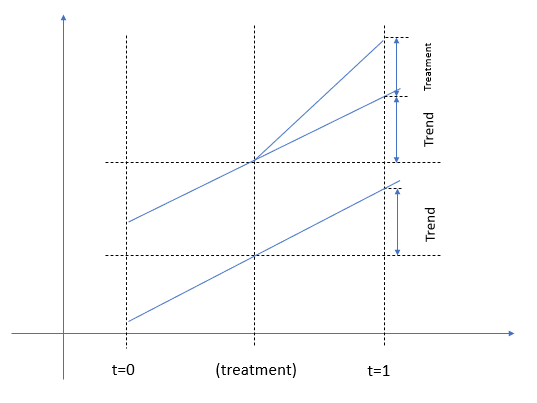
## Day 5: DID Analysis

Learn how to conduct DID Analysis in STATA

1. **Basic Idea**



Two time periods (), two groups (treatment group and control group).

Parallel trend assumption: the outcomes of both groups would have followed the same trend over time if treatment were not in place.

----------------------------------------------------

webuse hospdd

Several hospitals are using a new procedure for registration, we would like to see if the new procedure affects patients’ outcomes, measured by *satis*.

didregress (satis) (procedure), group(hospital) time(month)

-----*The model specifications should be enclosed in parentheses*

estat trendplots

estat ptrend

-----*Cannot reject null hypothesis, therefore parallel trend is satisfied*

estat granger

-----*‘Placebo’ test, cannot reject null hypothesis, therefore the effects exists*

*OR*

reghdfe satis procedure, absorb(hospital month)

-------------------------------------------------------

1. **Two-way fixed static DID model**

Beck et.al (2010), Big bad banks? The winners and losers from bank deregulation in the United States. *The Journal of Finance*, 65(5): 1637-1667.

is a dummy variable of our interest, in equals one in the years after unit *s* is treated.

Fixed effects: and

--------------------------------

replace procedure=0 if hospital<6 & month==4

----- *Assume hospitals from 1 to 5 are treated in May rather than April*

didregress (satis) (procedure), group(hospital) time(month)

----- *However, the pre-trend plot function are no longer optional*

reghdfe satis procedure, absorb(hospital month)

1. **Event study DID model**

Model from my own paper:

A black and white math symbols

Description automatically generated with medium confidence

With multiple treatment timing, we would like to know how treatment effects vary across time with an event study model.

-----*The canonical DID model might be biased in multiple time treatment settings. Event study model might also be biased but that’s another story.*

The basic idea is to compare the outcomes with base time ( here).

--------------------------------

webuse hospdd, clear

replace procedure=0 if hospital<10 & month==4

replace procedure=0 if hospital>5 &hospital<10 &month==5

----*Same as above, we load the data again, hospitals 1 to 5 are treated in May, hospitals 6 to 9 are treated in June, hospitals 10 to 18 are treated in April.*

gen a=month if procedure==1

bysort hospital: egen t\_month=min(a)

-----*We generate a variable describing when a hospital is treated.*

gen rel\_month=month-t\_month

forvalues i=1/4{

gen post`i’=1 if rel\_month==`i’

replace post`i’=0 if missing(post`i’)

}

forvalues i=1/7{

gen pre`i’=1 if rel\_month==`i’

replace pre`i’=0 if missing(pre`i’)

}

gen current=1 if rel\_month==0

replace current=0 if missing(current)

-----*Creating the D variables in the model*

(replace post7=1 if rel\_month>7)

(replace pre4=1 if rel\_month)

reghdfe satis pre7 pre6 pre5 pre4 pre3 pre2 current post1 post2 post3 post4, absorb(hospital month)

-----*As baseline, pre1 is not included here*

coefplot, baselevels keep(pre5 pre4 pre3 pre2 pre1 current post1 post2 post3) vertical coeflabels(pre5 = "-5" pre4 = "-4" pre3 = "-3" pre2 = "-2" current = "0" post1 = "1" post2 = "2" post3 = "3") yline(0) ylabel(-1(0.1)1) xline(4.5, lwidth(vthin) lpattern(dash) lcolor(teal)) ylabel(,labsize(\*0.75)) xlabel(,labsize(\*0.75)) ytitle("Dynamic Effect", size(small)) xtitle("Rel\_month", size(small)) addplot(line @b @at) ciopts(lpattern(dash) recast(rcap) msize(medium)) msymbol(circle\_hollow) scheme(s1mono)

---------------------