Assignment

483670

September 15, 2022

“‘{ INFO, include=FALSE, results=‘asis’, echo=FALSE} Code Book for New Jersey-Pennsylvania Data Set

Note: there are 410 observations in the data set

```
Column Location
```

Name: Start End Format Explanation ffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffffff SHEET 1 3 3.0 sheet number (unique store id) CHAIN 5 5 1.0 chain 1=bk; 2=kfc; 3=roys; 4=wendys CO_OWNED 7 7 1.0 1 if company owned STATE 9 9 1.0 1 if NJ; 0 if Pa

Dummies for location: SOUTHJ 11 11 1.0 1 if in southern NJ CENTRALJ 13 13 1.0 1 if in central NJ NORTHJ 15 15 1.0 1 if in northern NJ PA1 17 17 1.0 1 if in PA, northeast suburbs of Phila PA2 19 19 1.0 1 if in PA, Easton etc SHORE 21 21 1.0 1 if on NJ shore

First Interview NCALLS 23 24 2.0 number of call-backs* EMPFT 26 30 5.2 # full-time employees EMPPT 32 36 5.2 # part-time employees NMGRS 38 42 5.2 # managers/ass’t managers WAGE_ST 44 48 5.2 starting wage ($/hr)INCTIME 50545.1 monthstousualfirstraiseFIRSTINC 56605.2 usualamountoffirstraise$(/hr) BONUS 62 62 1.0 1 if cash bounty for new workers PCTAFF 64 68 5.1 % employees affected by new minimum MEALS 70 70 1.0 free/reduced price code (See below) OPEN 72 76 5.2 hour of opening HRSOPEN 78 82 5.2 number hrs open per day PSODA 84 88 5.2 price of medium soda, including tax PFRY 90 94 5.2 price of small fries, including tax PENTREE 96 100 5.2 price of entree, including tax NREGS 102 103 2.0 number of cash registers in store NREGS11 105 106 2.0 number of registers open at 11:00 am

Second Interview TYPE2 108 108 1.0 type 2nd interview 1=phone; 2=personal STATUS2 110 110 1.0 status of second interview: see below DATE2 112 117 6.0 date of second interview MMDDYY format NCALLS2 119 120 2.0 number of call-backs* EMPFT2 122 126 5.2 # full-time employees EMPPT2 128 132 5.2 # part-time employees NMGRS2 134 138 5.2 # managers/ass’t managers WAGE_ST2 140 144 5.2 starting wage ($/hr)INCTIME2 1461505.1 monthstousualfirstraiseFIRSTIN2 1521565.2 usualamountoffirstraise$(/hr) SPECIAL2 158 158 1.0 1 if special program for new workers MEALS2 160 160 1.0 free/reduced price code (See below) OPEN2R 162 166 5.2 hour of opening HRSOPEN2 168 172 5.2 number hrs open per day PSODA2 174 178 5.2 price of medium soda, including tax PFRY2 180 184 5.2 price of small fries, including tax PENTREE2 186 190 5.2 price of entree, including tax NREGS2 192 193 2.0 number of cash registers in store NREGS112 195 196 2.0 number of registers open at 11:00 am

Codes:

Free/reduced Meal Variable: 0 = none 1 = free meals 2 = reduced price meals 3 = both free and reduced price meals

Second Interview Status 0 = refused second interview (count = 1) 1 = answered 2nd interview (count = 399) 2 = closed for renovations (count = 2) 3 = closed “permanently” (count = 6) 4 = closed for highway construction (count = 1) 5 = closed due to Mall fire (count = 1)

*Note: number of call-backs = 0 if contacted on first call

```
# Q1
The mean number of employees in NJ are 'r tostring(NJemp1)' and "NJemp2" for the first and second wave r


'''r
NJemp1a <- mean(data$EMPFT[data$STATE == 1]) + mean(data$EMPPT[data$STATE ==
    1]) + mean(data$NMGRS[data$STATE == 1])
NJemp2a <- mean(data$EMPFT2[data$STATE == 1]) + mean(data$EMPPT2[data$STATE ==
    1]) + mean(data$NMGRS2[data$STATE == 1])
PAemp1a <- mean(data$EMPFT[data$STATE == 0]) + mean(data$EMPPT[data$STATE ==
    0]) + mean(data$NMGRS[data$STATE == 0])
PAemp2a <- mean(data$EMPFT2[data$STATE == 0]) + mean(data$EMPPT2[data$STATE ==
    0]) + mean(data$NMGRS2[data$STATE == 0])
EffNJa <- NJemp2a - NJemp1a
EffPAa <- PAemp2a - PAemp1a
data2 <- subset(data, STATUS2 == 1)
NJemp1b <- mean(data2$EMPFT[data$STATE == 1]) + mean(data2$EMPPT[data$STATE ==
    1]) + mean(data2$NMGRS[data$STATE == 1])
NJemp2b <- mean(data2$EMPFT2[data$STATE == 1]) + mean(data2$EMPPT2[data$STATE ==
    1]) + mean(data2$NMGRS2[data$STATE == 1])
PAemp1b <- mean(data2$EMPFT[data$STATE == 0]) + mean(data2$EMPPT[data$STATE ==
    0]) + mean(data2$NMGRS[data$STATE == 0])
PAemp2b <- mean(data2$EMPFT2[data$STATE == 0]) + mean(data2$EMPPT2[data$STATE ==
    0]) + mean(data2$NMGRS2[data$STATE == 0])
EffNJb <- NJemp2b - NJemp1b
EffPAb <- PAemp2b - PAemp1b
data <- data2
```

Q2

As we are looking at the difference in each shop we restrict the sample to those who responded in both waves. The effect of the minimum wage in the simple diff-in-diff is "dif_eff1". Once we add controls for the amount of time the shop is open and a proxy for its capacity (the number of registers) the magnitude of effect increases slightly to "dif_eff2". CAN WE CONTROL FOR THE NUMBER OF EMPLOYEES IMPACTED BY THE WAGE SHARE? OR IS THAT A MECHANISM??

```
data$EMP1 <- data$EMPFT + data$EMPPT + data$NMGRS
data$EMP2 <- data$EMPFT2 + data$EMPPT2 + data$NMGRS2
data$difE <- data$EMP1 - data$EMP2
m1 <- lm(difE ~ STATE, data)
m2 <- lm(difE ~ STATE + HRSOPEN + NREGS, data)
dif_eff1 <- m1$coefficients[2]
dif_eff2 <- m2$coefficients[2]
stargazer(m1, m2, column.labels = c("", "Controls"), type = "text",
    title = "Minimum Wage", header = FALSE, label = "tab:reg1")
```

```
##
```

```
## Minimum Wage
## ================================================================
##                                 Dependent variable:
##                     ---------------------------------------
##                                        difE
##                                               Controls
##                              (1)                 (2)
## ----------------------------------------------------------------
## STATE                      -2.327*             -2.391*
##                            (1.299)             (1.297)
##
## HRSOPEN                                        0.374**
##                                                (0.187)
##
## NREGS                                           0.088
##                                                (0.425)
##
## Constant                    1.838              -3.902
##                            (1.165)             (2.948)
##
## ----------------------------------------------------------------
## Observations                 378                 373
## R2                          0.008               0.021
## Adjusted R2                 0.006               0.014
## Residual Std. Error  10.022 (df = 376)     9.928 (df = 369)
## F Statistic         3.208* (df = 1; 376) 2.697** (df = 3; 369)
## ================================================================
## Note:                                *p<0.1; **p<0.05; ***p<0.01
```

Q3

The status variable provides problems to the command and has been left out of the analysis. It looks like there are potential differences between in the states on the number of employees. This difference seems to originate in a difference in the number of full-time employees in the first wave.

Comparing stepped dummy variables does not make an awful lot of sense.

```
data3 <- subset(data, select = -c(STATUS2))
balance <- balance_table(data3, "STATE")
knitr::kable(balance, caption = "Balance Table")
```

Q4

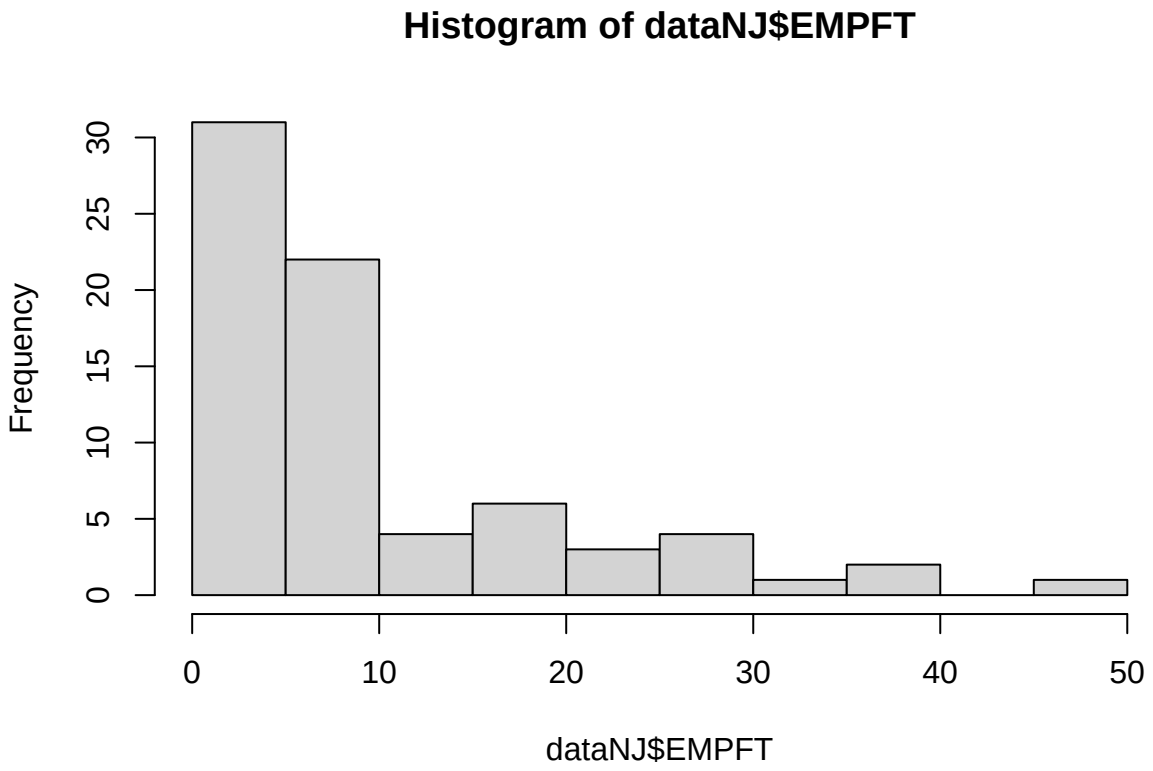
To estimate the propensity score we run a probit model estimating the probability of being in a given state. Table **??** calculates the probability of each restaurant being in NJ as opposed to Pa.

```
dataNJ <- subset(data, !STATE == 1)
dataPA <- subset(data, !STATE == 0)
# graph1 <- cdfCompare(dataNJ$EMPFT[0:79], dataPA$EMPFT[0:331])
```

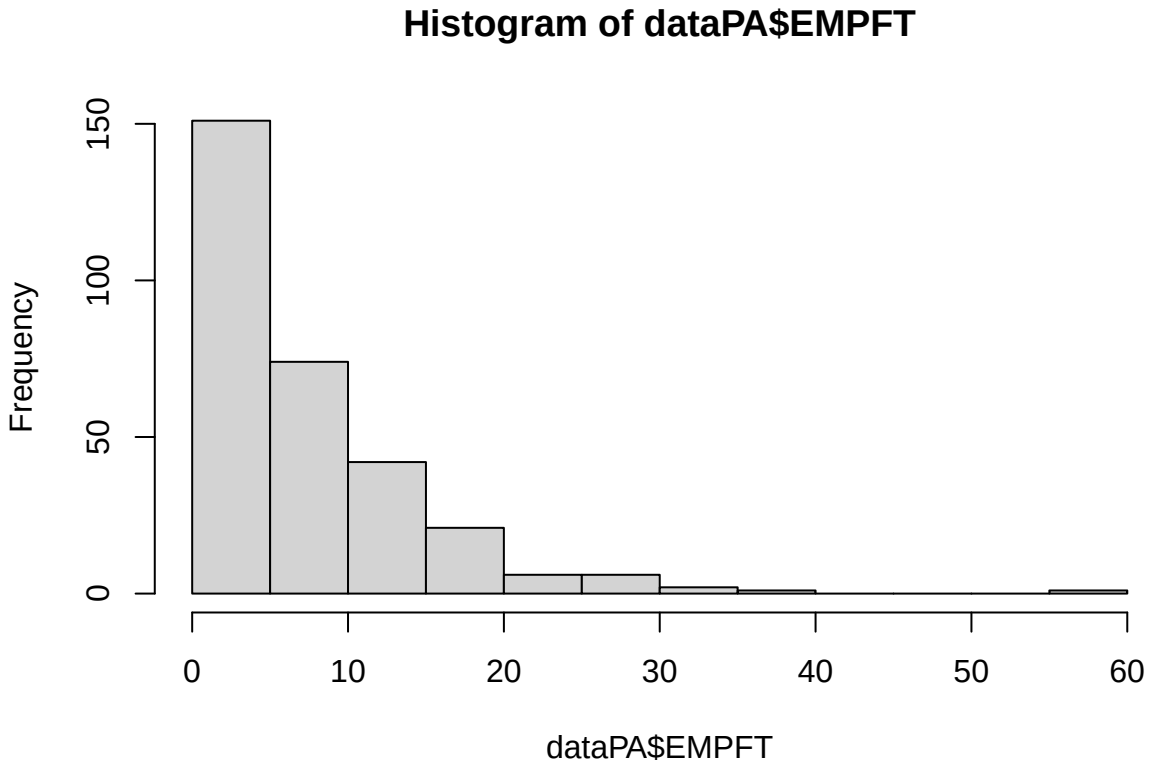
Table 1: Balance Table

variables1	Media_control1	Media_trat1	p_value1
BONUS	2.972973e-01	2.434211e-01	0.3624466
CENTRALJ	0.000000e+00	1.875000e-01	0.0000000
CHAIN	2.121622e+00	2.098684e+00	0.8782334
CO_OWNED	3.513514e-01	3.453947e-01	0.9238736
DATE2	1.121285e+05	1.120950e+05	0.9261648
difE	1.837838e+00	-4.891447e-01	0.1128536
EMP1	3.313514e+01	3.000428e+01	0.0960222
EMP2	3.129730e+01	3.049342e+01	0.5578162
EMPFT	1.020270e+01	7.751645e+00	0.0708302
EMPFT2	7.777027e+00	8.496711e+00	0.5118295
EMPPT	1.937838e+01	1.884211e+01	0.6749107
EMPPT2	1.983108e+01	1.850493e+01	0.3034505
FIRSTIN2	1.844444e-01	2.246032e-01	0.0012576
FIRSTINC	2.079688e-01	2.297091e-01	0.1133117
HRSOPEN	1.453378e+01	1.439638e+01	0.7197910
HRSOPEN2	1.463514e+01	1.438696e+01	0.5044687
INCTIME	1.936957e+01	1.793214e+01	0.4053026
INCTIME2	2.115152e+01	2.235135e+01	0.4757115
MEALS	2.027027e+00	1.871711e+00	0.0076043
MEALS2	1.905405e+00	1.733553e+00	0.0035400
NCALLS	7.567568e-01	1.213816e+00	0.0013335
NCALLS2	1.677419e+00	2.281818e+00	0.0262003
NMGRS	3.554054e+00	3.410526e+00	0.3077894
NMGRS2	3.689189e+00	3.491776e+00	0.2040121
NORTHJ	0.000000e+00	5.263158e-01	0.0000000
NREGS	3.378378e+00	3.705686e+00	0.0296376
NREGS11	2.821918e+00	2.710884e+00	0.2833660
NREGS112	2.575343e+00	2.683849e+00	0.2735462
NREGS2	3.479452e+00	3.671186e+00	0.2301576
OPEN	7.797297e+00	8.110197e+00	0.2704106
OPEN2R	7.871622e+00	8.172442e+00	0.2879269
PA1	4.594595e-01	0.000000e+00	0.0000000
PA2	5.405405e-01	0.000000e+00	0.0000000
PCTAFF	4.478261e+01	4.866407e+01	0.4289599
PENTREE	1.236389e+00	1.366122e+00	0.1278442
PENTREE2	1.190986e+00	1.413265e+00	0.0061747
PFRY	8.418056e-01	9.414533e-01	0.0000000
PFRY2	8.598551e-01	9.594178e-01	0.0000000
PSODA	9.741667e-01	1.062349e+00	0.0000000
PSODA2	9.740541e-01	1.062925e+00	0.0000000
SHEET	3.759459e+02	2.163125e+02	0.0000000
SHORE	0.000000e+00	1.052632e-01	0.0000000
SOUTHJ	0.000000e+00	2.861842e-01	0.0000000
SPECIAL2	2.328767e-01	2.046980e-01	0.6097328
TYPE2	1.067568e+00	1.065789e+00	0.9566654
WAGE_ST	4.632254e+00	4.612222e+00	0.6726108
WAGE_ST2	4.617246e+00	5.081200e+00	0.0000000

```
hEMP1 <- hist(dataNJ$EMPFT)
```




```
hEMP2 <- hist(dataPA$EMPFT)
```



```
# plot(ha,col = 'blue' , xlim = c(0, 50), ylim = c(0,100)) plot
# (hb, col = 'red', add = TRUE) probit <- glm(STATE ~ HRSOPEN +
# INCTIME + NREGS+ CHAIN, family = binomial(link = 'probit'),
# data = data) stargazer(probit, column.labels = c('', ''),type
# = 'text', title = 'Minimum Wage', header = FALSE, label =
# 'tab:probit')
```

Q5

https://www.r-bloggers.com/2022/04/propensity-score-matching/

```
# THIS DOESN'T WORK YET!!

# prop <- matchit(STATE ~ HRSOPEN + INCTIME + NREGS, data =
# data, method = 'nearest', distance ='glm',ratio = 1, replace =
# FALSE)
```