LAB on DiD - Simple DiD

Michael Kummer

Minimum Wages and Employment: A case study of the fast-food industry in New Jersey and Pennsylvania

By David Card and Alan Krueger, *American Economic Review*, 1993

David Edward Card is a Canadian labor economist and Professor of Economics at the University of California, Berkeley.

Alan Bennett Krueger is an American economist, Bendheim Professor of Economics and Public Affairs at Princeton University and Research Associate at the National Bureau of Economic Research.

Introduction

- April 1st 1992: New Jersey (USA) raises minimum wage from $4.25/hour to $5.05/hour.
- This increase gave New Jersey the highest state minimum wage in the US and was strongly opposed by the state's business.
- Conventional economic theory unambiguously predicts that an increase in the minimum wage will lead perfectly competitive employers to cut down employment.

Research Question

What was the impact of the increase in minimum wage on employment?

How would you go about answering this question?

- What do you think would be an appropriate sample to address the research question?
  - Who works for minimum wage?
  - Which industries have a majority of minimum wage workers?
- What would be a good control and a good treatment group?
  - Remember that both groups should follow a common trend prior to treatment (the minimum wage increase).
- At what point(s) in time would you collect the data?
- What data (variables for analysis) would you collect?
  - What are the key variables?
  - What other factors could be impacted by this change besides employment?
- How would you collect the data?
  - Survey the companies in person, by telephone, or by mail. (Card and Krueger)
  - Get payroll records from each company. (Neumark and Wascher)
  - Use Bureau of Labor Statistics' (BLS) data. (Card and Krueger)

Method

Sample: Fast-food restaurants

Why are fast-food restaurants a good choice for this analysis?

- They are a leading employer of low-wage workers;
- They comply with minimum wage regulations and thus are expected to comply with the change in legislation;
- Their job-requirements and products are relatively homogeneous and thus it is easy to obtain reliable measures of employment, wages, and product prices;
- They are known to have a high response rate to telephone surveys.

The authors surveyed fast-food stores in New Jersey and its neighbor eastern Pennsylvania (where the minimum wage law did not suffer any changes during this period) before and after the time of New Jersey's change in minimum wage.

- **Treatment group**: New Jersey fast-food stores.
- **Control group**: Eastern Pennsylvania fast-food stores.

Telephone survey:

- First wave: a little over one month before the increase in NJ's minimum wage.
- Second wave: eight months after the minimum wage increase.

Data

```r
setwd("C:/TopicsDig/Labs/Lab_08W_DiD_Simple")
# change the file's path to your own

require(data.table) # you can use library or require
require(ggplot2)    # you can use library or require
require(stargazer)  # you can use library or require
load("fastfood3.RData")
load("fastfood4.RData")
load("fastfood.RData")
```

Analysis and Results

Explore the data.

```r
head(dt.fastfood)
```

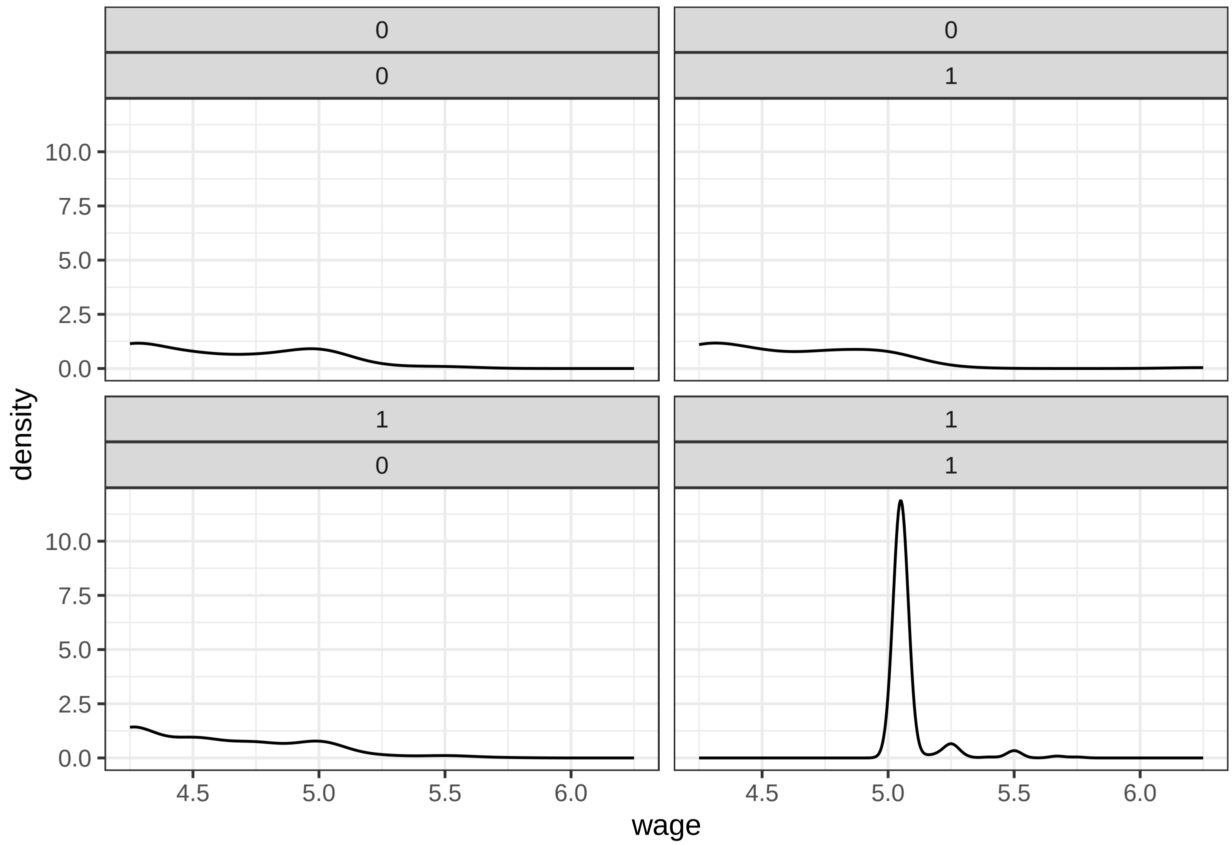
```
##    emptot gap   demp state chain co_owned atmin meals wage hrsopen pmeal
## 1:  40.50   0 -16.50     0     1        0    NA     2   NA    16.5  2.58
## 2:  13.75   0  -2.25     0     2        0    NA     2   NA    13.0  4.26
## 3:   8.50   0   2.00     0     2        1    NA     2   NA    10.0  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

Change in wages:

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()
```

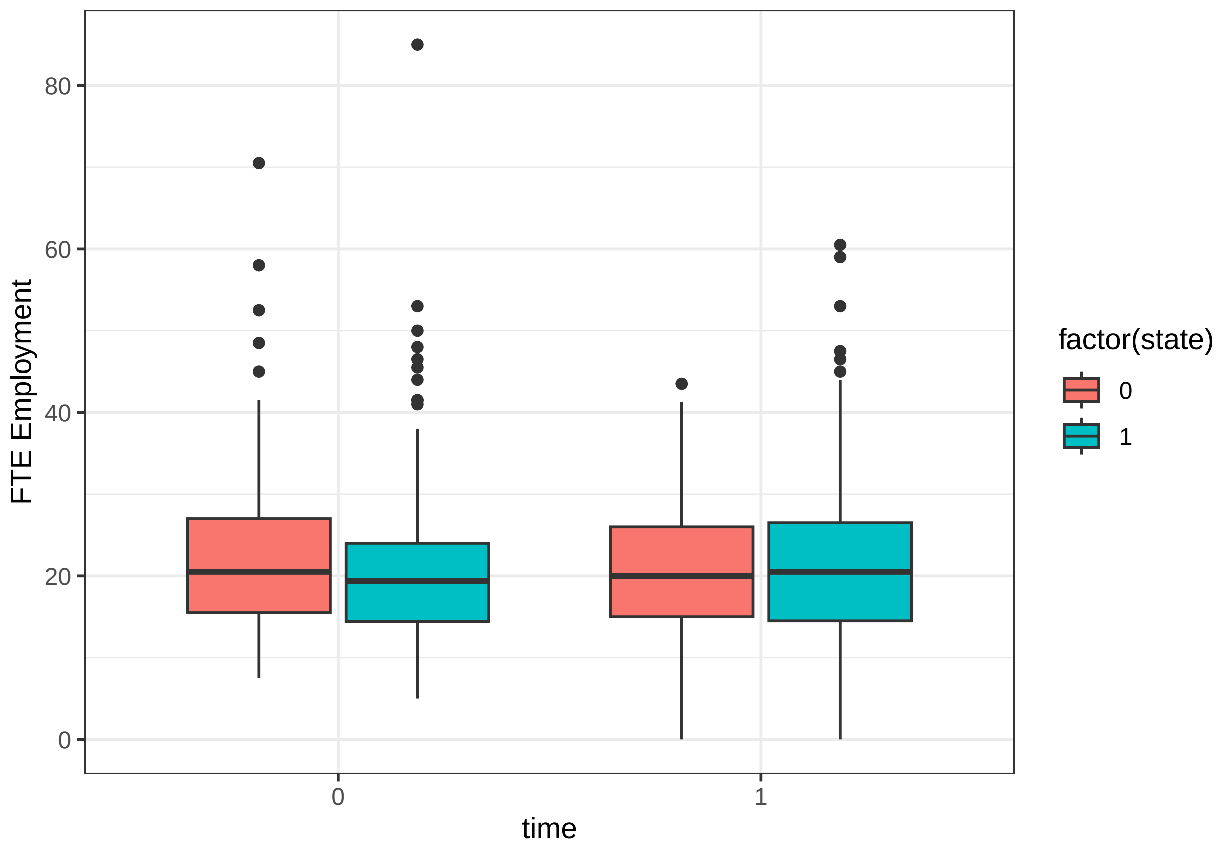


Please note that the plot above is **not testing the assumption of a common trend prior to treatment**. In order to test the common trend assumption, one needs to have data for the control and treatment groups at more than one time point before the treatment takes place. This data set does not allow 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")
```



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).

Means of key 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
## 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
## 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
```

Differences 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.

Treatment Effect $= (\bar{Y}_{NJ,T2} - \bar{Y}_{NJ,T1}) - (\bar{Y}_{PA,T2} - \bar{Y}_{PA,T1})$

Using all data:

```r
(21.02743-20.44557) - (21.16558-23.33117) # 2.74745
```

```
## [1] 2.74745
```

Using the clean clean data (balanced sub sample):

```r
(20.71293-20.51397) - (21.50000-23.62687) # 2.32583
```

```
## [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 standard errors).
- 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?

```r
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
```

```
coeffs <- coefficients(lm1)
coeffs
```

```
## (Intercept)        time       state  time:state
##   23.626866   -2.126866   -3.112900    2.325831
```

Are the coefficients statistically significant?

How do we interpret the regression coefficients?

- emptot00: average FTE employment at T1 in PA (β_0)
- emptot01: average FTE employment at T1 in NJ (β_0 + β_2)
- emptot10: average FTE employment at T2 in PA (β_0 + β_1)
- emptot11: average FTE employment at T2 in PA (β_0 + β_1 + β_2 + β_3)

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
```

(β_0) = 23.62687

(β_1) = emptot10 - emptot00 =

```
21.50000 - 23.62687
```

```
## [1] -2.12687
```

(β_2) = emptot01 - emptot00 =

```
21.50000 - 23.62687
```

```
## [1] -2.12687
```

(β_3) = (emptot11 - emptot10)-(emptot01 - emptot00) =

```r
(20.71293 - 20.51397) - (21.50000 - 23.62687)
```

```
## [1] 2.32583
```

Add controls for chain and ownership

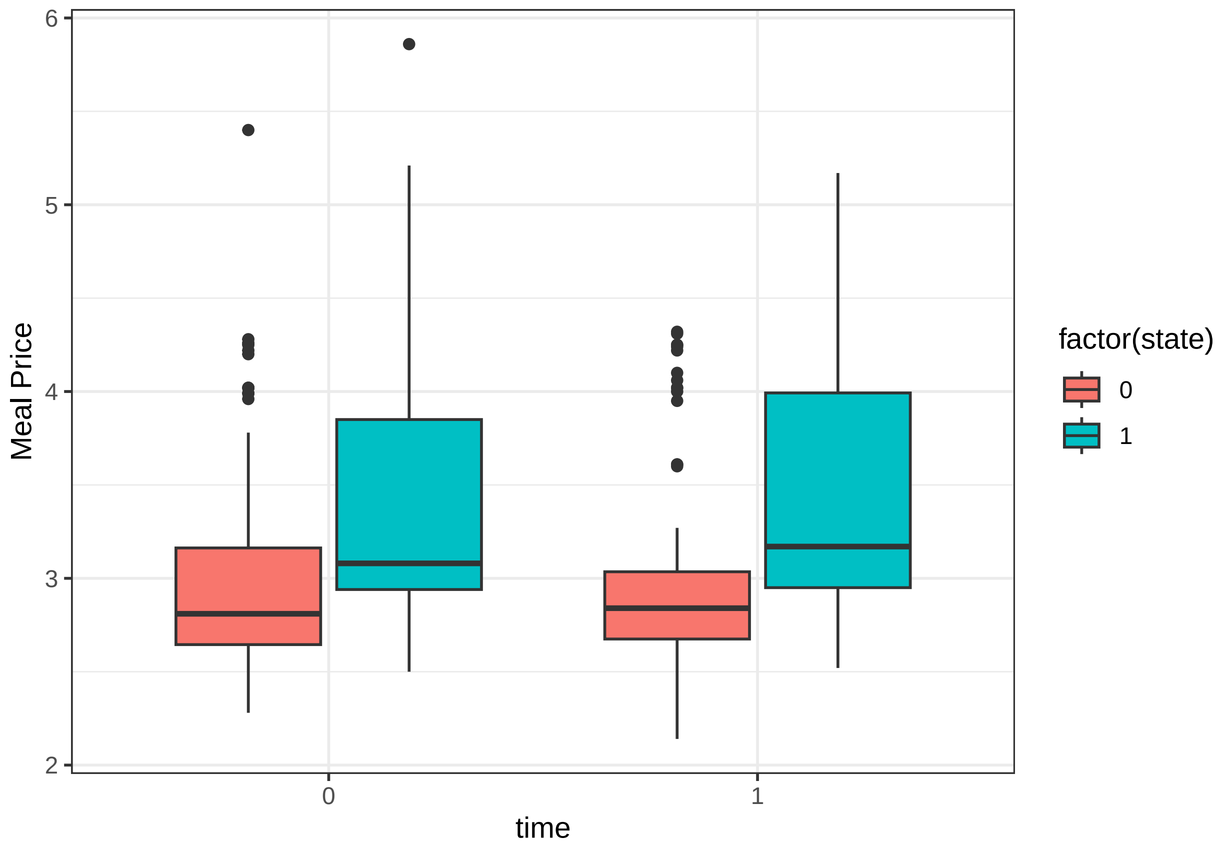
```r
lm <- lm(   emptot ~ time + state + time*state + factor(chain) + co_owned
          , data = dt.fastfood.clean)
stargazer(lm, type = "text")
```

```
##
## ===============================================
##                         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.

Change in meal prices:

```
qplot(   data = dt.fastfood, x = factor(time), y = pmeal
       , fill = factor(state)
       , geom = 'boxplot') + theme_bw() + xlab("time") + ylab("Meal Price")
```



Effect on meal prices

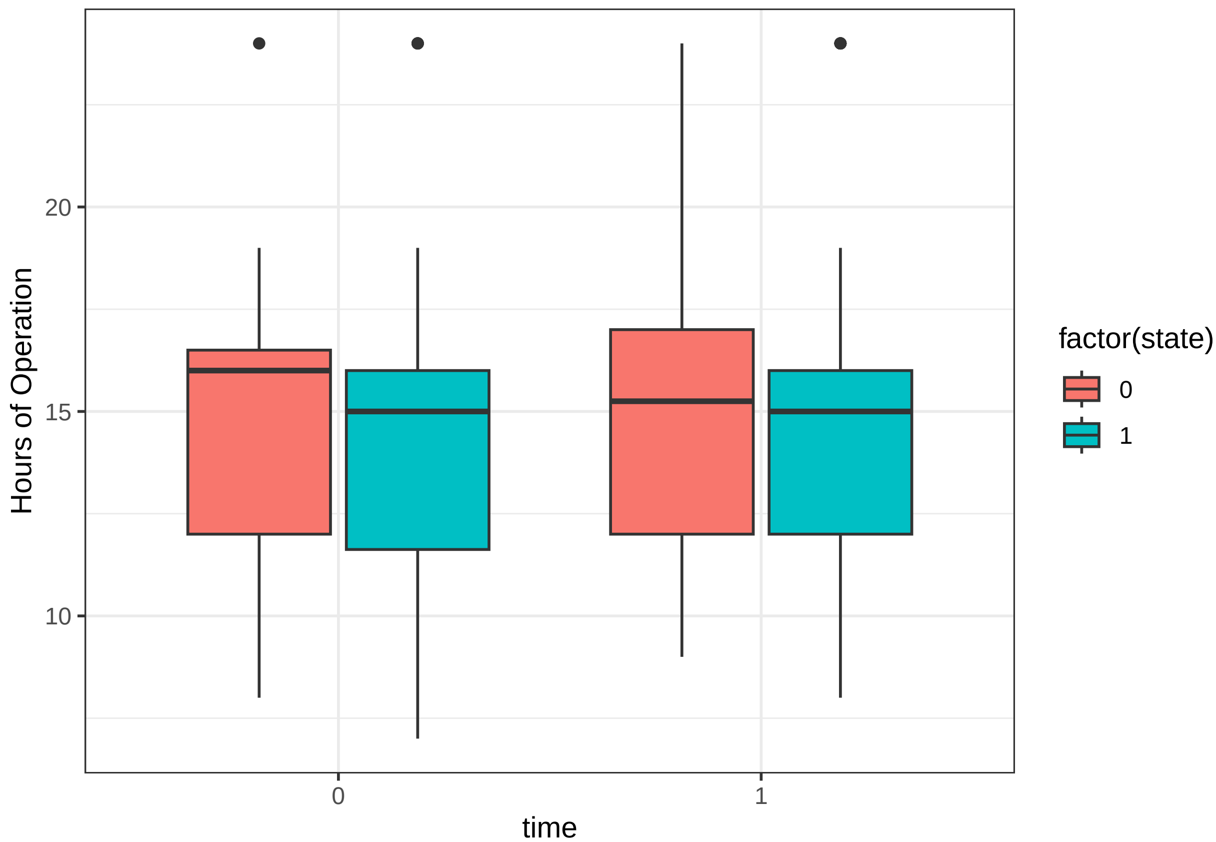
```
lm <- lm(   pmeal ~ time + state + time*state
          , data = dt.fastfood.clean)
stargazer(lm, type = "text")
```

```
##
## ===============================================
##                         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")
```



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

```r
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

The variable *Gap*, measures the intensity of treatment - it is the percent increase in wages that fastfood restaurants had to incur in order to meet the new minimum wage requirements.

```r
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
```

```r
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 equi- librium 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.

Extras: Data preparation

Create variables for analysis.

```
dt.fastfood.orig <- data.table(dt.fastfood.orig)
dt.fastfood.orig <- dt.fastfood.orig[,
           `:=` (
             emptot  = emppt  * 0.5 + empft  + nmgrs
            ,emptot2 = emppt2 * 0.5 + empft2 + nmgrs2
            ,demp    = (emppt2 * 0.5 + empft2 + nmgrs2) - (emppt  * 0.5 + empft + nmgrs)
            ,gap     = ifelse( state == 0, 0
                             , ifelse(  wage_st >= 5.05
                                      , 0
                                      , (5.05-wage_st)/wage_st))
            ,nj      = state
            ,bk      = as.double(chain == 1)
            ,kfc     = as.double(chain == 2)
            ,roys    = as.double(chain == 3)
            ,wendys  = as.double(chain == 4)
            ,pmeal   = psoda + pfry + pentree
            ,pmeal2  = psoda2 + pfry2 + pentree2
            ,dpmeal  = (psoda2 + pfry2 + pentree2) - (psoda + pfry + pentree)
            ,closed  = as.double(status2 == 3)
            ,fracft  = empft / (emppt  * 0.5 + empft  + nmgrs)
            ,fracft2 = ifelse(  emppt2 * 0.5 + empft2 + nmgrs2 > 0
                            , empft2 / (emppt2  * 0.5 + empft2 + nmgrs2)
                            , NA)
            ,atmin   = as.double(wage_st == 4.25)
            ,newmin  = as.double(wage_st == 5.05)
            ,icode   = ifelse( state == 0, 'PA STORE',
                             ifelse(  is.na(wage_st)
                                    , 'NJ Store Bad Wage'
                                    , ifelse(wage_st == 4.25
                                             , 'NJ Store Low-Wage'
                                             , ifelse(wage_st >= 5.0
                                                      , 'NJ Store High-Wage'
                                                      , 'NJ Store Med-Wage')))))]
dt.fastfood.orig <- dt.fastfood.orig[ order(nj) ]
dt.fastfood.orig <- dt.fastfood.orig[!status2 %in% c(0,2,4,5)]
head(dt.fastfood.orig)
```

```
##    sheet chain co_owned state southj centralj northj pa1 pa2 shore ncalls empft
## 1:    46     1        0     0      0        0      0   1   0     0      0  30.0
## 2:    49     2        0     0      0        0      0   1   0     0      0   6.5
## 3:   506     2        1     0      0        0      0   1   0     0      0   3.0
## 4:    56     4        1     0      0        0      0   1   0     0      0  20.0
```

```
## 5:    61     4        1     0      0        0      0   1   0      0      0   6.0
## 6:    62     4        1     0      0        0      0   1   0      0      2   0.0
##    emppt nmgrs wage_st inctime firstinc bonus pctaff meals open hrsopen psoda
## 1:  15.0     3      NA      19       NA     1     NA     2  6.5    16.5  1.03
## 2:   6.5     4      NA      26       NA     0     NA     2 10.0    13.0  1.01
## 3:   7.0     2      NA      13     0.37     0     30     2 11.0    10.0  0.95
## 4:  20.0     4     5.0      26     0.10     1      0     2 10.0    12.0  0.87
## 5:  26.0     5     5.5      52     0.15     1      0     3 10.0    12.0  0.87
## 6:  31.0     5     5.0      26     0.07     0     45     2 10.0    12.0  0.87
##    pfry pentree nregs nregs11 type2 status2  date2 ncalls2 empft2 emppt2 nmgrs2
## 1: 1.03    0.52     3       3     1       1 111792       1    3.5     35      3
## 2: 0.90    2.35     4       3     1       1 111292      NA    0.0     15      4
## 3: 0.74    2.33     3       3     1       1 111292      NA    3.0      7      4
## 4: 0.82    1.79     2       2     1       1 111492      NA    0.0     36      2
## 5: 0.77    1.65     2       2     1       1 111492      NA   28.0      3      6
## 6: 0.77    0.95     2       2     1       1 111492      NA     NA     NA     NA
##    wage_st2 inctime2 firstin2 special2 meals2 open2r hrsopen2 psoda2 pfry2
## 1:     4.30       26     0.08        1      2    6.5     16.5   1.03    NA
## 2:     4.45       13     0.05        0      2   10.0     13.0   1.01  0.89
## 3:     5.00       19     0.25       NA      1   11.0     11.0   0.95  0.74
## 4:     5.25       26     0.15        0      2   10.0     12.0   0.92  0.79
## 5:     4.75       13     0.15        0      2   10.0     12.0   1.01  0.84
## 6:       NA       26       NA        0      2   10.0     12.0     NA  0.84
##    pentree2 nregs2 nregs112 emptot emptot2   demp gap nj bk kfc roys wendys
## 1:     0.94      4        4  40.50    24.0 -16.50   0  0  1   0    0      0
## 2:     2.35      4        4  13.75    11.5  -2.25   0  0  0   1    0      0
## 3:     2.33      4        3   8.50    10.5   2.00   0  0  0   1    0      0
## 4:     0.87      2        2  34.00    20.0 -14.00   0  0  0   0    0      1
## 5:     0.95      2        2  24.00    35.5  11.50   0  0  0   0    0      1
## 6:     1.79      3        3  20.50      NA     NA   0  0  0   0    0      1
##    pmeal pmeal2      dpmeal closed    fracft   fracft2 atmin newmin    icode
## 1:  2.58     NA          NA      0 0.7407407 0.1458333    NA     NA PA STORE
## 2:  4.26   4.25 -0.00999999      0 0.4727273 0.0000000    NA     NA PA STORE
## 3:  4.02   4.02  0.00000000      0 0.3529412 0.2857143    NA     NA PA STORE
## 4:  3.48   2.58 -0.89999992      0 0.5882353 0.0000000     0      0 PA STORE
## 5:  3.29   2.80 -0.49000001      0 0.2500000 0.7887324     0      0 PA STORE
## 6:  2.59     NA          NA      0 0.0000000        NA     0      0 PA STORE
```

Remove missing data from key variables:

- Remove observations with missing data in the starting wage variable at T1 and at T2
- Remove observations with missing data in the dependent variable
- Remove observations for which the status at T2 is 3 (=closed)

```
dt.fastfood.orig <- dt.fastfood.orig[!is.na(wage_st) & !is.na(demp) & !(is.na(wage_st2)  & status2!=3)]
```

Create panel:

```
dt.fastfood.panel <- rbind(
   dt.fastfood.orig[,list(  emptot
                          , gap
                          , demp
                          , state
                          , chain
                          , co_owned
                          , atmin
```

```
                            , meals
                            , wage = wage_st
                            , hrsopen
                            , pmeal
                            , fracft
                            , time = 0
                            , id = 1:nrow(dt.fastfood.orig))]
  ,dt.fastfood.orig[,list(  emptot = emptot2
                            , gap
                            , demp
                            , state
                            , chain
                            , co_owned
                            , atmin
                            , meals = meals2
                            , wage = wage_st2
                            , hrsopen = hrsopen2
                            , pmeal = pmeal2
                            , fracft = fracft2
                            , time = 1
                            , id = 1:nrow(dt.fastfood.orig))])
head(dt.fastfood.panel)
```

```
##    emptot gap  demp state chain co_owned atmin meals wage hrsopen pmeal
## 1:   34.0   0 -14.0     0     4        1     0     2 5.00      12  3.48
## 2:   24.0   0  11.5     0     4        1     0     3 5.50      12  3.29
## 3:   70.5   0 -41.5     0     1        0     0     2 5.00      18  2.86
## 4:   23.5   0  13.0     0     1        0     0     2 5.00      24  2.85
## 5:   11.0   0   0.0     0     2        1     0     1 5.25      10  3.78
## 6:    9.0   0  -0.5     0     2        1     0     1 5.00      10  3.99
##       fracft time id
## 1: 0.5882353    0  1
## 2: 0.2500000    0  2
## 3: 0.7092199    0  3
## 4: 0.4255319    0  4
## 5: 0.1818182    0  5
## 6: 0.2222222    0  6
```


Acknowledgements and Thanks:

This lab is based on material by M Godinho de Matos, R Belo and F Reis. Gratefully acknowledged!