Differences-in-Differences

Richard L. Sweeney

(based on notes from Anant Nyshadham)

Setup

$$Y_i = a + bT_i + cX_i + e_i$$

- We are interested in learning about the **causal impact** of some treatment (T) on an outcome of interest (Y)
- Some common Y's in this class are health outcomes, expenditure on housing, wages, and costs (ie electricity bills, gas prices ,etc)
- What's a "treatment"?
  - Sometimes we are interested in the impact of some continuous variable. As a first pass, it is easiest to reduce this to "high" and "low" levels
  - In policy work, we often want to evaluate some specific intervention: ie ban fracking, cap carbon, etc
- Important outcomes are generally a function of MANY variables X (not just the treatment)

Simple comparison is biased

If we just compare treated to untreated units, we will typically not recover the true causal effect (b)

(1) $Yi = a + bTi + cXi + ei$

(2) $E(Yi \mid Ti=1) - E(Yi \mid Ti=0)$

$= [a + b + cE(Xi \mid Ti=1) + E(ei \mid Ti=1)]$
$\quad - [a + 0 + cE(Xi \mid Ti=0) + E(ei \mid Ti=0)]$

$= b \qquad + \qquad c\,[E(Xi \mid Ti=1) - E(Xi \mid Ti=0)]$

True effect "Omitted variable/selection bias" term

What if we had data from before the program?

What if we estimated this equation using data from before any one got the treatment?

(1) $Yi = a + bTi + cXi + ei$

Specifically, what would our estimate of b be?

What if we had data from before the program?

What if we estimated this equation using data from before any one got the treatment?

(1) $Yi = a + bTi + cXi + ei$

(2) $E(Y_{i0}| T_{i1}=1) – E(Y_{i0}| T_{i1}=0)$

$= [a + \mathbf{0} + cE(X_{i0} | T_{i1}=1) + E(e_{i0}| T_{i1}=1)]$
$– [a + 0 + cE(X_{i0}| T_{i1}=0) + E(e_{i0}| T_{i1}=0)]$

$= c [E(Xi | Ti=1) – E(Xi | Ti=0)]$

"Omitted variable/selection bias" term

ALL THAT'S LEFT IS THE PROBLEMATIC TERM – HOW COULD THIS BE HELPFUL TO US?

Differences-in-Differences (just what it sounds like)

- Use two periods of data
  - add second subscript to denote time

$= \{E(Y_{i1} \mid T_{i1}=1) - E(Y_{i1} \mid T_{i1}=0)\}$ (difference btwn T&C, post)

$- \{E(Y_{i0} \mid T_{i1}=1) - E(Y_{i0} \mid T_{i1}=0)\}$ – *(difference btwn T&C, pre)*

$= b + c\,[E(X_{i1} \mid T_{i1}=1) - E(X_{i1} \mid T_{i1}=0)]$

$- c\,[E(X_{i0} \mid T_{i1}=1) - E(X_{i0} \mid T_{i1}=0)]$

Differences-in-Differences (just what it sounds like)

- Use two periods of data
  - add second subscript to denote time

$= \{E(Y_{i1} \mid T_{i1}=1) - E(Y_{i1} \mid T_{i1}=0)\}$ (difference btwn T&C, post)

$- \{E(Y_{i0} \mid T_{i1}=1) - E(Y_{i0} \mid T_{i1}=0)\}$ *– (difference btwn T&C, pre)*

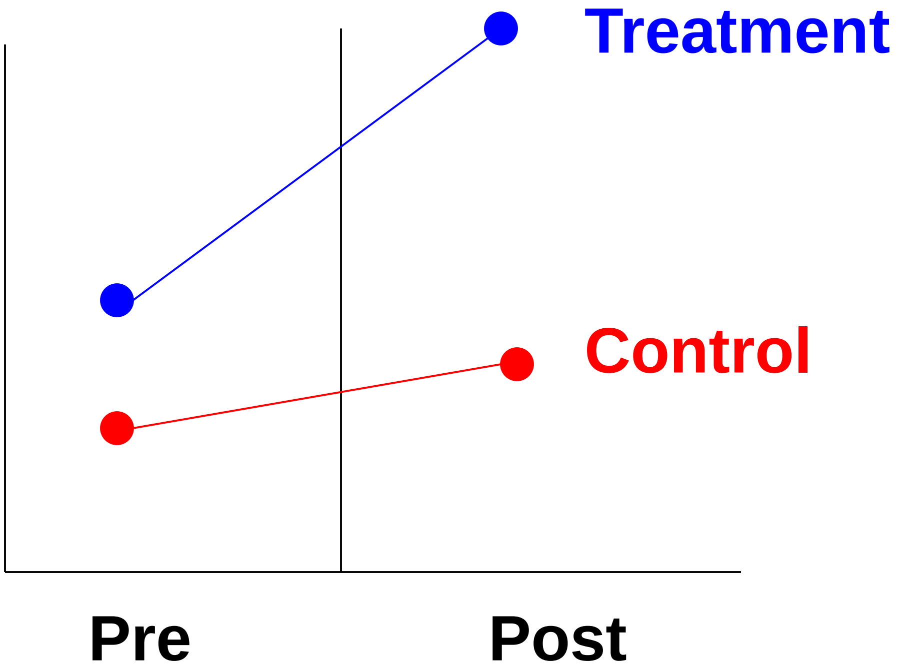
$= b + \cancel{c\,[E(X_{i1} \mid T_{i1}=1) - E(X_{i1} \mid T_{i1}=0)]}$

$\cancel{- c\,[E(X_{i0} \mid T_{i1}=1) - E(X_{i0} \mid T_{i1}=0)]}$

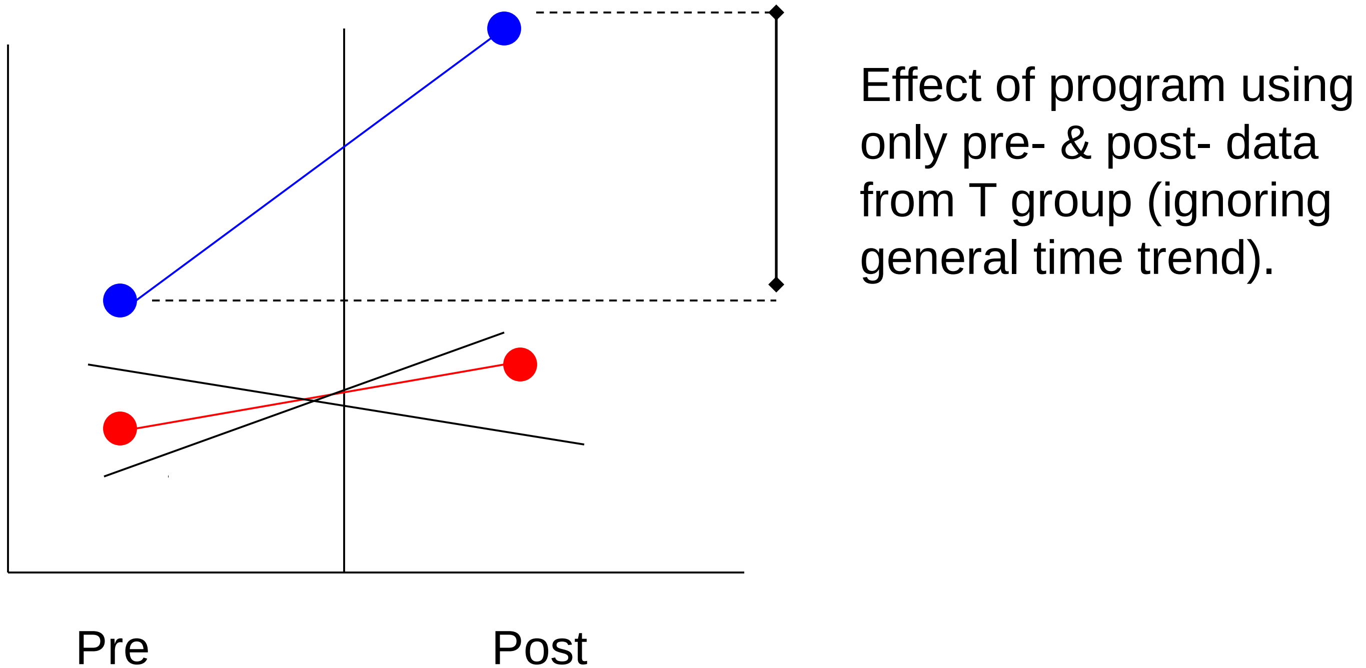
$= b$ *YAY!*

- Assume differences between X don't change over time.

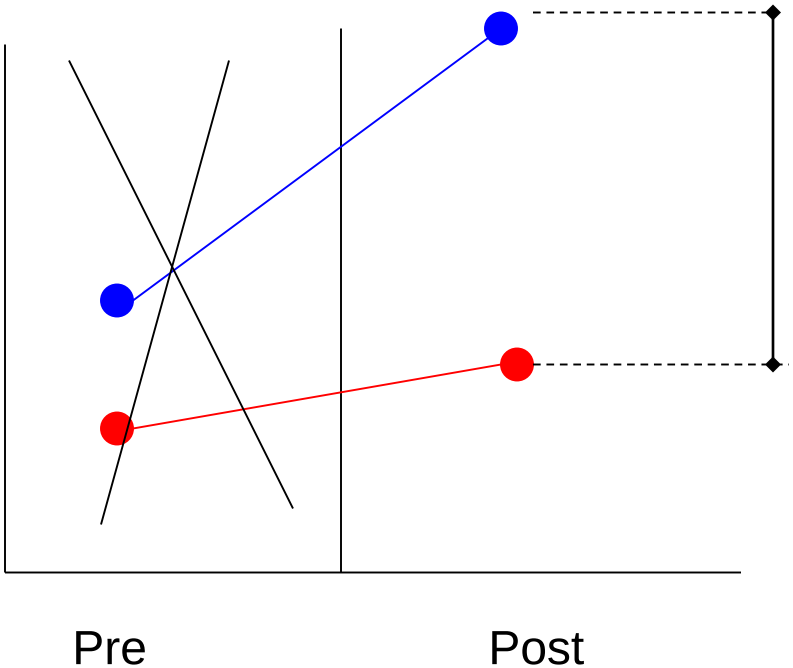
Differences-in-Differences, Graphically



Differences-in-Differences, Graphically

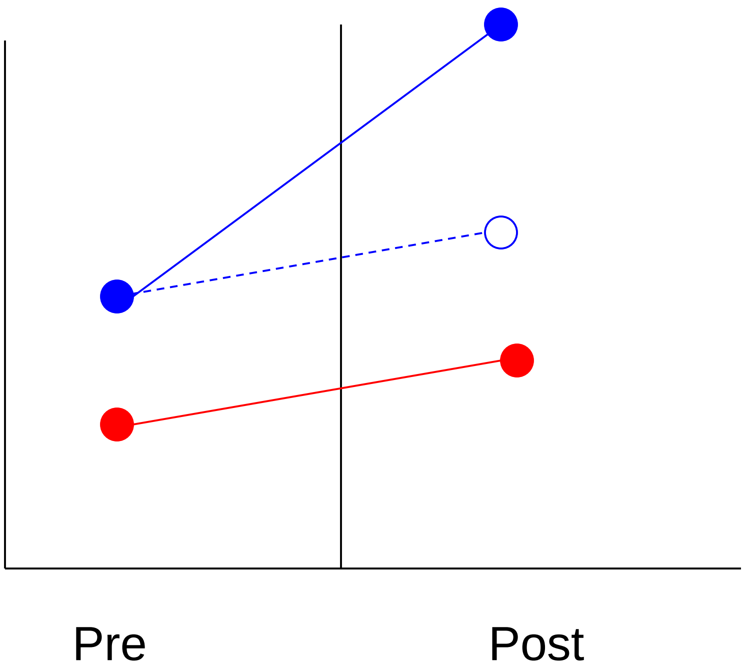


Differences-in-Differences, Graphically

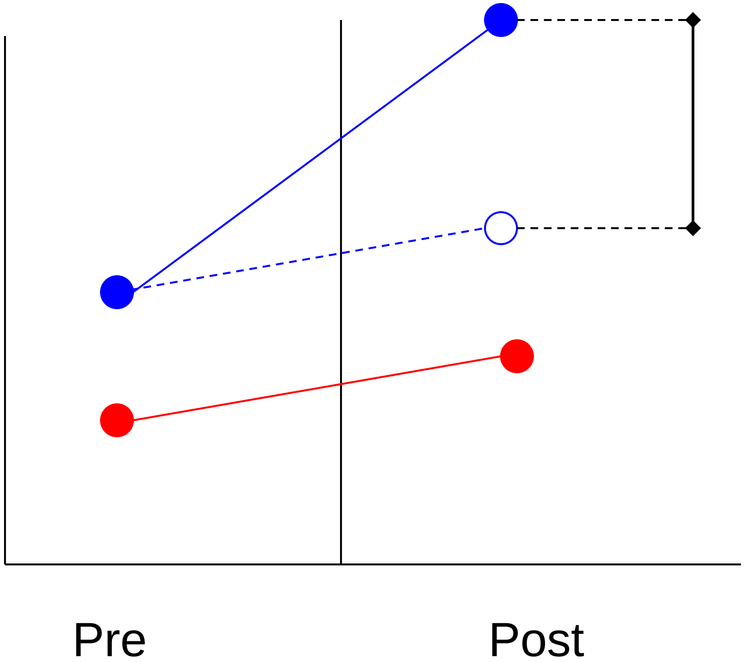


Effect of program using only T & C comparison from post-intervention (ignoring pre-existing differences between T & C groups).

Differences-in-Differences, Graphically



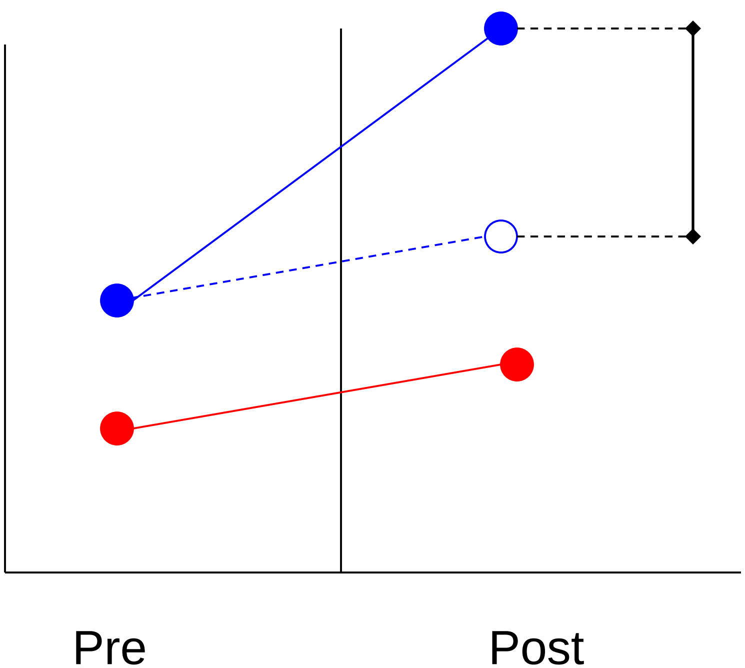
Differences-in-Differences, Graphically



Effect of program difference-in-difference (taking into account pre-existing differences between T & C and general time trend).

Identifying Assumption

- Whatever happened to the control group over time is what would have happened to the treatment group in the absence of the program.



Effect of program difference-in-difference (taking into account pre-existing differences between T & C and general time trend).

Uses of Diff-in-Diff

- Simple two-period, two-group comparison
  - very useful in combination with other methods
    - Randomization
    - Regression Discontinuity
    - Matching (propensity score)

- Can also do much more complicated "cohort" analysis, comparing many groups over many time periods

The (Simple) Regression

$$Y_{i,t} = a + bTreat_{i,t} + cPost_{i,t} + d(Treat_{i,t}Post_{i,t}) + e_{i,t}$$

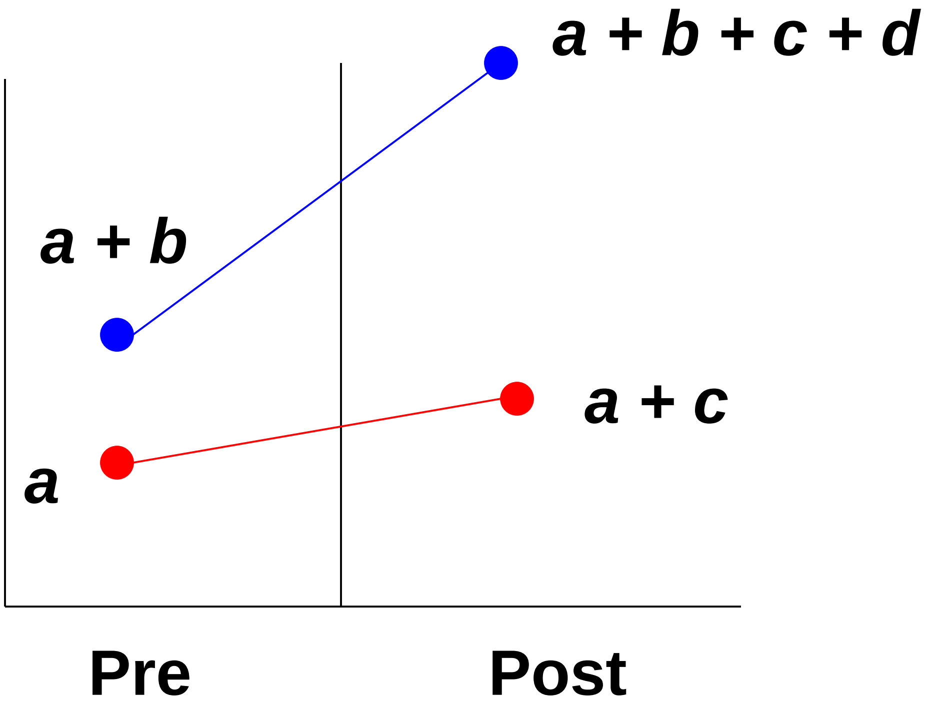
- $Treat_{i,t}$ is a binary indicator ("turns on" from 0 to 1) for being in the treatment group
  - Time invariant
- $Post_{i,t}$ is a binary indicator for the period after treatment begins
  - Regardless of if you get it
- and $Treat_{i,t}Post_{i,t}$ is the interaction (product)

Interpretation of *a, b, c, d* is "holding all else constant"

Putting Graph & Regression Together

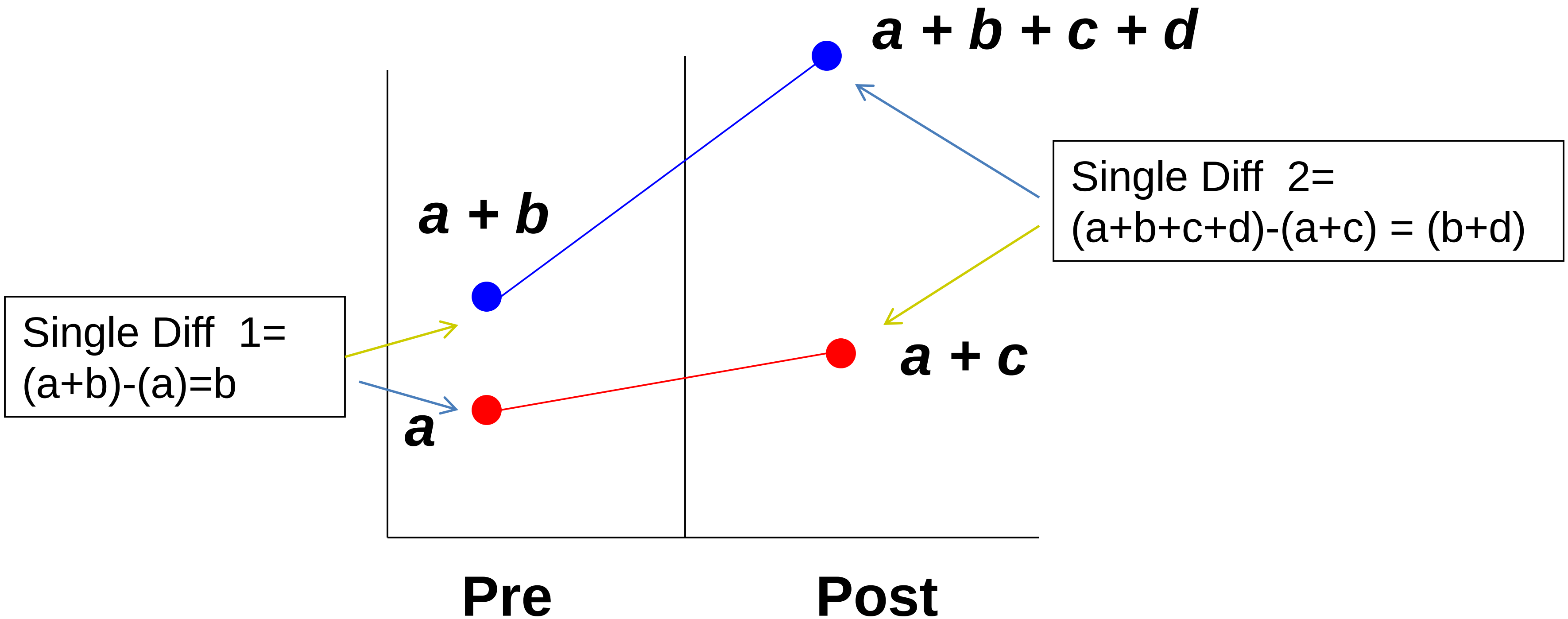
$$Y_{i,t} = a + bTreat_{i,t} + cPost_{i,t} + d(Treat_{i,t}Post_{i,t}) + e_{i,t}$$

d **is the causal effect of treatment**



Putting Graph & Regression Together

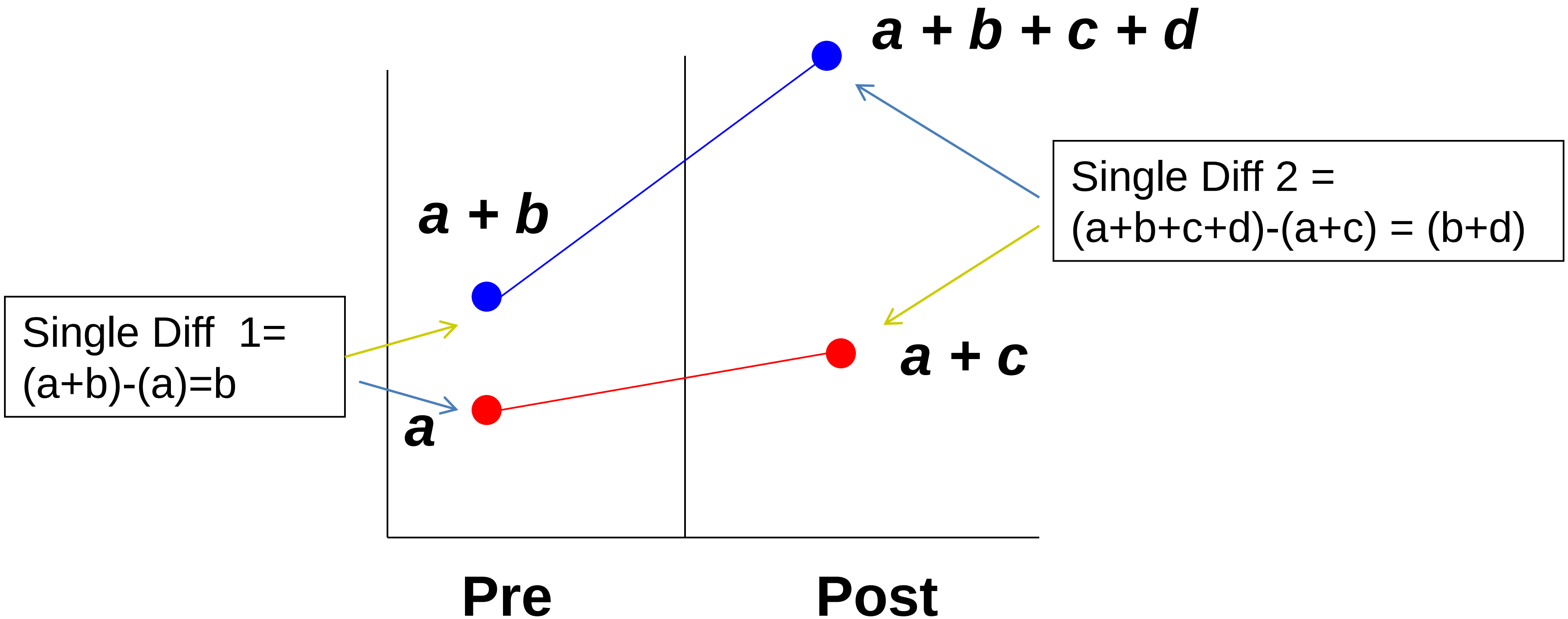
$$Y_{i,t} = a + bTreat_{i,t} + cPost_{i,t} + d(Treat_{i,t}Post_{i,t}) + e_{i,t}$$



Putting Graph & Regression Together

$$Y_{i,t} = a + bTreat_{i,t} + cPost_{i,t} + d(Treat_{i,t}Post_{i,t}) + e_{i,t}$$

Diff-in-Diff=(Single Diff 2-Single Diff 1)=(b+d)-b=d



Cohort Analysis

- When you've got richer data, it's not as easy to draw the picture or write the equations
  - cross-section (lots of individuals at one point in time)
  - time-series (one individual over lots of time)
  - repeated cross-section (lots of individuals over several times)
  - ★ panel (lots of individuals, multiple times for each) ★

- Basically, control for each time period and each "group" (fixed effects) – the coefficient on the treatment dummy is the effect you're trying to estimate

DiD Data Requirements

- Either repeated cross-section or panel
- Treatment can't happen for everyone at the same time
- If you believe the identifying assumption, then you can analyze policies ex post
  - Let's us tackle really big questions that we're unlikely to be able to randomize

Robustness Checks

- If possible, use data on multiple pre-program periods to show that difference between treated & control is stable
  - Not necessary for trends to be parallel, just to know function for each
- If possible, use data on multiple post-program periods to show that unusual difference between treated & control occurs only concurrent with program
- Alternatively, use data on multiple indicators to show that response to program is only manifest for those we expect it to be (e.g. the diff-in-diff estimate of the impact of ITN distribution on diarrhea should be zero)