Ch 18 the effects

Emilio Horner

2023-03-31

1. The assumption is that the treatment was the variable that led to the difference over time. The difference in difference provides a control group as a comparative. For example, it could have been another factor that led to the change over time.

2. C There may be differences in typical outcome levels between the treated group and the untreated group.

3. a. Parallel trends could be violated even with the same outcome if they had been trending differently leading up to the point of being measured. For example, I could make the same amount of money one year as someone else but over a longer pre-treatment period it could be shown that the slopes of the line regarding income over many years is not the same.

b. The difference in difference method will be inaccurate if the parallel trends are violated in the pre-treatment period. The results will be more innacurate the more the lines are not parallel with each other.

4.

a. the parallel trends assumption is violated in the graph
b. we would underestimate the impact of the treatment

5.

This could work but the research designer needs to make sure that the rates of covid were paralel in the United States and Canada before the treatment. It is hard to find a comparative control group because the United States also did lockdowns. And the strength of the lockdowns ebbed and flowed throughout the pandemic.

6.

(9-5)-(7.5-6)= 2.5

How is it performed?

1. a. iv. A set of fixed effects for state, and for year, and an interaction between “is 2016” and “is a treated state”.

b. Voter Turnout = α + β1(Treated State) + β2(Year 2016) + β3(Treated State × Year 2016) + State Fixed Effects + Year Fixed Effects + ε

β3 would give the DiD.

2. Assuming that the parallel trends assumption is met, the effect of laptops on test scores was 5.034, and this effect was statistically significant at the 95% level.

3. The standard priors test exists to know whether the outcome variable was changing at a constant rate before the treatment is introduced. It is similar to the parallel trends assumption.

4. a. There’s a such initial spike but that dips back down at the 6th time interval. This could bias the results of the treatment.

b. Between periods 4 and 5 the treatment appears to be 3.

5. It is not proper to use an already treated group as a control group because it makes it so we can't really know what the effect of the treatment is since the group we would supposedly be comparing the new treatment group to has already also recieved the treatment.

6.

```
sr <- read.csv("https://raw.githubusercontent.com/NickCH-K/TheEffectAssignments/main/sourdough_trends.csv")
```

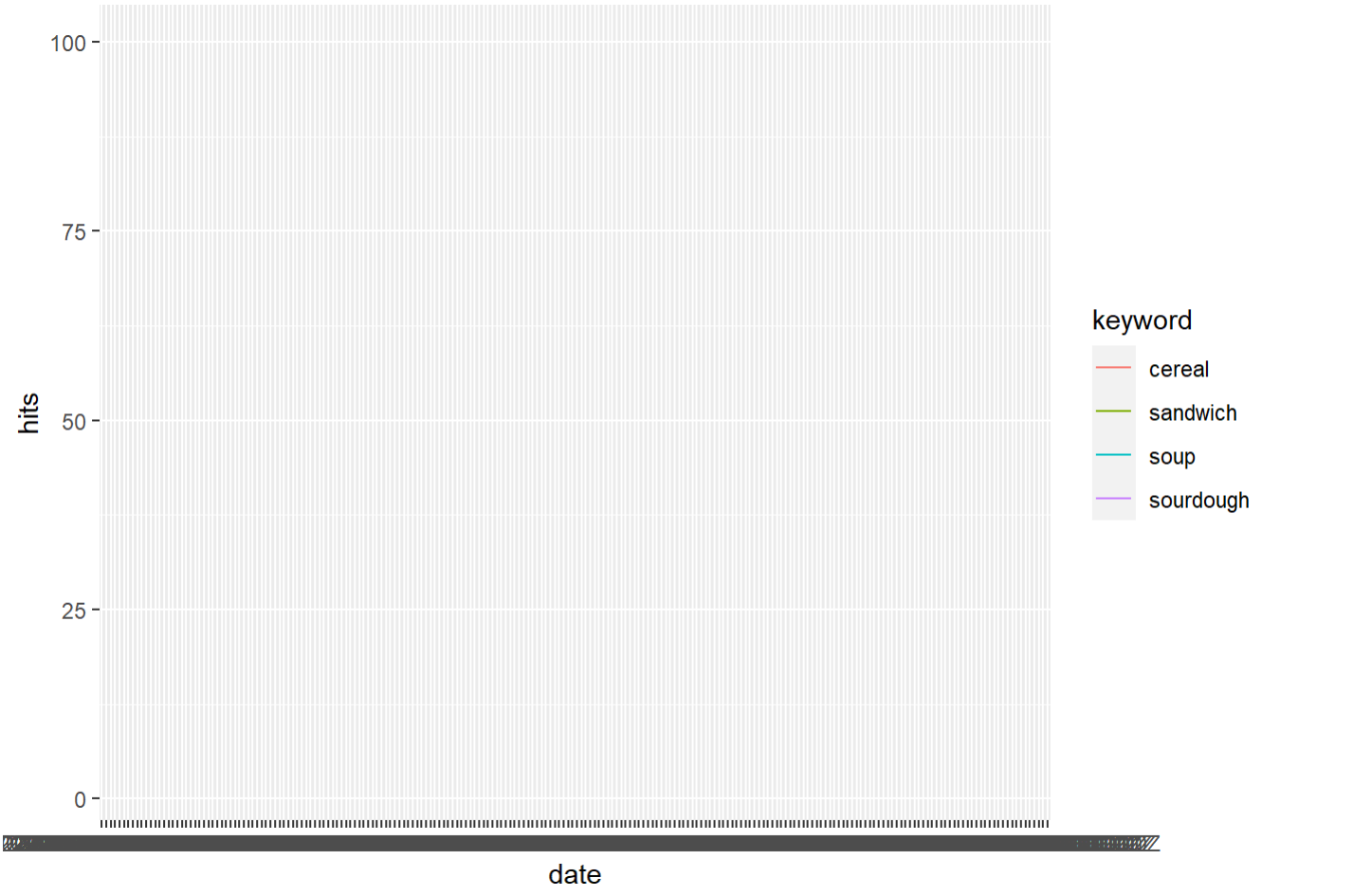
```
sr_subset <- sr[, c("date", "hits", "keyword")]
```

2.

```
library(ggplot2)
```

```
ggplot(sr_subset, aes(x = date, y = hits, color = keyword)) +
  geom_line() +
  geom_vline(aes(xintercept = as.numeric(as.Date("2020-03-15"))), linetype = "dashed")
```

```
## `geom_line()`: Each group consists of only one observation.
## ℹ Do you need to adjust the group aesthetic?
```



3. Looking at your graph from problem 2, comment on (a) whether it looks like the lockdown had an effect on the popularity of sourdough, (b) the shape that effect takes (i.e. is it a permanent increase in popularity? Temporary?), (c) whether you might be concerned about any of the control groups we've chosen

a. My graph is not showing it, but just looking at the data it does look like the lockdowns increased the popularity of sourdough.
b. The popularity is temporary
c. there could be unobserved confounding variables.

4.

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
sr_subset <- sr_subset %>%
  mutate(Treated = ifelse(keyword == "sourdough", 1, 0))
```

```
library(tidyverse); library(modelsummary); library(fixest)
```

```
## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.2 ──
## ✓ tibble  3.1.8     ✓ purrr   0.3.5
## ✓ tidyr   1.2.1     ✓ stringr 1.5.0
## ✓ readr   2.1.3     ✓ forcats 0.5.2
## ── Conflicts ────────────────────────────────────────── tidyverse_conflicts() ──
## ✖ dplyr::filter() masks stats::filter()
## ✖ dplyr::lag()    masks stats::lag()
```

```
## Warning: package 'fixest' was built under R version 4.2.3
```

```
model <- feols(hits ~ date + Treated,
           data = sr_subset)
msummary(model, stars = c('*' = .1, '**' = .05, '***' = .01))
```

	(1)
(Intercept)	32.187***
	(9.046)
date2020-01-02T00 × 00 × 00Z	0.000
	(12.783)
date2020-01-03T00 × 00 × 00Z	−1.500
	(12.783)
date2020-01-04T00 × 00 × 00Z	2.750
	(12.783)
date2020-01-05T00 × 00 × 00Z	6.250
	(12.783)
date2020-01-06T00 × 00 × 00Z	−1.750
	(12.783)
date2020-01-07T00 × 00 × 00Z	−0.500
	(12.783)
date2020-01-08T00 × 00 × 00Z	−3.500
	(12.783)
date2020-01-09T00 × 00 × 00Z	−2.500
	(12.783)
date2020-01-10T00 × 00 × 00Z	−2.250
	(12.783)
date2020-01-11T00 × 00 × 00Z	1.500
	(12.783)
date2020-01-12T00 × 00 × 00Z	3.750
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-01-13T00 × 00 × 00Z	−1.750
	(12.783)
date2020-01-14T00 × 00 × 00Z	−3.750
	(12.783)
date2020-01-15T00 × 00 × 00Z	−3.000
	(12.783)
date2020-01-16T00 × 00 × 00Z	−3.250
	(12.783)
date2020-01-17T00 × 00 × 00Z	−2.750
	(12.783)
date2020-01-18T00 × 00 × 00Z	4.000
	(12.783)
date2020-01-19T00 × 00 × 00Z	2.750
	(12.783)
date2020-01-20T00 × 00 × 00Z	3.500
	(12.783)
date2020-01-21T00 × 00 × 00Z	−2.500
	(12.783)
date2020-01-22T00 × 00 × 00Z	−1.500
	(12.783)
date2020-01-23T00 × 00 × 00Z	−3.250
	(12.783)
date2020-01-24T00 × 00 × 00Z	−4.250
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-01-25T00 × 00 × 00Z	0.000
	(12.783)
date2020-01-26T00 × 00 × 00Z	2.750
	(12.783)
date2020-01-27T00 × 00 × 00Z	−3.000
	(12.783)
date2020-01-28T00 × 00 × 00Z	−3.750
	(12.783)
date2020-01-29T00 × 00 × 00Z	−2.500
	(12.783)
date2020-01-30T00 × 00 × 00Z	−4.000
	(12.783)
date2020-01-31T00 × 00 × 00Z	−4.000
	(12.783)
date2020-02-01T00 × 00 × 00Z	−0.250
	(12.783)
date2020-02-02T00 × 00 × 00Z	−1.750
	(12.783)
date2020-02-03T00 × 00 × 00Z	−7.500
	(12.783)
date2020-02-04T00 × 00 × 00Z	−3.250
	(12.783)
date2020-02-05T00 × 00 × 00Z	−4.750
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-02-06T00 × 00 × 00Z	−4.500
	(12.783)
date2020-02-07T00 × 00 × 00Z	−5.750
	(12.783)
date2020-02-08T00 × 00 × 00Z	−2.000
	(12.783)
date2020-02-09T00 × 00 × 00Z	3.250
	(12.783)
date2020-02-10T00 × 00 × 00Z	−5.000
	(12.783)
date2020-02-11T00 × 00 × 00Z	−4.750
	(12.783)
date2020-02-12T00 × 00 × 00Z	−6.250
	(12.783)
date2020-02-13T00 × 00 × 00Z	−6.750
	(12.783)
date2020-02-14T00 × 00 × 00Z	−9.000
	(12.783)
date2020-02-15T00 × 00 × 00Z	−5.000
	(12.783)
date2020-02-16T00 × 00 × 00Z	0.250
	(12.783)
date2020-02-17T00 × 00 × 00Z	−3.500
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-02-18T00 × 00 × 00Z	−5.500
	(12.783)
date2020-02-19T00 × 00 × 00Z	−6.750
	(12.783)
date2020-02-20T00 × 00 × 00Z	−4.500
	(12.783)
date2020-02-21T00 × 00 × 00Z	−6.750
	(12.783)
date2020-02-22T00 × 00 × 00Z	−3.000
	(12.783)
date2020-02-23T00 × 00 × 00Z	−0.500
	(12.783)
date2020-02-24T00 × 00 × 00Z	−5.500
	(12.783)
date2020-02-25T00 × 00 × 00Z	−4.000
	(12.783)
date2020-02-26T00 × 00 × 00Z	−3.500
	(12.783)
date2020-02-27T00 × 00 × 00Z	−5.750
	(12.783)
date2020-02-28T00 × 00 × 00Z	−6.250
	(12.783)
date2020-02-29T00 × 00 × 00Z	−3.500
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-03-01T00 × 00 × 00Z	0.000
	(12.783)
date2020-03-02T00 × 00 × 00Z	−5.250
	(12.783)
date2020-03-03T00 × 00 × 00Z	−7.500
	(12.783)
date2020-03-04T00 × 00 × 00Z	−7.000
	(12.783)
date2020-03-05T00 × 00 × 00Z	−8.250
	(12.783)
date2020-03-06T00 × 00 × 00Z	−7.500
	(12.783)
date2020-03-07T00 × 00 × 00Z	−3.250
	(12.783)
date2020-03-08T00 × 00 × 00Z	−3.000
	(12.783)
date2020-03-09T00 × 00 × 00Z	−5.250
	(12.783)
date2020-03-10T00 × 00 × 00Z	−6.750
	(12.783)
date2020-03-11T00 × 00 × 00Z	−7.750
	(12.783)
date2020-03-12T00 × 00 × 00Z	−8.500
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-03-13T00 × 00 × 00Z	−6.250
	(12.783)
date2020-03-14T00 × 00 × 00Z	−2.500
	(12.783)
date2020-03-15T00 × 00 × 00Z	−0.250
	(12.783)
date2020-03-16T00 × 00 × 00Z	−7.000
	(12.783)
date2020-03-17T00 × 00 × 00Z	−6.250
	(12.783)
date2020-03-18T00 × 00 × 00Z	−5.250
	(12.783)
date2020-03-19T00 × 00 × 00Z	−6.250
	(12.783)
date2020-03-20T00 × 00 × 00Z	−7.500
	(12.783)
date2020-03-21T00 × 00 × 00Z	−1.250
	(12.783)
date2020-03-22T00 × 00 × 00Z	1.000
	(12.783)
date2020-03-23T00 × 00 × 00Z	−4.250
	(12.783)
date2020-03-24T00 × 00 × 00Z	−4.250
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-03-25T00 × 00 × 00Z	−4.750
	(12.783)
date2020-03-26T00 × 00 × 00Z	−4.750
	(12.783)
date2020-03-27T00 × 00 × 00Z	−5.750
	(12.783)
date2020-03-28T00 × 00 × 00Z	−0.250
	(12.783)
date2020-03-29T00 × 00 × 00Z	2.000
	(12.783)
date2020-03-30T00 × 00 × 00Z	−4.750
	(12.783)
date2020-03-31T00 × 00 × 00Z	−3.500
	(12.783)
date2020-04-01T00 × 00 × 00Z	−5.750
	(12.783)
date2020-04-02T00 × 00 × 00Z	−5.250
	(12.783)
date2020-04-03T00 × 00 × 00Z	−5.250
	(12.783)
date2020-04-04T00 × 00 × 00Z	1.000
	(12.783)
date2020-04-05T00 × 00 × 00Z	2.750
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-04-06T00 × 00 × 00Z	−3.750
	(12.783)
date2020-04-07T00 × 00 × 00Z	−3.750
	(12.783)
date2020-04-08T00 × 00 × 00Z	−3.500
	(12.783)
date2020-04-09T00 × 00 × 00Z	−3.000
	(12.783)
date2020-04-10T00 × 00 × 00Z	−1.250
	(12.783)
date2020-04-11T00 × 00 × 00Z	0.500
	(12.783)
date2020-04-12T00 × 00 × 00Z	−1.000
	(12.783)
date2020-04-13T00 × 00 × 00Z	0.000
	(12.783)
date2020-04-14T00 × 00 × 00Z	−2.750
	(12.783)
date2020-04-15T00 × 00 × 00Z	−2.250
	(12.783)
date2020-04-16T00 × 00 × 00Z	−4.000
	(12.783)
date2020-04-17T00 × 00 × 00Z	−3.250
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-04-18T00 × 00 × 00Z	2.000
	(12.783)
date2020-04-19T00 × 00 × 00Z	2.500
	(12.783)
date2020-04-20T00 × 00 × 00Z	−4.500
	(12.783)
date2020-04-21T00 × 00 × 00Z	−4.250
	(12.783)
date2020-04-22T00 × 00 × 00Z	−4.000
	(12.783)
date2020-04-23T00 × 00 × 00Z	−4.500
	(12.783)
date2020-04-24T00 × 00 × 00Z	−6.000
	(12.783)
date2020-04-25T00 × 00 × 00Z	−2.500
	(12.783)
date2020-04-26T00 × 00 × 00Z	0.250
	(12.783)
date2020-04-27T00 × 00 × 00Z	−5.000
	(12.783)
date2020-04-28T00 × 00 × 00Z	−6.000
	(12.783)
date2020-04-29T00 × 00 × 00Z	−6.500
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-04-30T00 × 00 × 00Z	−5.000
	(12.783)
date2020-05-01T00 × 00 × 00Z	−7.000
	(12.783)
date2020-05-02T00 × 00 × 00Z	−5.250
	(12.783)
date2020-05-03T00 × 00 × 00Z	−2.250
	(12.783)
date2020-05-04T00 × 00 × 00Z	−6.250
	(12.783)
date2020-05-05T00 × 00 × 00Z	−7.250
	(12.783)
date2020-05-06T00 × 00 × 00Z	−6.500
	(12.783)
date2020-05-07T00 × 00 × 00Z	−7.750
	(12.783)
date2020-05-08T00 × 00 × 00Z	−6.250
	(12.783)
date2020-05-09T00 × 00 × 00Z	−3.000
	(12.783)
date2020-05-10T00 × 00 × 00Z	−4.500
	(12.783)
date2020-05-11T00 × 00 × 00Z	−6.500
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-05-12T00 × 00 × 00Z	−6.500
	(12.783)
date2020-05-13T00 × 00 × 00Z	−7.500
	(12.783)
date2020-05-14T00 × 00 × 00Z	−7.750
	(12.783)
date2020-05-15T00 × 00 × 00Z	−8.750
	(12.783)
date2020-05-16T00 × 00 × 00Z	−4.500
	(12.783)
date2020-05-17T00 × 00 × 00Z	−2.750
	(12.783)
date2020-05-18T00 × 00 × 00Z	−6.500
	(12.783)
date2020-05-19T00 × 00 × 00Z	−7.500
	(12.783)
date2020-05-20T00 × 00 × 00Z	−8.000
	(12.783)
date2020-05-21T00 × 00 × 00Z	−8.500
	(12.783)
date2020-05-22T00 × 00 × 00Z	−10.250
	(12.783)
date2020-05-23T00 × 00 × 00Z	−6.500
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-05-24T00 × 00 × 00Z	−5.500
	(12.783)
date2020-05-25T00 × 00 × 00Z	−8.500
	(12.783)
date2020-05-26T00 × 00 × 00Z	−10.000
	(12.783)
date2020-05-27T00 × 00 × 00Z	−9.500
	(12.783)
date2020-05-28T00 × 00 × 00Z	−9.000
	(12.783)
date2020-05-29T00 × 00 × 00Z	−11.500
	(12.783)
date2020-05-30T00 × 00 × 00Z	−9.000
	(12.783)
date2020-05-31T00 × 00 × 00Z	−8.500
	(12.783)
date2020-06-01T00 × 00 × 00Z	−11.750
	(12.783)
date2020-06-02T00 × 00 × 00Z	−11.750
	(12.783)
date2020-06-03T00 × 00 × 00Z	−10.750
	(12.783)
date2020-06-04T00 × 00 × 00Z	−10.750
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-06-05T00 × 00 × 00Z	−12.000
	(12.783)
date2020-06-06T00 × 00 × 00Z	−8.250
	(12.783)
date2020-06-07T00 × 00 × 00Z	−7.000
	(12.783)
date2020-06-08T00 × 00 × 00Z	−10.750
	(12.783)
date2020-06-09T00 × 00 × 00Z	−10.500
	(12.783)
date2020-06-10T00 × 00 × 00Z	−10.750
	(12.783)
date2020-06-11T00 × 00 × 00Z	−11.000
	(12.783)
date2020-06-12T00 × 00 × 00Z	−11.500
	(12.783)
date2020-06-13T00 × 00 × 00Z	−9.500
	(12.783)
date2020-06-14T00 × 00 × 00Z	−7.250
	(12.783)
date2020-06-15T00 × 00 × 00Z	−10.000
	(12.783)
date2020-06-16T00 × 00 × 00Z	−11.250
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-06-17T00 × 00 × 00Z	−11.250
	(12.783)
date2020-06-18T00 × 00 × 00Z	−10.500
	(12.783)
date2020-06-19T00 × 00 × 00Z	−11.250
	(12.783)
date2020-06-20T00 × 00 × 00Z	−9.000
	(12.783)
date2020-06-21T00 × 00 × 00Z	−9.500
	(12.783)
date2020-06-22T00 × 00 × 00Z	−11.000
	(12.783)
date2020-06-23T00 × 00 × 00Z	−12.250
	(12.783)
date2020-06-24T00 × 00 × 00Z	−11.250
	(12.783)
date2020-06-25T00 × 00 × 00Z	−11.250
	(12.783)
date2020-06-26T00 × 00 × 00Z	−12.000
	(12.783)
date2020-06-27T00 × 00 × 00Z	−9.000
	(12.783)
date2020-06-28T00 × 00 × 00Z	−7.500
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-06-29T00 × 00 × 00Z	−11.000
	(12.783)
date2020-06-30T00 × 00 × 00Z	−11.500
	(12.783)
date2020-07-01T00 × 00 × 00Z	−12.000
	(12.783)
date2020-07-02T00 × 00 × 00Z	−11.750
	(12.783)
date2020-07-03T00 × 00 × 00Z	−10.750
	(12.783)
date2020-07-04T00 × 00 × 00Z	−11.250
	(12.783)
date2020-07-05T00 × 00 × 00Z	−10.500
	(12.783)
date2020-07-06T00 × 00 × 00Z	−10.750
	(12.783)
date2020-07-07T00 × 00 × 00Z	−11.000
	(12.783)
date2020-07-08T00 × 00 × 00Z	−11.250
	(12.783)
date2020-07-09T00 × 00 × 00Z	−10.500
	(12.783)
date2020-07-10T00 × 00 × 00Z	−11.250
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-07-11T00 × 00 × 00Z	−9.500
	(12.783)
date2020-07-12T00 × 00 × 00Z	−8.500
	(12.783)
date2020-07-13T00 × 00 × 00Z	−10.500
	(12.783)
date2020-07-14T00 × 00 × 00Z	−11.000
	(12.783)
date2020-07-15T00 × 00 × 00Z	−12.250
	(12.783)
date2020-07-16T00 × 00 × 00Z	−10.750
	(12.783)
date2020-07-17T00 × 00 × 00Z	−11.250
	(12.783)
date2020-07-18T00 × 00 × 00Z	−8.250
	(12.783)
date2020-07-19T00 × 00 × 00Z	−6.750
	(12.783)
date2020-07-20T00 × 00 × 00Z	−11.500
	(12.783)
date2020-07-21T00 × 00 × 00Z	−11.250
	(12.783)
date2020-07-22T00 × 00 × 00Z	−10.500
	(12.783)

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
date2020-07-23T00 × 00 × 00Z	−10.000
	(12.783)
date2020-07-24T00 × 00 × 00Z	−9.750
	(12.783)
date2020-07-25T00 × 00 × 00Z	−9.250
	(12.783)
date2020-07-26T00 × 00 × 00Z	−7.750
	(12.783)
date2020-07-27T00 × 00 × 00Z	−10.000
	(12.783)
date2020-07-28T00 × 00 × 00Z	−11.000
	(12.783)
date2020-07-29T00 × 00 × 00Z	−10.500
	(12.783)
date2020-07-30T00 × 00 × 00Z	−10.750
	(12.783)
date2020-07-31T00 × 00 × 00Z	−11.750
	(12.783)
date2020-08-01T00 × 00 × 00Z	−9.250
	(12.783)
Treated	−18.749***
	(1.427)
Num.Obs.	856
R2	0.253

* p < 0.1, ** p < 0.05, *** p < 0.01

	(1)
R2 Adj.	0.004
AIC	7567.3
BIC	8589.0
RMSE	15.64
Std.Errors	IID

* p < 0.1, ** p < 0.05, *** p < 0.01

5.

```
library(lubridate)
```

```
## Loading required package: timechange
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```
library(fixest)
```

```
sr_subset <- sr_subset %>%
  mutate(date = as.Date(date),
         month = month(date - days(15)),
         After = ifelse(date >= as.Date('2020-03-15'), 1, 0))
```

```
sr_subset <- sr_subset %>%
  mutate(month = ifelse(month == 1 & day(date) < 15, 12, month - 1),
         month = ifelse(month >= 2 & month <= 12, month - 2, month + 10))
```

est <- feols(hits ~ After*Treated | keyword + month | 0 | keyword, data = data, cluster = keyword)

^This is the formula I think, but I can't get it to work without an error message.