Linear Models Lecture 13: Instrumental Variables

Robert Gulotty

University of Chicago

February 26, 2026

Structural vs. Projection Parameters

- Recall from our earlier lectures: OLS estimates the **best linear predictor** (projection).
- A **structural model** posits a causal data-generating process:

$$Y = \boldsymbol{x}'\beta + e$$

- We can always define a projection: $\beta^* = (\mathbb{E}[\boldsymbol{x}\boldsymbol{x}'])^{-1}\mathbb{E}[\boldsymbol{x}Y]$, with $\mathbb{E}[\boldsymbol{x}e^*] = 0$.
- When $\mathbb{E}[\boldsymbol{x}e] \neq 0$:

$$\beta^* = \beta + (\mathbb{E}[\boldsymbol{x}\boldsymbol{x}'])^{-1}\mathbb{E}[\boldsymbol{x}e] \neq \beta$$

- OLS is inconsistent for the structural parameter: $\hat{\beta} \xrightarrow{p} \beta^* \neq \beta$.
- We call $\boldsymbol{x}$ **endogenous** when this occurs.

Endogeneity Source 1: Measurement Error

- The true structural model is $\mathbb{E}[Y \mid \boldsymbol{z}] = \boldsymbol{z}'\beta$, but we don't observe $\boldsymbol{z}$.
- Our measured variables: $\boldsymbol{x} = \boldsymbol{z} + \boldsymbol{u}$, where $\boldsymbol{u}$ is measurement error.
- Assumptions on the error:
  - $\text{plim}\ \frac{\boldsymbol{z}'\boldsymbol{u}}{n} = 0$: measurement error uncorrelated with truth
  - $\text{plim}\ \frac{\boldsymbol{e}'\boldsymbol{u}}{n} = 0$: measurement error uncorrelated with structural disturbance
- **Political science examples:**
  - Survey-reported ideology (ANES self-placement on liberal–conservative scale)
  - GDP in developing countries used in aid allocation studies
  - Self-reported voter turnout (overreported by $\sim$10–15%)

Measurement Error: Rewriting in Observables

- Substitute $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{u}$ into the structural equation:

$$\begin{aligned} Y &= \boldsymbol{z}'\beta + e \\ &= (\boldsymbol{x} - \boldsymbol{u})'\beta + e \\ &= \boldsymbol{x}'\beta + \underbrace{(e - \boldsymbol{u}'\beta)}_{\equiv \nu} \end{aligned}$$

- But $\boldsymbol{x}$ and ν are correlated:

$$\mathbb{E}[\boldsymbol{x}\nu] = \mathbb{E}[(\boldsymbol{z} + \boldsymbol{u})(e - \boldsymbol{u}'\beta)] = -\mathbb{E}[\boldsymbol{u}\boldsymbol{u}']\beta \neq 0$$

- The measurement error in regressors creates endogeneity, even though the true model is correctly specified.

Measurement Error: Attenuation Bias

- The OLS probability limit:

$$
\begin{aligned}
\text{plim } \hat{\beta} &= \beta + \left(\text{plim } \frac{\boldsymbol{X}'\boldsymbol{X}}{n}\right)^{-1} \text{plim } \frac{\boldsymbol{X}'\boldsymbol{u}}{n}\beta \\
&= \beta - \Sigma_X^{-1}\Sigma_{\boldsymbol{u}}\beta \\
&= \left(\boldsymbol{I} - \underbrace{\Sigma_X^{-1}}_{\text{signal}}\underbrace{\Sigma_{\boldsymbol{u}}}_{\text{noise}}\right)\beta
\end{aligned}
$$

- In the scalar case: plim $\hat{\beta} = \frac{\sigma_z^2}{\sigma_z^2+\sigma_u^2}\beta$
- OLS is biased **toward zero** — this is attenuation bias.
- Even if only one variable has measurement error, it affects **all** slope coefficients.

Measurement Error Example: Ideology and Voting

- **Question:** Does ideological extremism reduce electoral support?

$$\text{VoteShare}_i = \beta_0 + \beta_1 \,\text{Ideology}_i^* + \boldsymbol{x}_i'\gamma + e_i$$

- $\text{Ideology}^* =$ true ideological position (unobserved).
- We measure ideology with error:

$$\text{Ideology}_i = \text{Ideology}_i^* + u_i$$

  (e.g., survey responses, roll-call scores like DW-NOMINATE).
- Attenuation bias: OLS **underestimates** the penalty for extremism.
- **IV solution:** Use a second measure of ideology (e.g., campaign finance scores, CFscores) as an instrument — it shares the true signal but has independent measurement error.

Endogeneity Source 2: Simultaneity

- Two equations are jointly determined — each variable is both cause and effect.
- **Classic example:** Arms races (Richardson model).

$$\text{MilSpend}_A = \alpha_0 + \alpha_1 \, \text{MilSpend}_B + e_A$$
$$\text{MilSpend}_B = \gamma_0 + \gamma_1 \, \text{MilSpend}_A + e_B$$

- MilSpend_B appears on the RHS of equation 1, but is determined in equation 2 (which depends on MilSpend_A).
- OLS on either equation is inconsistent: the regressor is correlated with the error by construction.
- **Other examples:** Trade policy and trade flows; campaign spending and vote share; policing levels and crime rates.

Endogeneity Source 3: Endogenous Choice (Selection)

- Agents **choose** their treatment based on expected gains (Roy 1951).
- Potential outcomes: $Y_i(1) = \boldsymbol{x}_i'\beta_1 + e_{1i}, \quad Y_i(0) = \boldsymbol{x}_i'\beta_0 + e_{0i}$.
- Selection rule: individual chooses treatment if net benefit exceeds threshold:

$$D_i = \mathbb{1}\{\boldsymbol{z}_i'\gamma + \eta_i > 0\}, \qquad \eta_i = e_{1i} - e_{0i}$$

- If η_i and e_{1i} are correlated:

$$\mathbb{E}[e_{1i} \mid D_i = 1] = \mathbb{E}[e_{1i} \mid \boldsymbol{z}_i'\gamma + \eta_i > 0] \neq 0$$

- OLS on the treated sample is biased — we only observe outcomes for those who chose treatment.
- If $\boldsymbol{z}$ includes variables excluded from $\boldsymbol{x}$, we have an **exclusion restriction**.

Endogenous Choice: Political Selection

- **Question:** Does holding office increase personal wealth?

$$\text{Wealth}_i = \beta_0 + \beta_1 \,\text{HeldOffice}_i + \boldsymbol{x}_i'\gamma + e_i$$

- Problem: Who runs for office? Who wins?
  - Wealthier individuals may be more likely to run and win
  - More politically connected individuals may both win and accumulate wealth
  - Selection on gains: those who expect to profit most from office seek it out
- Observed difference = ATT + Selection bias:

$$\mathbb{E}[\text{Wealth} \mid \text{Office} = 1] - \mathbb{E}[\text{Wealth} \mid \text{Office} = 0]$$

- **IV solution:** Use close election outcomes (regression discontinuity / Lee 2008) as an instrument — winning a close race is quasi-random.

Selection Bias Decomposition

- The observed difference in outcomes:

$$\mathbb{E}[Y \mid D=1] - \mathbb{E}[Y \mid D=0] = \underbrace{\mathbb{E}[Y(1) - Y(0) \mid D=1]}_{\text{ATT}} + \underbrace{\mathbb{E}[Y(0) \mid D=1] - \mathbb{E}[Y(0) \mid D=0]}_{\text{Type I: Selection on levels}}$$

- If treatment effects are heterogeneous ($\tau_i = Y_i(1) - Y_i(0)$ varies):

$$\mathbb{E}[\tau_i \mid D = 1] \neq \mathbb{E}[\tau_i] \quad \Rightarrow \quad \text{Type II: Selection on gains}$$

- **Type I:** Officeholders would have been wealthier anyway (baseline differences).
- **Type II:** Those who gain most from office are most likely to seek it (differential returns).

The Common Thread

In all three cases — measurement error, simultaneity, selection — the core problem is $\mathbb{E}[\boldsymbol{x}e] \neq 0$. The solution: find an instrument Z.

Source	**Why $\mathbb{E}[\boldsymbol{x}e] \neq 0$**	**IV strategy**
Measurement error	Noise in $\boldsymbol{x}$ enters error	Second measure of $\boldsymbol{x}^*$
Simultaneity	Y and X jointly determined	Exogenous shifter of one eq.
Selection/omitted var.	Choice correlated with e	Excluded exogenous variable

Notation: Structural Equation (Hansen Ch. 12)

- Y is a linear function of exogenous variables $\boldsymbol{x}_1$ and endogenous variables $\boldsymbol{y}_2$:

$$Y_1 = \boldsymbol{x}_1'\beta_1 + \boldsymbol{y}_2'\beta_2 + e, \qquad \mathbb{E}[\boldsymbol{y}_2 e] \neq 0$$

- **Instruments:** $\boldsymbol{z} = (\boldsymbol{z}_1', \boldsymbol{z}_2')'$, dimension $l \times 1$:
  - $\boldsymbol{z}_1 = \boldsymbol{x}_1$: included instruments (the exogenous regressors), dimension k_1
  - $\boldsymbol{z}_2$: excluded instruments, dimension $l_2 \geq k_2$
- Writing $\boldsymbol{x} = (\boldsymbol{x}_1', \boldsymbol{y}_2')'$ of dimension $k = k_1 + k_2$:

$$Y_1 = \boldsymbol{x}'\beta + e$$

Instrumental Variable Conditions

Three conditions for valid instruments $\boldsymbol{z}$:

1. **Exogeneity:** $\mathbb{E}[\boldsymbol{z}e] = 0$ (instrument uncorrelated with structural error)
2. **Relevance:** rank $\mathbb{E}[\boldsymbol{z}\boldsymbol{x}'] = k$ (instruments predict endogenous regressors)
3. **Order condition:** $l \geq k$ (at least as many instruments as regressors)

- When $l = k$: **just identified** (exactly as many instruments as regressors)
- When $l > k$: **overidentified** ($q = l - k$ overidentifying restrictions)
- When $l < k$: **underidentified** (cannot estimate β)

Example: Labeling the IV Setup

Do democratic institutions cause economic growth? (Acemoglu, Johnson & Robinson 2001)

- **Structural equation** ($Y_1 = \boldsymbol{x}_1'\beta_1 + \boldsymbol{y}_2'\beta_2 + e$):

$$\text{GDP/capita}_i = \beta_1 \text{ Latitude}_i + \beta_2 \underbrace{\text{Institutions}_i}_{\boldsymbol{y}_2\text{: endogenous}} + e_i$$

- **Endogeneity:** Richer countries may adopt better institutions (reverse causality); omitted factors (geography, culture) affect both.
- **Excluded instrument** ($\boldsymbol{z}_2$): Colonial settler mortality.
  - **Relevance:** High settler mortality $\Rightarrow$ extractive colonies $\Rightarrow$ weak institutions today.
  - **Exogeneity:** Historical disease environment affects current GDP only through institutions (exclusion restriction — debatable!).

Reduced Form: Definitions

- The **reduced form** expresses all endogenous variables as functions of instruments only.
- For the endogenous regressors $\boldsymbol{y}_2$ ($k_2 \times 1$):

$$\boldsymbol{y}_2 = \Gamma'\boldsymbol{z} + \boldsymbol{u}_2 = \Gamma_{12}'\boldsymbol{z}_1 + \Gamma_{22}'\boldsymbol{z}_2 + \boldsymbol{u}_2, \qquad \mathbb{E}[\boldsymbol{z}\boldsymbol{u}_2'] = 0$$

  where Γ is $l \times k_2$, defined by $\Gamma = \mathbb{E}[\boldsymbol{z}\boldsymbol{z}']^{-1}\mathbb{E}[\boldsymbol{z}\boldsymbol{y}_2']$.
- The full projection of all regressors $\boldsymbol{x} = (\boldsymbol{x}_1', \boldsymbol{y}_2')'$ on $\boldsymbol{z}$:

$$\bar{\Gamma} = \begin{bmatrix} \boldsymbol{I}_{k_1} & \Gamma_{12} \\ \boldsymbol{0} & \Gamma_{22} \end{bmatrix} = \begin{bmatrix} \boldsymbol{I}_{k_1} & \Gamma \\ \boldsymbol{0} & \end{bmatrix} = \mathbb{E}[\boldsymbol{z}\boldsymbol{z}']^{-1}\mathbb{E}[\boldsymbol{z}\boldsymbol{x}']$$

- Key: OLS consistently estimates Γ and $\bar{\Gamma}$ because $\boldsymbol{z}$ is exogenous.

Reduced Form for Y

- Plugging the reduced form for $\boldsymbol{y}_2$ into the structural equation:

$$
\begin{aligned}
Y_1 &= \boldsymbol{z}_1'\beta_1 + (\Gamma_{12}'\boldsymbol{z}_1 + \Gamma_{22}'\boldsymbol{z}_2 + \boldsymbol{u}_2)'\beta_2 + e \\
&= \boldsymbol{z}_1' \underbrace{(\beta_1 + \Gamma_{12}\beta_2)}_{\lambda_1} + \boldsymbol{z}_2' \underbrace{\Gamma_{22}\beta_2}_{\lambda_2} + \underbrace{(\boldsymbol{u}_2'\beta_2 + e)}_{u_1} \\
&= \boldsymbol{z}'\lambda + u_1
\end{aligned}
$$

- The **structural** parameters are β_1, β_2.
- The **reduced form** parameters are λ, Γ.
- Relationship: $\lambda = \bar{\Gamma}\beta$.

Reduced Form: The AJR Example

- **Structural equation** (what we want):

$$\text{GDP}_i = \beta_1 \text{ Lat}_i + \beta_2 \text{ Institutions}_i + e_i$$

- **First stage** (reduced form for $\boldsymbol{y}_2$):

$$\text{Institutions}_i = \underbrace{\Gamma_{12}}_{\text{Lat coeff}} \text{ Lat}_i + \underbrace{\Gamma_{22}}_{\text{Mortality coeff}} \text{ SettlerMort}_i + u_{2i}$$

- **Reduced form for** Y:

$$\text{GDP}_i = \underbrace{\lambda_1}_{=\beta_1+\Gamma_{12}\beta_2} \text{ Lat}_i + \underbrace{\lambda_2}_{=\Gamma_{22}\beta_2} \text{ SettlerMort}_i + u_{1i}$$

- λ_2 is the "reduced form effect" of settler mortality on GDP.
- The structural parameter: $\beta_2 = \lambda_2 / \Gamma_{22}$ (ratio of reduced form to first stage).

Identification

- β is **identified** if it is the unique solution to the moment conditions:

$$\mathbb{E}[\boldsymbol{z}(Y_1 - \boldsymbol{x}'\beta)] = 0$$

- This is a system of l equations in k unknowns.
- **Just identified** ($l = k$): unique solution $\beta = (\mathbb{E}[\boldsymbol{z}\boldsymbol{x}'])^{-1}\mathbb{E}[\boldsymbol{z}Y_1]$
- Equivalently: if $\bar{\Gamma}$ has rank k, then $\beta = (\bar{\Gamma}'\bar{\Gamma})^{-1}\bar{\Gamma}'\lambda$.
- **Overidentified** ($l > k$): system is overdetermined.
  - No exact solution in general — we need a method to combine the moment conditions.
  - Foreshadowing: this is exactly the problem GMM solves (Lecture 15).
- The **relevance condition** (rank $\mathbb{E}[\boldsymbol{z}\boldsymbol{x}'] = k$) is what makes the solution exist.

IV Estimator: Just-Identified Case

- When $l = k$, the **IV estimator** is the sample analogue of the moment condition:

$$\hat{\beta}_{IV} = (Z'X)^{-1}Z'Y$$

- Decompose:

$$\begin{aligned}\hat{\beta}_{IV} &= (Z'X)^{-1}Z'(X\beta + e) \\ &= \beta + (Z'X)^{-1}Z'e\end{aligned}$$

- Consistency: plim $\hat{\beta}_{IV} = \beta + (E[ZX'])^{-1}E[Ze] = \beta \quad$ under exogeneity.

Indirect Least Squares

- **ILS**: Estimate reduced forms by OLS, then recover β.
- Reduced form estimates:

$$\hat{\Gamma} = (Z'Z)^{-1}Z'X, \qquad \hat{\lambda} = (Z'Z)^{-1}Z'Y$$

- When $l = k$: $\hat{\beta}_{ILS} = \hat{\Gamma}^{-1}\hat{\lambda}$
- Show equivalence:

$$\begin{aligned}
\hat{\beta}_{ILS} &= [(Z'Z)^{-1}Z'X]^{-1}(Z'Z)^{-1}Z'Y \\
&= (Z'X)^{-1}(Z'Z)(Z'Z)^{-1}Z'Y \\
&= (Z'X)^{-1}Z'Y = \hat{\beta}_{IV}
\end{aligned}$$

- ILS = IV in the just-identified case.

The Wald Estimator

- Special case: single endogenous X, single binary instrument $Z \in \{0, 1\}$.
- The IV estimator simplifies to the **Wald estimator**:

$$\hat{\beta}_{Wald} = \frac{\bar{Y}_{Z=1} - \bar{Y}_{Z=0}}{\bar{X}_{Z=1} - \bar{X}_{Z=0}} = \frac{\text{Reduced form effect on } Y}{\text{First stage effect on } X}$$

- Intuition: scale the **intent-to-treat** (ITT) effect by the first-stage compliance rate.
- This is a ratio of two consistent estimators — a **ratio estimator**.

The Wald/IV estimator is the ratio of the reduced form to the first stage. This "ratio" structure is central to understanding IV.

Wald Estimator: Draft Lottery and Civic Participation

- **Question:** Does military service increase political participation?
- Y = voter turnout, X = served in military, Z = draft lottery number (low = drafted).
- **Reduced form:** $\bar{Y}_{Z=\text{low}} - \bar{Y}_{Z=\text{high}}$ = effect of lottery on turnout (ITT).
- **First stage:** $\bar{X}_{Z=\text{low}} - \bar{X}_{Z=\text{high}}$ = effect of lottery on service rate.

$$\hat{\beta}_{Wald} = \frac{\text{ITT on turnout}}{\text{First stage compliance}} = \frac{\text{Reduced form}}{\text{First stage}}$$

- Not everyone drafted actually serves (deferments, exemptions).
- The Wald estimator rescales the ITT by the share who comply with the draft.

Two-Stage Least Squares

- When $l > k$ (overidentified), we need 2SLS.
- **Stage 1:** Regress X on Z:

$$\hat{X} = P_Z X, \qquad P_Z = Z(Z'Z)^{-1}Z'$$

- **Stage 2:** Regress Y on $\hat{X}$:

$$\hat{\beta}_{2SLS} = (\hat{X}'\hat{X})^{-1}\hat{X}'Y = (X'P_Z X)^{-1}X'P_Z Y$$

- The projection P_Z extracts the **exogenous variation** in X — the part predicted by instruments.

2SLS = IV When Just Identified

- When $l = k$, $P_Z = Z(Z'Z)^{-1}Z'$ and:

$$
\begin{aligned}
\hat{\beta}_{2SLS} &= (\hat{X}'\hat{X})^{-1}\hat{X}'Y \\
&= (X'P_Z X)^{-1}X'P_Z Y \\
&= (X'Z(Z'Z)^{-1}Z'X)^{-1}X'Z(Z'Z)^{-1}Z'Y \\
&= (Z'X)^{-1}(Z'Z)(X'Z)^{-1}X'Z(Z'Z)^{-1}Z'Y \qquad ((ABC)^{-1} = C^{-1}B^{-1}A^{-1}) \\
&= (Z'X)^{-1}Z'Y = \hat{\beta}_{IV}
\end{aligned}
$$

- 2SLS generalizes IV to the overidentified case.

R Implementation

```
library(estimatr)

# 2SLS estimation
# Y ~ endogenous + exogenous | excluded_instruments + exogenous
fit_iv <- iv_robust(Y ~ D + X | Z + X, data = dat)
summary(fit_iv)
```

- Variables before "|" are the structural equation regressors.
- Variables after "|" are the instruments (excluded + included).
- `iv_robust` uses heteroskedasticity-robust standard errors by default.

Consistency of 2SLS (Hansen Thm. 12.1)

Theorem (Consistency)

Under the conditions: (1) $E[Ze] = 0$, *(2) rank* $E[ZX'] = k$, *(3)* $E[ZZ'] > 0$, *the 2SLS estimator is consistent:* $\hat{\beta}_{2SLS} \xrightarrow{p} \beta$.

Proof sketch:

$$\hat{\beta}_{2SLS} - \beta = \left(\frac{X'P_Z X}{n}\right)^{-1} \frac{X'P_Z e}{n}$$

Define $Q_{XZ} = E[XZ']$, $Q_{ZZ} = E[ZZ']$, $Q_{ZX} = E[ZX']$. Then:

$$\xrightarrow{p} (Q_{XZ} Q_{ZZ}^{-1} Q_{ZX})^{-1} Q_{XZ} Q_{ZZ}^{-1} \underbrace{E[Z'e]}_{=0} = 0$$

Asymptotic Distribution (Hansen Thm. 12.2)

Theorem (Asymptotic Normality)

Under regularity conditions:

$$\sqrt{n}(\hat{\beta}_{2SLS} - \beta) \xrightarrow{d} N(0, V_\beta)$$

where:

$$V_\beta = (Q_{XZ} Q_{ZZ}^{-1} Q_{ZX})^{-1} Q_{XZ} Q_{ZZ}^{-1} \Omega Q_{ZZ}^{-1} Q_{ZX} (Q_{XZ} Q_{ZZ}^{-1} Q_{ZX})^{-1}$$

with $Q_{XZ} = E[XZ']$, $Q_{ZZ} = E[ZZ']$, $Q_{ZX} = E[ZX']$, $\Omega = E[Z'Zee'Z]$.

- Under homoskedasticity ($E[e^2 \mid Z] = \sigma^2$), this simplifies to:

$$V_\beta = \sigma^2 (Q_{XZ} Q_{ZZ}^{-1} Q_{ZX})^{-1}$$

- Under heteroskedasticity, use robust covariance estimation (as in Lecture 8).

Variance Estimation

- **Homoskedastic** variance estimator:

$$\hat{V}_\beta = \hat{\sigma}^2 (X' P_Z X)^{-1}, \qquad \hat{\sigma}^2 = \frac{\hat{e}'\hat{e}}{n-k}$$

  where $\hat{e} = Y - X\hat{\beta}_{2SLS}$ (use **original** X, not $\hat{X}$).

- **Heteroskedastic-robust** (analogous to HC):

$$\hat{V}_\beta = (X' P_Z X)^{-1} \left(\sum_{i=1}^{n} \hat{e}_i^2 \hat{X}_i \hat{X}_i' \right) (X' P_Z X)^{-1}$$

- Important: Standard errors from the "naive" second-stage regression (regressing Y on $\hat{X}$ and reading off SEs) are **incorrect** — they use $\hat{X}$ instead of X in the residuals.

IV as a Method of Moments

- The IV estimator solves the **sample moment condition**:

$$\frac{1}{n}\sum_{i=1}^{n} Z_i(Y_i - X_i'\beta) = 0$$

- This is a system of l equations in k unknowns.
- When $l = k$: exactly identified, unique solution.
- When $l > k$: overidentified, no exact solution.
- **2SLS resolves this** by using a specific weighting matrix:

$$\hat{\beta}_{2SLS} = \arg\min_{\beta} \left[\frac{1}{n}\sum Z_i(Y_i - X_i'\beta)\right]' (Z'Z/n)^{-1} \left[\frac{1}{n}\sum Z_i(Y_i - X_i'\beta)\right]$$

Foreshadowing GMM

- 2SLS uses the weighting matrix $\hat{W} = (Z'Z/n)^{-1}$.
- Is this the **optimal** weighting?
  - Under homoskedasticity: **yes**.
  - Under heteroskedasticity: **no** — there exists a more efficient choice.
- The **Generalized Method of Moments** (GMM) chooses $\hat{W}$ optimally:

$$\hat{W}_{opt} = \hat{\Omega}^{-1}, \qquad \hat{\Omega} = \frac{1}{n}\sum_{i=1}^{n} \hat{e}_i^2 Z_i Z_i'$$

- GMM also extends beyond linear IV to **any** set of moment conditions.
- Next week: we develop the full GMM framework.

Summary

1. Endogeneity ($E[Xe] \neq 0$) makes OLS inconsistent for structural parameters.
2. Instruments Z satisfying exogeneity and relevance allow consistent estimation.
3. The IV estimator $\hat{\beta}_{IV} = (Z'X)^{-1}Z'Y$ works when just identified.
4. 2SLS $\hat{\beta}_{2SLS} = (X'P_Z X)^{-1}X'P_Z Y$ generalizes to overidentification.
5. Both are consistent and asymptotically normal under standard conditions.
6. IV is a **method of moments** — 2SLS is a specific weighting of moment conditions.

Next lecture: Finite sample properties, testing, LATE, and the MTE framework.