Lecture11_NikeLab

February 23, 2025

1 Lecture 11 - EEP 118 Spring 2025¶

Nike Zion Williamsons' shoe explosion

This is the notebook for Lecture 11 where we will estimate the effect of a Nike show destruction during an NCAA basketball game on Nike prices, before and after the shoe falls apart on national television.

Can we reject the null hypothesis that Zion's shoe explosion was not correlated with a change in Nike Stock prices at 5 % significance level, or at 1 % significance?

Research Strategy: Collect Data on hourly Nike Stock Prices on Feb 20th (pre shoe explosion) and on Feb 21 (next day after shoe explosion) of 2019

To run, hit the `i>|Run` button on top middle bar and keep hitting and it will run line by line,

OR

To run a line that starts with In [ ]: highlight the content and hit CONTROL ENTER at same time

[ ]:
```
# Load the 'pacman' package
install.packages("pacman")
library(pacman)
#packages to use load them now using the pacman "manager"
p_load(dplyr, haven, reader,psych)
#Another great feature of p_load(): if you try to load a package that is not␣
↪installed on your machine, p_load() install the package for you, rather than␣
↪throwing an error. For instance, let's install and load one final package␣
↪named ggplot2.
p_load(ggplot2)
pacman::p_load(lfe, lmtest, haven, sandwich, tidyverse)
# lfe for running fixed effects regression
# lmtest for displaying robust SE in output table
# haven for loading in dta files

#set scientific display off, thank you Roy
options(scipen=999)
```

[9]:
```
#-------------------------------------------
#1. Read in data and see the top rows to see column names etc
```

```
#-------------------------------------------
#read in a Stata dataset
StockData <- read_dta("Lecture11_stockPricesn.dta")
head(StockData)
```

A tibble: 6 × 4

counter <dbl>	price1 <dbl>	price2 <dbl>	price3 <dbl>
1	84.63	114.15	21.40
2	84.93	114.49	21.40
3	84.46	114.64	21.37
4	84.70	115.00	21.27
5	84.58	115.00	21.12
6	84.89	115.30	21.38

[10]:
```
##Renaming first four columns columns
colnames(StockData) <- c("Hour", "Nike", "Adidas", "UnderArmour")
```

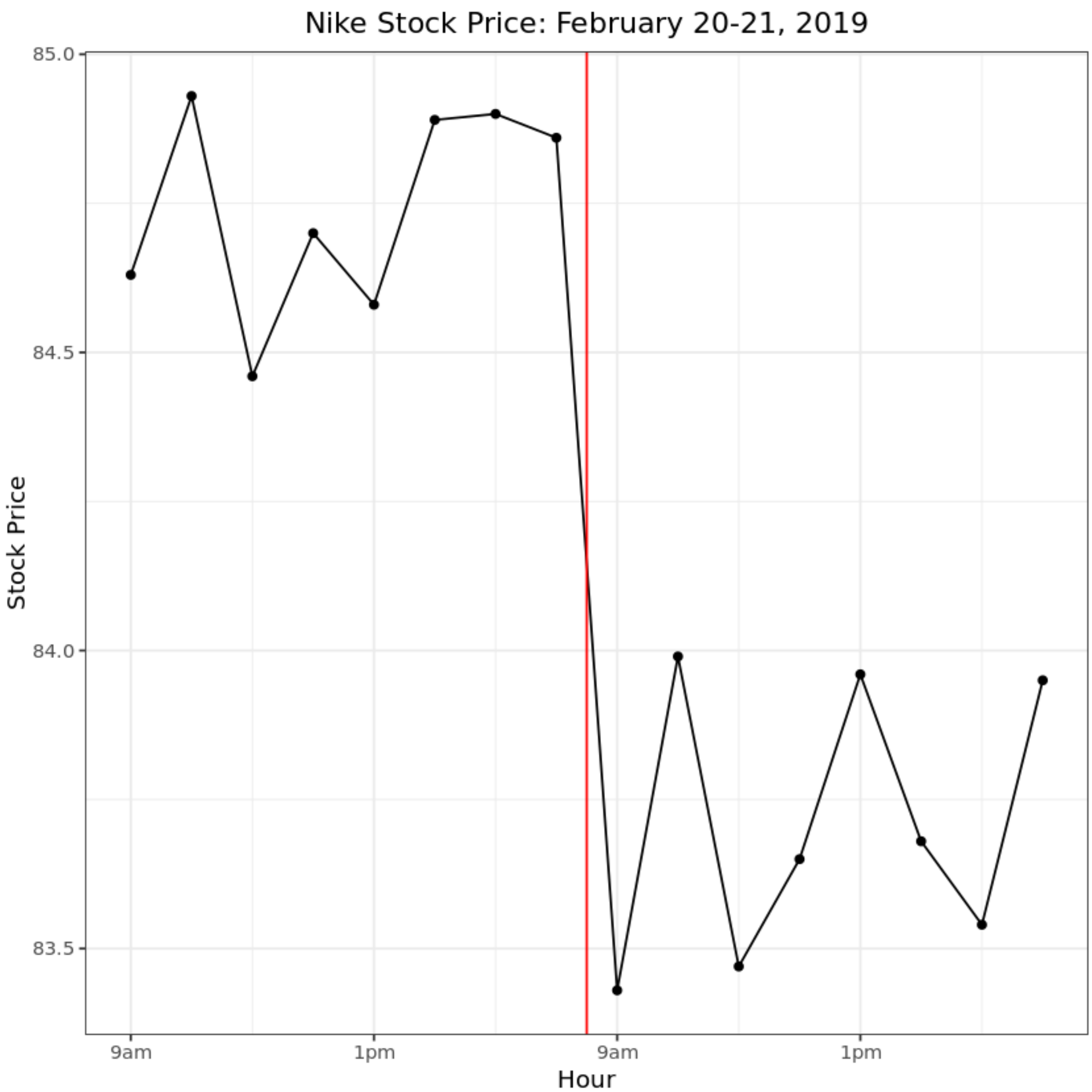
[11]:
```
##Defining Post "Shoe Explodes" Dummy
StockData$PostExplosion <- as.numeric(StockData$Hour >= 9)
#see the bottom rows of the dataset
tail(StockData)
```

A tibble: 6 × 5

Hour <dbl>	Nike <dbl>	Adidas <dbl>	UnderArmour <dbl>	PostExplosion <dbl>
11	83.47	114.93	21.14	1
12	83.65	115.25	21.50	1
13	83.96	114.90	21.58	1
14	83.68	115.23	21.50	1
15	83.54	115.08	21.39	1
16	83.95	115.11	21.42	1

scatter plot for nike stock prices with explosion marked

[12]:
```
NikePlot <- ggplot(StockData, aes(x = Hour, y = Nike))
#NikePlot <- NikePlot + geom_smooth(se=FALSE) + geom_point() +␣
↪geom_line(color="black") + geom_vline(xintercept=8.5, colour="Red")
NikePlot <- NikePlot + geom_point() + geom_line(color="black") +␣
↪geom_vline(xintercept=8.5, colour="Red")
NikePlot <- NikePlot + theme_bw() + theme(plot.title = element_text(hjust = 0.
↪5)) +
 scale_x_continuous(breaks=(seq(1,16,4)), labels=c("9am", "1pm", "9am", "1pm"))
NikePlot <- NikePlot + xlab("Hour") + ylab("Stock Price") + ggtitle("Nike Stock␣
↪Price: February 20-21, 2019")
NikePlot
```



Summary Stats

[13]:
```
#one way
library(psych)
describeBy(StockData$Nike, StockData$PostExplosion)
#library(doBy)
#summaryBy(Nike ~ PostExplosion, data = StockData,
#          FUN = function(x) { c(m = mean(x), s = sd(x)) } )
# produces mpg.m wt.m mpg.s wt.s for each
#  level of PostExplosion
```

```
 Descriptive statistics by group
group: 0
```

```
   vars n  mean   sd median trimmed mad   min   max range  skew kurtosis   se
X1    1 8 84.74 0.18  84.78   84.74 0.2 84.46 84.93  0.47 -0.32    -1.71 0.06
------------------------------------------------------------
group: 1
   vars n  mean   sd median trimmed  mad   min   max range skew kurtosis   se
X1    1 8 83.71 0.23  83.67   83.71 0.32 83.43 83.99  0.56 0.12    -1.93 0.08
```

Regression of stock price on hourly trend and explosion indicator

[14]:
```
#regression
regLecture11 <- lm(Nike ~ Hour+PostExplosion,StockData)
#show output
summary(regLecture11)
```

```
Call:
lm(formula = Nike ~ Hour + PostExplosion, data = StockData)

Residuals:
    Min      1Q  Median      3Q     Max
-0.2426 -0.1762 -0.0197  0.1109  0.3551

Coefficients:
              Estimate Std. Error t value             Pr(>|t|)
(Intercept)   84.61089    0.11994 705.452 < 0.0000000000000002 ***
Hour           0.02952    0.02163   1.365                0.195
PostExplosion -1.27119    0.19941  -6.375            0.0000244 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.1982 on 13 degrees of freedom
Multiple R-squared:  0.8951,        Adjusted R-squared:  0.8789
F-statistic: 55.45 on 2 and 13 DF,  p-value: 0.0000004321
```

What do we see above? Interpretation, controlling for an hourly trend (Hour), the explosion is correlated with a significant drop of Nike Stock prices (coefficient on post explosion is -1.27, and is statistically significant from zero, at 5%, 1%, 0.001 % signif the p value is 0.0000244=2.44e-05.

If you want to know if -1.27 dollars is a big drop, compare it to the baseline average of 84.61 (the constant estimate).

So -1.27/84.61 is about - 0.015, or -1.5% drop

How to get estimated parameters bata_hat b0_hat, b1_hat, b2_hat and also Generate Predicted stock price of nike using b0 and b1 and b2 estimates of the regression you estimated

[15]:
```
#how to extract the coeff from the regression above?
#R stores that in recLecture11
#as a variable called regLecture11$coefficients
```

```
#all of them into beta_hat vector:
beta_hat<-regLecture11$coefficients
beta_hat

#estimate of a constant is the first one in the coefficients
regLecture11$coefficients[1]

#estimate of explosion parameter is the third one in the coefficients:
regLecture11$coefficients[3]

#and it also has the fitted values and add them as an additional column to the␣
↪dataframe of data
StockData$Yhat<-regLecture11$fitted.values

#create predictions as a separate column of data called Yhat
Yhat<-regLecture11$fitted.values
#or Nikehat
Nikehat<-regLecture11$fitted.values
```

(Intercept) 84.6108918871198 **Hour** 0.0295239403134301 **PostExplosion** -1.27119136991955

(Intercept): 84.6108918871198

PostExplosion: -1.27119136991955

Another way to get predictions, multiply coeff and x's

[16]:
```
StockData$Nikeprediction<- regLecture11$coefficients[1]+␣
↪regLecture11$coefficients[2]*StockData$Hour+␣
↪regLecture11$coefficients[3]*StockData$PostExplosion
#look at the column Nikeprediction in StockData frame. It is equal to Yhat␣
↪defined above
```

Generate the Log of Nike stock price by hour from the original data and add as an additional column to the dataframe called StockData

[17]:
```
#generate log
StockData$NikeLog<-log(StockData$Nike)
```

We can get and interpret the regression estimates of the log price regression now. In fact, that way we get the percent change due to the explosion directly from the coefficient of the postExplosion variable, from the regression of log prices on a constant, hour, and post explosion

[18]:
```
#regression in logs
regLecture11Log <- lm(NikeLog ~ Hour+PostExplosion,StockData)
#show output
summary(regLecture11Log)
```

```
Call:
lm(formula = NikeLog ~ Hour + PostExplosion, data = StockData)

Residuals:
       Min         1Q     Median         3Q        Max 
-0.0028914 -0.0021055 -0.0002265  0.0013135  0.0042343 

Coefficients:
                Estimate Std. Error  t value             Pr(>|t|)    
(Intercept)    4.4380519  0.0014266 3110.889 < 0.0000000000000002 ***
Hour           0.0003507  0.0002573    1.363                0.196    
PostExplosion -0.0150956  0.0023719   -6.364            0.0000248 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.002358 on 13 degrees of freedom
Multiple R-squared:  0.8948,	Adjusted R-squared:  0.8786 
F-statistic: 55.26 on 2 and 13 DF,  p-value: 0.0000004407
```

What do we see above? Interpretation, controlling for an hourly trend (Hour), the explosion is correlated with a significant drop of log of Nike Stock prices (coefficient on post explosion is -0.015, and is statistically significant from zero, at 5%, 1%, 0.001 % signif the p value is 0.0000248=2.48e-05. A change in the log of -0.015 in the change in the Nike price of - 1.5%

beta_hat post explosion=-0.015 corresponds to a -1.5% drop in the Nike price.

THIS is a before and AFTER analysis, we can say if we find statistical evidence in favor of explosion being correlated with a change in Nike stock or not, all else constant. But we cannot say it CAUSED the change in price.

Later in class, where we are going towards is CAUSALITY!

Research Strategy: Collect Data on hourly Nike and UA and Adidas Stock Prices on Feb 20 and 21

That is, get data also for a control group that captures anything else that could have happened to stock prices for sports companies from 20 and 21 of Feb that had nothing to do with the explosion, e.g. get stock prices for Adidas and Under Armour.

FIRST : Check if Adidas and Under Armour, or average of (UA and Adidas), stock prices were trending similarly to Nike stock prices before the explosion

Then we have a good counterfactual, Of what would have happened in Feb 21 relative to Feb 20 to Nike if there was no explosion (using the change in average (UA and Adidas) stock prices as that control change)

Then the causal effect of the explosion on Nike on Feb 21 is change in Nike MINUS change in average(UA and Adidas). We call this Difference in Differences.

See the youtube of (NikeLab folder).

I will return at end of semester to do the rest with you all, and see in the next slide a blog post on my complete Econometric analysis Nike Lab.

With Scott Kaplan, who is starting this Fall as an Assistant Professor in Economics at the U. S. Naval academy. He graduated from EEP, then will graduate this year with a PhD in ARE

https://www.scottkaplan.org/post/the-explosion-heard-around-the-college-basketball-world

The video of this Lecture11 portion – Nike Lab, is here on Youtube, by yours truly https://youtu.be/eWE_vEEZhS0