

# SISCER Module 15 Practice 3: DID

Ting Ye and Qingyuan Zhao

July 2022

## 1

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --
## v ggplot2 3.3.6      v purrr  0.3.4
## v tibble  3.1.7      v dplyr  1.0.9
## v tidyr   1.2.0      v stringr 1.4.0
## v readr   2.1.2      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

library(clubSandwich)

## Registered S3 method overwritten by 'clubSandwich':
##   method      from
##   bread.mlm    sandwich

library(ggplot2)
library(did)
library(scales)

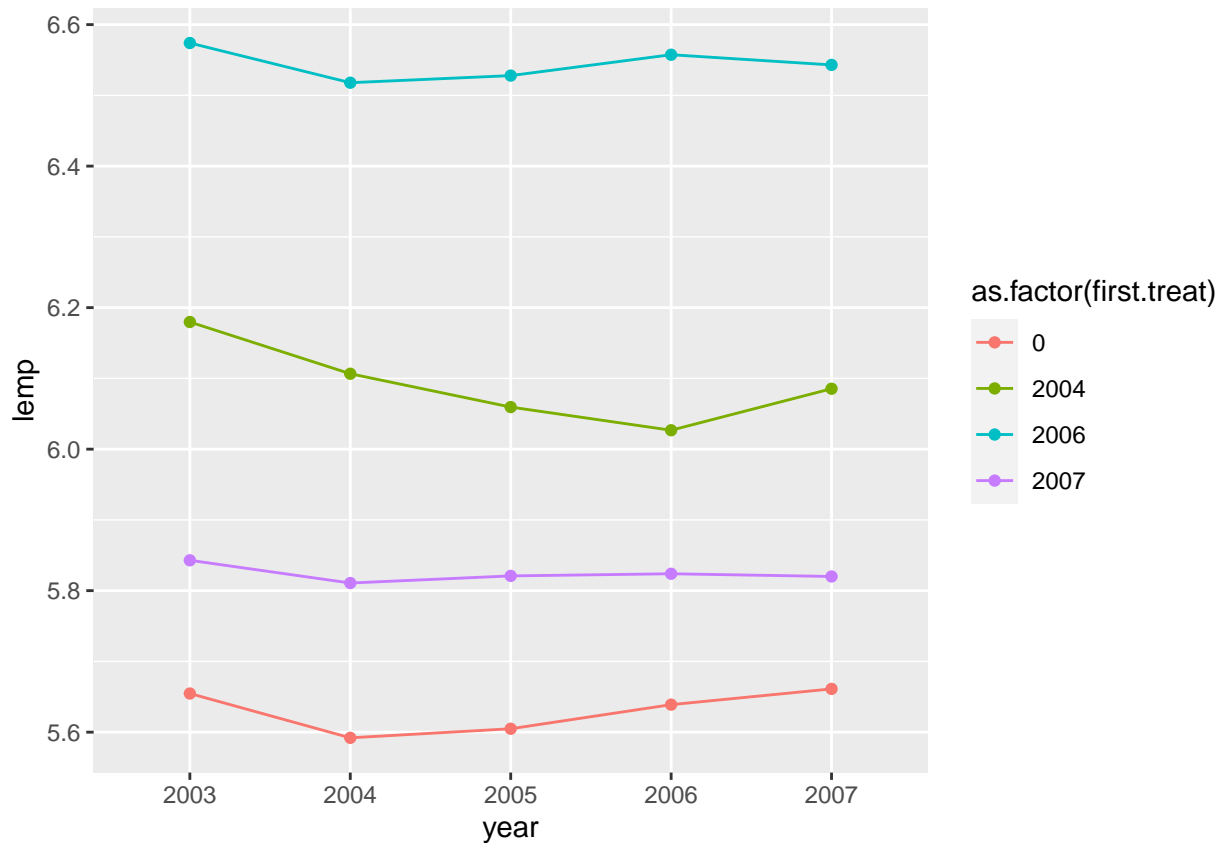
##
## Attaching package: 'scales'

## The following object is masked from 'package:purrr':
##
##   discard

## The following object is masked from 'package:readr':
##
##   col_factor

data("mpdta")

mpdta<-mpdta %>% mutate(after.ind=1*(year>=first.treat))
ggplot(data=mpdta, aes(x=as.factor(year),y=lemp,color=as.factor(first.treat),group=as.factor(first.treat)))
  stat_summary(fun = mean, geom = 'line')+stat_summary(fun = mean, geom = 'point')+xlab("year")+ylab("l"
```



```
# fixed TWFE
fit.twfe<-lm(lemp~after.ind+as.factor(year)+as.factor(countyreal),data=mpdta)
summary(fit.twfe)$coefficients[2,] # unadjusted
```

```
##      Estimate   Std. Error    t value   Pr(>|t|)
## -0.036548937  0.012646474  -2.890049673  0.003893424
```

```
# cluster-robust standard error with CR2 small-sample correction
```

```
coeftest.twfe <- coef_test(fit.twfe,
                           vcov = "CR2",
                           cluster = mpdta$countyreal)
coeftest.twfe[2,]
```

```
##      Coef. Estimate   SE t-stat d.f. (Satt) p-val (Satt) Sig.
## after.ind -0.0365 0.0133 -2.75      280      0.00627 **
```

```
fit.twfe.adj<-lm(lemp~after.ind+as.factor(year)+as.factor(countyreal)+lpop,data=mpdta)
```

```
# dynamic TWFE
```

```
gen_reg <- function(df){
  exposure <- c("eff_pre_3", "eff_pre_2", "eff_pre_1",
              "eff_0", "eff_post_1", "eff_post_2", "eff_post_3")

  regs <- c(exposure, "factor(countyreal)", "factor(year)")

  covs <- paste0(regs, collapse = " + ")
  form <- formula(sprintf("%s ~ %s", "lemp", covs))
```

```

mod <- lm(form, data = df)

# using arguments from here:
# https://jepusto.github.io/clubSandwich/articles/panel-data-CRVE.html
coefptest <- coef_test(mod,
                       vcov = "CR2",
                       cluster = df$countyreal,
                       test = "Satterthwaite")

# adding confidence interval
coefptest <- coefptest %>%
  mutate(lower_ci = beta - (1.96 * SE),
         upper_ci = beta + (1.96 * SE))

list(form = form,
     mod = mode,
     coefptest = coefptest)
}

graph_event <- function(res){

# choose relevant variables
plot_df <- res$coefptest[2:8,] %>%
  mutate(t = seq(-3, 3, 1),

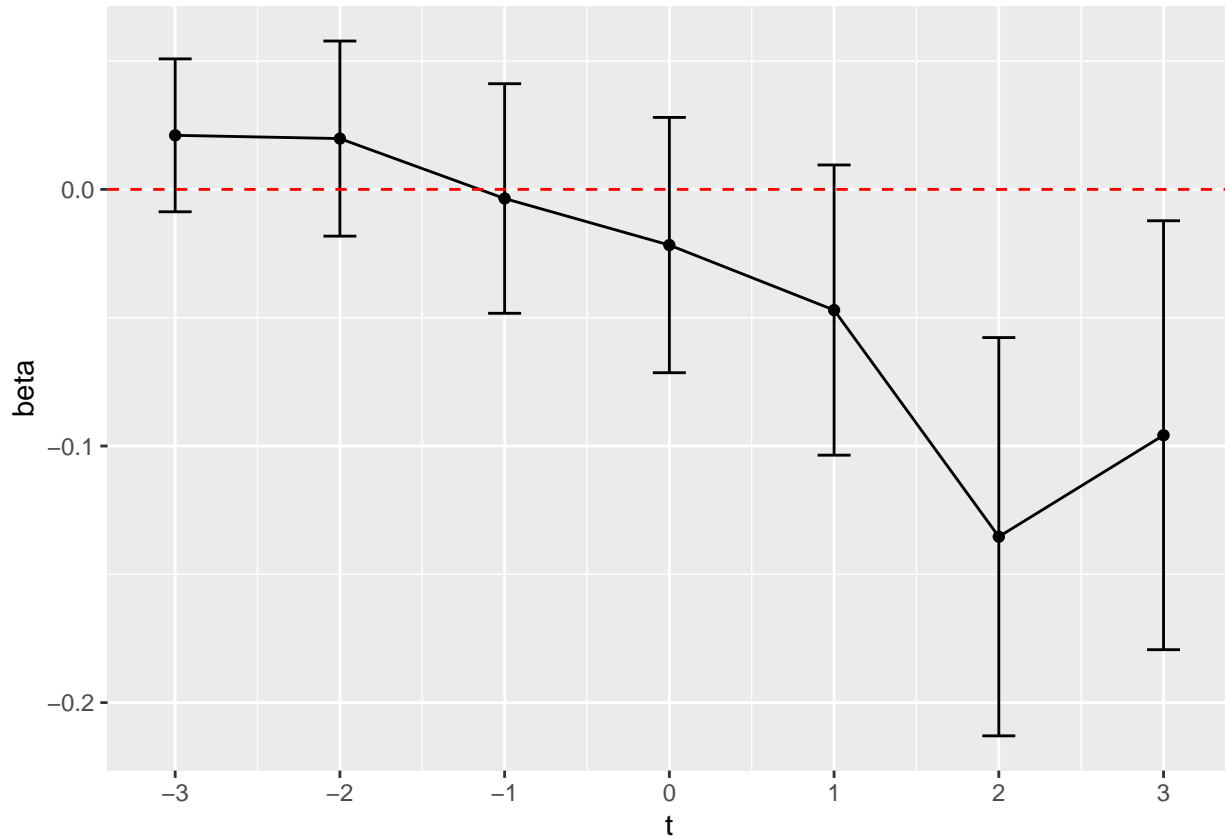
         # calculate upper and lower confidence intervals
         lower_ci = beta - (1.96 * SE),
         upper_ci = beta + (1.96 * SE)) %>%
  data.frame()

# draw plot
ggplot(plot_df, aes(x = t, y = beta)) + geom_point() + geom_line() +
  geom_errorbar(aes(ymin = lower_ci, ymax = upper_ci), width=.2,
               position=position_dodge(0.05)) +
  scale_x_continuous(breaks= pretty_breaks()) +
  geom_hline(yintercept=0, linetype="dashed", color = "red")
}

mpdta<- mpdta %>%
  mutate(
    eff_pre_3 = ifelse(year == first.treat - 3, 1, 0),
    eff_pre_2 = ifelse(year == first.treat - 2, 1, 0),
    eff_pre_1 = ifelse(year == first.treat - 1, 1, 0),
    eff_0 = ifelse(year == first.treat, 1, 0),
    eff_post_1 = ifelse(year == first.treat + 1, 1, 0),
    eff_post_2 = ifelse(year == first.treat + 2, 1, 0),
    eff_post_3 = ifelse(year == first.treat + 3, 1, 0))

res<-gen_reg(mpdta)
graph_event(res)

```



```
fit.twfe.dynamic<-lm(lemp~eff_pre_3+eff_pre_2+eff_pre_1+eff_0+eff_post_1+eff_post_2+eff_post_3+
as.factor(year)+as.factor(countyreal),data=mpdta)
```

```
# cluster-robust standard error with CR2 small-sample correction
```

```
coeftest.twfe.dynamic <- coef_test(fit.twfe.dynamic,
vcov = "CR2",
cluster = mpdta$countyreal)
```

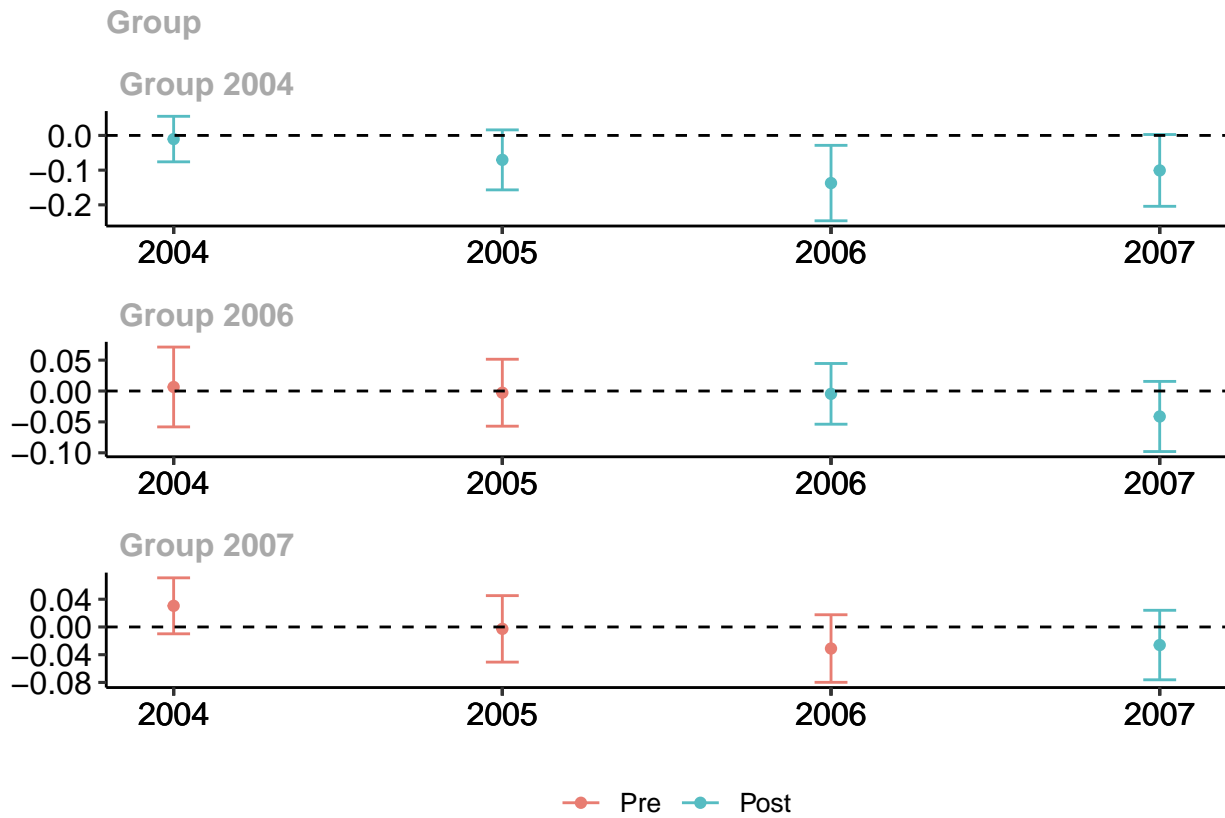
```
coeftest.twfe.dynamic[2:8,]
```

	Coef.	Estimate	SE	t-stat	d.f.	(Satt)	p-val	(Satt)	Sig.
##	eff_pre_3	0.02107	0.0152	1.387	231.8		0.16685		
##	eff_pre_2	0.01981	0.0194	1.021	257.3		0.30825		
##	eff_pre_1	-0.00355	0.0228	-0.156	267.5		0.87654		
##	eff_0	-0.02169	0.0254	-0.855	284.6		0.39340		
##	eff_post_1	-0.04702	0.0288	-1.630	177.8		0.10487		
##	eff_post_2	-0.13534	0.0396	-3.418	36.8		0.00155	**	
##	eff_post_3	-0.09580	0.0426	-2.247	38.3		0.03050	*	

```
## unadjusted
```

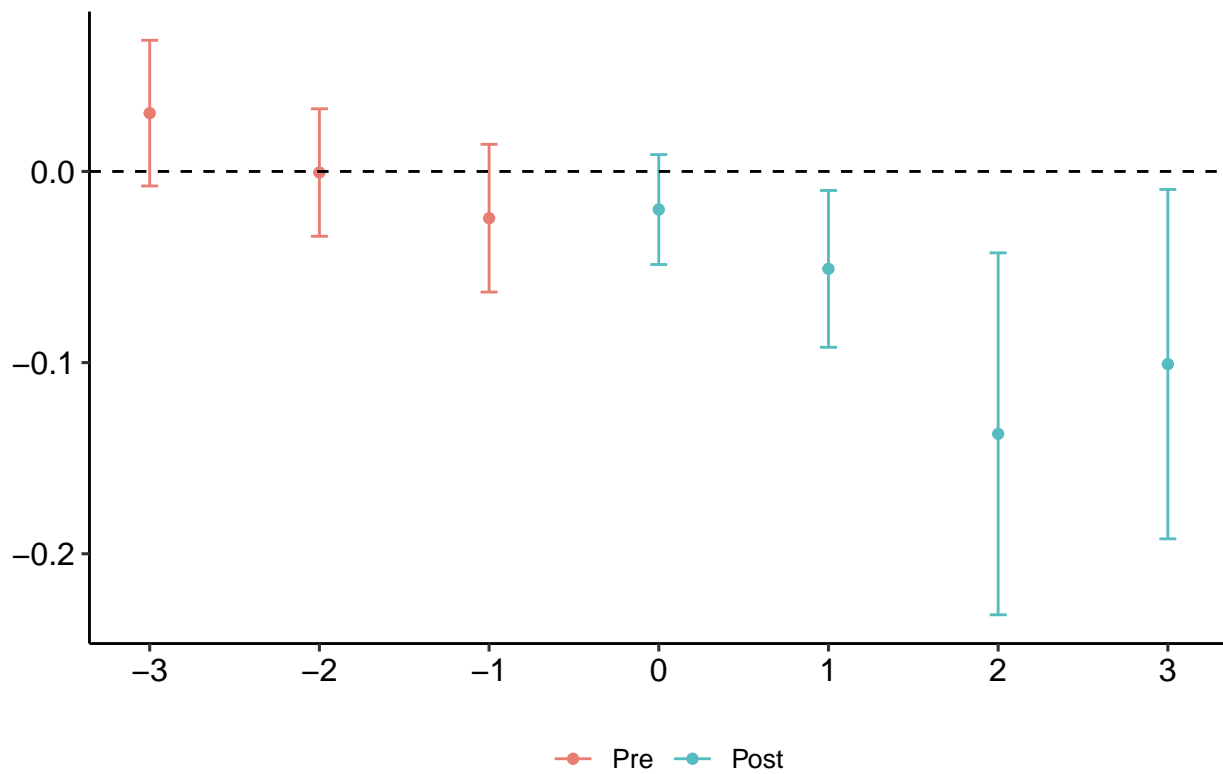
```
out <- att_gt(yname="lemp",
tname="year",
idname="countyreal",
gname="first.treat",
xformula=NULL,
data=mpdta)
```

```
ggdid(out) # group-specific treatment effect
```



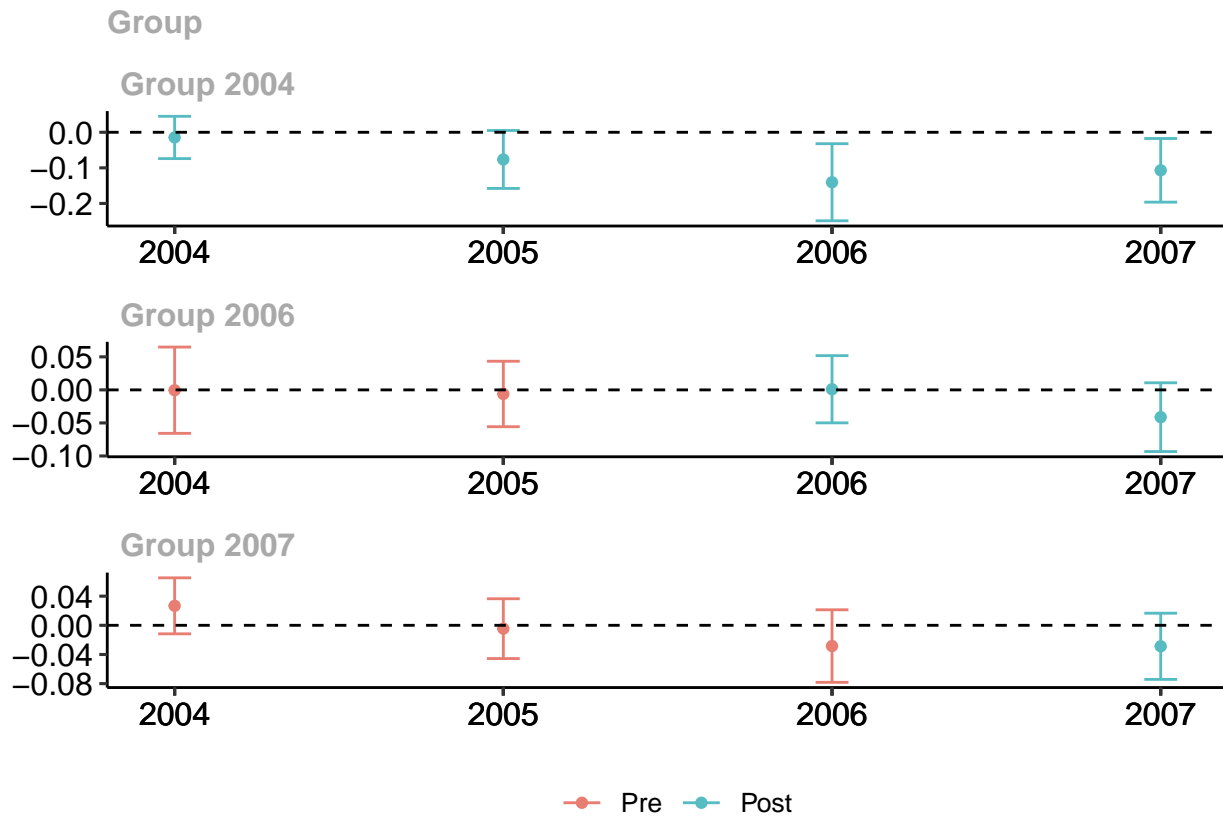
```
# average effect of treatment 0/1/2/3 time periods after adopting the treatment
agg_dyn <- aggte(out, type = "dynamic")
ggdid(agg_dyn)
```

Average Effect by Length of Exposure



```
## using covariates
out2 <- att_gt(yname="lemp",
  tname="year",
  idname="countyreal",
  gname="first.treat",
  xformula=~lpop,
  data=mpdta)

ggdid(out2) # group-specific treatment effect
```



```
# average effect of treatment 0/1/2/3 time periods after adopting the treatment
agg_dyn2 <- aggte(out2, type = "dynamic")
ggdid(agg_dyn2)
```

Average Effect by Length of Exposure

