Difference-in-difference causal effects designs

Difference-in-difference (*d-i-d*) designs receive growing popularity in accounting causal effect analyses. Panel data (time series and cross-sectional data) in which there is an intervention (possibly, regulatory) leading some firms to adopt treatment (perhaps, a different disclosure strategy) while other firms do not often prompt *d-i-d* experiments. Such *d-i-d* experiments are typically interpreted as providing evidence on the average treatment effect on the treated (ATT) during the post-intervention period ($t = 1$).

$$E\left[Y_1 - Y_0 \mid D = 1\right]$$

where Y_1 represents potential outcome with treatment, Y_0 represents potential outcome without treatment, and $D = 1$ identifies the subpopulation who adopt treatment while $D = 0$ identifies the subpopulation who don't adopt treatment. Hence, prior to the intervention date ($t = 0$) typically all firms are without treatment ($D = 0$). For completeness, other average treatment effects are average treatment effect on the untreated ($ATUT$)

$$E\left[Y_1 - Y_0 \mid D = 0\right]$$

and the unconditional average treatment effect (ATE) or via iterated expectations

$$E\left[Y_1 - Y_0\right] = \Pr\left(D = 1\right) E\left[Y_1 - Y_0 \mid D = 1\right] + \Pr\left(D = 0\right) E\left[Y_1 - Y_0 \mid D = 0\right]$$

As *d-i-d* is an ignorable treatment strategy, average treatment effect identification draws from conditional mean independence and common support. Conditional mean independence indicates the data generating process (DGP) is characterized by

$$E\left[Y_1 \mid X, D\right] = E\left[Y_1 \mid X\right]$$

and especially, for ATT

$$E\left[Y_0 \mid X, D\right] = E\left[Y_0 \mid X\right]$$

This condition is not verifiable via data tests as it is counterfactual in nature, $(Y_1 \mid D = 0)$ and $(Y_0 \mid D = 1)$ are unobservable. Rather, the analyst engages in thought experiments to query the logical consistency of the condition applied to the setting at-hand. On the other hand, common support applies to both the unknown DGP and the observed sample. Common support refers to the overlap of covariate distributions for the treated ($D = 1$) and untreated ($D = 0$) subpopulations/subsamples. Reliance on common support in the samples allows inference from the data rather than extrapolations outside the data. DGP common support is more subtle and will be discussed and illustrated via some stylized examples below.

For the examples, all DGPs have the following partially observable covariates (X_j, $j = 0, 1$) and unobservable (U_j, $j = 0, 1$) common structure.

$$\begin{aligned} Y_1 &= \alpha_1 X_1 + U_1 \\ Y_0 &= \alpha_0 X_0 + U_0 \end{aligned}$$

and observed outcomes and covariates are

$$\begin{aligned} Y &= DY_1 + (1 - D) Y_0 \\ X &= DX_1 + (1 - D) X_0 \end{aligned}$$

Both outcomes and covariates can be missing.

We explore three (linear) *d-i-d* designs from sparsest to most elaborate. Design one ignores covariates X where β_2, the coefficient on D, is interpreted as ATT.

$$Y = \beta_0 + \beta_1 t + \beta_2 D + \varepsilon_1 \tag{1}$$

Design two includes covariates, X, where γ_2 is interpreted as ATT.

$$Y = \gamma_0 + \gamma_1 t + \gamma_2 D + \gamma_3 X + \varepsilon_2 \tag{2}$$

Design three includes covariates and interactions, $X \times D$, where $\delta_2 + \delta_4 E[X]$ is interpreted as ATT.

$$Y = \delta_0 + \delta_1 t + \delta_2 D + \delta_3 X + \delta_4 X \times D + \varepsilon_3 \tag{3}$$

Design three is the only one of these designs capable of identifying conditional average treatment effects where $\delta_2 + \delta_4 (X = x)$ is interpreted as $ATT(X)$.

Example 1 (homogeneous outcome) *Suppose the DGP is*

	t	D	X_1	U_1	Y_1	X_0	U_0	Y_0	X	Y
	0	0	1	−1	0	1	1	0	1	0
	0	0	2	−1	1	2	1	−1	2	−1
	0	0	3	−1	2	3	1	−2	3	−2
	0	0	1	1	2	1	−1	−2	1	−2
	0	0	2	1	3	2	−1	−3	2	−3
	0	0	3	1	4	3	−1	−4	3	−4
	1	1	1	0	1	1	0	−1	1	1
	1	1	2	0	2	2	0	−2	2	2
	1	1	3	0	3	3	0	−3	3	3
	1	0	1	0	1	1	0	−1	1	−1
	1	0	2	0	2	2	0	−2	2	−2
	1	0	3	0	3	3	0	−3	3	−3
means	0.5	0.25	2	0	2	2	0	−2	2	−1

Conditional average treatment effects vary with the covariate.

	$X = 1$	$X = 2$	$X = 3$
$E[Y_1 \mid X_1, D = 1]$ $-E[Y_0 \mid X_0, D = 1]$	$1 - (-1) = 2$	$2 - (-2) = 4$	$3 - (-3) = 6$
$E[Y_1 \mid X_1, D = 0]$ $-E[Y_0 \mid X_0, D = 0]$	$1 - (-1) = 2$	$2 - (-2) = 4$	$3 - (-3) = 6$
$E[Y_1 \mid X_1] - E[Y_0 \mid X_0]$	$1 - (-1) = 2$	$2 - (-2) = 4$	$3 - (-3) = 6$

Unconditional average treatment effects indicate homogeneity.

$$\begin{aligned}
ATT &= E_{X_1}[E[Y_1 \mid X_1, D = 1]] - E_{X_0}[E[Y_0 \mid X_0, D = 1]] \\
&= \frac{1}{3}(2) + \frac{1}{3}(4) + \frac{1}{3}(6) = 4 \\
ATUT &= E_{X_1}[E[Y_1 \mid X_1, D = 0]] - E_{X_0}[E[Y_0 \mid X_0, D = 0]] \\
&= \frac{1}{3}(2) + \frac{1}{3}(4) + \frac{1}{3}(6) = 4 \\
ATE &= E_{X_1}[E[Y_1 \mid X_1]] - E_{X_0}[E[Y_0 \mid X_0]] \\
&= \frac{1}{3}(2) + \frac{1}{3}(4) + \frac{1}{3}(6) = 4
\end{aligned}$$

Regression parameters are expected to be unbiased (and asymptotically consistent) as unobservables in each subpopulation regression are conditionally unbiased.

	$X_j = 1$	$X_j = 2$	$X_j = 3$	$t = 0$	$t = 1$
$E[U_1 \mid X_1, t]$	0	0	0	0	0
$E[U_0 \mid X_0, t]$	0	0	0	0	0

Importantly, conditional mean independence is satisfied.

	$D = 1$	$D = 0$
$E[Y_1 \mid X_1 = 1, D]$	1	1
$E[Y_1 \mid X_1 = 2, D]$	2	2
$E[Y_1 \mid X_1 = 3, D]$	3	3
$E[Y_0 \mid X_0 = 1, D]$	-1	-1
$E[Y_0 \mid X_0 = 2, D]$	-2	-2
$E[Y_0 \mid X_0 = 3, D]$	-3	-3

Design one yields

$$\begin{gathered} Y = -2 + 0t + 4D + \varepsilon_1, \\ \textit{suggested } ATT = 4 \end{gathered} \tag{1}$$

Design two yields

$$\begin{gathered} Y = -1 + 0t + 4D - \tfrac{1}{2}X + \varepsilon_2, \\ \textit{suggested } ATT = 4 \end{gathered} \tag{2}$$

Design three yields

$$\begin{array}{c} Y = 0 + 0t + 0D - X + 2X \times D + \varepsilon_3, \\ \textit{suggested } ATT\,(X = 1) = 0 + 2 \times 1 = 2 \\ \textit{suggested } ATT\,(X = 2) = 0 + 2 \times 2 = 4 \\ \textit{suggested } ATT\,(X = 3) = 0 + 2 \times 3 = 6 \\ \textit{suggested } ATT = 0 + 2 \times 2 = 4 \end{array} \tag{3}$$

As suggested the parameters are consistent with the DGP and,[1] *most significantly, all three designs effectively identify* ATT*. However, only design three identifies the conditional average treatment effects. This is a prototype setting for d-i-d designs. The next example is a slight variation in which the regression condition* $E[U_j \mid X_j, t] = 0$ *is not satisfied but conditional mean independence is satisfied by the DGP.*

Example 2 (homogeneous outcome but $E[U_j \mid X_j, t] \neq 0$**)** *Suppose the DGP is a slight variation of example 1*

	t	D	X_1	U_1	Y_1	X_0	U_0	Y_0	X	Y
	0	0	1	−1	0	1	1	0	1	0
	0	0	2	0	2	2	0	−2	2	−2
	0	0	3	1	4	3	−1	−4	3	−4
	0	0	1	−1	0	1	1	0	1	0
	0	0	2	0	2	2	0	−2	2	−2
	0	0	3	1	4	3	−1	−4	3	−4
	1	1	1	−1	0	1	1	0	1	0
	1	1	2	0	2	2	0	−2	2	2
	1	1	3	1	4	3	−1	−4	3	4
	1	0	1	−1	0	1	1	0	1	0
	1	0	2	0	2	2	0	−2	2	−2
	1	0	3	1	4	3	−1	−4	3	−4
means	0.5	0.25	2	0	2	2	0	−2	2	−1

Conditional average treatment effects vary with the covariate.

	$X = 1$	$X = 2$	$X = 3$
$E[Y_1 \mid X_1, D = 1]$ $-E[Y_0 \mid X_0, D = 1]$	$0 - (0) = 0$	$2 - (-2) = 4$	$4 - (-4) = 8$
$E[Y_1 \mid X_1, D = 0]$ $-E[Y_0 \mid X_0, D = 0]$	$0 - (0) = 0$	$2 - (-2) = 4$	$4 - (-4) = 8$
$E[Y_1 \mid X_1] - E[Y_0 \mid X_0]$	$0 - (0) = 0$	$2 - (-2) = 4$	$4 - (-4) = 8$

[1] This is most transparent for design three where the coefficient on X for the $D = 0$ subpopulation is -1 and for the $D = 1$ subpopulation is $-1 + 2 = 1$, and other parameters are zero.

Unconditional average treatment effects indicate homogeneity.

$$\begin{aligned}
ATT &= E_{X_1}\left[E\left[Y_1 \mid X_1, D=1\right]\right] - E_{X_0}\left[E\left[Y_0 \mid X_0, D=1\right]\right] \\
&= \frac{1}{3}(0) + \frac{1}{3}(4) + \frac{1}{3}(8) = 4 \\
ATUT &= E_{X_1}\left[E\left[Y_1 \mid X_1, D=0\right]\right] - E_{X_0}\left[E\left[Y_0 \mid X_0, D=0\right]\right] \\
&= \frac{1}{3}(0) + \frac{1}{3}(4) + \frac{1}{3}(8) = 4 \\
ATE &= E_{X_1}\left[E\left[Y_1 \mid X_1\right]\right] - E_{X_0}\left[E\left[Y_0 \mid X_0\right]\right] \\
&= \frac{1}{3}(0) + \frac{1}{3}(4) + \frac{1}{3}(8) = 4
\end{aligned}$$

Regression parameters may not be unbiased (or asymptotically consistent) as unobservables in each subpopulation regression are not conditionally unbiased (in particular, $X_j = 1, 3$).

	$X_j = 1$	$X_j = 2$	$X_j = 3$	$t = 0$	$t = 1$
$E[U_1 \mid X_1, t]$	-1	0	1	0	0
$E[U_0 \mid X_0, t]$	1	0	-1	0	0

However, conditional mean independence is satisfied.

	$D = 1$	$D = 0$
$E[Y_1 \mid X_1 = 1, D]$	0	0
$E[Y_1 \mid X_1 = 2, D]$	2	2
$E[Y_1 \mid X_1 = 3, D]$	4	4
$E[Y_0 \mid X_0 = 1, D]$	0	0
$E[Y_0 \mid X_0 = 2, D]$	-2	-2
$E[Y_0 \mid X_0 = 3, D]$	-4	-4

Design one yields

$$\begin{gathered} Y = -2 + 0t + 4D + \varepsilon_1, \\ \textit{suggested } ATT = 4 \end{gathered} \tag{1}$$

Design two yields

$$\begin{gathered} Y = 0 + 0t + 4D - X + \varepsilon_2, \\ \textit{suggested } ATT = 4 \end{gathered} \tag{2}$$

Design three yields

$$\begin{gathered} Y = 2 + 0t - 4D - 2X + 4X \times D + \varepsilon_3, \\ \textit{suggested } ATT\,(X = 1) = -4 + 4 \times 1 = 0 \\ \textit{suggested } ATT\,(X = 2) = -4 + 4 \times 2 = 4 \\ \textit{suggested } ATT\,(X = 3) = -4 + 4 \times 3 = 8 \\ \textit{suggested } ATT = -4 + 4 \times 2 = 4 \end{gathered} \tag{3}$$

Even though the regression parameters might be biased the symmetry in (X_j, U_j) is such that bias does not emerge (in particular, positive corr (X_1, U_1) is offset by negative corr (X_0, U_0)). All three designs effectively identify unconditional ATT *and conditional $ATT(X)$ as the DGP exhibits full common or balanced support. The next example maintains homogeneous outcome but begins to explore the impact of unbalanced covariate support in a d-i-d design.*

Example 3 (homogeneous outcome with unbalanced covariates) *Suppose the DGP is*

	t	D	X_1	U_1	Y_1	X_0	U_0	Y_0	X	Y
	0	0	1	-1	0	1	1	0	1	0
	0	0	2	0	2	3	0	-3	3	-3
	0	0	3	1	4	1	-1	-2	1	-2
	0	0	1	-1	0	1	1	0	1	0
	0	0	2	0	2	3	0	-3	3	-3
	0	0	3	1	4	1	-1	-2	1	-2
	1	1	1	-1	0	1	1	0	1	0
	1	1	2	0	2	3	0	-3	2	2
	1	1	3	1	4	1	-1	-2	3	4
	1	0	1	-1	0	1	1	0	1	0
	1	0	2	0	2	3	0	-3	3	-3
	1	0	3	1	4	1	-1	-2	1	-2
means	0.5	0.25	2	0	2	$1\frac{2}{3}$	0	$-1\frac{2}{3}$	$1\frac{3}{4}$	$-\frac{3}{4}$

Conditional average treatment effects vary with the covariate.

	$X=1$	$X=2$	$X=3$
$E[Y_1 \mid X_1, D=1]$ $-E[Y_0 \mid X_0, D=1]$	$0-(-1)=1$	$2-(NA)=NA$	$4-(-3)=7$
$E[Y_1 \mid X_1, D=0]$ $-E[Y_0 \mid X_0, D=0]$	$0-(-1)=1$	$2-(NA)=NA$	$4-(-3)=7$
$E[Y_1 \mid X_1] - E[Y_0 \mid X_0]$	$0-(-1)=1$	$2-(NA)=NA$	$4-(-3)=7$

Unconditional average treatment effects indicate homogeneity.

$$\begin{aligned}
ATT &= E_{X_1}[E[Y_1 \mid X_1, D=1]] - E_{X_0}[E[Y_0 \mid X_0, D=1]] \\
&= \frac{1}{3}(0) + \frac{1}{3}(2) + \frac{1}{3}(4) - \left[\frac{2}{3}(-1) + \frac{1}{3}(-3)\right] = 3\frac{2}{3} \\
ATUT &= E_{X_1}[E[Y_1 \mid X_1, D=0]] - E_{X_0}[E[Y_0 \mid X_0, D=0]] \\
&= \frac{1}{3}(0) + \frac{1}{3}(2) + \frac{1}{3}(4) - \left[\frac{2}{3}(-1) + \frac{1}{3}(-3)\right] = 3\frac{2}{3} \\
ATE &= E_{X_1}[E[Y_1 \mid X_1]] - E_{X_0}[E[Y_0 \mid X_0]] \\
&= \frac{1}{3}(0) + \frac{1}{3}(2) + \frac{1}{3}(4) - \left[\frac{2}{3}(-1) + \frac{1}{3}(-3)\right] = 3\frac{2}{3}
\end{aligned}$$

Regression parameters are not expected to be unbiased (or asymptotically consistent) as unobservables in each subpopulation regression are not conditionally unbiased.

	$X_j=1$	$X_j=2$	$X_j=3$	$t=0$	$t=1$
$E[U_1 \mid X_1, t]$	-1	0	1	0	0
$E[U_0 \mid X_0, t]$	0	NA	0	0	0

However, conditional mean independence is satisfied.

	$D=1$	$D=0$
$E[Y_1 \mid X_1=1, D]$	0	0
$E[Y_1 \mid X_1=2, D]$	2	2
$E[Y_1 \mid X_1=3, D]$	4	4
$E[Y_0 \mid X_0=1, D]$	-1	-1
$E[Y_0 \mid X_0=2, D]$	NA	NA
$E[Y_0 \mid X_0=3, D]$	-3	-3

Design one yields

$$\begin{array}{c} Y=-1\tfrac{2}{3}+0t+3\tfrac{2}{3}D+\varepsilon_1, \\ \textit{suggested } ATT=3\tfrac{2}{3} \end{array} \tag{1}$$

Design two yields

$$\begin{array}{c} Y=-1+0t+3\tfrac{4}{5}D-\tfrac{2}{5}X+\varepsilon_2, \\ \textit{suggested } ATT=3\tfrac{4}{5} \end{array} \tag{2}$$

Design three yields[2]

$$\begin{array}{c} Y=0+0t-2D-X+3X\times D+\varepsilon_3, \\ \textit{suggested } ATT(X=1)=-2+3\times 1=1 \\ \textit{suggested } ATT(X=2)=-2+3\times 2=4 \\ \textit{suggested } ATT(X=3)=-2+3\times 3=7 \\ \textit{suggested } ATT=-2+3\times 1\tfrac{3}{4}=3\tfrac{1}{4} \end{array} \tag{3}$$

As suggested the parameters are inconsistent with the DGP (in particular, positive corr (X_1, U_1) is not offset by zero corr (X_0, U_0)). Only design one effectively identifies unconditional ATT presumably because of the lack of common support (no overlap at $X=2$). Surprisingly, design three effectively identifies conditional average treatment effects for $X=1,3$. To explore common support, we focus on the data excluding $(X=2)$. The average treatment effects conditional on $(X=1,3)$ limited support exhibit the same outcome homogeneity.

$$\begin{aligned} ATT(X=1,3) &= E[Y_1 \mid X=1,3, D=1]-E[Y_0 \mid X=1,3, D=1] \\ &= \frac{1}{2}(0)+\frac{1}{2}(4)-\left[\frac{2}{3}(-1)+\frac{1}{3}(-3)\right]=3\frac{2}{3} \\ ATUT(X=1,3) &= E[Y_1 \mid X=1,3, D=0]-E[Y_0 \mid X=1,3, D=0] \\ &= \frac{1}{2}(0)+\frac{1}{2}(4)-\left[\frac{2}{3}(-1)+\frac{1}{3}(-3)\right]=3\frac{2}{3} \\ ATE(X=1,3) &= E_{X_1}[E[Y_1 \mid X_1=1,3]]-E_{X_0}[E[Y_0 \mid X_0=1,3]] \\ &= 3\frac{2}{3} \end{aligned}$$

[2] Here suggested ATT involves iteration of X over the entire observed sample. If we employ only the $D=1$ subsample suggested ATT is $-2+3\times 2=4\neq 3\frac{2}{3}$.

Although conditional mean independence is satisfied, the covariate distribution for Y_1 differs between the two subpopulations, $D = 0, 1$. Design one yields

$$\begin{array}{c} Y = -1\tfrac{2}{3} + 0t + 3\tfrac{2}{3}D + \varepsilon_1, \\ \textit{suggested } ATT\,(X = 1, 3) = 3\tfrac{2}{3} \end{array} \tag{1}$$

Design two yields

$$\begin{array}{c} Y = -1 + 0t + 3\tfrac{4}{5}D - \tfrac{2}{5}X + \varepsilon_2, \\ \textit{suggested } ATT\,(X = 1, 3) = 3\tfrac{4}{5} \end{array} \tag{2}$$

Design three yields[3]

$$\begin{array}{c} Y = 0 + 0t - 2D - X + 3X \times D + \varepsilon_3, \\ \textit{suggested } ATT\,(X = 1) = -2 + 3 \times 1 = 1 \\ \textit{suggested } ATT\,(X = 3) = -2 + 3 \times 3 = 7 \\ \textit{suggested } ATT\,(X = 1, 3) = -2 + 3 \times 1\tfrac{8}{11} = 3\tfrac{2}{11} \end{array} \tag{3}$$

Only design one identifies $ATT\,(X = 1, 3)$ because of covariate imbalance in the two subpopulations ($D = 0$ and $D = 1$). Again, design three effectively identifies conditional average treatment effects for $X = 1, 3$ and we're spared the embarrassment of suggesting $ATT\,(X = 2)$. Next, we consider a slight variation on this example where the DGP reflects $E[U_j \mid X_j, t] = 0$.

Example 4 (homogeneous outcome with unbalanced covariates and $E[U_j \mid X_j] = 0$)

Suppose the DGP is

	t	D	X_1	U_1	Y_1	X_0	U_0	Y_0	X	Y
	0	0	1	-1	0	1	1	0	1	0
	0	0	2	-1	1	1	1	0	1	0
	0	0	3	-1	2	3	1	-2	3	-2
	0	0	1	1	2	1	-1	-2	1	-2
	0	0	2	1	3	1	-1	-2	1	-2
	0	0	3	1	4	3	-1	-4	3	-4
	1	1	1	0	1	1	0	-1	1	1
	1	1	2	0	2	1	0	-1	2	2
	1	1	3	0	3	3	0	-3	3	3
	1	0	1	0	1	1	0	-1	1	-1
	1	0	2	0	2	1	0	-1	1	-1
	1	0	3	0	3	3	0	-3	3	-3
means	0.5	0.25	2	0	2	$1\frac{2}{3}$	0	$-1\frac{2}{3}$	$1\frac{3}{4}$	$-\frac{3}{4}$

[3] Here suggested ATT involves iteration of X over the entire observed sample. If we employ only the $D = 1$ subsample suggested ATT is $-2 + 3 \times 2 = 4 \neq 3\frac{2}{3}$.

Conditional average treatment effects vary with the covariate.

	$X=1$	$X=2$	$X=3$
$E[Y_1 \mid X_1, D=1]$ $-E[Y_0 \mid X_0, D=1]$	$1-(-1)=2$	$2-(NA)=NA$	$3-(-3)=6$
$E[Y_1 \mid X_1, D=0]$ $-E[Y_0 \mid X_0, D=0]$	$1-(-1)=2$	$2-(NA)=NA$	$3-(-3)=6$
$E[Y_1 \mid X_1]-E[Y_0 \mid X_0]$	$1-(-1)=2$	$2-(NA)=NA$	$3-(-3)=6$

Unconditional average treatment effects indicate homogeneity.

$$\begin{aligned}
ATT &= E_{X_1}[E[Y_1 \mid X_1, D=1]] - E_{X_0}[E[Y_0 \mid X_0, D=1]] \\
&= \frac{1}{3}(1)+\frac{1}{3}(2)+\frac{1}{3}(3)-\left[\frac{2}{3}(-1)+\frac{1}{3}(-3)\right]=3\frac{2}{3} \\
ATUT &= E_{X_1}[E[Y_1 \mid X_1, D=0]] - E_{X_0}[E[Y_0 \mid X_0, D=0]] \\
&= \frac{1}{3}(1)+\frac{1}{3}(2)+\frac{1}{3}(3)-\left[\frac{2}{3}(-1)+\frac{1}{3}(-3)\right]=3\frac{2}{3} \\
ATE &= E_{X_1}[E[Y_1 \mid X_1]] - E_{X_0}[E[Y_0 \mid X_0]] \\
&= \frac{1}{3}(1)+\frac{1}{3}(2)+\frac{1}{3}(3)-\left[\frac{2}{3}(-1)+\frac{1}{3}(-3)\right]=3\frac{2}{3}
\end{aligned}$$

Regression parameters are expected to be unbiased (and asymptotically consistent) as unobservables in each subpopulation regression are conditionally unbiased.

	$X_j=1$	$X_j=2$	$X_j=3$	$t=0$	$t=1$
$E[U_1 \mid X_1,t]$	0	0	0	0	0
$E[U_0 \mid X_0,t]$	0	NA	0	0	0

Also, conditional mean independence is satisfied.

	$D=1$	$D=0$
$E[Y_1 \mid X_1=1, D]$	1	1
$E[Y_1 \mid X_1=2, D]$	2	2
$E[Y_1 \mid X_1=3, D]$	3	3
$E[Y_0 \mid X_0=1, D]$	-1	-1
$E[Y_0 \mid X_0=2, D]$	NA	NA
$E[Y_0 \mid X_0=3, D]$	-3	-3

Design one yields

$$\begin{gathered} Y=-1\tfrac{2}{3}+0t+3\tfrac{2}{3}D+\varepsilon_1, \\ \textit{suggested } ATT=3\tfrac{2}{3} \end{gathered} \tag{1}$$

Design two yields

$$\begin{gathered} Y=-\tfrac{2}{3}+0t+3\tfrac{13}{15}D-\tfrac{3}{5}X+\varepsilon_2, \\ \textit{suggested } ATT=3\tfrac{13}{15} \end{gathered} \tag{2}$$

Design three yields[4]

$$
\begin{array}{c}
Y = 0 + 0t + 0D - X + 2X \times D + \varepsilon_3, \\
\textit{suggested } ATT\,(X = 1) = 0 + 2 \times 1 = 2 \\
\textit{suggested } ATT\,(X = 2) = 0 + 2 \times 2 = 4 \\
\textit{suggested } ATT\,(X = 3) = 0 + 2 \times 3 = 6 \\
\textit{suggested } ATT = 0 + 2 \times 1\tfrac{3}{4} = 3\tfrac{1}{2}
\end{array}
\tag{3}
$$

As suggested the parameters are consistent with the DGP (see design three). Only design one effectively identifies ATT because of the lack of common support (no overlap at $X = 2$). Design three effectively identifies conditional average treatment effects for $X = 1, 3$. To explore common support, we again focus on the data excluding $(X = 2)$. The average treatment effects conditional on $(X = 1, 3)$ limited support exhibit the same outcome homogeneity.

$$
\begin{aligned}
ATT\,(X = 1, 3) &= E\left[Y_1 \mid X_1 = 1, 3, D = 1\right] - E\left[Y_0 \mid X_0 = 1, 3, D = 1\right] \\
&= \frac{1}{2}(1) + \frac{1}{2}(3) - \left[\frac{2}{3}(-1) + \frac{1}{3}(-3)\right] = 3\frac{2}{3} \\
ATUT\,(X = 1, 3) &= E\left[Y_1 \mid X_1 = 1, 3, D = 0\right] - E\left[Y_0 \mid X_0 = 1, 3, D = 0\right] \\
&= \frac{1}{2}(1) + \frac{1}{2}(3) - \left[\frac{2}{3}(-1) + \frac{1}{3}(-3)\right] = 3\frac{2}{3} \\
ATE\,(X = 1, 3) &= E_{X_1}\left[E\left[Y_1 \mid X_1 = 1, 3\right]\right] - E_{X_0}\left[E\left[Y_0 \mid X_0 = 1, 3\right]\right] \\
&= 3\frac{2}{3}
\end{aligned}
$$

Although conditional mean independence is satisfied, the covariate distribution for Y_1 differs between the two subpopulations, $D = 0, 1$. Design one yields

$$
\begin{array}{c}
Y = -1\tfrac{2}{3} + 0t + 3\tfrac{2}{3}D + \varepsilon_1, \\
\textit{suggested } ATT\,(X = 1, 3) = 3\tfrac{2}{3}
\end{array}
\tag{1}
$$

Design two yields

$$
\begin{array}{c}
Y = -\tfrac{2}{3} + 0t + 3\tfrac{13}{15}D - \tfrac{3}{5}X + \varepsilon_2, \\
\textit{suggested } ATT\,(X = 1, 3) = 3\tfrac{13}{15}
\end{array}
\tag{2}
$$

Design three yields[5]

$$
\begin{array}{c}
Y = 0 + 0t + 0D - 1X + 2X \times D + \varepsilon_3, \\
\textit{suggested } ATT\,(X = 1) = 0 + 2 \times 1 = 2 \\
\textit{suggested } ATT\,(X = 3) = 0 + 2 \times 3 = 6 \\
\textit{suggested } ATT\,(X = 1, 3) = 0 + 2 \times 1\tfrac{8}{11} = 3\tfrac{5}{11}
\end{array}
\tag{3}
$$

As in example 3, only design one identifies the local $ATT\,(X = 1, 3)$. Design three effectively identifies conditional average treatment effects for $X = 1, 3$.

[4] Here suggested ATT involves iteration of X over the entire observed sample. If we employ only the $D = 1$ subsample suggested ATT is $0 + 2 \times 2 = 4 \neq 3\frac{2}{3}$.

[5] Here suggested ATT involves iteration of X over the entire observed sample. If we employ only the $D = 1$ subsample suggested ATT is $0 + 2 \times 2 = 4 \neq 3\frac{2}{3}$.

Apparently, lack of common support in the DGP confounds d-i-d's ability to identify average treatment effects even if sample adjustments are made in attempt to work with common support evidenced in the sample. Notice, if the covariate distribution in the two subpopulations ($D = 0$ and $D = 1$) are balanced as in example 1 average treatment effects are identified by all three designs. Next, we explore a case in which covariates are balanced but conditional mean independence fails.

Example 5 (homogeneous outcome with balanced covariates but conditional mean independence fails) *Suppose the DGP is*

	t	D	X_1	U_1	Y_1	X_0	U_0	Y_0	X	Y
	0	0	1	−1	0	1	1	0	1	0
	0	0	2	0	2	2	0	−2	2	−2
	0	0	3	1	4	3	−1	−4	3	−4
	0	0	1	−1	0	1	1	0	1	0
	0	0	2	0	2	2	0	−2	2	−2
	0	0	3	1	4	3	−1	−4	3	−4
	1	1	1	1	2	3	1	−2	1	2
	1	1	2	0	2	2	0	−2	2	2
	1	1	3	−1	2	1	−1	−2	3	2
	1	0	1	1	2	3	1	−2	3	−2
	1	0	2	0	2	2	0	−2	2	−2
	1	0	3	−1	2	1	−1	−2	1	−2
means	0.5	0.25	2	0	2	2	0	−2	2	−1

Conditional average treatment effects vary with the covariate.

	$X = 1$	$X = 2$	$X = 3$
$E[Y_1 \mid X_1, D = 1]$ $-E[Y_0 \mid X_0, D = 1]$	$2 - (-2) = 4$	$2 - (-2) = 4$	$2 - (-2) = 4$
$E[Y_1 \mid X_1, D = 0]$ $-E[Y_0 \mid X_0, D = 0]$	$\frac{2}{3} - \left(-\frac{2}{3}\right) = \frac{4}{3}$	$2 - (-2) = 4$	$3\frac{1}{3} - \left(-3\frac{1}{3}\right) = 6\frac{2}{3}$
$E[Y_1 \mid X_1] - E[Y_0 \mid X_0]$	$2 - (-2) = 4$	$2 - (-2) = 4$	$2 - (-2) = 4$

Unconditional average treatment effects indicate homogeneity.

$$\begin{aligned}
ATT &= E_{X_1}[E[Y_1 \mid X_1, D = 1]] - E_{X_0}[E[Y_0 \mid X_0, D = 1]] \\
&= \frac{1}{3}(4) + \frac{1}{3}(4) + \frac{1}{3}(4) = 4 \\
ATUT &= E_{X_1}[E[Y_1 \mid X_1, D = 0]] - E_{X_0}[E[Y_0 \mid X_0, D = 0]] \\
&= \frac{1}{3}\left(\frac{4}{3}\right) + \frac{1}{3}(4) + \frac{1}{3}\left(6\frac{2}{3}\right) = 4 \\
ATE &= E_{X_1}[E[Y_1 \mid X_1]] - E_{X_0}[E[Y_0 \mid X_0]] \\
&= \frac{1}{3}(2) + \frac{1}{3}(4) + \frac{1}{3}(6) = 4
\end{aligned}$$

Regression parameters are expected to be unbiased (and asymptotically consistent) as unobservables in each subpopulation regression are conditionally unbiased.

	$X_j = 1$	$X_j = 2$	$X_j = 3$	$t = 0$	$t = 1$
$E[U_1 \mid X_1,t]$	0	0	0	0	0
$E[U_0 \mid X_0,t]$	0	0	0	0	0

Further, conditional mean independence is not satisfied for Y_0 *or* Y_1*.*

	$D = 1$	$D = 0$
$E[Y_1 \mid X_1 = 1, D]$	2	$\frac{2}{3}$
$E[Y_1 \mid X_1 = 2, D]$	2	2
$E[Y_1 \mid X_1 = 3, D]$	2	$3\frac{1}{3}$
$E[Y_0 \mid X_0 = 1, D]$	-2	$-\frac{2}{3}$
$E[Y_0 \mid X_0 = 2, D]$	-2	-2
$E[Y_0 \mid X_0 = 3, D]$	-2	$-3\frac{1}{3}$

Design one yields

$$\begin{array}{c} Y = -2 + 0t + 4D + \varepsilon_1, \\ \textit{suggested } ATT = 4 \end{array} \tag{1}$$

Design two yields

$$\begin{array}{c} Y = 0 + 0t + 4D - 1X + \varepsilon_2, \\ \textit{suggested } ATT = 4 \end{array} \tag{2}$$

Design three yields

$$\begin{array}{c} Y = \frac{2}{3} + 0t + 1\frac{1}{3}D - 1\frac{1}{3}X + 1\frac{1}{3}X \times D + \varepsilon_3, \\ \textit{suggested } ATT\,(X = 1) = 1\frac{1}{3} + 1\frac{1}{3} \times 1 = 2\frac{2}{3} \\ \textit{suggested } ATT\,(X = 2) = 1\frac{1}{3} + 1\frac{1}{3} \times 2 = 4 \\ \textit{suggested } ATT\,(X = 3) = 1\frac{1}{3} + 1\frac{1}{3} \times 3 = 5\frac{1}{3} \\ \textit{suggested } ATT = 1\frac{1}{3} + 1\frac{1}{3} \times 2 = 4 \end{array} \tag{3}$$

The parameters are inconsistent with the DGP (as the DGP does not exhibit conditional mean independence), but remarkably, all three designs effectively identify unconditional ATT*. Not surprisingly as conditional mean independence is violated by the DGP, design three does not identify the conditional average treatment effects for* $X = 1, 3$*. Nonetheless, apparently balanced covariates can overcome other identification pitfalls regarding unconditional average treatment effects.*[6]

Example 6 (heterogeneous outcome with unbalanced covariates) *Suppose*

[6] This result is due to covariate balance and not symmetry of the covariate, X_j, distributions. For instance, replacing $X_j = 3$ with $X_j = 6$ for $j = 0, 1$, such that X_j is evenly distributed between $1, 2$, and 6 — clearly asymmetric and $E[X] = 3$, yields similar identification results.

the DGP is

	t	D	X_1	U_1	Y_1	X_0	U_0	Y_0	X	Y
	0	0	1	-1	0	1	-1	-2	1	-2
	0	0	1	1	2	1	1	0	1	0
	0	0	3	0	3	3	0	-3	3	-3
	0	0	1	1	2	1	-1	-2	1	-2
	0	0	1	-1	0	1	1	0	1	0
	0	0	3	0	3	3	0	-3	3	-3
	1	1	3	-1	2	1	-1	-2	3	2
	1	1	1	0	1	1	1	0	1	1
	1	1	3	1	4	3	0	-3	3	4
	1	0	1	1	2	1	-1	-2	1	-2
	1	0	1	-1	0	1	1	0	1	0
	1	0	3	0	3	3	0	-3	3	-3
means	0.5	0.25	$1\frac{5}{6}$	0	$1\frac{5}{6}$	$1\frac{2}{3}$	0	$-1\frac{2}{3}$	$1\frac{5}{6}$	$-\frac{2}{3}$

Conditional average treatment effects vary with the covariate.

	$X=1$	$X=3$
$E[Y_1 \mid X_1, D=1]$ $-E[Y_0 \mid X_0, D=1]$	$1-(-1)=2$	$3-(-3)=6$
$E[Y_1 \mid X_1, D=0]$ $-E[Y_0 \mid X_0, D=0]$	$1-(-1)=2$	$3-(-3)=6$
$E[Y_1 \mid X_1] - E[Y_0 \mid X_0]$	$1-(-1)=2$	$3-(-3)=6$

Unconditional average treatment effects indicate outcome heterogeneity.

$$\begin{aligned}
ATT &= E_{X_1}[E[Y_1 \mid X_1, D=1]] - E_{X_0}[E[Y_0 \mid X_0, D=1]] \\
&= \frac{1}{3}(1) + \frac{2}{3}(3) - \left[\frac{2}{3}(-1) + \frac{1}{3}(-3)\right] = 4 \\
ATUT &= E_{X_1}[E[Y_1 \mid X_1, D=0]] - E_{X_0}[E[Y_0 \mid X_0, D=0]] \\
&= \frac{2}{3}(1) + \frac{1}{3}(3) - \left[\frac{2}{3}(-1) + \frac{1}{3}(-3)\right] = 3\frac{1}{3} \\
ATE &= E_{X_1}[E[Y_1 \mid X_1]] - E_{X_0}[E[Y_0 \mid X_0]] \\
&= \frac{7}{12}(1) + \frac{5}{12}(3) - \left[\frac{2}{3}(-1) + \frac{1}{3}(-3)\right] \\
&= \Pr(D=1)\, ATT + \Pr(D=0)\, ATUT \\
&= \frac{1}{4}(4) + \frac{3}{4}\left(3\frac{1}{3}\right) = 3\frac{1}{2}
\end{aligned}$$

Regression parameters are expected to be unbiased (and asymptotically consistent) as unobservables in each subpopulation regression are conditionally unbiased.

	$X_j=1$	$X_j=2$	$X_j=3$	$t=0$	$t=1$
$E[U_1 \mid X_1, t]$	0	0	0	0	0
$E[U_0 \mid X_0, t]$	0	0	0	0	0

Further, conditional mean independence is satisfied for both Y_0 *or* Y_1*.*

	$D=1$	$D=0$
$E[Y_1 \mid X_1 = 1, D]$	1	1
$E[Y_1 \mid X_1 = 3, D]$	3	3
$E[Y_0 \mid X_0 = 1, D]$	-1	-1
$E[Y_0 \mid X_0 = 3, D]$	-3	-3

Design one yields

$$\begin{array}{c} Y = -1\frac{2}{3} + 0t + 4D + \varepsilon_1, \\ \textit{suggested } ATT = 4 \end{array} \tag{1}$$

Design two yields

$$\begin{array}{c} Y = -\frac{5}{6} + 0t + 4\frac{1}{3}D - \frac{1}{2}X + \varepsilon_2, \\ \textit{suggested } ATT = 4\frac{1}{3} \end{array} \tag{2}$$

Design three yields[7]

$$\begin{array}{c} Y = 0 + 0t + 0D - 1X + 2X \times D + \varepsilon_3, \\ \textit{suggested } ATT\,(X=1) = 0 + 2 \times 1 = 2 \\ \textit{suggested } ATT\,(X=3) = 0 + 2 \times 3 = 6 \\ \textit{suggested } ATT = 0 + 2 \times 1\frac{5}{6} = 3\frac{2}{3} \end{array} \tag{3}$$

The parameters are consistent with the DGP, but because of covariate imbalance only design one effectively identifies the unconditional ATT*. However, design three identifies the conditional average treatment effects for* $X = 1, 3$*. Apparently, a d-i-d design typically does not effectively identify average treatment effects for a DGP with unbalanced covariates.*

[7] Here suggested ATT involves iteration of X over the entire observed sample. If we employ only the $D = 1$ subsample suggested ATT is $0 + 2 \times 2\frac{1}{3} = 4\frac{2}{3} \neq 4$.