**DECOMPOSITION OF DID WITH NO ANTICIPATION VIOLATION**

No Anticipation means that Y=Y(0) at baseline for the treatment group

Step 1: Write down our DID equation (2x2):

Step 2: “**Switching Equation”** means to replace Y with either Y(1) or Y(0) depending on whether it is in that period treated (Y=Y(1)) or not treatd (Y=Y(0)).

Step 3: **“the econometric trick – add in zeroes for every Y(1)”.**

***Hint: Our goal is two fold: 1) we want to see an ATT parameter. 2) isolate a ‘parallel trends bias term’ which remember is ‘four average Y(0) and three subtractions”***

DID = ( E[Y(1)\_k|Post] – E[Y(1)\_k|Pre] ) – (E[Y(0)\_U|Post] – E[Y(0)\_U|Pre] )

+ E[Y(0)\_k|Post] – E[Y(0)\_k|Post]

+ E[Y(0)\_k|Pre] – E[Y(0)\_k|Pre]

Step 4: Rearrange the above expression into (1) an ATT term for the post-treatment period and the treatment group, *k*, and (2) a parallel trends term expressed **only** in Y(0) terms. When you do this, whatever is left over that is neither the ATT nor the parallel trends bias term **is the additional bias associated with losing No Anticipation.**

DID = E[Y(1)\_k|Post] – E[Y(0)\_k|Post] 🡨 ATT for group *k* in *post*. This is the target parameter

+ ( E[Y(0)\_k|Post] – E[Y(0)\_k|Pre] ) – (E[Y(0)\_U|Post] – E[Y(0)\_U|Pre] )

|-> “parallel trends bias term”

* ( E[Y(1)\_k|Pre] ) - E[Y(0)\_k|Pre] )

|-> A NA violation means that DID equals:  
  
ATT for group *k* in the post period (which is the goal of DD)

+ PT bias term (all expressed in Y(0))

* **ATT for group *k* at baseline.**

Let’s say for the sake of argument, that this treatment when applied to group k causes outcomes to rise by 10 points. No matter the period. So that means that the ATT in the pre period is 10 and the ATT in the post period is also 10.

Now assume parallel trends. And work out what the DID coefficient would be:

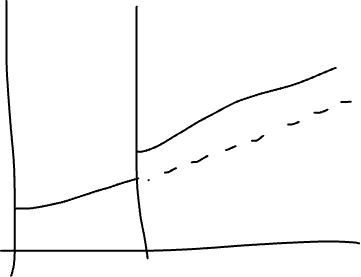
DD coefficient = ATT\_k(post) + PT – ATT\_k(pre)

DD coefficient = 10 + 0 – 10 = 0

Now assume that at baseline, the ATT is 5, and at post the ATT is 10 and PT=0.

DD coefficient = 10 + 0 **– 5 = 5**. Still biased.

Y(1)



ATT=10 in year 1

Y(0)

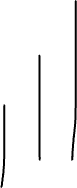
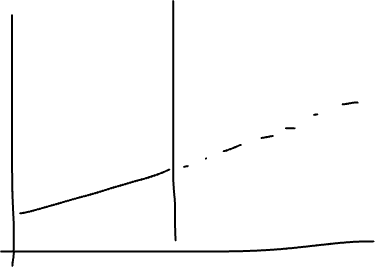
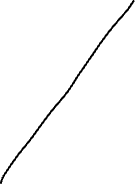


ATT = 10 in year 2

“Constant treatment effects” means treatment effect is the same over time.

So what is “dynamic treatment effecs”. That means the ATT is changing over time.

Y(1)



Y(0)

What happens if you use in diff-in-diff an already treated group as a control?

We know from that decomposition that the diff-in-diff coefficient **always equals the following with NA and already treated groups as a control:**

**DD = ATT\_k,post + PT bias – Change in ATT\_U**

**10 + 0 – 12= - 2**

**ATT\_k,post = 10**

**PT = 0**

**ATT\_U, pre = 5**

**ATT\_U, post = 17**

**Change in ATT\_U = ATT\_U,post – ATT\_U,pre = 17-5 = 12.**