Notes on Lab Session 5

Giulia Tani
https://tanigiulia.github.io/

TSE - MSc course in Program Evaluation
February 2026

Setting: French Enterprise Zone Program

- Program initiated in 1997 to encourage employment in disadvantaged areas.
- The program offered significant wage-tax exemptions to firms hiring at least 20% of their labor force locally.
- We have a **panel** dataset covering 271 municipalities in the Paris region from 1993 to 2003. The program was implemented only in 13 municipalities, selected on the basis of previous unemployment, youth share, fiscal potential, etc.
- Theory predicts two possible effects of the program on local unemployment. As the program makes employment of local residents less costly:
  1) unemployed people should find a job more easily
     $\longrightarrow$ $\uparrow$ probability to exit unemployment,
  2) non-employed people become more hopeful to find a job, so they start looking for one
     $\longrightarrow$ $\uparrow$ probability to enter unemployment.
- Our question: do enterprise zones help locals to find a job?
  $\longrightarrow$ Treatment: participation to the program. Outcome: probability to exit unemployment.

Example: Bondy (treated) vs Le Raincy (untreated)

Example

- Suppose there are only 2 municipalities $R \in \{0, 1\}$ and 2 periods $t \in \{0, 1\}$.
  Treatment occurs only in municipality $R = 1$ at time $t = 1$.

- The probability of exit from unemployment is the following:

	1996	1998
Bondy ($R = 1$)	0.10	0.16
Le Raincy ($R = 0$)	0.12	0.15

- First approach: compare outcomes before vs after treatment.
  But: other things might be changing between 1996 and 1998 (e.g. national economic shock, changes in demographics, housing...).

- Second approach: compare outcomes in treated vs control municipality.
  But: municipalities might differ for reasons unrelated to treatment (e.g. education, firm density, public services...).

- **Solution**: use **both** dimensions to identify the effect.

Principle of difference-in-differences

- Assumptions:
  - No migration in/out each municipality.
  - **Equal (or parallel) trend**:
    In the absence of treatment the evolution of potential outcomes would have been the same in both municipalities:

$$E(Y_1(0) - Y_0(0) \mid R = 0) = E(Y_1(0) - Y_0(0) \mid R = 1)$$

    . Equivalently, differences across municipalities are kept constant over time:

$$E(Y_1(0) \mid R = 1) - E(Y_1(0) \mid R = 0) = E(Y_0(0) \mid R = 1) - E(Y_0(0) \mid R = 0)$$

- The average treatment on the treated $ATT = E(Y_1(1) - Y_1(0) \mid R = 1)$ can be written as:

$$ATT = E(Y_1 - Y_0 \mid R = 1) - E(Y_1 - Y_0 \mid R = 0) = \text{diff-in-diff}$$

Example

- We use as outcome Y the logarithm of the probability of exit from unemployment:

	1996	1998	Δ (1998–1996)
Bondy ($R = 1$)	-2.3026	-1.8326	0.4700
Le Raincy ($R = 0$)	-2.1203	-1.8971	0.2231
Δ (Bondy – Le Raincy)	-0.1823	0.0645	0.2469

- When the outcome is in logs, a change in the outcome corresponds to a proportional change in the original probability:

$$\Delta \log(p) = \log(p_2) - \log(p_1) = \log\left(\frac{p_2}{p_1}\right) \quad \Longrightarrow \quad \exp\big(\Delta \log(p)\big) = \frac{p_2}{p_1}.$$

So a DiD estimate of 0.2469 in logs implies a relative increase of $\exp(0.2469) - 1 \approx 0.28$.

$\longrightarrow$ The program increased the probability of leaving unemployment by approximately 28% relative to the counterfactual trend captured by the control municipality.

- Parallel trends assumption: in the absence of the program, Bondy's exit-to-job probability would have evolved like Le Raincy's, even if Bondy started at a different level.

Linear model

- In general, we can write the observed outcome as

$$Y_{t,m} = \alpha + \gamma\ R_m + \delta\ 1(t \geq t^*) + \beta\ d_{m,t} + \epsilon_{m,t}, \qquad E[\epsilon_{m,t} \mid R_m, t] = 0$$

where the treatment occurs at time t^* in the municipalities where $R_m = 1$ and it is defined as

$$d_{m,t} = R_m\ 1(t \geq t^*)$$

- The conditional mean outcomes $E[Y_{m,t} \mid R_m, t]$ are:

	$t < t^*$	$t \geq t^*$	Δ
$R_m = 1$	$\alpha + \gamma$	$\alpha + \gamma + \delta + \beta$	$\delta + \beta$
$R_m = 0$	α	$\alpha + \delta$	δ
Δ	γ	$\gamma + \beta$	β

- The coefficient β captures the ATT.