DiD

DiD strategies take into account the timing of the treatment (e.g., a policy shift) to obtain an estimate for the ATET.

Remember that we can write

$$\begin{aligned} Y_i &= D_i Y_{i1} + (1 - D_i) Y_{i0} \\ &= Y_{i0} + D_i (Y_{i1} - Y_{i0}) \\ &= \alpha + \Delta_i D_i + \epsilon_i \end{aligned}$$

where $\alpha = \mu_0$ and $\epsilon_i = Y_{i0} - \mu_0$.

DiD

To incorporate timing, let t_0 denote the time period before the assignement and t_1, the time period after. We then append time subscripts to the above relation:

$$Y_{it} = \alpha + \Delta_i D_{it} + \epsilon_{it}, t = t_1, t_0$$

where $\Delta_i = Y_{i1t_1} - Y_{i0t_1}$ and $D_{it} = 0$ if $t = t_0$.

We can also decompose the shock ϵ_{it} as

$$m_t + \eta_i + v_{it}$$

where m_t is a macro shock affecting all equally at time t and η_i is a time-invariant individual effect.

DiD

The identifying assumption we will maintain here is that

$$E[\epsilon_{it}|D_{it}] = m_t + E[\eta_i|D_{it}]$$

which is equivalent to no selection on untreated outcomes (in first differences):

$$E[Y_{i0t_1} - Y_{i0t_0}|D_{it_1} = 1] = E[Y_{i0t_1} - Y_{i0t_0}].$$

This does not rule out selection on unobservables (since this may still depend on η_i).

It does exclude the possibility that selection is based on transitory individual-specific unobservables (v_{it}).

DiD

- Suppose for example that we want to evaluate a training programme for those with low earnings. Let the threshold for eligibility be B.
- We have a panel of individuals and those with low earnings qualify for training, forming the treatment group.
- Those with higher earnings form the comparison group.
- The treatment group may have low earnings for two reasons:
  1. They have low permanent earnings: η_i is low. This is handled by DiD.
  2. They have a negative transitory shock: v_{it} is low. This is *not* handled by DiD.

DiD

- To see this note that participants are such that $Y_{i0t_0} < B$. Assume that the shocks v_{it} are *iid*. Then, $v_{it_0} < B - \alpha - m_{t_0} - \eta_i$. this implies that

  $$E[Y_{i0t_1} - Y_{i0t_0} | D_{it} = 1] = m_{t_1} - m_{t_0} - E[v_{it_0} | v_{it_0} < B - \alpha - m_{t_0} - \eta_i]$$

  and for the comparison group

  $$E[Y_{i0t_1} - Y_{i0t_0} | D_{it} = 0] = m_{t_1} - m_{t_0} - E[v_{it_0} | v_{it_0} > B - \alpha - m_{t_0} - \eta_i].$$

- This implies that

  $$E[Y_{i0t_1} - Y_{i0t_0} | D_{it} = 1] - E[Y_{i0t_1} - Y_{i0t_0} | D_{it} = 0] =$$
  $$E[v_{it_0} | v_{it_0} > B - \alpha - m_{t_0} - \eta_i] - E[v_{it_0} | v_{it_0} < B - \alpha - m_{t_0} - \eta_i]$$
  $$> 0$$

DiD

- Ashenfelter (1978, REStat) was one of the first to consider DiD to evaluate training programs.

TABLE 1.—MEAN EARNINGS PRIOR, DURING, AND SUBSEQUENT TO TRAINING FOR 1964 MDTA CLASSROOM TRAINEES AND A COMPARISON GROUP

	White Males		Black Males		White Females		Black Females	
	Trainees	Comparison Group	Trainees	Comparison Group	Trainees	Comparison Group	Trainees	Comparison Group
1959	$1,443	$2,588	$ 904	$1,438	$ 635	$ 987	$ 384	$ 616
1960	1,533	2,699	976	1,521	687	1,076	440	693
1961	1,572	2,782	1,017	1,573	719	1,163	471	737
1962	1,843	2,963	1,211	1,742	813	1,308	566	843
1963	1,810	3,108	1,182	1,896	748	1,433	531	937
1964	1,551	3,275	1,273	2,121	838	1,580	688	1,060
1965	2,923	3,458	2,327	2,338	1,747	1,698	1,441	1,198
1966	3,750	4,351	2,983	2,919	2,024	1,990	1,794	1,461
1967	3,964	4,430	3,048	3,097	2,244	2,144	1,977	1,678
1968	4,401	4,955	3,409	3,487	2,398	2,339	2,160	1,920
1969	$4,717	$5,033	$3,714	$3,681	$2,646	$2,444	$2,457	$2,133
Number of Observations	7,326	40,921	2,133	6,472	2,730	28,142	1,356	5,192

DiD

- With covariates, it is possible that

$$E[Y_{i0t_1} - Y_{i0t_0}|D_{it_1} = 1, X] = E[Y_{i0t_1} - Y_{i0t_0}|X]$$

even if

$$E[Y_{i0t_1} - Y_{i0t_0}|D_{it_1} = 1] \neq E[Y_{i0t_1} - Y_{i0t_0}].$$

- This is just as in matching where the outcome now is the change in Ys!

DiD

Under the assumption that $E[Y_{i0t_1} - Y_{i0t_0}|D_{it_1} = 1] = E[Y_{i0t_1} - Y_{i0t_0}]$,

$$\begin{aligned} ATET &= E[\Delta_i|D_{it_1} = 1] \\ &= E[Y_{it_1}|D_{it_1} = 1] - E[Y_{it_0}|D_{it_1} = 1] \\ &\quad -\{E[Y_{it_1}|D_{it_1} = 0] - E[Y_{it_0}|D_{it_1} = 0]\}. \end{aligned}$$

Which suggests the empirical analog estimator

$$\widehat{ATET} = \frac{\sum_i d_{it_1}(y_{it_1} - y_{it_0})}{\sum_i d_{it_1}} - \frac{\sum_i (1 - d_{it_1})(y_{it_1} - y_{it_0})}{\sum_i (1 - d_{it_1})}$$

Notice this is just the within groups panel data estimator.

DiD

- If treatment and control groups can be separated before the policy change **and** composition of the groups with respect to the individual fixed effects does not change, repeated cross-section can be used to estimate the $ATET$.
- Differential macro trends would also invalidate the estimator.
- Selection on idiosyncratic shocks will also invalidate the estimator. One example is "Ashenfelter's dip". If enrolment to a training program is more likely when earnings are temporarily decreased, a faster earnings growth should be expected among the treated *even without program participation*.

Nonlinear DiD

- To extend DiD to a nonlinear setting, suppose that

$$Y_{it} = 1[\alpha + \Delta_i D_{it} + \epsilon_{it} \geq 0].$$

- As before, assume that

$$\epsilon_{it} = m_t + \eta_i - v_{it}$$

  and

$$E[\epsilon_{it} | D_{it}] = m_t + E[\eta_i | D_{it}].$$

- Now, assume also that $v_{it} \sim F$ where F is known and $\eta_i = \eta_d$ if $D_{it} = d$.

Nonlinear DiD

- The above model corresponds to a binary response model with non-independent shocks.
- Under those parametric assumptions, letting $Y_{it}^{d}, d = 0, 1$ denote treated and untreated outcomes for individual i as before,

$$E[Y_{it}^{0}|D_{it_1} = d] = F(\alpha + \eta_d + m_t)$$
$$\Rightarrow F^{-1}(E[Y_{it}^{0}|D_{it_1} = d]) = \alpha + \eta_d + m_t$$

for $d = 0, 1$ and $t = t_0, t_1$ since $E[\epsilon_{it}|D_{it}] = m_t + E[\eta_i|D_{it}]$.

Nonlinear DiD

- Then,

$$F^{-1}(E[Y^0_{it_1}|D_{it_1} = d]) - F^{-1}(E[Y^0_{it_0}|D_{it_1} = d]) = m_{t_1} - m_{t_0}$$

  for $d = 0, 1$.
- This leads to

$$E[Y^0_{it_1}|D_{it_1} = 1] = F\Big(F^{-1}(E[Y^0_{it_0}|D_{it_1} = 1])$$
$$+\{F^{-1}(E[Y^0_{it_1}|D_{it_1} = 0]) - F^{-1}(E[Y^0_{it_0}|D_{it_1} = 0])\}\Big)$$

Nonlinear DiD

- The *ATET* $(= E[Y^1_{it_1} - Y^0_{it_1} | D_i = 1])$ can then be expressed as

$$\begin{aligned} ATET &= E[Y^1_{it_1}|D_{it_1} = 1] - F\Big(F^{-1}(E[Y^0_{it_0}|D_{it_1} = 1]) \\ &\quad + \{F^{-1}(E[Y^0_{it_1}|D_{it_1} = 0]) - F^{-1}(E[Y^0_{it_0}|D_{it_1} = 0])\}\Big) \end{aligned}$$

- An estimator for this quantity is

$$\begin{aligned} \widehat{ATET} &= \frac{\sum d_i y_{it_1}}{\sum d_i} - F\left[F^{-1}\left(\frac{\sum d_i y_{it_0}}{\sum d_i}\right)\right. \\ &\quad \left. + \left\{F^{-1}\left(\frac{\sum(1-d_i)y_{it_1}}{\sum(1-d_i)}\right) - F^{-1}\left(\frac{\sum(1-d_i)y_{it_0}}{\sum(1-d_i)}\right)\right\}\right] \end{aligned}$$