ECONOMETRICS III

CAUSAL INFERENCE

Štěpán Jurajda

CERGE-EI

September 2, 2026

Part I

INTRO

INTRODUCTION

What has been covered in previous courses:

- LS asymptotics, Maximum Likelihood, GMM
- non-parametric density and regression estimation
- discrete choice: probit/logit, MNL, conditional and nested MNL
- panel data: static, dynamic, non-linear
- causal inference: **Potential Outcomes**, RCTs, matching, **IV - LATE**, RDD, DinDs

We will cover some of the same topics in more detail, and add a few extensions: IPW 'matching' (IPW LATE), DML (DML IPW), weak IVs, errors in variables, Synth. Cohort/Stacked DinDs, semi-parametric qualitative choice and sample selection, and **duration**.

Suggested Reading

- Heckman (2000) “Causal parameters and policy analysis in econometrics,”*QJE*
- Athey and Imbens (2017) “The State of Applied Econometrics: Causality and Policy Evaluation,” *JEP*
- Heckman (2010) “Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy,” *JEL*
- A unified framework for IV and selection models: Manning (2003) “Instrumental Variables for Binary Treatments with Heterogeneous Treatment Effects” link
- Imbens, G., and Y. Xu (2024) “LaLonde (1986) after Nearly Four Decades: Lessons Learned” link
- Mogstad and Torgovitsky (2024) “Instrumental Variables with Unobserved Heterogeneity in Treatment Effects” link
- Baker et al. (2025) Difference-in-Differences Designs: A Practitioner’s Guide
- Abadie et al. (2025) Differences-in-Differences and Event-Study Evidence
- Ahrens et al. (2025) An Introduction to Double/Debiased Machine Learning

Introduction

Policy Relevant Parameters

What do we mean by "x affects y "?

- $E[y|x]$ minimizing the MSE?
- $E[y|x] = \int_{-\infty}^{\infty} ydF(y|x)$ has no behavioral meaning.
- OLS is the BLUE (within linear approximations) if $E[\epsilon|x] = 0$.

What are the parameters we care about with experimental variation?
But first, what is experimental variation?

- The CIA, the Conditional Independence Assumption (unconfoundedness and ignorability): treatment assignment is independent of potential outcomes, at least conditional on observables.[1]
- SUTVA, the Stable Unit Treatment Value Assumption: treatment is the same thing for each unit and there are no spillovers between units.[2]

x is a *causal determinant* of y when we have a model and exogenous (experimental) variation in x.[3]
With an identified structural model, one can answer an array of policy-relevant counterfactual questions.

---

[1] In the IV framework, assume the exclusion restriction and monotonicity ($E[X|Z = 1] \geq E[X|Z = 0]$).
[2] The potential outcome of unit i depends only on the treatment of unit i and not on other units.
[3] Bin-scatter: Frisch-Waugh-Lovell theorem and Cattaneo et al.: $y = f(g(x) + W\gamma)$, where f and g can be non-smooth, non-linear functions, but with no interactions between x and W (only through f).

POTENTIAL OUTCOMES MODEL

- $Y_{1i} = \mu_1 + U_{1i}$ and $Y_{0i} = \mu_0 + U_{0i}$ where $T \in \{0, 1\}$ denotes treatment.
- $E[Y_{1i} - Y_{0i}] = \mu_1 - \mu_0$, the *average treatment effect* (ATE)
- $E[Y_{1i} - Y_{0i}|T_i = 1]$, the *average effect of treatment on the treated* (ATT)
- $ATT = ATE + E[U_{1i} - U_{0i}|T_i = 1]$

The fundamental problem with estimating the ATT is that the data only provides $E[Y_{1i}|T_i = 1]$ and $E[Y_{0i}|T_i = 0]$ but not the "what-if" counterfactual $E[Y_{0i}|T_i = 1]$.

$$\begin{aligned} E[Y_{1i}|T_i = 1] - E[Y_{0i}|T_i = 0] &= \text{ATT} + \{E[Y_{0i}|T_i = 1] - E[Y_{0i}|T_i = 0]\} \\ &= \mu_1 - \mu_0 + E[U_{1i} - U_{0i}|T_i = 1] + \\ &\quad \{E[U_{0i}|T_i = 1] - E[U_{0i}|T_i = 0]\} . \end{aligned}$$

The sample selection bias (the term in curly brackets) comes from the fact that the treatments and controls may have a different outcome if neither got treated. This model suggests that there is heterogeneity in treatment effects that should be at the center of our thinking.

A HOMOGENEOUS EFFECT OR NOT?

In a simple world, there is one β effect of T on Y, for everybody. There may be correlation between T and unobservables, so we IV for T in estimating $Y_i = \alpha + X_i\theta + \beta T_i + U_i$.
But what if individuals know (guess) their *person-specific* return from choosing T and *act upon it*?

$$\begin{aligned} Y_{i0} &= \alpha_0 + X_i\theta_0 + U_{i0} \\ Y_{i1} &= \alpha_1 + X_i\theta_1 + U_{i1}, \end{aligned}$$

The person-specific (causal) effect is then

$$\beta_i = Y_{i1} - Y_{i0} = \alpha_1 - \alpha_0 + X_i(\theta_1 - \theta_0) + U_{i1} - U_{i0}. \tag{1}$$

There is a *distribution* of returns (ex post causal effects, random coefficients).

When $\beta_i = \beta$ for all i, we worry about (i) $COV(T, U) \neq 0$, i.e., unobservable heterogeneity bias and measurement error, and (ii) the weak IV bias, where $COV(U, T)/COV(IV, T)$ is large because $COV(IV, T)$ is small. With person-specific β_i, there are more problems: $COV(T, U_0) \neq 0$ as before, but also $COV(\beta, U_0) \neq 0$ and crucially $COV(\beta, T) \neq 0$.

A HOMOGENEOUS EFFECT OR NOT? CONTD.

To complete the model (similar to a Heckman's λ switching regressions), introduce an instrument Z and postulate the first-stage selection equation:

$$\begin{aligned} T_i^* &= \theta_T Z_i + U_{iT} \\ T_i &= 1 \text{ if } T_i^* \geqq 0. \end{aligned}$$

We can now define three more treatment parameters: LATE, MTE, and ITT.

Imbens and Angrist (1994) prove the Local Average Treatment Effect (LATE) theorem, which provides an interpretation of IV coefficients in a binary-treatment, binary-instrument 2SLS. LATE IV allows for heterogeneous treatment effects (under the monotonicity assumption[4]) and classifies the population into distinct, non-overlapping but unobserved groups: compliers, always-takers, never-takers, and defiers (see also Sigstad, AER, 2026). The LATE treatment effect can be identified for the complier subpopulation.

Now, define the marginal treatment effect $\text{MTE}(x, u_T) = E[\beta | X = x, U_T = u_T]$ as the effect of T on a person with values (x, u_T) who is just indifferent between choosing $T = 1$ and 0.[5] It is a willingness to pay measure for those at the margin of indifference given X and U_D.

[4] Which is equivalent to additive separability of the error term and the explanatory variables in the first-stage equation; see Mogstad and Torgovitsky (2024). How can one square up multiple instruments with the monotonicity assumption? For example, in the case of multiple judges (Dahl et al. 2014), instead of using judge-specific IVs, use a binary IV consisting of strict/lenient judges whilst dropping moderate judges.

[5] Start with writing down the LATE for the case of a binary IV (values z and z') and let z' go to z. This is MTE. The people who are affected by such a small change in IV are indifferent between the two choices.

A HOMOGENEOUS EFFECT OR NOT? CONTD.

Running the outcome on the instrument instead of the treatment (the so-called reduced form) gives the *intention to treat* (ITT) parameter.[6]

The ultimate goal is to provide evidence for policy making. The typical range of parameters of policy interest (ATT, ATE) can be expressed as differentially weighted averages (integrals over population) of the MTE. IV and OLS can also be expressed as weighted averages of MTE, but the weights are not those of ATT, ATE, etc. IV (LATE) weights are driven by the instrument: whose treatment assignment is predicted by the instrument? LATE estimates the mean return at the margin defined by the IV manipulation.[7]

Treatment effect heterogeneity is at the core of the recent difference-in-differences and IV literature. One may be able to assume away unobservable heterogeneity in treatment effects if there are no interactions of observables with treatment, and if treatment assignment is made in absence of agents being able to guess about the size of treatment effects.[8]

[6] LATE equals ITT divided by the *compliance rate* (the first stage associated with the IV) based on the indirect least squares argument (see Example **??**).

[7] See the discussion between Deaton (2010), Imbens (2010), and Heckman and Urzua (2010) on whether LATE is a policy relevant parameter. It turns out that linear IV actually often does not estimate LATE (Mogstad and Torgovitsky, 2024).

[8] With binary treatment and binary outcome, there must be treatment heterogeneity, as otherwise the average treatment effect across the population (or within any sub-group) would have to be -1, 0, or +1.

A HOMOGENEOUS EFFECT OR NOT? CONTD.

In the Moving to Opportunity (MTO) experiment in Boston, vouchers supporting re-location of families were assigned at random, but compliance with the randomized experiment was imperfect. One uses the random assignment as IV for actual treatment to estimate LATE—to zoom in on the 'compliers' and to avoid the 'defiers' (and this also omits the 'always takers' and the 'never takers'). If we run the outcome on the randomized voucher assignment instead, we estimate the ITT parameter.

Angrist et al. (NBER WP 33635) revisit the landmark results from the Women's Health Initiative trial that showed that random assignment to menopausal hormone therapy (MHT) elevated risks of breast cancer. Recent analyses argued that MHT risks are small based on ITT effects that ignored non-compliance. IV methods estimate effects of MHT on compliers (with MHT due to random assignment) and show larger risks and benefits.