PS5

Christian Birchler   Fenqi Guo   Mingrui Zhang   Wenjie Tu   Zunhan Zhang

Names are listed in alphabetical order

Difference-in-Difference

1. Identification in the DiD model with constant effects

$$y_{it} = \delta_0 + \delta_1 \mathbf{1}(T_t = 2) + \delta_2 \mathbf{1}(D_i = 1) + \delta_3 \mathbf{1}(T_t = 2)\mathbf{1}(D_i = 1) + u_{it}$$

$$u_{it} = \lambda_t + \theta_i + \epsilon_{it}$$

1(a)

Before treatment

$$\begin{cases} \mathbb{E}(y_{it}|T_t = 1, D_i = 0) = \delta_0 + \mathbb{E}(\lambda_t|T_t = 1) + \mathbb{E}(\theta_i|D_i = 0) & \text{period 1 before treatment} \\ \mathbb{E}(y_{it}|T_t = 2, D_i = 0) = \delta_0 + \delta_1 + \mathbb{E}(\lambda_t|T_t = 2) + \mathbb{E}(\theta_i|D_i = 0) & \text{period 2 before treatment} \end{cases}$$

$$\mathbb{E}(y_{it}|T_t = 2, D_i = 0) - \mathbb{E}(y_{it}|T_t = 1, D_i = 0) = \delta_1 + \mathbb{E}(\lambda_t|T_t = 2) - \mathbb{E}(\lambda_t|T_t = 1)$$

After treatment

$$\begin{cases} \mathbb{E}(y_{it}|T_t = 1, D_i = 1) = \delta_0 + \delta_2 + \mathbb{E}(\lambda_t|T_t = 1) + \mathbb{E}(\theta_i|D_i = 1) & \text{period 1 after treatment} \\ \mathbb{E}(y_{it}|T_t = 2, D_i = 1) = \delta_0 + \delta_1 + \delta_2 + \delta_3 + \mathbb{E}(\lambda_t|T_t = 2) + \mathbb{E}(\theta_i|D_i = 1) & \text{period 2 after treatment} \end{cases}$$

$$\mathbb{E}(y_{it}|T_t = 2, D_i = 1) - \mathbb{E}(y_{it}|T_t = 1, D_i = 1) = \delta_1 + \delta_3 + \mathbb{E}(\lambda_t|T_t = 2) - \mathbb{E}(\lambda_t|T_t = 1)$$

DiD estimator

$$\begin{aligned} \Delta^{DiD} &= \{\mathbb{E}(y_{it}|T_t = 2, D_i = 1) - \mathbb{E}(y_{it}|T_t = 1, D_i = 1)\} - \{\mathbb{E}(y_{it}|T_t = 2, D_i = 0) - \mathbb{E}(y_{it}|T_t = 1, D_i = 0)\} \\ &= \delta_1 + \delta_3 + \mathbb{E}(\lambda_t|T_t = 2) - \mathbb{E}(\lambda_t|T_t = 1) - (\delta_1 + \mathbb{E}(\lambda_t|T_t = 2) - \mathbb{E}(\lambda_t|T_t = 1)) \\ &= \delta_3 \end{aligned}$$

We see that taking first difference eliminates θ_i across groups and taking difference in differences eliminates λ_t across periods.

1(b)

In order to identify the DiD estimator, we need to assume the error term ϵ_{it}

- $\mathbb{E}(\epsilon_{it}) = 0$. The error term has a zero mean.
- $Cov(\epsilon_{it}, T_t) = 0$. The error term is uncorrelated with the period-specific term.
- $Cov(\epsilon_{it}, D_i) = 0$. The error term is uncorrelated with the group-specific term.
- $Cov(\epsilon_{it}, T_t, D_i) = 0$. This ensures that period-specific term and group-specific term are uncorrelated with each other.

2. Replication of Card and Krueger (1994)

2(a)

```
# read csv file
raw.data <- read.csv('did_card_krueger.csv')

# drop rows with sample equal to 0
data <- raw.data[raw.data$sample==1, ]
```

```
# check classes of all dataframe columns
str(data)
```

$$\text{state} \begin{cases} 1 & \text{NJ} \\ 0 & \text{PA} \end{cases}$$

```
# first interview wave
ave.fte.NJ <- mean(data$fte[data$state==1])
ave.fte.PA <- mean(data$fte[data$state==0])

# second interview wave
ave.fte2.NJ <- mean(data$fte2[data$state==1])
ave.fte2.PA <- mean(data$fte2[data$state==0])

cat(sprintf('The average employment in NJ for the first interview wave is %.2f
           \nThe average employment in PA for the first interview wave is %.2f
           \nThe average employment in NJ for the second interview wave is %.2f
           \nThe average employment in PA for the second interview wave is %.2f',
           ave.fte.NJ, ave.fte.PA, ave.fte2.NJ, ave.fte2.PA))
```

The average employment in NJ for the first interview wave is 17.28

The average employment in PA for the first interview wave is 20.11

The average employment in NJ for the second interview wave is 17.56

The average employment in PA for the second interview wave is 18.10

```r
# employment difference in NJ (treatment)
ave.dfte.NJ <- ave.fte2.NJ - ave.fte.NJ

# employment difference in PA (comparison)
ave.dfte.PA <- ave.fte2.PA - ave.fte.PA

cat(sprintf(
'Difference in average employment between second and first interviews in NJ is %.2f
\nDifference in average employment between second and first interviews in PA is %.2f',
ave.dfte.NJ, ave.dfte.PA))
```

Difference in average employment between second and first interviews in NJ is 0.29

Difference in average employment between second and first interviews in PA is -2.02

```r
didfte <- ave.dfte.NJ - ave.dfte.PA

cat(sprintf('The difference between NJ and PA of the time differences is %.2f',
            didfte))
```

The difference between NJ and PA of the time differences is 2.30

In order to obtain a valid estimate of the employment effect, the following conditions must be fulfilled:

- Treatment (minimum wage increase) unrelated to outcome (employment) at baseline. In other words, the allocation of treatment was not determined by outcome.
- Parallel Trend Assumption. In the absence of treatment, the difference between the treatment group (NJ) and the comparison group (PA) is constant over time.
- Stable Unit Treatment Value Assumption. The potential outcomes for any unit do not vary with the treatments assigned to other units. For each unit, there is no different forms or versions of each treatment level, which lead to different potential outcomes.

From the DiD estimator, we see an increase in employment in New Jersey where the minimum wage was increased.

2(b)

```r
# subset data in NJ
data.NJ <- data[data$state==1, ]

# high wage in NJ (comparison)
data.NJ.hw <- data.NJ[data.NJ$wage_st>=5, ]

# low wage in NJ (treatment)
data.NJ.lw <- data.NJ[data.NJ$wage_st<5, ]

# DiD estimator of starting wages obtained within NJ data
didw.NJ <- mean(data.NJ.lw$dw)-mean(data.NJ.hw$dw)

# DiD estimator of employment obtained within NJ data
```

```r
didfte.NJ <- mean(data.NJ.lw$dfte)-mean(data.NJ.hw$dfte)

cat(sprintf('The relative impact of the minimum wage on starting wages is %.2f
            \nThe relative impact of the minimum wage on employment is %.2f',
            didw.NJ, didfte.NJ))
```

```
## The relative impact of the minimum wage on starting wages is 0.62
##
## The relative impact of the minimum wage on employment is 3.30
```

2(c)

```r
library(stargazer)
```

```r
# regress change of employment on state
dfte.did1 <- lm(dfte ~ as.factor(state), data=data)

# regress change of starting wages on state
dw.did1 <- lm(dw ~ as.factor(state), data=data)

# regress change of employment on state and other control variables
dfte.did2 <- lm(dfte ~ as.factor(state) +
                  as.factor(co_owned) + as.factor(chain), data=data)

# regress change of starting wages on state and other control variables
dw.did2 <- lm(dw ~ as.factor(state) +
                as.factor(co_owned) + as.factor(chain), data=data)

stargazer(dfte.did1, dw.did1, dfte.did2, dw.did2, keep='state',
          header=F, font.size='small', keep.stat=c('n', 'rsq'),
          title='Comparison between regressions with and without control variables',
          column.labels=c('No controls', 'No controls', 'Controls', 'Controls'))
```

Table 1: Comparison between regressions with and without control variables

	Dependent variable:			
	dfte No controls	dw No controls	dfte Controls	dw Controls
	(1)	(2)	(3)	(4)
as.factor(state)1	2.302^{**}	0.504^{***}	2.297^{*}	0.504^{***}
	(1.167)	(0.048)	(1.174)	(0.047)
Observations	351	351	351	351
R^2	0.011	0.243	0.021	0.279
Note:				*p<0.1; **p<0.05; ***p<0.01

- As we can see from column(1) to column(3), the DiD estimate of the employment slightly decreases while the standard errors slightly increase. Both estimates are consistent with the estimate (2.30) obtained in 2(a).

- As we can see from column(2) to column(4), the DiD estimate of the wage remains the same while the standard errors slightly decrease. Both estimates are smaller than the estimate (0.62) obtained in 2(b).

2(d)

```r
# subset PA data
data.PA <- data[data$state==0, ]

# create a new variable like gap for restaurants in PA
data.PA$gap4PA <- ifelse(data.PA$wage_st<5.05, (5.05-data.PA$wage_st)/data.PA$wage_st, 0)
```

```r
# regress change in employment on gap4PA (PA)
dfte.PA <- lm(dfte ~ gap4PA, data=data.PA)

# regress change in wages on gap4PA (PA)
dw.PA <- lm(dw ~ gap4PA, data=data.PA)

# regress change in employment on gap (NJ)
dfte.NJ <- lm(dfte ~ gap, data=data[data$state==1, ])

# regress change in wages on gap (NJ)
dw.NJ <- lm(dw ~ gap, data=data[data$state==1, ])

stargazer(dfte.PA, dw.PA, dfte.NJ, dw.NJ, header=F,
          keep.stat=c('n', 'rsq'), omit='Constant',
          column.labels=c('PA', 'PA', 'NJ', 'NJ'),
          title='Placebo test')
```

Table 2: Placebo test

	Dependent variable:			
	dfte PA	dw PA	dfte NJ	dw NJ
	(1)	(2)	(3)	(4)
gap4PA	1.121 (18.142)	2.544*** (0.540)		
gap			14.259** (6.468)	4.379*** (0.111)
Observations	66	66	285	285
R^2	0.0001	0.258	0.017	0.846
Note:		*p<0.1; **p<0.05; ***p<0.01		

- Column(1) and column(3) show that the coefficient on gap for Pennsylvania is small and not statistically different from zero while the coefficient on gap for New Jersey is large and statistically different from zero at 5% significance level. We cannot reject the null hypothesis in column(1).

- Column(2) and column(4) show that the coefficient on gap for Pennsylvania is smaller than the coefficient on gap for New Jersey. We can see that coefficient from $\text{dw} = \beta \cdot \text{gap} + \epsilon$ captures the impact of minimum wage increase on starting wage in those restaurants that were required to raise their wage when the policy was put in place. Such effects are stronger in New Jersey than in Pennsylvania.

We know that there was no minimum wage increase in Pennsylvania during that period. We perform a **placebo test** using a "fake" treatment group. In this setting, the "fake" treatment group is the Pennsylvania where the minimum wage was assumed to be increased but actually remained at the federal level. Since the "fake" treatment group should not be affected by the program, we should expect small and insignificant coefficient in column(1) compared to the coefficient in column(3) and smaller coefficient in column(2) than in column(4). The regression results are consistent with our expectation and support the parallel trend assumption.