Econometrics of Evaluation
Difference-in-Differences (DID)

Jean-Noël Senne

Université Paris-Saclay

Problématiques Contemporaines du Développement
M2EIED
Université de Pau et des Pays de l'Adour
2023 - 2024

The false counterfactuals

- Remember that the basic principle of an evaluation method is to find or build an unobserved counterfactual
- There are two 'naive' but generally **false counterfactuals :**
  1. **Pre-treatment vs. post-treatment** comparisons (before/after)
     $\Rightarrow$ **Time-trend** bias
  2. **Treated vs. controls** comparisons (with/without)
     $\Rightarrow$ **Selection** bias

$\Rightarrow$ The **Difference-in-Differences estimation** ("diff-in-diff" or DID) combines these two flawed approaches !

- **Identification assumption : Parallel or common trend** in the outcomes

$$E(Y_{0t_1} - Y_{0t_0} | T = 1) = E(Y_{0t_1} - Y_{0t_0} | T = 0)$$

$\rightarrow$ In the absence of treatment, the evolution of the outcomes would have been the same between treated and controls (between t_0 and t_1)

$\rightarrow$ The counterfactual **levels** of Y_0 can be different but their **time variation** is similar

When two false make it right

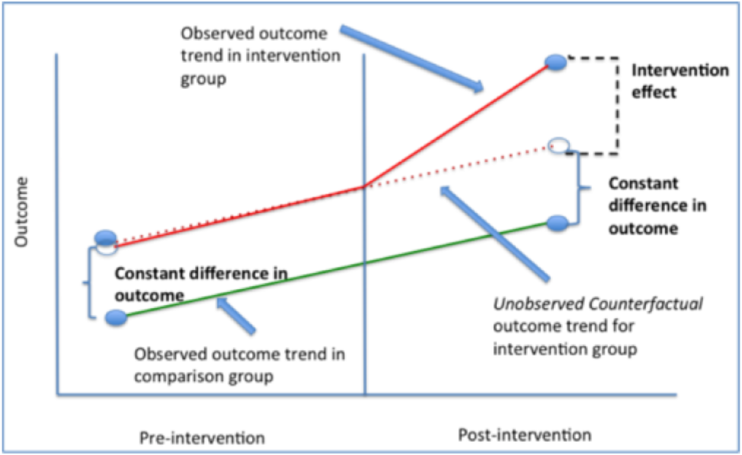
- **DID estimation :** Compare the differences in outcomes **pre and post** treatment between **treated and controls**

$$
\begin{aligned}
ATT &= E(Y_{1t_1} - Y_{0t_1} | T = 1) = E(Y_{1t1} - Y_{0t_0} + Y_{0t_0} - Y_{0t_1} | T = 1) \\
&= \underbrace{E(Y_{1t_1} - Y_{0t_0} | T = 1)}_{observed} - \underbrace{E(Y_{0t_1} - Y_{0t_0} | T = 1)}_{unobserved} \\
&= \underbrace{E(Y_{1t_1} - Y_{0t_0} | T = 1)}_{observed} - \underbrace{E(Y_{0t_1} - Y_{0t_0} | T = 0)}_{observed}
\end{aligned}
$$

→ 1^{st} **diff. (treated/controls)** eliminates the selection bias (systematic differences between the 2 groups)

→ 2^{nd} **diff. (before/after)** eliminates the time-trend bias (if similar for the 2 groups)

DID on a graph



Data for a DID framework

To use DID as a robust identification strategy to estimate causal treatment effects, basic inputs are needed :

1. You need data with a **time dimension** (panel or repeated cross-section)
2. You have **both treated and untreated** individuals, i.e. the treatment (policy, program, shock) does not affect all the sample
3. You observe individuals (at least 1 period) **before** and (at least 1 period) **after** the treatment

$\Rightarrow$ **How does DID estimation work in practice ?**

$\Rightarrow$ **What can we learn from DID estimations ?**

$\Rightarrow$ **To what extent are DID estimations valid and robust ?**

Plan

The basic DID principle

DID in regression framework

The validity of DID

The simple case

Simplest DID framework with **2 groups and 2 periods**

- The 2 groups are **treated ($T = 1$) and untreated ($T = 0$)** individuals
- The 2 periods are **before and after** the treatment :
  - $t = t_0$ before individuals are treated
  - $t = t_1$ after individuals are treated
- We can decompose the **potential outcomes** into :
  ($\simeq$ **separability hypothesis** in panel data) :

$$Y_{0it} = \gamma_i + \lambda_t + \epsilon_{0it}$$
$$Y_{1it} = Y_{0it} + \beta$$

where : γ_i is the time-invariant unobserved heterogeneity
λ_t is the time-trend
ϵ_{0it} are individual-specific shocks (with zero mean)
β is the treatment effect

- The **observed outcome** writes :

$$Y_{it} = \gamma_i + \lambda_t + \beta T_{it} + \epsilon_{0it}$$

Formalizing the DID identification assumption

- The assumption underlying DID estimation is that, **in the absence of treatment,** individual i's outcome at time t is given by :

$$E(Y_{it}|T = 0, t = \tau) = \gamma_i + \lambda_\tau$$

- There are actually **2 implicit assumptions :**
  1. Selection bias is related to **fixed characteristics** of individuals (γ_i) (the magnitude of the selection bias is stable over time)
  2. Time-trend (λ_t) is the **same for treated and controls**

$\Rightarrow$ **Parallel or common trend assumptions**

Back to the false counterfactuals

- **Treated/controls comparison** (post-treatment, i.e. at time t_1)

$$
\begin{aligned}
E(Y^{treated}_{post}) - E(Y^{controls}_{post}) &= E(Y_{1it}|T=1, t=t_1) - E(Y_{0it}|T=0, t=t_1) \\
&= [E(\gamma_i|T=1) + \beta + \lambda_{t_1}] - [E(\gamma_i|T=0) + \lambda_{t_1}] \\
&= \beta + \underbrace{E(\gamma_i|T=1) - E(\gamma_i|T=0)}_{\text{Selection bias}}
\end{aligned}
$$

$\Rightarrow$ Treatment effect is confounded with the **selection bias**

- **Pre/post comparison** (among the treated)

$$
\begin{aligned}
E(Y^{treated}_{post}) - E(Y^{treated}_{pre}) &= E(Y_{1it}|T=1, t=t_1) - E(Y_{0it}|T=1, t=t_0) \\
&= [E(\gamma_i|T=1) + \beta + \lambda_{t_1}] - [E(\gamma_i|T=1) + \lambda_{t_0}] \\
&= \beta + \underbrace{\lambda_{t_1} - \lambda_{t_0}}_{\text{Time-trend bias}}
\end{aligned}
$$

$\Rightarrow$ Treatment effect is confounded with the **time-trend**

The unbiased DID estimator

- Let's write the pre/post-treatment outcome the control group

$$E(Y_{pre}^{controls}) = E(Y_{0it}|T=0, t=t_0) = E(\gamma_i|T=0) + \lambda_{t_0}$$
$$E(Y_{post}^{controls}) = E(Y_{0it}|T=0, t=t_1) = E(\gamma_i|T=0) + \lambda_{t_1}$$

- The control group allows us to **estimate the (common) time trend !**

$$\begin{aligned} E(Y_{post}^{controls}) - E(Y_{pre}^{controls}) &= [E(\gamma_i|T=0) + \lambda_{t_1}] - [E(\gamma_i|T=0) + \lambda_{t_0}] \\ &= \lambda_{t_1} - \lambda_{t_0} \end{aligned}$$

- The **unbiased DID estimator** of the treatment effect therefore writes :

$$\begin{aligned} \beta_{DID} &= [E(Y_{post}^{treated}) - E(Y_{pre}^{treated})] - [E(Y_{post}^{controls}) - E(Y_{pre}^{controls})] \\ &= [\beta + \lambda_{t_1} - \lambda_{t_0}] - [\lambda_{t_1} - \lambda_{t_0}] \\ &= \beta \end{aligned}$$

$\Rightarrow$ **DID = ATT** (as long as the common trend assumption holds)

Note – *Treatment effect does not depend on the time period or i's characteristics (homogeneity). When treatment effect changes over time, DID estimate may depend on the choice of evaluation window.*

What is a good control group ?

- A plausible control group is likely to show time trends similar to the treated group
- An untreated group when there is **self-selection into treatment** (e.g a free-entry program) is unlikely to work...
  - → The fact that they didn't enter the program may be related to their anticipated benefits from the program
  - → The change in their (anticipated) outcomes might be different from the change in outcomes among the treated

⇒ **DID cannot solve classical self-selection bias...**

- **Unanticipated treatments** (shocks, policy reforms or interventions) usually provide the best framework
  - → Control group is defined *ex-ante*
  - → The targeting is not directly determined by the expected benefits
  - → The targeting does not depend on (macroeconomic) cycles
  - → Or the targeting is based on fixed characteristics that are not related to expected benefits (e.g. region, age, income groups)

⇒ **The best DID scenarii are provided by natural or quasi-experiments**

Some famous DID examples

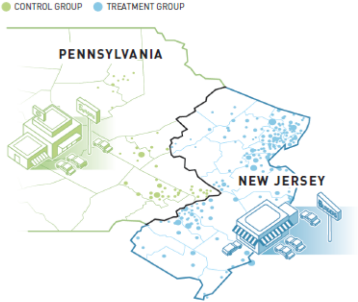
- **David Card, Joshua Angrist :** Nobel Prize in Economics 2021 (with Guido Imbens) awarded for the use of natural experiments in labor market analysis
- **Card (1990)** : effect of immigration on the labor market
  - Mariel exodus (1980) : Following riots, Fidel Castro allowed Cubans to leave the island if the wished. Nearly 125,000 people migrated to the U.S., most of them from the Mariel Bay facing the coasts of Florida and Miami.
  - Compares Miami's wage and unemployment trends to those of four other comparable cities (DID).
  - Assumption : without this migration shock, the evolution of wages and unemployment would have been similar in Miami and in the other cities.

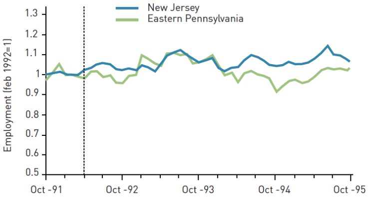
  $\rightarrow$ No effect on wages or unemployment.

Some famous DID examples

- **Card and Krueger (1993)** : impact of minimum wage increases on employment.
  - Increase in the minimum wage in the State of New Jersey : from \$4.25 to \$5.05
  - Compare the evolution of employment in NJ fast-food restaurants (unskilled employment) with that of the neighboring state of Pennsylvania.
  - Assumption : without wage increases, employment trends would have been similar in fast food restaurants in both states.

  → No effect on employment.





Some famous DID examples

- **Duflo (2001)** : Impact of school construction on education and earnings
  - Huge school construction program in Indonesia : 61,000 schools were built between 1973 and 1979
  - Compares school-age cohorts that benefited from the program to older cohorts between areas where school construction has been more/less intensive
  - Assumption : without this intensive construction, trends in schooling would have been similar in high and low building areas.

  → No effect on education and subsequent earnings.

Plan

The basic DID principle

DID in regression framework

The validity of DID

DID estimation through a simple regression

- To implement DID in a simple **linear regression** framework (with 2 groups and 2 periods), we can estimate :

$$Y_{it} = \alpha + \rho T_i + \lambda t + \beta(T_i \times t) + \epsilon_{it}$$

where : T_i is a treatment dummy and t is dummy equal to 1 if $t = t_1$
β is the treatment effect of interest
$\alpha = E(Y_{it}|T_i = 0, t = t_0)$ (pre-treatment mean in the control group)
$\rho = E(Y_{it}|T_i = 1, t = t_0) - E(Y_{it}|T_i = 0, t = t_0)$ (selection bias)
$\lambda = E(Y_{it}|T_i = 0, t = t_1) - E(Y_{it}|T_i = 0, t = t_0)$ (time-trend)

- Therefore :

$$\begin{aligned}
E(Y_{post}^{treated}) - E(Y_{pre}^{treated}) &= E(Y_{it_1} - Y_{it_0}|T = 1) = (\alpha + \rho + \lambda + \beta) - (\alpha + \rho) \\
&= \beta + \lambda \\
E(Y_{post}^{controls}) - E(Y_{pre}^{controls}) &= E(Y_{it_1} - Y_{it_0}|T = 0) = (\alpha + \lambda) - \alpha \\
&= \lambda \\
\text{DID estimator} &= E(Y_{it_1} - Y_{it_0}|T = 1) - E(Y_{it_1} - Y_{it_0}|T = 0) \\
&= \beta
\end{aligned}$$

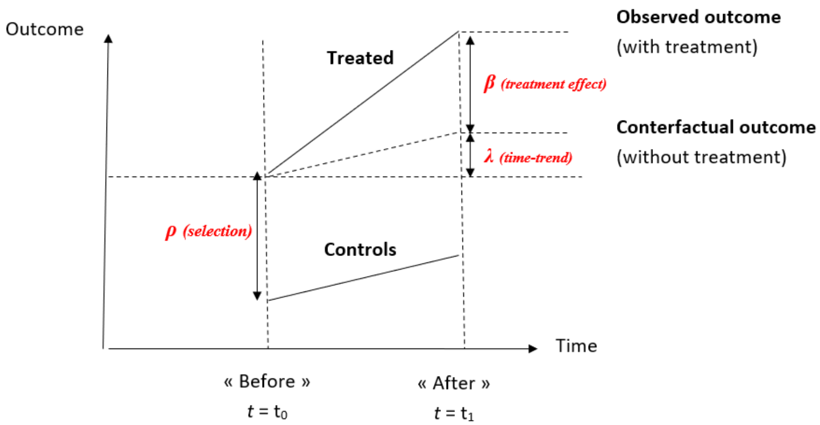
$\Rightarrow$ The variables T_i and t account for the **(time-invariant) selection bias** and the **(common) time-trend**

DID regression intuition

	Treatment	Controls
Pre-Treatment	$\bar{Y}^{treatment}_{pre}$	$\bar{Y}^{Controls}_{pre}$
Post-Treatment	$\bar{Y}^{treatment}_{post}$	$\bar{Y}^{Controls}_{post}$

- Intuitively, DID estimation is just a comparison between 4 cell-levels means
- Only one cell is treated : **Treatment** $\times$ **Post-treatment**

DID regression on a graph



DID estimation with panel-data

- The DID simple regression (with 2 groups and 2 periods) is equivalent to a **first-difference equation** :

$$\Delta Y_{it} = Y_{it_1} - Y_{it_0} = \lambda + \beta T_i + \epsilon_{it}$$

where : ΔY_{it} is the difference (change) in outcome between the 2 periods
β is the treatment effect of interest
λ is the time-trend

- With panel data, the **two-way fixed-effects regression** can be seen as a generalization of DID with multiple groups and multiple time periods :

$$Y_{it} = \eta_i + \nu_t + \beta T_{it} + \epsilon_{it}$$

where : T_{it} is a dummy equal to 1 if individual i is treated at time t
β is the treatment effect of interest
η_i is an individual fixed-effect (individual dummies)
λ_t is a time fixed-effect (period dummies)

⇒ The fixed-effects η_i and λ_t account for the **(time-invariant) unobserved heterogeneity** (selection bias) and the **(common) cyclical effects** (time-trend).

⇒ Identification comes from **within-group** time variation

Assumptions and warnings with panel-data DID

- **Identification assumption :** Treatment T_i is strictly **exogenous**, i.e uncorrelated to individual-specific shocks ϵ_{it} :

$$cov(\epsilon_{it}, T_{it}) = 0 \quad \forall t$$
$$cov(\epsilon_{it}, \lambda_t) = 0 \quad \forall t$$
$$cov(\epsilon_{it}, \eta_i) = 0 \quad \forall t$$

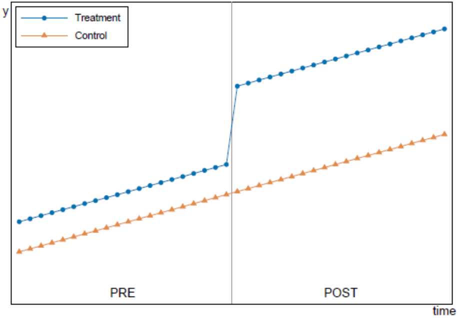
$\Rightarrow$ Selection into treatment does not depend on **past and/or future shocks** on the outcome (expected benefits)

- **Fixed-effects regression pros :** we can add time-varying observable characteristics X_{it} to increase precision and/or if the common trend assumption holds conditional on X_{it} and/or other shocks over the period :

$$Y_{it} = \eta_i + \nu_t + \phi X_{it} + \beta T_{it} + \epsilon_{it}$$

- **Fixed-effects regression cons :**
  - Usual panel-data issues with standard-errors = **serial correlation** (use panel-robust clustered standard-errors)
  - Treatment effect might **vary over time**
  - Treatment timing might **vary across groups**

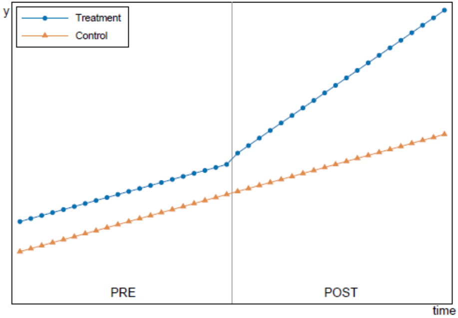
DID with a constant treatment effect over time



- Treatment effect **does not depend** on the time period

⇒ DID estimate **does not depend** on the choice of the time window

DID with varying treatment effects over time



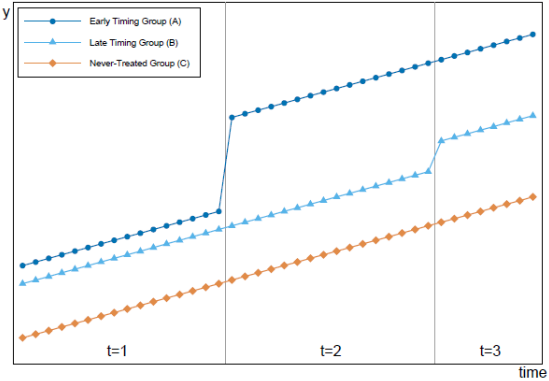
- Treatment effect **does depend** on the time period

⇒ DID estimate **does depend** on the choice of the time window

- We can estimate a fixed-effects regression with an **event study framework** (i.e. dummies for each post treatment period) :

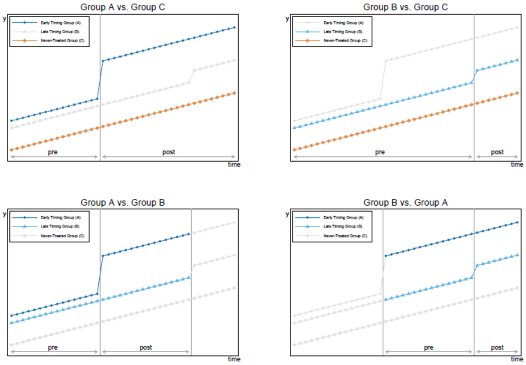
$$Y_{it} = \eta_i + \nu_t + \phi X_{it} + \sum_{l=1}^{k} \beta_l T_{it+l} + \epsilon_{it}$$

DID with varying treatment timing across groups



- With 3 timing groups (one of which is never treated), we can decompose 3 timing windows ($t = 1, 2, 3$)

Decomposition into 2 $\times$2 standard DID



- When treatment effect is **homogenous**, β_{DID} is the ATT
- When treatment effect is **heterogenous (across units)**, β_{DID} is a weighted-average of each ATT
- When treatment effect **changes over time**, we need to mix with an event study framework (complex)

Note – *See Roth & alii (2022) for recent DID developments*

DID estimation with repeated cross-sections

- **Aggregated data on multiple groups** observed at different periods may be enough **("repeated cross-sections or "pseudo-panel")** if :
  - Only some groups are treated (group-level treatment)
  - Cyclical effects are similar across groups (common trend)
  - Group composition is stable over time
  - Representative data for each group is available
- **Two-way fixed-effects group regression :**

$$Y_{gt} = \eta_g + \nu_t + \phi X_{gt} + \beta T_{gt} + \epsilon_{gt}$$

where : T_{gt} is a dummy equal to 1 if group g is treated at time t
β is the treatment effect of interest
η_g is a group fixed-effect (group dummies)
λ_t is a time fixed-effect (period dummies)

- Panel data (if available) bring additional information and allow to better control for **individual heterogeneity :**

$$Y_{it} = \eta_i + \nu_t + \phi X_{it} + \beta T_{it} + \epsilon_{it}$$

- Careful with the standard-errors : **serial + within-group correlation** (use panel-robust clustered standard errors or (better) block bootstrap)

Note – *See Menéndez & Gignoux (2016) and Bertand & alii (2004) for more details on inference*

Plan

How convincing are DID estimations ?

- DID does not identify the treatment effect if treated and controls were on different trajectories prior to the treatment (and then possibly after...)

⇒ **Fundamental common trend assumption**

- This assumption **cannot be tested** (it would imply that we can observe the counterfactual evolution of the treated group outcome in the absence of treatment)

⇒ **You need to defend the common trend !**

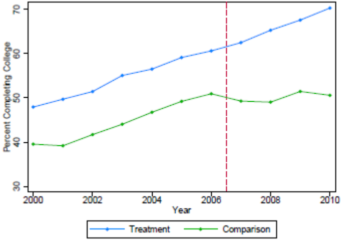
- Some tests allow to **judge its plausibility :**
    1. Check that treated and controls have **similar observable characteristics** (balance check)
    2. Control for **unbalanced observable characteristics** between treated and controls (common trend assumption conditional on X)
    3. Compare the **trends in past outcomes** (if more than 2 periods)
    4. Run **falsification or "placebo" tests** (if more than 2 periods)

- **If it is not enough... :**
    1. Run **triple differences** (if several control groups)
    2. Combine **DID with matching** (if selection on observables)

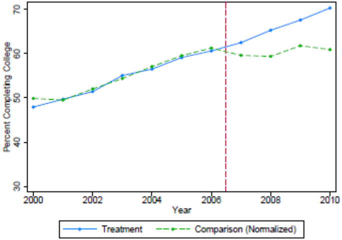
Comparison of past income trends

- We can check that **over the past periods (before treatment)**, treated and controls have indeed experienced similar evolutions
- Requires data for at least **two periods before treatment** occurred
- A **graphical examination of observed outcome trends** before treatment can provide insight into the credibility of the identification assumption.

Valid common trend assumption

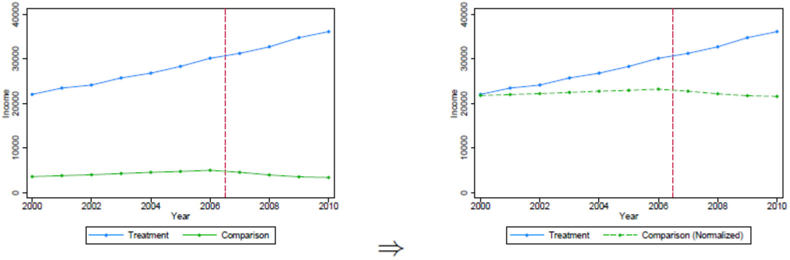


⇒



Sometimes, the common trend assumption is clearly OK !

Invalid common trend assumption



Sometimes, the common trend assumption clearly fails...

The famous Ashenfelter's Dip

- **A textbook case** where the assumption is not verified (Ashenfelter, 1978).
- Impact of training on the income of low wage earners.
- Detailed analysis of **earnings changes prior to training** shows that in the year prior to training, trained workers had lower earnings growth than before training

→ Likely a **transitory shock to become eligible for the program**

- Not clear whether the effect on wages that we observe is the impact of training or the impact of "catching up" to an average growth path after a negative shock.

Falsification or "placebo" tests

- The same DID estimation can be applied around a **date when nothing happened :**
  - ⇒ If the effect is **significant :** there is a problem, we can fear that there is a significant difference in trajectories between the treated and controls
  - ⇒ If the impact is **not significant :** this is a strong indicator of parallel trajectories.

- In practice, you can estimate a DID regression within an event study framework but adding **"future" treatment dummies** (i.e dummies for each pre- et post-treatment period) :

$$Y_{it} = \eta_i + \nu_t + \phi X_{it} + \sum_{l=0}^{j} \alpha_j T_{it-j} + \sum_{l=1}^{k} \beta_l T_{it+l} + \epsilon_{it}$$

⇒ We hope the α_j not to be significant

- You can also **replace** Y **by an alternative outcome** Y' that is not supposed to be affected by the treatment and hope to find a zero treatment effect

What can we do if the common trend assumption is violated ?

- In the recent literature, 2 ways to introduce more flexibility
  1. Run **triple differences** (if several control groups)
  2. Combine **DID with matching** (if selection on observables)

Note – *See Roth & alii (2022) for recent DID developments*

Triple differences

- If there are **several possible control groups**, you can check that the treatment effect is similar according to whether you use alternative control groups

- The treatment effect can also be estimated with **triple differences**

⇒ **The DDD estimator** is the difference between the DID of interest and the placebo DID (that is supposed to be 0, but it isn't...)

Example – The impact of a labor market integration program targeted at unemployed people over 50 in a region :

- 2 possible control groups :
  - A group of younger unemployed people from the same region
  - A group of unemployed people of the same age but from a neighboring region
- These groups can be combined into a **triple differences** : we subtract from the before/after evolution of the treated (1), the before/after evolution of the younger unemployed people from the same region (2), and the before/after evolution of the unemployed people from the neighboring region (3)

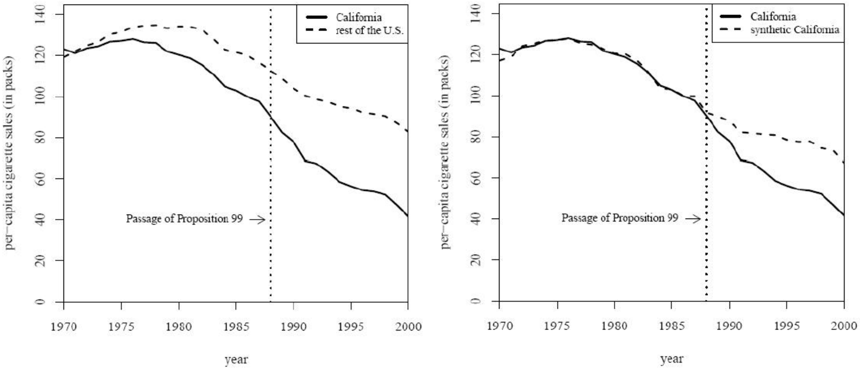
DID combined with matching

- Abadie (2005) suggests a **semi-parametric DID estimation method** that matches observations on observable characteristics X.
- The identification assumption is transformed into a conditional assumption : conditional on observed characteristics X, treated and controls individuals evolve on parallel trends
- This involves comparing **previously matched individuals** according to their observable characteristics or according to a **propensity score**

$$E(Y_{1t_1} - Y_{0t_1}|T = 1) = E(Y_{1t_1} - Y_{0t_0}|T = 1) - E[(Y_{0t_1} - Y_{0t_0})\frac{P(D = 1|X)P(D = 0)}{P(D = 0|X)P(D = 1)}|T = 0]$$

- Abadie & alii (2010) alternatively suggest to build a **synthetic control group** (weighted-average of all untreated units where weights minimize the distance with the treated unit)

Synthetic control graph



To sum up

Steps for implementing DID estimations :

- **Find consistent data pre and post-treatment**
  - Longitudinal data (preferably panel data).
  - Data should be relatively close in time, to avoid capturing other changes that would not be related to the treatment.
- **Select a control group :**
  - It will only be "credible" if you can assume that the changes in outcome in this control group are the same as those that would have occurred in the treated group in the absence of treatment.
  - Caution, it cannot be made up of untreated units if there is self-selection, because selection into treatment is then likely be related to expected benefits (trajectories)
- **Estimate the treatment effect** with non-parametric methods (i.e. differences in means) or through a fixed-effects regression, but :
  - Correctly computing the standard-errors of the estimate
  - When using individual data, take into account the serial and/or the within-group correlation between individual-specific shocks that bias the estimation of the variance
  - Bertrand, Duflo & Mullainathan (2004) show that this can have a significant impact on the results (significance). Use bootstrap procedures.

References

Abadie (2005) “Semiparametric difference-in-differences estimators” The Review of Economic Studies, 72(1), 1-19

Abadie, Diamond, Hainmueller (2010) “Synthetic control methods for comparative case studies : Estimating the effect of California’s tobacco control program” Journal of the American statistical Association, 105(490), 493-505.

Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. The Review of Economics and Statistics, 47-57.

Bertrand, Duflo, Mullainathan (2004) “How much should we trust differences-in-differences estimates ?” The Quarterly Journal of Economics, 119(1), 249-275.

Card (1990) “The impact of the Mariel boatlift on the Miami labor market” ILR Review, 43(2), 245-257.

Card, Krueger (1993) “Minimum wages and employment : A case study of the fast food industry in New Jersey and Pennsylvania” NBER working paper.

Gignoux, Menéndez (2016) “Benefit in the wake of disaster : long-run effects of earthquakes on welfare in rural Indonesia” Journal of Development Economics, 118, 26-44.

Roth, Sant’Anna, Bilinski, & Poe, J. (2022). “What’s Trending in Difference-in-Differences ? A Synthesis of the Recent Econometrics Literature” arXiv preprint arXiv :2201.01194.