Baker, et al (2025) “Practitioner’s Guide”

2x2 – four averages and three subtractions. You can identify the ATT with only two groups and only two periods if parallel trends holds from the pre period to the post period expressed as E[Y(0)].

* Two groups, two periods
* No anticipation, Parallel trends and untreated comparison group
* DID = ATT
* The challenge though is that with the 2x2, you cannot easily support the parallel trends claim.
* Every 2x2 **always is equal to**:
  + ATT + PT bias term
  + Under NA, all pre-treatment outcomes are Y(0) or don’t have a treatment effect. Which would mean, if you calculated a 2x2 on the pre-treatment period, under NA, those 2x2s equal PT because ATT=0 in the pre-period under NA.

2xT – Event study

* Two groups, but now you have T time periods (T>2).
* No anticipation, parallel trends, untreated comparison group
* Event study coefficients are each 2x2s and therefore each event study coefficient *no matter where it is found in the regression* is always equal to:
  + ATT + PT (PT bias terms).
  + Under NA, ATT=0, so therefore all pre-treatment event study coefficients are *only* equal to PT.
* When you are estimating an event study in a 2xT using OLS, each of the coefficients is a *long difference 2x2*.
  + **Pre-treatment:** When you run a standard OLS “interaction model” of the 2xT where you interact the treatment indicator with calendar time indicators, you have to drop a time period, and therefore all coefficients will be measured relative to that omitted time period. And this is called “long differences”. OLS cannot do it any other way.
  + **Pre-treatment:** But since each coefficient is a 2x2, we could *manually* have calculated pre-treatment coefficients as 2x2s but used the neighbor, not a fixed baseline, in which case you have a *rolling comparison*. This called sometimes “short gaps”. OLS cannot do this – we can do it manually, but OLS is not manual.
  + **Post-treatment:** Everyone, no matter the method, cannot calculate those post-treatment event study coefficients any other way than using a fixed baseline. So all post-treatment coefficients in an event study are always calculated using *fixed dropped omitted baseline time period*. Because it’s literally impossible to use “the neighbor” because the neighbor is treated and we need no anticipation (an untreated comparison).

2x2xT

GxT

DxT

Unconditional parallel trends

Males earnings grows +2 a year

Female earnings grows +1 a year

* Randomly assign the treatment to this population. What is the average earnings growth in the treatment group versus the control group.
* If it was truly random, then the treatment is assigned independent of Y(0) and it is also assigned independent of *changes* in Y(0) which would mean that the average change in Y(0) is the same in the treatment as it is in the control group because the covariates that cause Y(0) trends are balanced.
* Treatment group will be half male; control group will be half female.
* So what is the average change in Y(0) in the treatment group:
  + Treatment group: Change in E[Y(0)] = ( 0.5 x 2) + (0.5 x 1) = 1.5
  + Control group: Change in E[Y(0)] = ( 0.5 x 2) + (0.5 x 1) = 1.5

What if it was not an RCT? What if your observational study was like this:

* Treatment group: 75% male, 25% female
* Control group: 25% male, 75% female

What is the average change in Y(0) for the treatment group vs control group?

* Treatment group: (0.75\*2) + (0.25\*1) = 1.75. On average, untreated potential outcomes in the treatment group grow by 1.75 a year.
* Control group: (0.25\*2) + (0.75\*1) = 1.25. On average, in the control group, untreated potential outcomes grow by 1.25 a year.

THEREFORE, UNCONDITIONAL PARALLEL TRENDS DOES NOT HOLD**. And it does not hold because the covariates that cause trends in untreated potential outcomes are *imbalanced across the two groups***.

**Abadie IPW (inverse probability weighting)**

Weighted 2x2=

E [ ( **long difference**/average number of units treated ) x (D-p(x))/(1-p(x)) ]

**Long difference is “after minus before” average outcomes**.

Abadie IPW 2x2 = w1 x long diff + w0 x long diff

= ( **w1**(Y\_post, D=1) - w1 (Y\_Pre, D=1) ) - ( w0(Y0\_post, D=0) – w0 Y(\_Pre, D=0) )

What is **w1**? That is the weight on treated units (D=1).

W1 = 1/Pr(D=1)