Experimental Analysis of Public Policies

Session 4

Compliance, IV, LATE, and External Validity

Etienne Dagorn - Universite de Lille

Session Objectives

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

- → distinguish assignment Z_i from treatment received D_i;
- → compute ITT, first stage, and the Wald estimator;
- → define always-takers, never-takers, compliers, and defiers;
- → interpret LATE as an effect for compliers;
- → state relevance, independence, exclusion, and monotonicity;
- → diagnose weak instruments and exclusion restrictions;
- → discuss why IV estimates can be internally credible but externally local.

Roadmap

Assignment Versus Receipt

Who Is Moved by the Instrument?

IV Assumptions and Threats

Estimation, Inference, and External Validity

Roadmap

Assignment Versus Receipt

Who Is Moved by the Instrument?

IV Assumptions and Threats

Estimation, Inference, and External Validity

From RCT Design to Compliance

→ Session 3 studied credible assignment.

→ Many programmes randomise access, encouragement, or eligibility.

→ Actual take-up may differ from assignment.

→ We now distinguish the effect of being assigned from the effect of receiving treatment.

Central question

What can we learn when assignment changes treatment probability but does not perfectly determine treatment?

Assignment vs Receipt

$$Z_i = \text{assignment or instrument}, \qquad D_i = \text{treatment received}$$

- $\rightarrow$ Z_i is controlled or plausibly exogenous.
- $\rightarrow$ D_i may be chosen by the unit or constrained by implementation.
- $\rightarrow$ Outcomes respond to D_i, but D_i is not randomly chosen.

Notation reflex: when $Z_i \neq D_i$, never call the assigned group simply "the treated group" without clarifying the margin.

Encouragement Example

- → A job centre randomly sends invitation letters to a training programme.
- → Some invited workers do not attend.
- → Some non-invited workers find another way to attend.
- → The invitation is random; attendance is not.

Interpretation: the letter is a candidate instrument: it is randomly assigned and shifts attendance, but exclusion and monotonicity still need defending.

Compliance Workshop: What Is the Policy Lever?

Scenario: a city randomly sends childcare-voucher letters to eligible parents.

1. Define Z_i, D_i, and Y_i.
2. Which effect does the city directly control: offer or childcare use?
3. Why might D_i remain endogenous?
4. Who are the likely compliers?

Debrief

The first policy-relevant estimate is often the ITT: the effect of offering the policy as actually implemented.

Why this matters: policymakers can usually send letters, open slots, or change eligibility; they cannot force every eligible person to comply.

Roadmap

Assignment Versus Receipt

Who Is Moved by the Instrument?

IV Assumptions and Threats

Estimation, Inference, and External Validity

Compliance Types

Type	$D_i(1)$	$D_i(0)$	**Behaviour**
Always-takers	1	1	treated regardless of Z
Never-takers	0	0	untreated regardless of Z
Compliers	1	0	treated only when $Z = 1$
Defiers	0	1	move opposite to Z

Before any ratio

Classify which group the instrument actually moves.

Four Useful Quantities

Quantity	Meaning
ITT / reduced form	effect of assignment on outcome
First stage	effect of assignment on treatment received
TOT	effect for actual treatment recipients; not generally identified by assignment alone
LATE	effect for compliers moved by the instrument

TOT Is Not Automatically LATE

$$TOT = E[Y_i(1) - Y_i(0) \mid D_i = 1]$$

- → TOT is the average effect for actual treatment recipients.
- → LATE is the average effect for compliers, $D_i(1) > D_i(0)$.
- → With two-sided non-compliance, actual recipients include always-takers as well as compliers.
- → In one-sided non-compliance, a treatment-on-the-treated reading may be defensible for the induced recipients.

Interpretive rule: do not call a Wald estimate "the effect on the treated" unless the compliance structure justifies it.

Intent-to-Treat

$$ITT_Y = E[Y_i \mid Z_i = 1] - E[Y_i \mid Z_i = 0]$$

- → Under random assignment, this assigned-arm contrast identifies the causal effect of being assigned or offered treatment.
- → It is policy-relevant when policymakers control offers, not take-up.
- → It includes noncompliance and implementation frictions.

Interpretation: a small ITT may reflect a weak treatment effect, low take-up, or both.

First Stage

$$ITT_D = E[D_i \mid Z_i = 1] - E[D_i \mid Z_i = 0]$$

- → First stage measures how much assignment changes treatment received.
- → If the first stage is zero, assignment cannot identify a treatment-received effect.
- → Weak first stages produce imprecise and fragile IV estimates.

Reporting habit: always show the first-stage coefficient in treatment-receipt units that are easy to understand, such as percentage-point take-up.

Wald Estimand and Estimator

$$LATE = \frac{E[Y_i \mid Z_i = 1] - E[Y_i \mid Z_i = 0]}{E[D_i \mid Z_i = 1] - E[D_i \mid Z_i = 0]}$$

$$\widehat{LATE} = \frac{\bar{Y}_{Z=1} - \bar{Y}_{Z=0}}{\bar{D}_{Z=1} - \bar{D}_{Z=0}}$$

→ The first line is the population identifying contrast under IV assumptions.

→ The second line is the sample analogue.

→ The ratio is the outcome effect per induced treatment.

Logic: the Wald ratio rescales the offer effect by the share of people whose treatment status actually changed.

Numerical Example

→ Invitation raises training attendance from 20% to 60%.

→ Invitation raises employment from 45% to 53%.

→ First stage: $0.60 - 0.20 = 0.40$.

→ ITT: $0.53 - 0.45 = 0.08$.

→ Wald estimate: $0.08/0.40 = 0.20$.

Interpretation

Employment rises by 20 percentage points for compliers, under IV assumptions.

Check: the estimate is not 8 percentage points because only 40 percentage points of the sample were induced into treatment.

Local Average Treatment Effect

LATE

The IV estimand is the average treatment effect for compliers: units whose treatment status changes because of the instrument.

→ LATE is local to a behavioural margin.

→ It may differ from ATE or ATT.

→ It is often exactly the group policy can move.

Plain words: IV tells us about people persuaded by the instrument, not people who would always or never take treatment.

Roadmap

Assignment Versus Receipt

Who Is Moved by the Instrument?

IV Assumptions and Threats

Estimation, Inference, and External Validity

Assumption 1: Relevance

Relevance

The instrument must change the probability of treatment:

$$E[D_i \mid Z_i = 1] \neq E[D_i \mid Z_i = 0].$$

→ Check the first stage.

→ Report its magnitude, not just significance.

→ Ask whether the instrument moves the treatment meaningfully.

Assumption 2: Independence

Independence

The instrument is as good as randomly assigned with respect to potential outcomes and potential treatment behaviour:

$$Z_i \perp \{Y_i(1), Y_i(0), D_i(1), D_i(0)\}.$$

→ Random encouragement can satisfy this by design.

→ Natural instruments require institutional justification.

→ Balance checks can support but not prove the assumption.

Assumption 3: Exclusion

Exclusion restriction

The instrument affects the outcome only through treatment received.

- → Invitation letters may affect motivation even without attendance.
- → Distance to college may affect local labour markets, not only schooling.
- → Exclusion is often the most fragile IV assumption.

Diagnostic habit: write the most credible direct path from Z to Y that does not use D; then ask whether the paper rules it out.

Assumption 4: Monotonicity

No defiers

The instrument moves treatment in one direction for every unit.

→ No one attends only because they were not invited.

→ Plausibility depends on the institution and behavioural response.

→ Monotonicity is what makes the local group interpretable as compliers.

IV Assumptions: Debrief Grid

Assumption	Question for the empirical setting
Relevance	Does the instrument visibly move treatment received?
Independence	Who or what determines the instrument?
Exclusion	Could the instrument affect outcomes without treatment?
Monotonicity	Is it plausible that nobody moves opposite to the instrument?

Most common exam mistake

Reporting a Wald ratio without saying which assumption is most fragile.

When LATE Is Useful

- → The policy lever is an encouragement or eligibility rule.
- → The complier group is policy-relevant.
- → Full compliance is unrealistic.
- → The instrument is transparent enough to defend.

Limit

LATE does not automatically describe always-takers, never-takers, or a future universal rollout.

Weak First Stages

- → A weak instrument barely changes treatment received.
- → The Wald ratio divides by a small number.
- → Estimates become unstable and standard errors can become large.
- → Weak instruments can also bias IV toward OLS and make conventional t-tests unreliable.
- → A strong theoretical instrument still needs an empirical first stage.

Exam phrase: "The exclusion story may be plausible, but the instrument has too little predictive power to be useful."

Exclusion Threats

Instrument	**Possible exclusion problem**
Training invitation	information or motivation effect
Distance to school	neighborhood differences
Judge assignment	courtroom practices beyond sentencing
Eligibility threshold	direct benefit of eligibility status

Roadmap

Assignment Versus Receipt

Who Is Moved by the Instrument?

IV Assumptions and Threats

Estimation, Inference, and External Validity

Two-Stage Least Squares

$$D_i = \pi_0 + \pi_1 Z_i + \eta_i$$
$$Y_i = \alpha + \tau \widehat{D}_i + \varepsilon_i$$

- → First stage predicts treatment using the instrument.
- → Second stage relates outcomes to instrument-induced treatment variation.
- → The causal interpretation still comes from IV assumptions.

Implementation warning: use an IV/2SLS estimator command for correct standard errors; these equations explain the logic, not a recipe for naive two-step OLS inference.

Multiple Instruments

- → Multiple instruments can improve precision.
- → They may identify a weighted average of different complier groups.
- → Overidentification tests do not prove validity.
- → Interpretability matters as much as statistical convenience.

Example: distance, eligibility, and encouragement may each move different people into training, so the resulting IV estimand may mix margins.

Inference

- → Use standard errors appropriate to assignment or instrument variation.
- → Cluster when instruments vary at group level.
- → First-stage uncertainty is part of IV estimation.
- → Report confidence intervals, not only point estimates.

Reading habit: interpret the first-stage table and the reduced-form table before trusting the second-stage coefficient.

External Validity

- → Who are the compliers?
- → Would a different instrument move a different group?
- → Would treatment effects change at scale?
- → Is the study population similar to the policy population?

Question

The result may be internally valid and still local to a narrow margin.

Example: a reminder letter identifies effects for forgetful or low-information individuals, not necessarily for everyone eligible.

Scale-Up

- → Treatment quality can change when a programme expands.
- → Participants at scale may differ from pilot participants.
- → General equilibrium effects may appear.
- → Administrative capacity can become a binding constraint.

Policy example: a voucher may work for parents with nearby childcare slots but not after expansion if slots become scarce.

Reporting IV Results

1. State the instrument and treatment received.
2. Report first stage and ITT.
3. State IV assumptions explicitly.
4. Interpret the complier group.
5. Discuss exclusion and external-validity threats.

Minimum table package: first stage, reduced form, IV estimate, control mean, sample size, and clustering level.

Common Mistakes

Watch for these

- → Calling IV an ATE without discussing compliers.
- → Reporting only the second-stage coefficient.
- → Treating a significant first stage as proof of exclusion.
- → Ignoring how the instrument changes behaviour.
- → Confusing assignment effects with treatment-received effects.

Practice Bridge

- → Theory sheet: `exercises/theory/session_04_theory.md`.
- → QCM: `assessments/qcm/session_04_qcm.md`.
- → Code task: compute ITT, first stage, and Wald estimate from a small dataset.
- → Source support: `../S1_experimental_method/code/r/02_power_rct.R`.

Student deliverable: a three-line interpretation: offer effect, take-up effect, and treatment effect for compliers.

Mini Scenario

- → A city randomly offers childcare vouchers.
- → Voucher offers increase childcare use by 25 percentage points.
- → Voucher offers increase employment by 5 percentage points.
- → What is the ITT? What is the first stage? What is the LATE?
- → Which exclusion threat worries you most?

Solution cue: $ITT_Y = 0.05$, $ITT_D = 0.25$, so $LATE = 0.20$ for parents induced to use childcare by the voucher offer.

Formula Sheet

$$ITT_Y = E[Y \mid Z = 1] - E[Y \mid Z = 0]$$
$$ITT_D = E[D \mid Z = 1] - E[D \mid Z = 0]$$
$$LATE = ITT_Y / ITT_D$$

→ These formulas are simple.

→ The hard part is defending the assumptions.

Reading an IV Paper

1. What is the endogenous treatment?
2. What is the instrument, and who assigns it?
3. How large is the first stage?
4. What is the exclusion story?
5. Who are the compliers?
6. Would another instrument identify another effect?

Interpretive discipline

An IV coefficient is not just "the effect of treatment". It is an effect for the group whose treatment is shifted by the instrument.

Key Takeaways

- → ITT estimates the effect of assignment or offer.
- → IV uses assignment-induced treatment variation.
- → LATE applies to compliers.
- → Relevance, independence, exclusion, and monotonicity are essential.
- → External validity requires asking who the compliers are.

One-sentence memory aid: IV replaces direct treatment variation with instrument-induced treatment variation.

Transition to Session 5

- → IV uses a credible instrument Z_i to shift treatment received D_i.
- → Some policies have no individual-level instrument.
- → They may still vary across places or institutions over time.
- → The next design asks whether untreated groups reveal the treated group's missing counterfactual trend.

Bridge question

When treatment timing differs across units, can changes in untreated units stand in for the treated counterfactual trend?

Appendix: From RCT to Natural IV

→ Some instruments are randomised by design.

→ Others come from rules, timing, distance, judges, or institutions.

→ Natural IV designs need stronger institutional explanation.

→ The same assumptions remain: relevance, independence, exclusion, monotonicity.

References:

→ Imbens and Angrist (1994), "Identification and Estimation of Local Average Treatment Effects".

→ Angrist, Imbens, and Rubin (1996), "Identification of Causal Effects Using Instrumental Variables".

→ Angrist and Pischke (2009), *Mostly Harmless Econometrics*, IV chapter.