

Práctica: Diferencias en Diferencias

Carlos Bautista

Mario Heredia

Mariana Lugo

2022-12-01

De acuerdo con el estudio de Card y Krueger (1994) analizaremos la relación entre un aumento del salario mínimo y el empleo.

La teoría económica sugiere que los aumentos del salario mínimo conducen a una reducción del empleo por al menos dos razones:

- Las empresas tienen menos probabilidades de contratar personal y prefieren invertir en otros recursos que ahora son más baratos debido al aumento salarial.
- Los salarios más altos inducen a las empresas a subir sus precios para compensar sus mayores costos; a medida que aumentan los precios, los compradores disminuyen, lo que conducirá a una menor demanda y empleo.

Uno de los primeros estudios que analizó este problema de política fue el de Card y Krueger, quienes aplicaron el método de diferencias en diferencias en dos grupos de restaurantes de comida rápida: restaurantes de comida rápida en Nueva Jersey donde el salario mínimo aumentó de \$4.25 a \$5.05 por hora (grupo de tratamiento) y restaurantes de comida rápida en Pensilvania donde el salario mínimo no aumentó (grupo de control).

Con ayuda del conjunto de datos anexo se pide analizar si los aumentos del salario mínimo afectan al empleo.

Se cargan las librerías y datos utilizados en la práctica, además se muestra un ejemplo de los datos:

```
library(tidyverse)
library(readr)
library(kableExtra)
library(dplyr)
library(stargazer)
library(car)
path = "/home/mariana/Documents/MCD/Diff_in_Diff_MCD/data/datadid.csv"
dta<- read.csv(path)
kable(head(dta), booktabs=TRUE) %>%
  kable_styling(latex_options = c("scale_down", "hold_position"))
```

ID	Chain	SouthJ	CentralJ	NorthJ	PA1	PA2	Shore	Group	Empl	C.Owned	Hours.Opening	Soda	Fries	Treatment
46	1	0	0	0	1	0	0	0	30.0	0	16.5	1.03	1.03	0
49	2	0	0	0	1	0	0	0	6.5	0	13.0	1.01	0.90	0
506	2	0	0	0	1	0	0	0	3.0	1	10.0	0.95	0.74	0
56	4	0	0	0	1	0	0	0	20.0	1	12.0	0.87	0.82	0
61	4	0	0	0	1	0	0	0	6.0	1	12.0	0.87	0.77	0
62	4	0	0	0	1	0	0	0	0.0	1	12.0	0.87	0.77	0

1. ¿El efecto en el grupo de control es significativo antes de implementar la política?

Se especifica el primer modelo 1 (m1) como:

$$Empleo_i = \beta_0 + \beta_1Tiempo + \beta_2Tratamiento + \beta_3TiempoTratamiento + u$$

. En nuestros datos la variable que se refiere al tiempo es la denominada como *Treatment* y la variable que indica el grupo de tratamiento es la denominada como *Group*.

```
m1 <- lm(Empl ~ Treatment + Group + (Treatment*Group) , data = dta)
stargazer(m1, type="text")
```

```
##
## =====
##                      Dependent variable:
##                      -----
##                      Empl
## -----
## Treatment              -2.640**
##                      (1.330)
##
## Group                  -2.481**
##                      (1.044)
##
## Treatment:Group         3.362**
##                      (1.481)
##
## Constant               10.205***
##                      (0.938)
##
## -----
## Observations              802
## R2                       0.008
## Adjusted R2              0.004
## Residual Std. Error      8.281 (df = 798)
## F Statistic              2.122* (df = 3; 798)
## =====
## Note:                    *p<0.1; **p<0.05; ***p<0.01
```

El coeficiente de la regresión $\beta_0 = 10.20513$, que corresponde al intercepto, resultó significativo en el modelo 1. Es decir, podemos afirmar que el número de empleados (variable de respuesta) promedio en Pensilvania (grupo de control) antes del aumento del salario mínimo en Nueva Jersey (tratamiento) es significativo.

```
kable(m1$coefficients[1])
```

	x
(Intercept)	10.20513

2. ¿Los grupos de tratamiento y de control muestran alguna diferencia significativa antes de implementar la política?

La diferencia entre el grupo de tratamiento y control antes de implementar la política se refiere al coeficiente β_2 . En el modelo 1, el coeficiente β_2 resultó significativo por lo que la diferencia del número de empleados en Pensilvania (PA), como grupo de control, y Nueva Jersey (NJ), como grupo de tratamiento, es significativa antes del aumento en el salario mínimo.

```
kable(m1$coefficients[3])
```

	x
Group	-2.481202

3. ¿En el grupo de tratamiento existe algún cambio significativo entre el primer y el segundo período de tiempo?

El cambio en el grupo de tratamiento del primer y segundo período está representado por $\beta_1 + \beta_3 = 0.7215565$.

```
b1_b3<-m1$coefficients[2]+m1$coefficients[4]
kable(b1_b3)
```

	x
Treatment	0.7215565

```
linearHypothesis(m1, c('Treatment + Treatment:Group = 0'))
```

```
## Linear hypothesis test
##
## Hypothesis:
## Treatment + Treatment:Group = 0
##
## Model 1: restricted model
## Model 2: Empl ~ Treatment + Group + (Treatment * Group)
##
##   Res.Df  RSS Df Sum of Sq    F Pr(>F)
## 1     799 54813
## 2     798 54728   1    84.209 1.2279 0.2682
```

De acuerdo con el *p-value* del estadístico F no podemos rechazar la hipótesis nula. Es decir, no podemos afirmar que la diferencia en el número de empleados NJ antes y después del aumento del salario mínimo es significativo.

##. 4. Después del tratamiento ¿Hay algún efecto significativo en el grupo de tratamiento sobre el de control? Es decir, ¿tiene impacto la política implementada?

El impacto del tratamiento en la variable de respuesta es el estimador β_3 de diferencias en diferencias. En el modelo 1, β_3 resultó significativo por lo que el incremento en el salario tuvo un impacto positivo en el empleo de 3.3 empleados en promedio.

```
kable(m1$coefficients[4])
```

	x
Treatment:Group	3.36175

Modelo con variables de control

Con el objetivo de obtener estimaciones más precisas, hacemos una mejor especificación del modelo incorporando variables de control al modelo, siguiendo lo establecido por Card y Krueger (1994). En ese sentido, se pretende controlar por diferencias entre restaurantes: por ejemplo, las variables de los precios de sodas y papas a la francesa ya que restaurantes con precios más elevados podrían contratar más empleados; o si algunos restaurantes abren por más tiempo que otros, esto también tendría un efecto en el número de empleados que contratan.

Por tanto, el modelo a estimar es el siguiente:

```
m2 <- lm(Empl ~ Treatment + Group + (Treatment*Group) +
          C.Owned + Hours.Opening + Soda + Fries +
          as.factor(Chain) + SouthJ + CentralJ + NorthJ +
          PA1 + PA2 + Shore, data = dta)

stargazer(m2, type="text")
```

```
##
## =====
##                               Dependent variable:
##                               -----
##                               Empl
## -----
## Treatment                    -1.316
##                               (1.293)
##
## Group                        -3.065**
##                               (1.285)
##
## C.Owned                      -0.129
##                               (0.681)
##
## Hours.Opening                0.550***
##                               (0.099)
##
## Soda                         2.727
##                               (4.260)
##
## Fries                        -0.646
##                               (4.065)
##
## as.factor(Chain)2            -0.595
##                               (1.089)
##
## as.factor(Chain)3            -1.798**
##                               (0.907)
##
## as.factor(Chain)4            2.230**
##                               (1.055)
##
## SouthJ                       -2.457***
##                               (0.829)
##
```

```
## CentralJ                -1.366
##                        (0.865)
##
## NorthJ
##
##
## PA1                     -2.405*
##                        (1.307)
##
## PA2
##
##
## Shore                   0.840
##                        (1.047)
##
## Treatment:Group        4.124***
##                        (1.414)
##
## Constant                1.617
##                        (4.658)
##
## -----
## Observations            757
## R2                     0.126
## Adjusted R2            0.110
## Residual Std. Error    7.670 (df = 742)
## F Statistic            7.670*** (df = 14; 742)
## =====
## Note:                  *p<0.1; **p<0.05; ***p<0.01
```

Respondiendo las mismas preguntas para el modelo 1 tenemos que:

1. ¿El efecto en el grupo de control es significativo antes de implementar la política?

El coeficiente b_0 no es estadísticamente significativo, es decir, no tenemos evidencia para afirmar que la media de PA (grupo de control) antes del aumento en el salario mínimo es diferente de cero

```
kable(m2$coefficients[1])
```

	x
(Intercept)	1.616835

2. ¿Los grupos de tratamiento y de control muestran alguna diferencia significativa antes de implementar la política?

Sí, el coeficiente β_2 es significativo en el modelo 2, por lo que podemos afirmar que los restaurantes de comida rápida de NJ tiene 3 empleados menos en promedio que los de PA, en el periodo previo al aumento en el salario mínimo

```
kable(m2$coefficients[3])
```

	x
Group	-3.065444

3. ¿En el grupo de tratamiento existe algún cambio significativo entre el primer y el segundo período de tiempo?

De acuerdo con el *p-value* del estadístico F tampoco podemos rechazar la hipótesis nula para el modelo 2. Es decir, no podemos afirmar que la diferencia en el número de empleados NJ antes y después del aumento del salario mínimo es significativo.

```
b1_b3<-m2$coefficients[2]+m2$coefficients[17]
kable(b1_b3)
```

	x
Treatment	2.807429

```
linearHypothesis(m1, c('Treatment + Treatment:Group = 0'))
```

```
## Linear hypothesis test
##
## Hypothesis:
## Treatment + Treatment:Group = 0
##
## Model 1: restricted model
## Model 2: Empl ~ Treatment + Group + (Treatment * Group)
##
##   Res.Df    RSS Df Sum of Sq    F Pr(>F)
## 1      799 54813
## 2      798 54728   1    84.209 1.2279 0.2682
```

##. 4. Después del tratamiento ¿Hay algún efecto significativo en el grupo de tratamiento sobre el de control? Es decir, ¿tiene impacto la política implementada?

En el modelo 2, β_3 , que representa el coeficiente de diferencias en diferencias, es significativo. Podemos afirmar que el aumento en el salario mínimo en NJ resultó en un incremento de 4 empleados en promedio.

```
kable(m2$coefficients[17])
```

	x
Treatment:Group	4.12365