

Experimental Analysis of Public Policies

Session 5

Difference-in-Differences: 2x2 and TWFE

Etienne Dagorn - Universite de Lille

Session Objectives

By the end of this session, students should be able to:

- explain why before-after and post-treatment level comparisons can fail;
- compute a 2x2 DiD estimator from group means;
- label the 2x2 estimand as an ATT for the treated group in the post period;
- state parallel trends in plain language and potential-outcomes language;
- write and interpret the basic DiD regression;
- explain why serial correlation motivates clustered standard errors;
- use pre-trends, placebo outcomes, and anticipation checks as diagnostics;
- know when simple TWFE has a clear interpretation.

Roadmap

Policy Timing And Failed Comparisons

The 2x2 Double Difference

Regression And Simple TWFE

Diagnostics And Inference

Practice And Bridge

Roadmap

Policy Timing And Failed Comparisons

The 2x2 Double Difference

Regression And Simple TWFE

Diagnostics And Inference

Practice And Bridge

From Instruments To Policy Timing

- Session 4 used an instrument Z to create exogenous variation in treatment received D .
- Some policy questions have no credible instrument at the individual level.
- They do have variation in when and where a policy begins.
- DiD asks whether untreated units reveal the missing trend for treated units.

Central question

Can untreated trends reveal the missing counterfactual trend?

Policy Example

- A region introduces a youth employment subsidy in 2025.
- Other similar regions do not introduce it.
- Youth employment rises in the treated region after 2025.
- Is the rise due to the subsidy or to a general recovery?

Counterfactual question

How would youth employment in the treated region have changed after 2025 without the subsidy?

Failed Comparison 1: Before And After

$$\bar{Y}_{T,1} - \bar{Y}_{T,0}$$

- This uses the treated group as its own comparison.
- It is tempting because it follows the same region over time.
- It fails if employment would have changed anyway because of recovery, recession, seasonality, or demographic change.

What is missing?

The treated region's untreated post-policy outcome, not just its pre-policy outcome.

Failed Comparison 2: Post-Treatment Levels

$$\bar{Y}_{T,1} - \bar{Y}_{C,1}$$

- This compares treated and untreated regions after the subsidy starts.
- It fails if treated and control regions had different baseline levels before the policy.
- A poorer treated region can still look worse after a successful policy.

What is missing?

The permanent treated-control difference that existed before treatment.

DiD Workshop: Build The Counterfactual

Scenario: a region introduces a youth employment subsidy in 2025.

1. Draw the treated-region employment path before and after 2025.
2. Draw the control-region path.
3. Use the control change to project the treated region's untreated post-2025 outcome.
4. State the exact trend assumption needed.

Debrief

The control group does not need the same level as the treated group. It must provide the untreated trend for the treated group.

Roadmap

Policy Timing And Failed Comparisons

The 2x2 Double Difference

Regression And Simple TWFE

Diagnostics And Inference

Practice And Bridge

The 2x2 Setup

	Before	After
Treated group	$\bar{Y}_{T,0}$	$\bar{Y}_{T,1}$
Control group	$\bar{Y}_{C,0}$	$\bar{Y}_{C,1}$

- One treated group, one control group.
- One pre-treatment period, one post-treatment period.
- Treatment starts between the two periods.
- In this 2x2 setting, treatment received is $D_{it} = G_i \times Post_t$.

Table Computation

	Before	After	Change
Treated	50	58	$58 - 50 = 8$
Control	55	59	$59 - 55 = 4$
DiD			$8 - 4 = 4$

Interpretation

Under parallel trends, the estimate identifies the ATT for the treated group in the post period.

The DiD Estimator

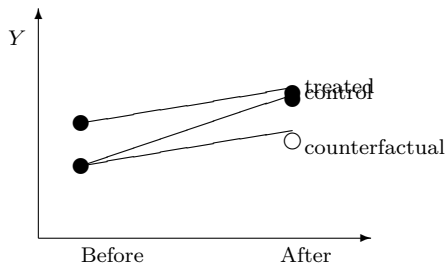
$$\hat{\tau}_{DiD} = (\bar{Y}_{T,1} - \bar{Y}_{T,0}) - (\bar{Y}_{C,1} - \bar{Y}_{C,0})$$

- Treated change: what happened to the treated group.
- Control change: the proposed untreated counterfactual trend.
- Double difference: the extra treated change after removing the common trend.

Estimand

The usual 2x2 causal target is $ATT_1 = E[Y_{i1}(1) - Y_{i1}(0) \mid G_i = 1]$.

Graphical Counterfactual



- The observed treated post outcome is factual.
- The open point is the missing untreated outcome for treated units.
- DiD is the vertical distance between those two post-period points.

Parallel Trends

Parallel trends

In the absence of treatment, the treated and control groups would have experienced the same average change in outcomes.

$$E[Y_{i1}(0) - Y_{i0}(0) \mid G_i = 1] = E[Y_{i1}(0) - Y_{i0}(0) \mid G_i = 0]$$

- Groups can have different levels.
- What matters is the untreated counterfactual trend.
- Pre-treatment trends are evidence about credibility, not the assumption itself.

Levels Can Differ

- Treated regions may be poorer than control regions.
- Treated schools may have lower baseline scores.
- DiD can allow fixed level differences.
- It cannot allow different untreated trends unless those trends are explicitly modelled and justified.

Shortcut

DiD removes permanent group differences and common shocks, not group-specific shocks.

Card And Krueger As A Design Audit

- Empirical question: did a higher minimum wage reduce fast-food employment?
- Treated group: restaurants in the state raising the minimum wage.
- Comparison group: similar restaurants in a nearby state without the increase.
- Identifying variation: one state changes policy while the other does not.
- Assumption: absent the policy, employment trends would have been comparable.

Student output

Name the treated group, comparison group, treatment date, estimand, and most credible threat to parallel trends.

Roadmap

Policy Timing And Failed Comparisons

The 2x2 Double Difference

Regression And Simple TWFE

Diagnostics And Inference

Practice And Bridge

Regression Specification

$$Y_{it} = \alpha + \beta G_i + \lambda Post_t + \tau(G_i \times Post_t) + u_{it}$$

- G_i : treated-group indicator.
- $Post_t$: post-treatment period indicator.
- $G_i \times Post_t = D_{it}$: treatment received in the 2x2 design.
- τ : the same DiD contrast computed from the table.

Interpretation

The interaction coefficient is causal only if the control trend identifies the treated group's untreated counterfactual trend.

Fixed Effects In The Simple Panel

$$Y_{it} = \alpha_i + \lambda_t + \tau D_{it} + u_{it}$$

- Unit fixed effects α_i absorb time-invariant differences across units.
- Time fixed effects λ_t absorb shocks common to all units in a period.
- Identification comes from within-unit changes relative to common calendar changes.
- Fixed effects do not absorb a shock that affects only treated units when treatment starts.

Adding Controls

$$Y_{it} = \alpha_i + \lambda_t + \tau D_{it} + X'_{it}\gamma + u_{it}$$

- Pre-treatment covariates can improve precision or define a conditional comparison.
- Clearly exogenous time-varying controls may support conditional parallel trends.
- Controls affected by treatment can change the estimand or bias the estimate.
- Controls do not replace the need to defend parallel trends.

TWFE In Simple Settings

$$Y_{it} = \alpha_i + \lambda_t + \tau D_{it} + u_{it}$$

- With one treatment date and valid untreated controls, TWFE has a transparent DiD interpretation.
- No anticipation, no interference, and stable group composition are still needed.
- Homogeneous effects make one coefficient easier to summarise, but they are not the only issue.
- With staggered timing and heterogeneous effects, Session 6 audits the comparison structure.

When TWFE Is Comfortable

- Treatment timing is simple.
- There are clear untreated comparison units.
- The post-period ATT target is clear.
- Treatment effects are plausibly stable or the reported average is policy-relevant.
- Event-study diagnostics do not show strong pre-trends.

Boundary

This is the comfortable case. Session 6 explains why staggered timing can break this simplicity.

Roadmap

Policy Timing And Failed Comparisons

The 2x2 Double Difference

Regression And Simple TWFE

Diagnostics And Inference

Practice And Bridge

Standard Errors And Serial Correlation

Inference problem

Outcomes close in time are often correlated. Ignoring this can make standard errors too small.

- Employment rates, school scores, and regional shocks are persistent.
- Policy treatment often varies at the group level.
- Cluster standard errors at the level of treatment assignment or panel dependence.
- With few clusters, inference is fragile even with many individual observations.

Pre-Trends

- Multiple pre-treatment periods allow a visual pre-trend check.
- Similar pre-trends support parallel trends but do not prove it.
- Diverging pre-trends weaken credibility.
- Very noisy pre-trends may be uninformative.

Exam phrase

The pre-trend evidence is supportive but cannot rule out a treated-specific shock after treatment.

Parallel Trends: Debrief Grid

Evidence

How to read it

Similar pre-trends

supportive but not a proof

Diverging pre-trends

serious threat to identification

Policy announced
early

possible anticipation

Treated region hit by
shock

parallel trends likely violated

Reflex

Ask whether the treated group would have followed the control trend without the policy, not whether the groups are identical.

Placebo Checks

- Pretend treatment happened before it actually did.
- Test outcomes that should not respond to the policy.
- Use groups that should not be affected.
- Placebos are diagnostic evidence, not automatic validation.

Good placebo

Choose a false timing, group, or outcome that would reveal the most plausible violation of parallel trends.

Threat	Question to ask
Different trends	Why would untreated trends be parallel?
Anticipation	Did behaviour change before treatment?
Spillovers	Did controls get indirectly treated?
Composition	Did the groups change over time?

Student output

Pick the most credible threat in the subsidy example and propose one diagnostic or design response.

Roadmap

Policy Timing And Failed Comparisons

The 2x2 Double Difference

Regression And Simple TWFE

Diagnostics And Inference

Practice And Bridge

Applied Reading Questions

- What is the treated group and treatment date?
- What comparison group is used?
- What ATT or treated-period target is being estimated?
- What is the evidence for parallel trends?
- What standard errors are reported?
- What threat remains most credible?

Code Mini-Lab

- Build a two-period panel.
- Compute DiD by hand from group means.
- Estimate the interaction regression.
- Plot treated and control means over time.
- Verify that the hand-computed DiD equals the interaction coefficient in the simple 2x2 case.

Exercise Bridge

- Theory sheet: `exercises/theory/session_05_theory.md`.
- QCM: `assessments/qcm/session_05_qcm.md`.
- Applied exercise: compute and critique a 2x2 DiD.
- Exam skill: defend or reject parallel trends in a policy scenario.

Exercise rhythm

Compute first, interpret second, then diagnose the assumption.

Transition To Session 6

- Today one DiD coefficient summarised one before-after contrast.
- Many policies have effects that appear slowly, fade, or change as institutions adapt.
- Many policies are adopted by different units in different years.

Bridge question

What changes when treatment effects evolve and units adopt at different times?

Design Checklist

1. Who is treated and when?
2. Who is the comparison group?
3. What ATT or treated-period estimand is targeted?
4. What is the identifying trend assumption?
5. What diagnostics support it?
6. What inference level is appropriate?
7. What remaining shock could break the design?

Key Takeaways

- DiD compares changes, not levels.
- The clean 2x2 estimand is usually the ATT for treated units in the post period.
- Parallel trends is the core identifying assumption.
- Fixed effects remove permanent unit differences and common time shocks.
- Pre-trends and placebos help diagnose credibility but do not prove identification.
- Standard errors must respect the panel and treatment structure.

Appendix: Equivalent 2x2 Regression

$$Y_{it} = \alpha + \beta G_i + \lambda Post_t + \tau(G_i \times Post_t) + u_{it}$$

- The interaction coefficient equals the 2x2 DiD.
- Adding fixed effects changes notation but not the core design logic.
- The estimate is causal only under parallel trends.

References and Further Reading

- Card and Krueger (1994), “Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania”.
- Bertrand, Duflo, and Mullainathan (2004), “How Much Should We Trust Differences-in-Differences Estimates?”.
- Angrist and Pischke (2009), *Mostly Harmless Econometrics*, DiD chapter.
- Cunningham (2021), *Causal Inference: The Mixtape*, DiD chapter.

Reading reflex

Before reading the table, find the treated group, comparison group, treatment date, pre-trend evidence, and clustering level.