

新北市垃圾費隨袋徵收效果- Synthetic Control Method

2020/06/11 邱奕勳

新北市

隨水費徵收

全面實施垃圾費
隨袋徵收

2010/12

全面實施垃圾費
隨袋徵收



Y

平均每人每日
一般廢棄物產生量 (公斤)

去除其他變數對Y的影響

DID

Pre Post

Treatment

Control

	20	25
	18	22

DID

Treatment

Control

Pre

Post

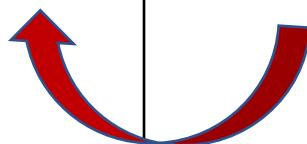
20

25

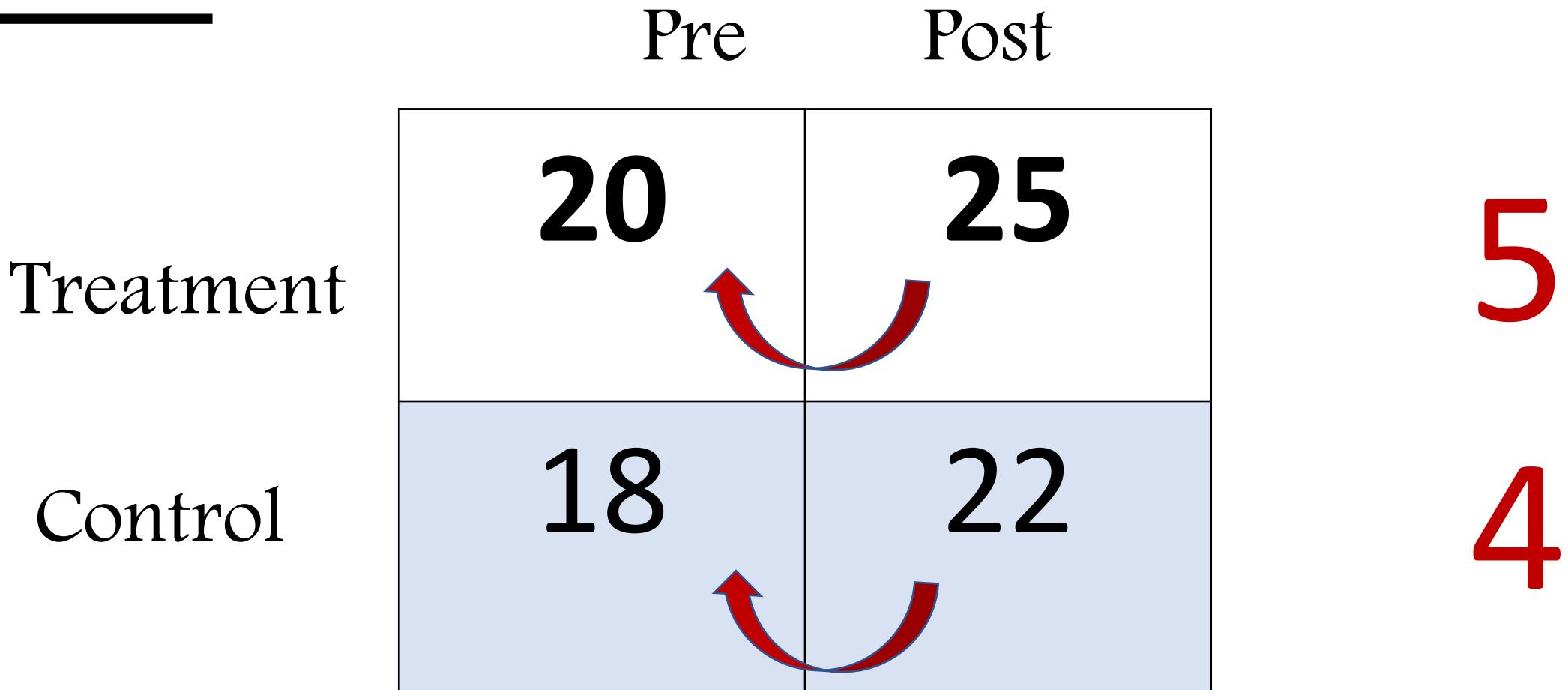
18

22

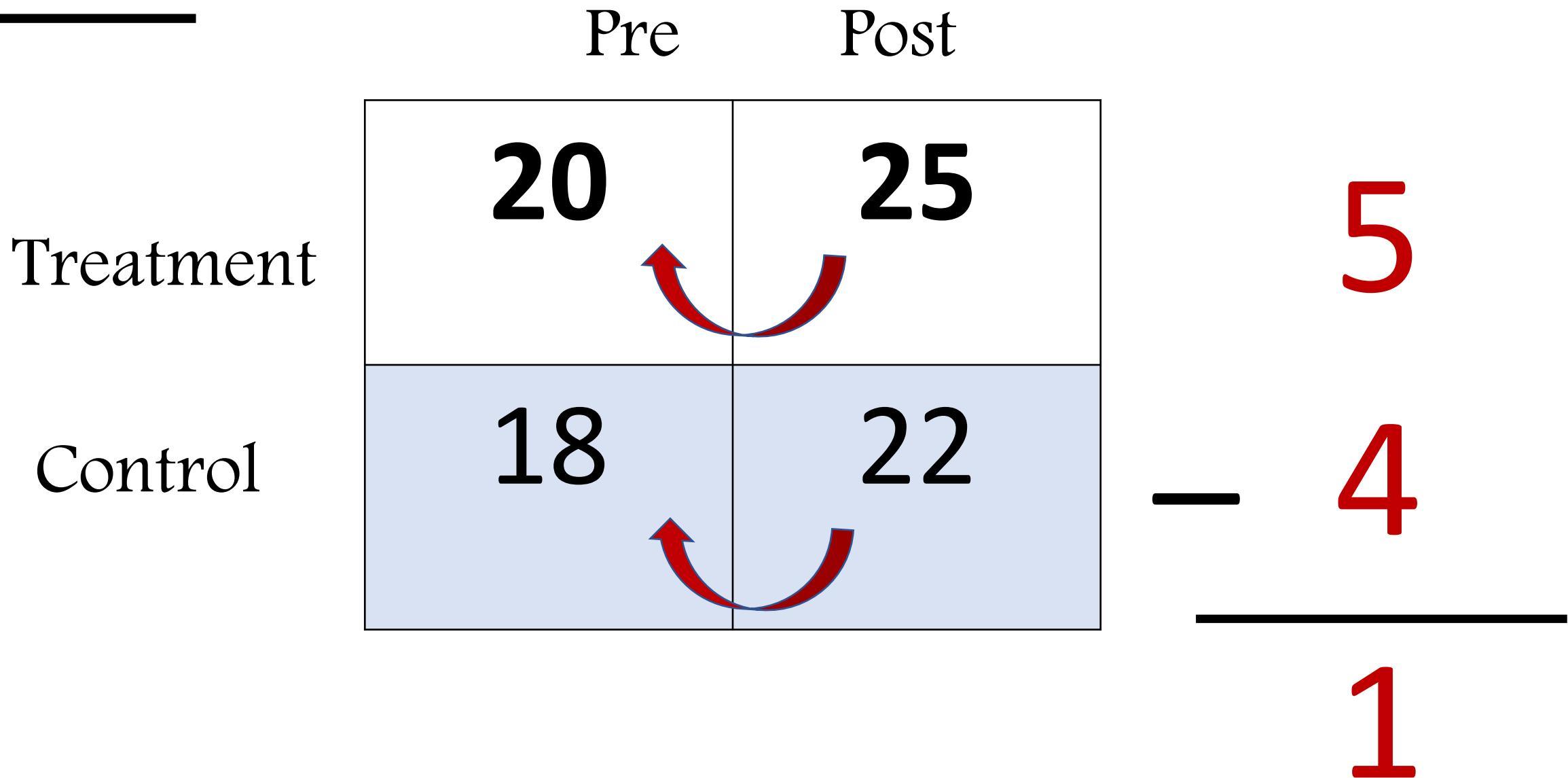
5



DID

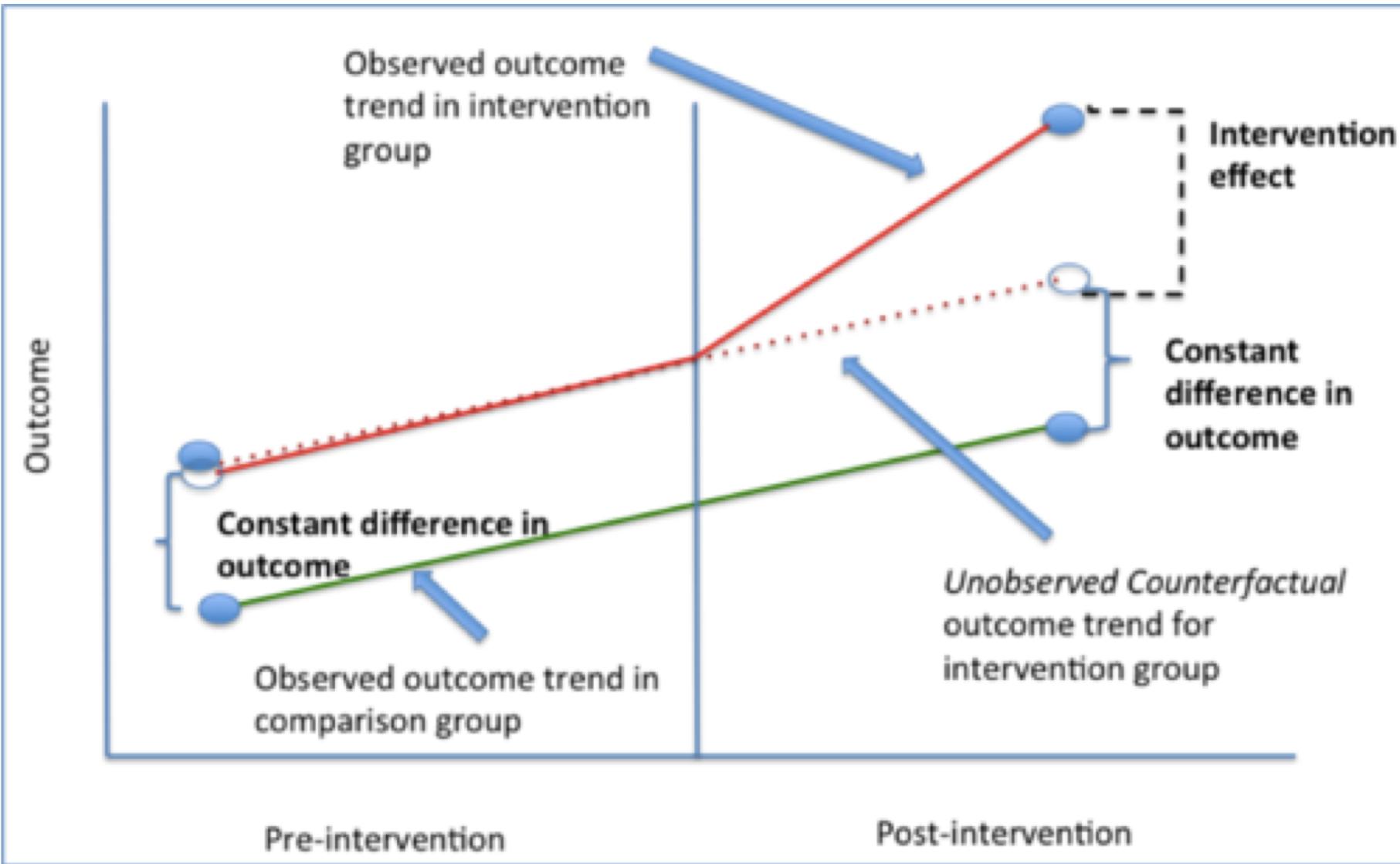


DID



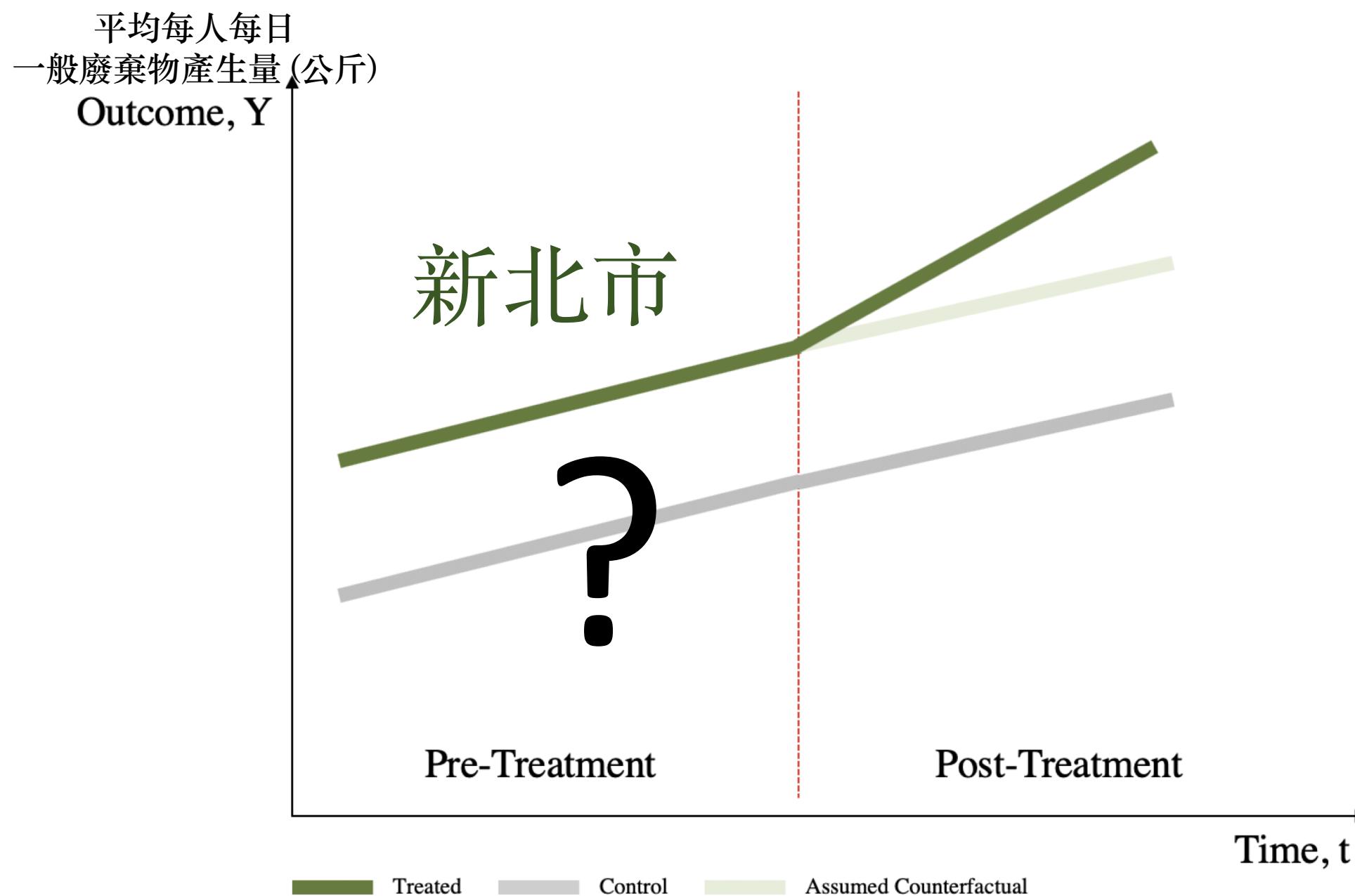
DID

parallel trend assumption

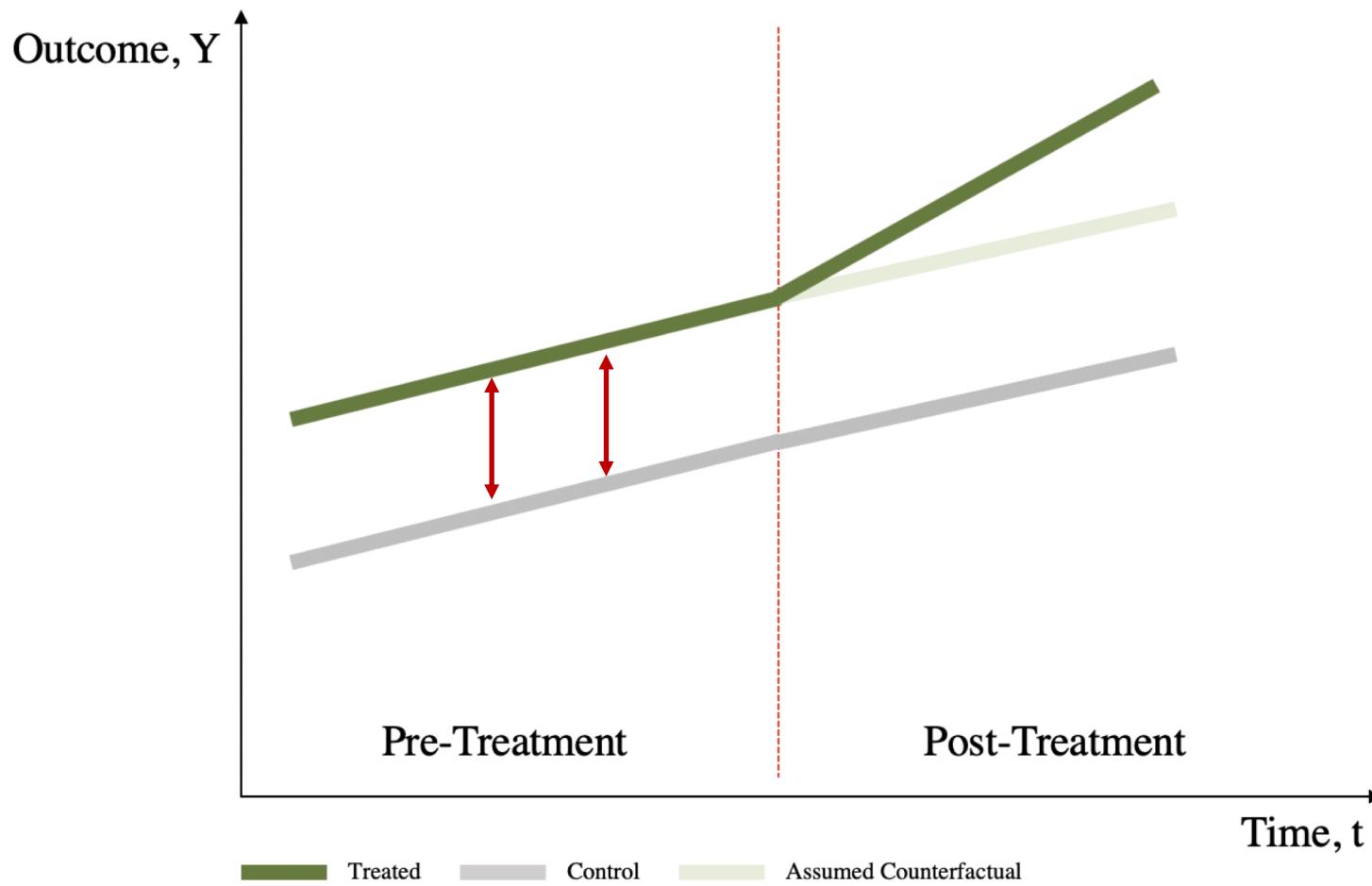


Treatment Group : 新北市
Control Group : 新嘉竹苗彰南雲...





Constant Difference



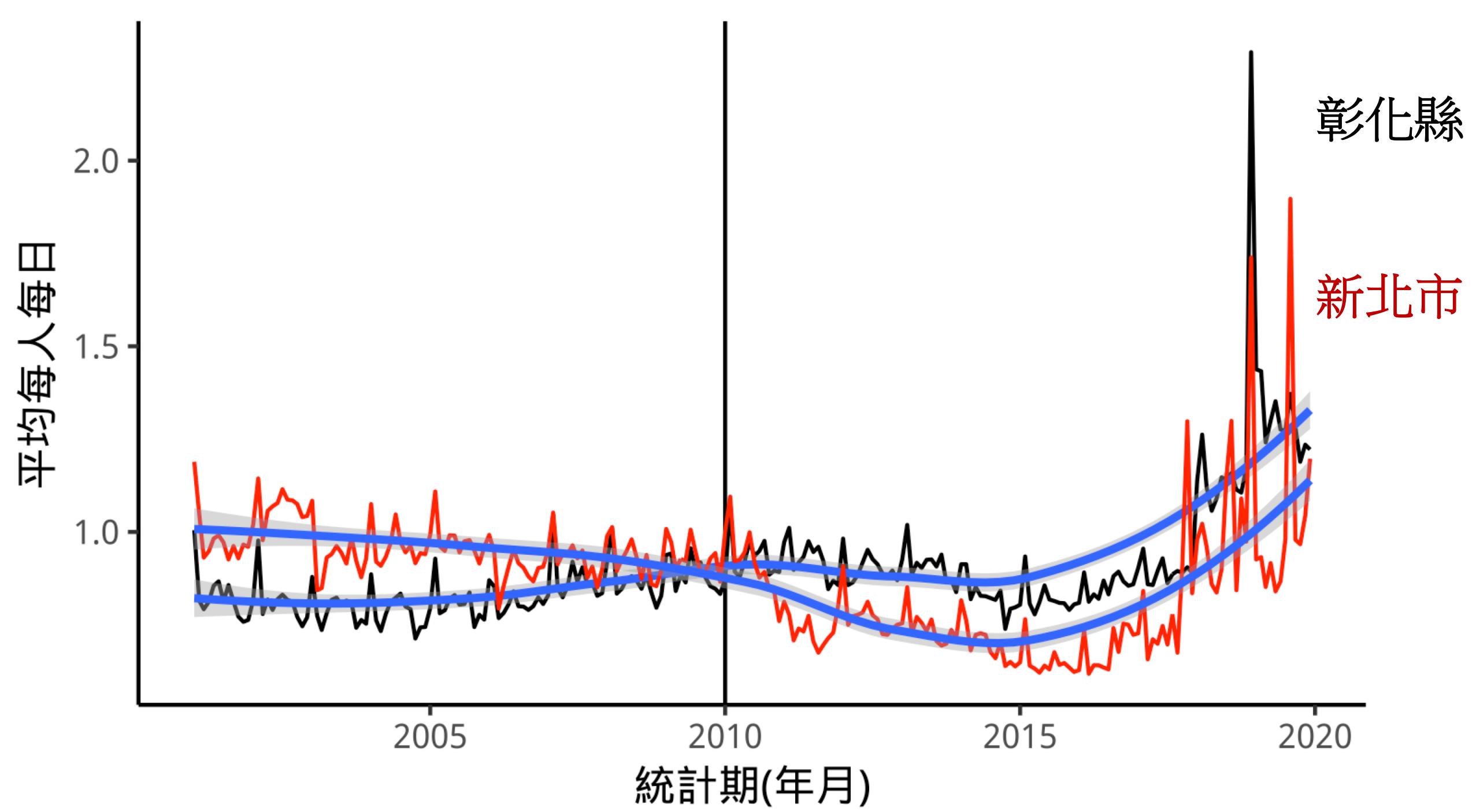
計算Differences 的標準差

▲	臺北市	桃園市	臺中市	高雄市	宜蘭縣	苗栗縣	彰化縣	雲林縣	嘉義縣	屏東縣	臺東縣	花蓮縣	澎湖縣
新北市	0.2413775	0.1866363	0.1638206	0.1678436	0.1662487	0.1891941	0.1592426	0.1904631	0.1909153	0.1807588	0.2248391	0.270326	0.2097177
排序	15.0000000	8.0000000	2.0000000	5.0000000	4.0000000	9.0000000	1.0000000	10.0000000	11.0000000	7.0000000	13.0000000	16.0000000	12.0000000

計算Differences 的標準差

▲	臺北市	桃園市	臺中市	高雄市	宜蘭縣	苗栗縣	彰化縣	雲林縣	嘉義縣	屏東縣	臺東縣	花蓮縣	澎湖縣
新北市	0.2413775	0.1866363	0.1638206	0.1678436	0.1662487	0.1891941	0.1592426	0.1904631	0.1909153	0.1807588	0.2248391	0.270326	0.2097177
排序	15.0000000	8.0000000	2.0000000	5.0000000	4.0000000	9.0000000	1.0000000	10.0000000	11.0000000	7.0000000	13.0000000	16.0000000	12.0000000

與新北市差異的變動幅度最小



```
lm(formula = DIDNtpChang2$`平均每人每日一般廢棄物產生量 (公斤)` ~  
  treat + post + effect, data = DIDNtpChang2)
```

Residuals:

Min	1Q	Median	3Q	Max
-0.192925	-0.046925	-0.007425	0.038925	0.236075

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.839158	0.006050	138.711	< 2e-16 ***
treatTRUE	0.113767	0.008556	13.297	< 2e-16 ***
postTRUE	0.034050	0.009879	3.447	0.000631 ***
effectTRUE	-0.272850	0.013971	-19.529	< 2e-16 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.06627 on 380 degrees of freedom

Multiple R-squared: 0.6119, Adjusted R-squared: 0.6088

F-statistic: 199.7 on 3 and 380 DF, p-value: < 2.2e-16

怎樣才足夠是一個好的控制組？

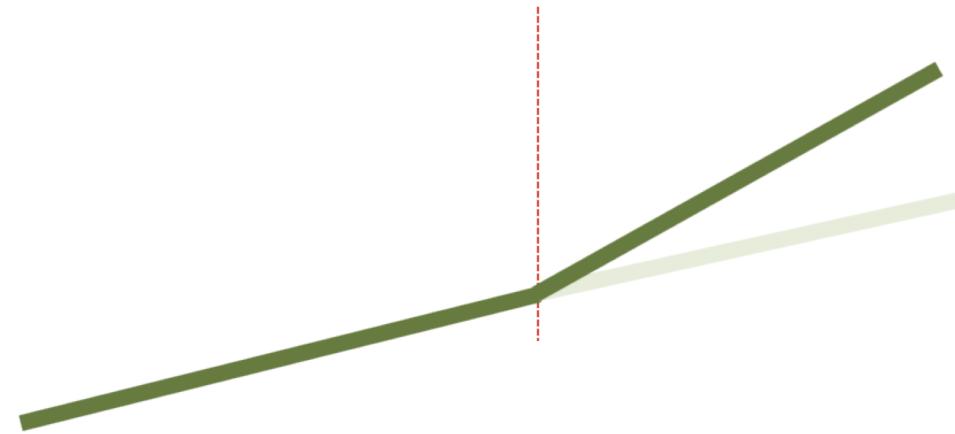
Parallel Trend Assumption Holds?

Difference SD 怎樣才算夠小？

極端：

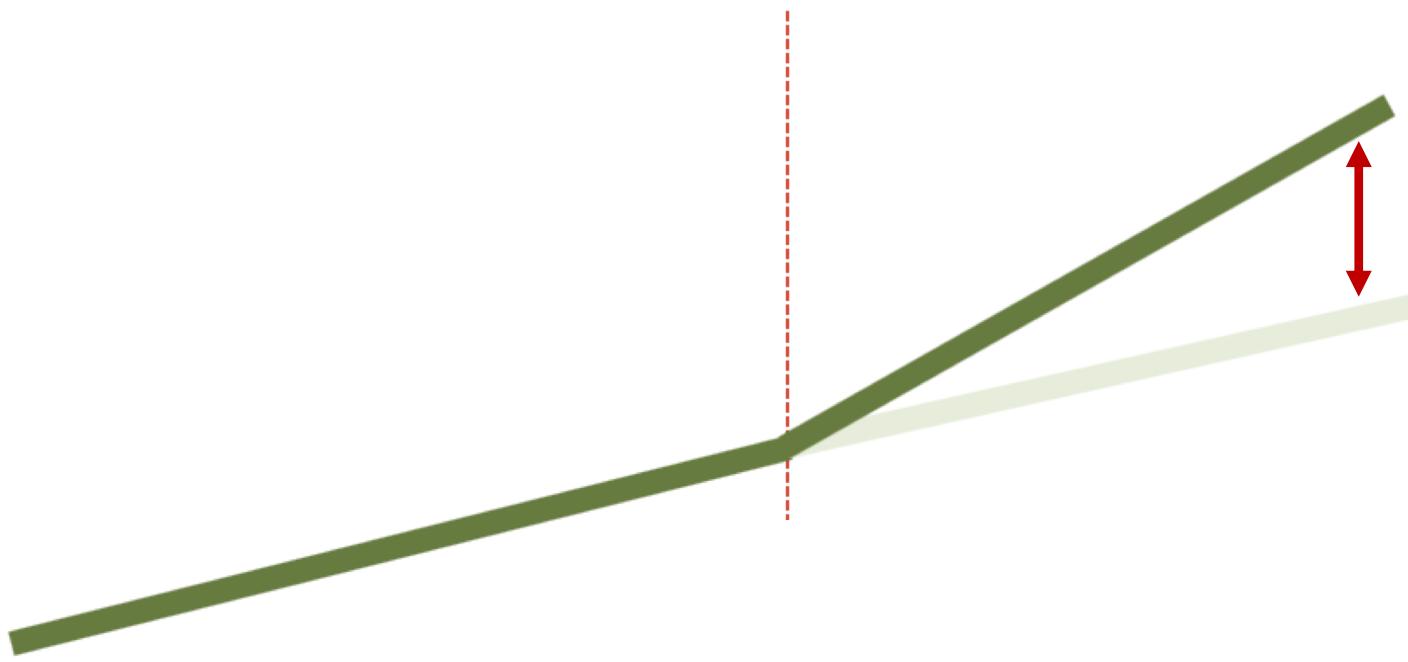
Pre-treatment Y of Control G = Pre-treatment Y of Treatment G

Difference SD = 0, parallel assumption holds perfectly



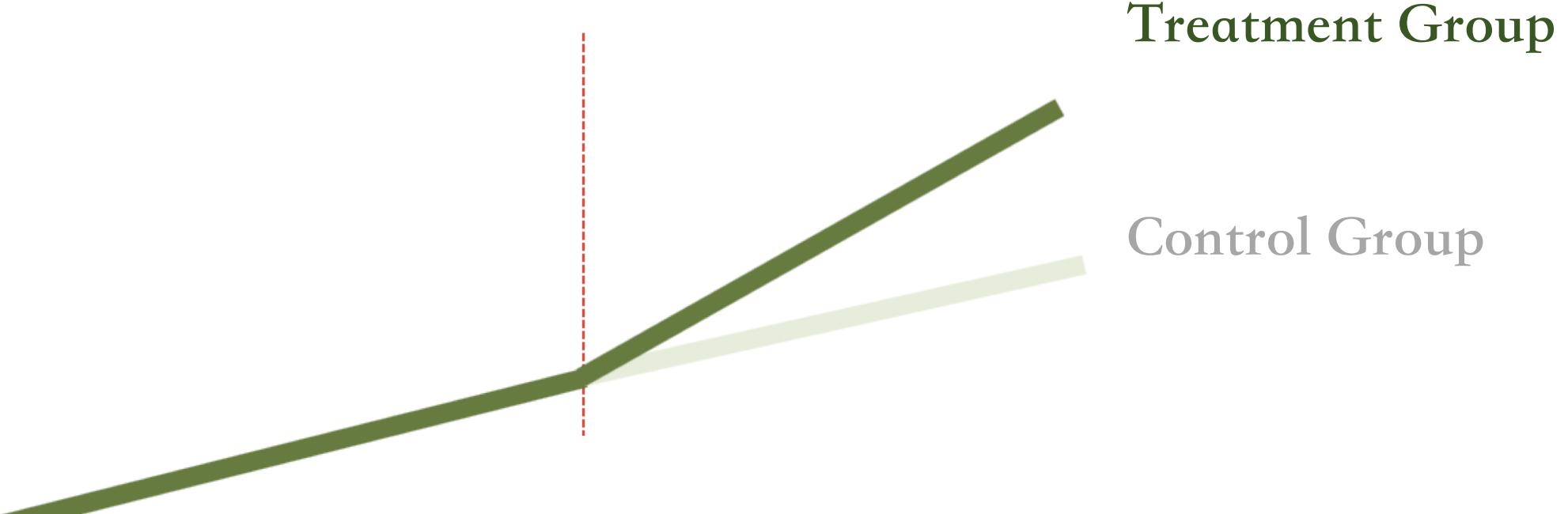
Treatment Group

Control Group



Objectives

找到一個近乎相似的control group，用control Group的結果，
估計實驗組如果沒有 實驗下的結果。



Control Group

新北市

Treatment Group

新北市

v.s



Control Group

Counterfactual
新北市

Treatment Group

新北市

v.s



Synthetic Control Method (SCM)

Control Group

合成
新北市

Treatment Group

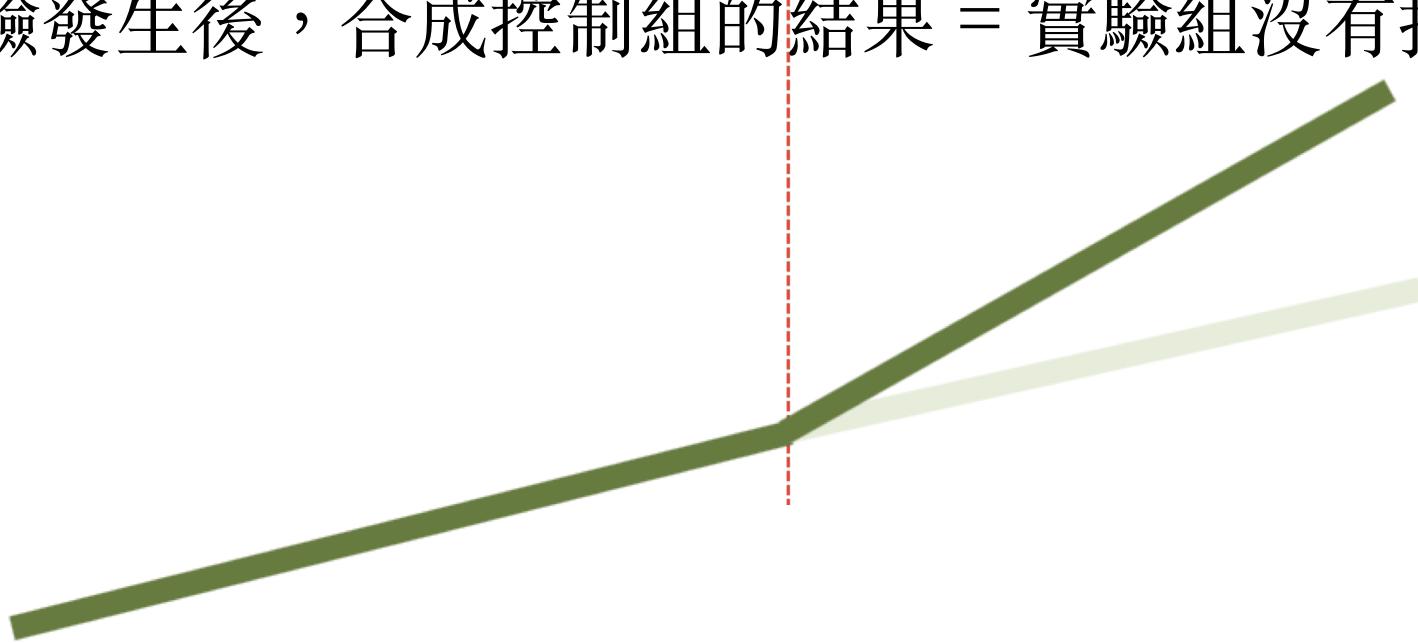
v.s
新北市

特徵	合成新北市	新北市
平均每戶消費支出	600,000	610,000
失業率	4.5	4.5
平均每戶人數	3.1	3.2
平均每人可支配所得	45,000	45,100
平均每人每日一般廢棄物產生量(公斤)	0.95	0.95

Synthetic Control Method (SCM)

給予數個控制組加權平均，
合成一個與實驗組特徵相像、Y變數相近的合成控制組

- > 在實驗發生前，合成控制組與實驗組有極高的相似度
- > 在實驗發生後，合成控制組的結果 = 實驗組沒有接受實驗的結果

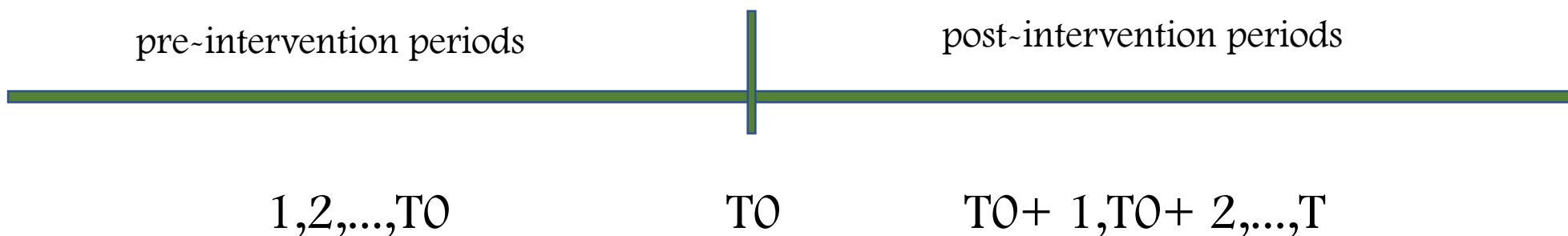


How SCM works

Observe units $j = 1, \dots, J+1$ for time periods $t = 1, \dots, T$

Treatment group : $J = 1$

Possible donor: $J = 2, \dots, J+1$



$$Y_{it}^N$$

Unit i 在t時刻沒有接收實驗，觀察到的結果

$$Y_{it}^I$$

Unit i 在t時刻接收實驗，觀察到的結果

Effect α_{1t} = $Y_{1t}^I - Y_{1t}^N$ for $t = T_0 + 1, \dots, T$

$$Y_{it}^N$$

Unit i 在t時刻沒有接收實驗，觀察到的結果

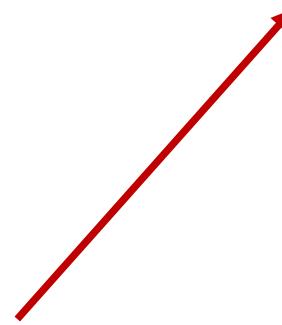
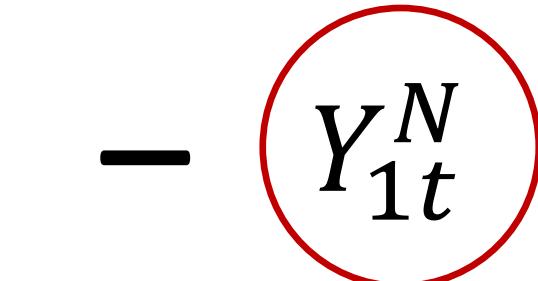
$$Y_{it}^I$$

Unit i 在t時刻接收實驗，觀察到的結果

Effect α_{1t} = $Y_{1t}^I - Y_{1t}^N$ for $t = T_0 + 1, \dots, T$

*Counterfactual
unobserved*

$$\text{Effect } \alpha_{1t} = Y_{1t}^I - Y_{1t}^N \quad \text{for } t = T_0 + 1, \dots, T$$



Reasonable Estimate of Y_{1t}^N

Reasonable Estimate of Y_{1t}^N

特徵

$U_i : (\gamma X 1)$ vector of observed covariates of each unit.

Pre-treatment

Outcome $(T_0 X 1)$ vector $K = (k_1, \dots, k_{T_0})'$ $\rightarrow Y_i^K = \sum_{s=1}^{T_0} k_s Y_{is}$

Reasonable Estimate of Y_{1t}^N

特徵

U_i : (γX_1) vector of observed covariates of each unit.

Pre-treatment

Outcome ($T_0 X_1$) vector $K = (k_1, \dots, k_{T_0})'$ $\rightarrow Y_i^K = \sum_{s=1}^{T_0} k_s Y_{is}$

$W = (w_2, \dots, w_{J+1})$ such that $w_j \geq 0$ for $j = 2, \dots, J+1$ and $\sum_{j=2}^{J+1} w_j = 1$

W^* such that 1) $\sum_{j=2}^{J+1} w_j^* Y_j^{K_1} = Y_1^{K_1}$... $\sum_{j=2}^{J+1} w_j^* Y_j^{K_M} = Y_1^{K_M}$

2) $\sum_{j=2}^{J+1} w_j^* U_j = U_1$

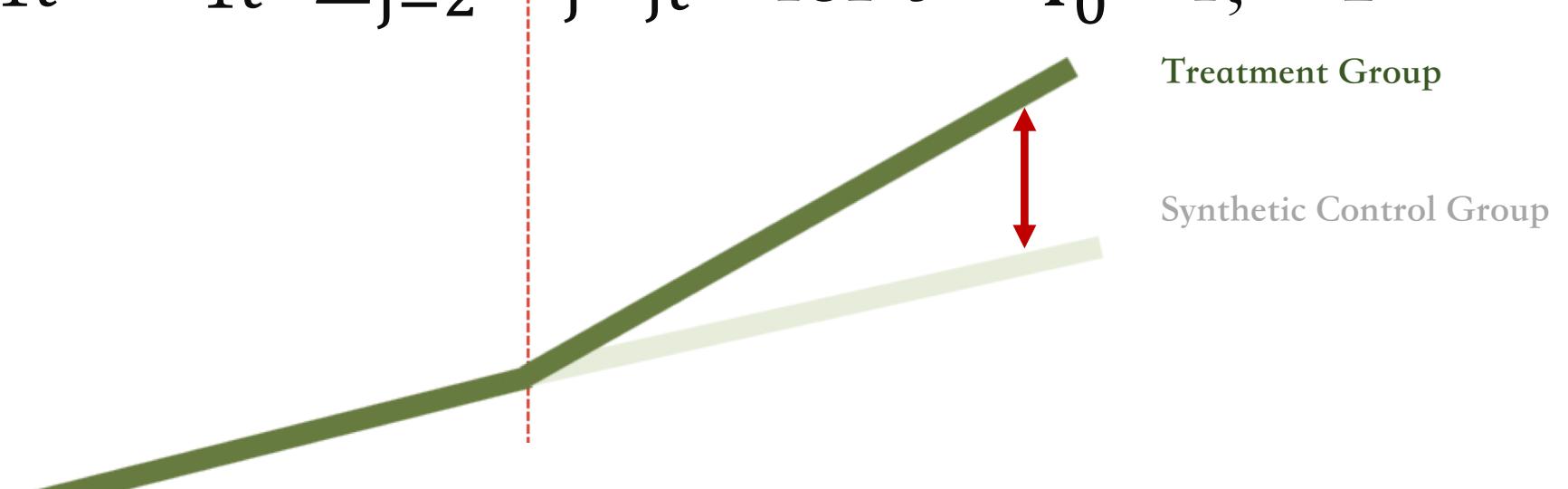
Reasonable Estimate of $\hat{\alpha}_{1t}^N$

$W = (w_2, \dots, w_{J+1})$ such that $w_j \geq 0$ for $j = 2, \dots, J+1$ and $\sum_{j=2}^{J+1} w_j = 1$

W^* such that 1) $\sum_{j=2}^{J+1} w_j^* Y_j^{K_1} = Y_1^{K_1}$... $\sum_{j=2}^{J+1} w_j^* Y_j^{K_M} = Y_1^{K_M}$

2) $\sum_{j=2}^{J+1} w_j^* U_j = U_1$

-> $\hat{\alpha}_{1t} = Y_{1t} - \sum_{j=2}^{J+1} w_j^* Y_{jt}$ for $t = T_0 + 1, \dots, T$



minimize the distance

$$\|X_1 - X_0 W\|_V = \sqrt{(X_1 - X_0 W)' V (X_1 - X_0 W)}$$

Treatment Group
Characteristics

$$(k \times 1) \quad X_1 = (U'_1, \bar{Y}_1^{K_1}, \dots, \bar{Y}_1^{K_M})'$$

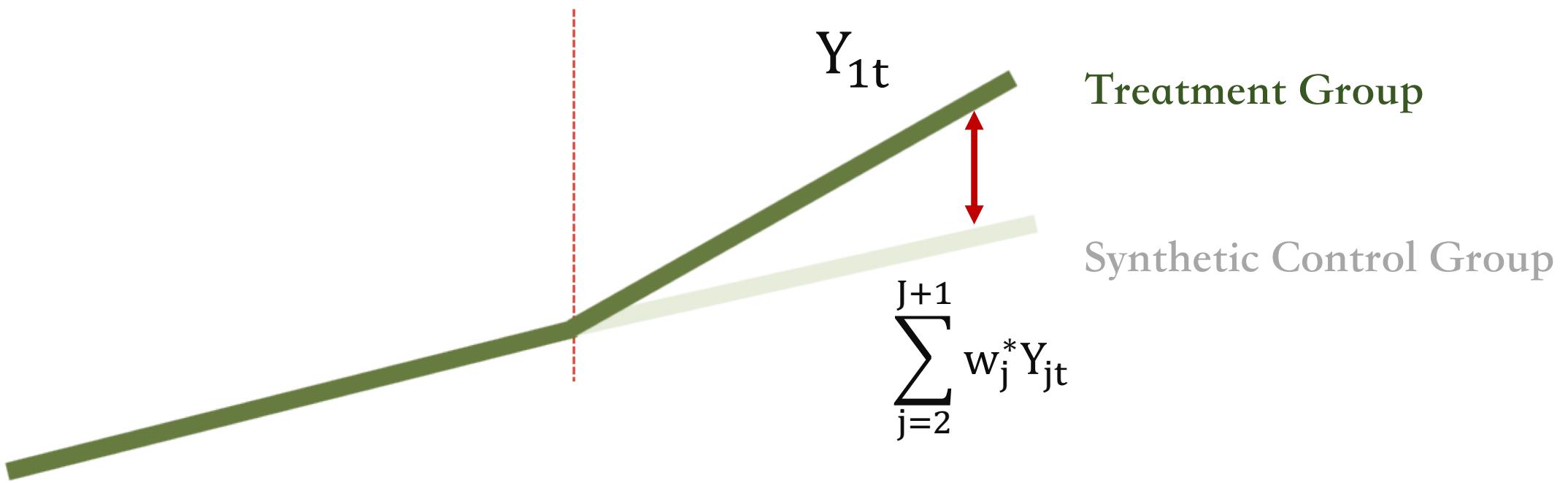
Control Group
Characteristics

($k \times J$) matrix X_0 = the values of the same characteristics of the control units

V is defined as some $(k \times k)$ symmetric and positive semidefinite matrix

Obtained by minimizing the mean square error of the synthetic control estimator,

$$(Y_1 - Y_0 W^*)' (Y_1 - Y_0 W^*)$$



Treatment Group

Synthetic Control Group

Synth

Y : 平均每人每日一般廢棄物產生量 (公斤)

Predictor (U):

平均每戶消費支出,

失業率,

平均每戶人數,

平均每人可支配所得,

平均每人每日一般廢棄物產生量 (公斤)

Data preparation

```
```{r}
trashprep.out <- dataprep(
 .foo = trash_by_year,
 .predictors = c("平均每戶消費支出", "失業率", "平均每戶人數", "平均每人可支配所得", "平均每人每日一般廢棄物產生量 (公斤)"),
 .predictors.op = "mean",
 .time.predictors.prior = 2001:2010,
 .dependent = "平均每人每日一般廢棄物產生量 (公斤)",
 .unit.variable = "regionno",
 .unit.names.variable = "統計區",
 .time.variable = "統計期",
 .treatment.identifier = 15,
 .controls.identifier = c(1:14, 16:20),
 .time.optimize.ssr = 2001:2010,
 .time.plot = 2001:2018
)
```

```

Synth

Optimization

```
```{r}
synth.out = synth(data.prep.obj = trashprep.out)
```
```

| w.weights | unit.names | unit.numbers |
|-----------|------------|--------------|
| 0.008 | 高雄市 | 1 |
| 0.010 | 花蓮縣 | 2 |
| 0.297 | 基隆市 | 3 |
| 0.007 | 嘉義市 | 4 |
| 0.000 | 嘉義縣 | 5 |
| 0.015 | 苗栗縣 | 6 |
| 0.019 | 南投縣 | 7 |
| 0.007 | 澎湖縣 | 8 |
| 0.009 | 屏東縣 | 9 |
| 0.000 | 臺北市 | 10 |
| 0.010 | 臺東縣 | 11 |
| 0.013 | 臺南市 | 12 |
| 0.177 | 臺中市 | 13 |
| 0.009 | 桃園市 | 14 |
| 0.275 | 新竹市 | 16 |
| 0.102 | 新竹縣 | 17 |
| 0.015 | 宜蘭縣 | 18 |
| 0.015 | 雲林縣 | 19 |
| 0.013 | 彰化縣 | 20 |

Synth

```
```{r}  
synth.tables$tab.pred
```
```

| | Treated | Synthetic | Sample Mean |
|---------------------|------------|------------|-------------|
| 平均每戶消費支出 | 734629.400 | 734371.533 | 628338.563 |
| 失業率 | 4.710 | 4.691 | 4.616 |
| 平均每戶人數 | 3.485 | 3.485 | 3.376 |
| 平均每人可支配所得 | 509251.900 | 509423.440 | 470422.116 |
| 平均每人每日一般廢棄物產生量 (公斤) | 0.953 | 0.953 | 0.937 |

Synth

```
```{r}
gaps = trashprep.out$Y1plot - .(trashprep.out$Y0plot %% synth.out$solution.w)
gaps
```
```

```
15
2001 -0.01569159
2002 0.11484778
2003 0.02489433
2004 0.01027955
2005 0.03043715
2006 -0.08098862
2007 -0.08329828
2008 0.01511565
2009 0.01074276
2010 -0.02656256
2011 -0.20460701
2012 -0.19940305
2013 -0.23910431
2014 -0.29349362
2015 -0.33784746
2016 -0.27667458
2017 -0.22621251
2018 -0.22185531
```

```
```{r}
gaps[1:10] %>% sd()
```
```

```
[1] 0.05738299
```

Gaps differences

SCM

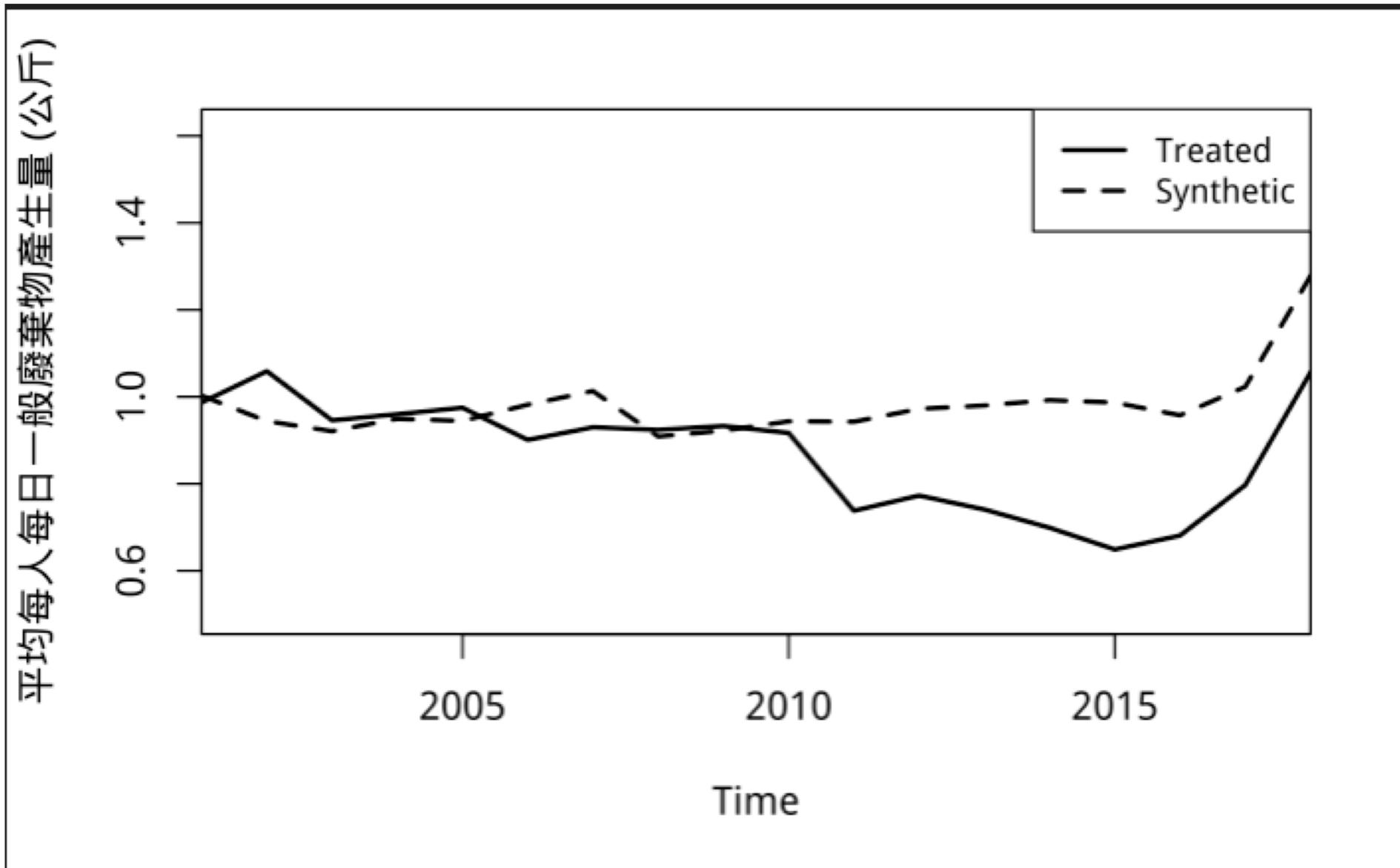
Synthetic Control Unit

0.05738299



DID

彰化縣



行政院環境保護署 106 年 10 月修正垃圾計算方式，將事業垃圾量加入生活廢棄物中，因此造成 106 年後之垃圾清運量增加。

Synth

| 統計
期 | 統計
區 | 每人每
日 | treat | post | effect |
|---------|-----------|-----------|-------|-------|--------|
| 2001 | synthetic | 1.0026916 | FALSE | FALSE | FALSE |
| 2001 | NTP | 0.9870000 | TRUE | FALSE | FALSE |
| 2002 | synthetic | 0.9441522 | FALSE | FALSE | FALSE |
| 2002 | NTP | 1.0590000 | TRUE | FALSE | FALSE |
| 2003 | synthetic | 0.9211057 | FALSE | FALSE | FALSE |
| 2003 | NTP | 0.9460000 | TRUE | FALSE | FALSE |
| 2004 | synthetic | 0.9497205 | FALSE | FALSE | FALSE |
| 2004 | NTP | 0.9600000 | TRUE | FALSE | FALSE |
| 2005 | synthetic | 0.9445628 | FALSE | FALSE | FALSE |
| 2005 | NTP | 0.9750000 | TRUE | FALSE | FALSE |
| 2006 | synthetic | 0.9819886 | FALSE | FALSE | FALSE |

Synth

```
Call:  
lm(formula = synthntp_ntp_lmdata$每人每日 ~ treat + post +  
    effect, data = synthntp_ntp_lmdata)  
  
Residuals:  
    Min      1Q  Median      3Q     Max  
-0.11775 -0.03629 -0.02170  0.01055  0.28925  
  
Coefficients:  
              Estimate Std. Error t value Pr(>|t|)  
(Intercept) 9.532e-01 2.644e-02 36.048 < 2e-16 ***  
treatTRUE   -2.238e-05 3.740e-02 -0.001     1.00  
postTRUE     6.343e-02 3.967e-02   1.599     0.12  
effectTRUE  -2.499e-01 5.609e-02  -4.455 9.62e-05 ***  
---  
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1  
  
Residual standard error: 0.08362 on 32 degrees of freedom  
Multiple R-squared:  0.5588,    Adjusted R-squared:  0.5175  
F-statistic: 13.51 on 3 and 32 DF,  p-value: 7.272e-06
```

平均每人每日一般廢棄物產生量 (公斤)

Synth

```
Call:  
lm(formula = synthntp_ntp_lmdata$每人每日 ~ treat + post +  
  effect, data = synthntp_ntp_lmdata)  
  
Residuals:  
    Min      1Q   Median     3Q     Max  
-0.11775 -0.03629 -0.02170  0.01055  0.28925  
  
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)  
(Intercept) 9.532e-01 2.644e-02 36.048 < 2e-16 ***  
treatTRUE   -2.238e-05 3.740e-02 -0.001    1.00  
postTRUE    6.343e-02 3.967e-02  1.599    0.12  
effectTRUE  -2.499e-01 5.609e-02 -4.455 9.62e-05 ***  
---  
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1  
  
Residual standard error: 0.08362 on 32 degrees of freedom  
Multiple R-squared:  0.5588,    Adjusted R-squared:  0.5175  
F-statistic: 13.51 on 3 and 32 DF,  p-value: 7.272e-06
```

DID

```
lm(formula = DIDNtpChang2$`平均每人每日一般廢棄物產生量 (公斤)` ~  
  treat + post + effect, data = DIDNtpChang2)  
  
Residuals:  
    Min      1Q   Median     3Q     Max  
-0.192925 -0.046925 -0.007425  0.038925  0.236075  
  
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)  
(Intercept) 0.839158  0.006050 138.711 < 2e-16 ***  
treatTRUE   0.113767  0.008556 13.297 < 2e-16 ***  
postTRUE    0.034050  0.009879  3.447 0.000631 ***  
effectTRUE -0.272850  0.013971 -19.529 < 2e-16 ***  
---  
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1  
  
Residual standard error: 0.06627 on 380 degrees of freedom  
Multiple R-squared:  0.6119,    Adjusted R-squared:  0.6088  
F-statistic: 199.7 on 3 and 380 DF,  p-value: < 2.2e-16
```

感動

於聽心

[Alberto A., Alexis D.,and Jens H.,\(2011\), “Synth: An R Package for Synthetic Control Methods in Comparative Case Studies” *Journal of Statistical Software*](#)

[Causal Inference Using Synthetic Control: The Ultimate Guide](#)

[“A Review on Factors affecting Municipal Solid Waste Generation”](#), 2nd International Engineering Conference.
Federal University of Technology, Minna, Niger State, Nigeria. Minna, Nigeria, October 17-19, 2017

[Prediction of municipal solid waste generation using artificial neural network approach enhanced by structural break analysis](#)