Observations 714   
## R2 0.197   
## Adjusted R2 0.189   
## Residual Std. Error 8.562 (df = 706)   
## F Statistic 24.769\*\*\* (df = 7; 706)   
## ===============================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

# One possible explanation for the plot above is that, instead of firing employees, fastfood stores found alternative ways to compensate for their cost increase. For instance, we can look at the price of meals and the hours of operation to see if these were impacted by the increase in minimum wage.

## On meal prices

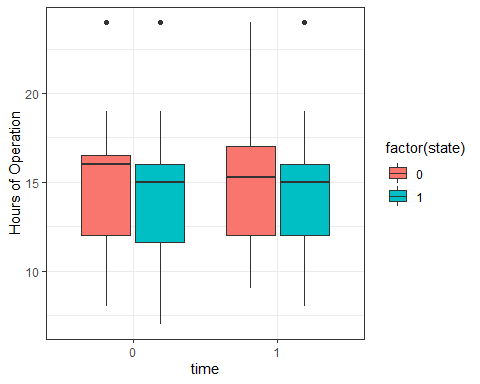
lm <- lm( pmeal ~ time + state + time\*state, data = dt.fastfood.clean)  
stargazer(lm, type = "text")

##   
## ===============================================  
## Dependent variable:   
## ---------------------------  
## pmeal   
## -----------------------------------------------  
## time -0.048   
## (0.113)   
##   
## state 0.323\*\*\*   
## (0.089)   
##   
## time:state 0.122   
## (0.126)   
##   
## Constant 3.054\*\*\*   
## (0.080)   
##   
## -----------------------------------------------  
## Observations 672   
## R2 0.055   
## Adjusted R2 0.051   
## Residual Std. Error 0.641 (df = 668)   
## F Statistic 13.058\*\*\* (df = 3; 668)   
## ===============================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## Change in number of hours of operation:

qplot( data = dt.fastfood, x = factor(time), y = hrsopen, fill = factor(state)  
, geom = "boxplot") + theme\_bw() + xlab("time") + ylab("Hours of Operation")

## Warning: Removed 7 rows containing non-finite values (stat\_boxplot).



## Effect on hours open

lm <- lm( hrsopen ~ time + state + time\*state, data = dt.fastfood.clean)  
stargazer(lm, type = "text")

##   
## ===============================================  
## Dependent variable:   
## ---------------------------  
## hrsopen   
## -----------------------------------------------  
## time 0.153   
## (0.490)   
##   
## state -0.163   
## (0.383)   
##   
## time:state -0.164   
## (0.544)   
##   
## Constant 14.575\*\*\*   
## (0.345)   
##   
## -----------------------------------------------  
## Observations 707   
## R2 0.001   
## Adjusted R2 -0.003   
## Residual Std. Error 2.825 (df = 703)   
## F Statistic 0.302 (df = 3; 703)   
## ===============================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

#Effect on the fraction of full-time employees

lm <- lm( fracft ~ time + state + time\*state, data = dt.fastfood.clean)  
stargazer(lm, type = "text")

##   
## ===============================================  
## Dependent variable:   
## ---------------------------  
## fracft   
## -----------------------------------------------  
## time -0.033   
## (0.042)   
##   
## state -0.021   
## (0.032)   
##   
## time:state 0.055   
## (0.046)   
##   
## Constant 0.355\*\*\*   
## (0.029)   
##   
## -----------------------------------------------  
## Observations 708   
## R2 0.003   
## Adjusted R2 -0.002   
## Residual Std. Error 0.239 (df = 704)   
## F Statistic 0.622 (df = 3; 704)   
## ===============================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

#Alternative Specifications

summary(dt.fastfood.clean$gap)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.00000 0.00000 0.06316 0.08553 0.18824 0.18824

lm <- lm( emptot ~ gap \* time , data = dt.fastfood.clean)  
stargazer(lm, type = "text")

##   
## ===============================================  
## Dependent variable:   
## ---------------------------  
## emptot   
## -----------------------------------------------  
## gap -20.193\*\*\*   
## (6.570)   
##   
## time -1.576   
## (1.064)   
##   
## gap:time 15.653\*   
## (9.291)   
##   
## Constant 22.825\*\*\*   
## (0.753)   
##   
## -----------------------------------------------  
## Observations 714   
## R2 0.014   
## Adjusted R2 0.010   
## Residual Std. Error 9.462 (df = 710)   
## F Statistic 3.346\*\* (df = 3; 710)   
## ===============================================  
## Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

## Conclusions

# Contrary to the central prediction of the textbook model of the minimum wage (. . . ) we find no evidence that the rise in New Jersey’s minimum wage reduced employment at fast-food restaurants in the state. Regardless of whether we compare stores in New Jersey that were affected by the $5.05 minimum to stores in eastern Pennsylvania (where the minimum wage was constant at $4.25 per hour) or to stores in New Jersey that were initially paying $5.00 per hour or more (and were largely unaffected by the new law), we find that the increase in the minimum wage increased employment. We present a wide variety of alternative specifications to probe the robustness of this conclusion. None of the alternatives shows a negative employment effect. (. . . ) We also find no evidence that minimum-wage increases negatively affect the number of McDonald’s outlets opened in a state. Finally, we find that prices of fast-food meals increased in New Jersey relative to Pennsylvania, suggesting that much of the burden of the minimum-wage rise was passed on to consumers. Within New Jersey, however, we find no evidence that prices increased more in stores that were most affected by the minimum-wage rise. Taken as a whole, these findings are difficult to explain with the standard competitive model or with models in which employers face supply constraints (e.g., monopsony or equilibrium search models).

## Limitations

# How could you do this better? \* Telephone survey may be a limitation - get data from more reliable sources. \* Common trend assumption - ideally, we would like to collect data in order to test this assumption. \* We could also consider the impact of the new minimum wage law on the number of store openings and closures