

Applied Economics Issues

Session 2: Endogeneity and a First Look at Instruments

Ewen Gallic

Master 2 DADEE – 2026

Today's session

- 1 **Reading debrief**: how do empirical economists write an introduction?
 - 2 **Econometrics, the endogeneity problem, and a first look at instruments**
 - 3 **Practice in R** (in class): does OLS recover the truth? Can an instrument help?
 - 4 **Next session**: the map of the course, and the reading assignment
- ▶ You read [Ashenfelter and Card \(1985\)](#) and [Angrist and Krueger \(1991\)](#): **you** will explain them when we get there.
 - ▶ Session 3: [Angrist and Krueger \(1991\)](#) and instrumental variables in depth. Session 4: [Ashenfelter and Card \(1985\)](#) and difference-in-differences.

1. Reading debrief: anatomy of an introduction

Part 1 of your assignment: what does each paragraph do?

- ▶ For each paragraph of the two introductions, you wrote one sentence describing its **function**.
- ▶ Let us build the map together, paragraph by paragraph:
 - 1 Angrist and Krueger (1991)
 - 2 Ashenfelter and Card (1985)
- ▶ For each paragraph, ask yourself:
 - Why did the authors put this paragraph **here**?
 - What would the reader lose if it were removed?

A standard structure: the “Introduction Formula”

Keith Head (UBC) summarizes the norm for introductions in economics in five steps (Head, n.d.):

- 1 **Hook**: attract the reader's interest (the paper relates to something that matters).
- 2 **Question**: tell the reader what the paper actually does.
- 3 **Antecedents**: the prior work needed to understand the contribution.
- 4 **Value-added**: how the paper adds to that work (typically about three contributions).
- 5 **Roadmap**: how the rest of the paper is organized.

Evans (2020) did your exercise (“what do the authors do in paragraph one?”) on 15 papers from the AER, the QJE and AEJ: Applied. Other guides he cites add two explicit steps: the **method** and the **main findings**.

Comparing the two introductions

Discussion

- ▶ Where is the **research question** stated? In which paragraph?
- ▶ Where do the authors tell us **how** they will identify the effect?
- ▶ Do they give the **main result** in the introduction? With a number?
- ▶ Which steps of the formula are missing, merged, or in a different order?

Why it matters for you

- ▶ Reading the introduction well is the **first pass** on any paper.
- ▶ If you cannot answer “what is the question, how do they answer it, what do they find?” after the introduction, read it again.
- ▶ The same structure applies when **you** write a synopsis or a report.

2. Core lenses of applied economics

From theory to application

- ▶ Applied work is rooted in economic **theory**.
- ▶ The **theories** provide simplified models of **behaviour**, **incentives**, and **interactions** among agents.
- ▶ Applied economists take **insights** from these theories and **test them against evidence**
 - using real-world data,
 - addressing policy questions.

With what tools?

- ▶ Applied work relies on multiple tools: different hammers for different nails.
- ▶ We will distinguish between two different questions applied economists usually have to answer:
 - 1 Does X cause Y ?
 - measuring causal effects.
 - 2 How much is X worth?
 - assigning economic value to things without explicit market prices.
- ▶ Econometrics is the common language across methods.

2.1. Econometrics

Econometrics: the root of applied work

- ▶ As you already know, econometrics consists in applying statistical models to economic data.
- ▶ In applied economics, econometrics can be used for different purposes:
 - 1 **description** ,
 - 2 **prediction** ,
 - 3 **causality** .

The econometric toolbox

Among the different existing econometric methods, the main are:

- ▶ **Regression analysis**: the jackknife of econometrics.
- ▶ **Panel data methods**: to exploit variation over two dimensions, usually over time and across individuals.
- ▶ **Instrumental variables** (IV): to handle endogeneity with external sources of variation.
- ▶ **Difference-in-Differences** (DiD): to compare treated vs. control before/after policy.
- ▶ **Regression Discontinuity** (RD): to exploit cutoffs or thresholds.

This list is obviously not exhaustive and will probably evolve with time.

Regression analysis example: Mincer's earnings function

- ▶ A nice example of **regression analysis** is that of [Mincer \(1974, NBER\)](#).
- ▶ The research question is: how do schooling and on the job training determine wages over working life?
- ▶ The paper mostly relies on a 1/1,000 sample from the U.S. Census of Population microdata (1950 and 1960).

Regression analysis example: Mincer's earnings function

Assumptions in [Mincer \(1974, NBER\)](#):

- ▶ Theoretically: investment in human capital leads to higher wages.
- ▶ Empirically: log wages increase linearly with years of schooling and concavely with work experience.

$$\ln w_i = \alpha + \rho S_i + \beta_1 X_i + \beta_2 X_i^2 + \varepsilon_i \quad (1)$$

Diagram illustrating the Mincer's earnings function regression model:

- The dependent variable is $\ln w_i$ (Wage/income).
- The independent variables are:
 - S_i (Years of schooling)
 - X_i (Potential experience)
 - X_i^2 (Sq. Potential experience)
- The error term is ε_i .

- ▶ The parameters are estimated with OLS:
 - ρ : average return to one extra year of schooling (each additional year of schooling increases wages by about $\rho \times 100\%$),
 - β_1, β_2 : returns to experience (experience increases earnings at a decreasing rate bc of concavity assumption).

2.2. The endogeneity problem

The Endogeneity Problem

We want to estimate the **causal effect** of some regressors on y . For observation i , we write the model:

$$y_i = \mathbf{x}_i' \beta + \varepsilon_i, \quad \mathbf{x}_i = (1, x_{i1}, \dots, x_{i,K-1})'$$

- ▶ β_k : what happens to y if we change x_k by one unit, *the other regressors and everything in ε held fixed*.
- ▶ ε_i : everything else that affects y_i (unobserved).
- ▶ OLS is **consistent** for β if and only if

$$E[\mathbf{x}_i \varepsilon_i] = \mathbf{0}.$$

- ▶ When $E[\mathbf{x}_i \varepsilon_i] \neq \mathbf{0}$, some regressors are said to be **endogenous**. As a consequence, OLS converges to something else than β .

Reminders: Convergence, Consistency, Unbiasedness

Let $\hat{\theta}_n$ be an estimator of a parameter θ_0 , computed on a sample of size n .

- **Convergence in probability** : a property of **any** sequence of random variables Z_n .

$$Z_n \xrightarrow{P} c \text{ (or } \text{plim } Z_n = c) \iff \forall \eta > 0, \Pr(|Z_n - c| > \eta) \xrightarrow[n \rightarrow \infty]{} 0.$$

- **Consistency** : the special case where the sequence is the estimator, and the limit is the true value:

$$\hat{\theta}_n \xrightarrow{P} \theta_0.$$

→ where does the estimator go *as n grows*?

- **Unbiasedness** : for a **given** n , the estimator is right on average across samples:

$$E[\hat{\theta}_n] = \theta_0.$$

→ where is the centre of the estimator *for a fixed n*

Unbiased vs. Consistent: Neither Implies the Other

Let $X_1, \dots, X_n$ be i.i.d. with mean μ and variance σ^2 .

	Consistent	Not consistent
Unbiased	$\bar{X}_n$	X_1 (first observation only)
Biased	$\bar{X}_n + \frac{1}{n}$; $\frac{1}{n} \sum_i (X_i - \bar{X}_n)^2$	$\bar{X}_n + 1$; OLS with an endogenous regressor

- ▶ X_1 : $E[X_1] = \mu$, but its distribution does **not** shrink when n grows.
- ▶ $\bar{X}_n + \frac{1}{n}$: biased by $\frac{1}{n}$, but the bias vanishes as $n \rightarrow \infty$. Same for the variance with $1/n$ (instead of $1/(n-1)$): $E = \frac{n-1}{n} \sigma^2$.
- ▶ $\bar{X}_n + 1$: the bias never vanishes.

Reminders: Two Tools to Prove Consistency

Law of large numbers (for i.i.d. $\mathbf{w}_i$ with finite mean): sample averages converge to expectations,

$$\frac{1}{n} \sum_{i=1}^n \mathbf{w}_i \xrightarrow{p} E[\mathbf{w}_i].$$

Rules for plim (continuous mapping / Slutsky): if $\mathbf{A}_n \xrightarrow{p} \mathbf{A}$ and $\mathbf{B}_n \xrightarrow{p} \mathbf{B}$, then

$$\mathbf{A}_n + \mathbf{B}_n \xrightarrow{p} \mathbf{A} + \mathbf{B}, \quad \mathbf{A}_n \mathbf{B}_n \xrightarrow{p} \mathbf{A} \mathbf{B}, \quad \mathbf{A}_n^{-1} \xrightarrow{p} \mathbf{A}^{-1} \text{ (}\mathbf{A} \text{ invertible)}.$$

- ▶ Expectations do not pass through products and inverses like that:
 $E[\mathbf{A}_n^{-1} \mathbf{B}_n] \neq E[\mathbf{A}_n^{-1}] E[\mathbf{B}_n]$ in general.
- ▶ plim is, therefore, much easier to manipulate than E: **consistency is often easy to establish when unbiasedness is not.**

OLS: the Sampling Error

Consider the linear model with n observations and K regressors (including the constant):

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

- ▶ The OLS estimator of $\boldsymbol{\beta}$ is:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}.$$

- ▶ Plugging $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, it can easily be shown that

$$\hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \underbrace{(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}}_{\text{sampling error}}$$

Both properties depend on how the sampling error behaves:

- ▶ Unbiasedness: what is its **mean**, for a given n ?
- ▶ Consistency: what is its **limit**, as $n \rightarrow \infty$?

OLS and Unbiasedness (Finite Sample)

Condition on $\mathbf{X}$ (treated as fixed) and take the expectation of $\hat{\beta} = \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon$:

$$E[\hat{\beta} \mid \mathbf{X}] = \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' E[\varepsilon \mid \mathbf{X}]$$

- Under the exogeneity condition $E[\varepsilon \mid \mathbf{X}] = \mathbf{0}$, the OLS coefficient is **unbiased**, for any sample size n (as long as $\mathbf{X}'\mathbf{X}$ is invertible):

$$E[\hat{\beta} \mid \mathbf{X}] = \beta \implies E[\hat{\beta}] = \beta.$$

- If X is endogenous, $E[\varepsilon \mid \mathbf{X}] \neq \mathbf{0}$, and thus

$$E[\hat{\beta} \mid \mathbf{X}] - \beta = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' E[\varepsilon \mid \mathbf{X}] \neq \mathbf{0} :$$

OLS is **biased**.

OLS and Consistency (Large Sample)

Now, let us turn to the consistency of the estimator. Recall $\hat{\beta} = \beta + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\varepsilon$ and divide by n inside each factor:

$$\hat{\beta} = \beta + \left(\frac{1}{n}\mathbf{X}'\mathbf{X}\right)^{-1}\left(\frac{1}{n}\mathbf{X}'\varepsilon\right)$$

- By the law of large numbers, writing $\mathbf{x}'_i$ for the i -th row of $\mathbf{X}$ (so that $\mathbf{X}'\mathbf{X} = \sum_i \mathbf{x}_i\mathbf{x}'_i$ and $\mathbf{X}'\varepsilon = \sum_i \mathbf{x}_i\varepsilon_i$):

$$\frac{1}{n}\mathbf{X}'\mathbf{X} \xrightarrow{p} \mathbf{Q} = \mathbb{E}[\mathbf{x}_i\mathbf{x}'_i] \text{ (invertible),} \quad \frac{1}{n}\mathbf{X}'\varepsilon \xrightarrow{p} \mathbb{E}[\mathbf{x}_i\varepsilon_i]$$

- Hence, by the rules for plim,

$$\text{plim } \hat{\beta} = \beta + \mathbf{Q}^{-1} \mathbb{E}[\mathbf{x}_i\varepsilon_i]$$

- OLS is **consistent** iff $\mathbb{E}[\mathbf{x}_i\varepsilon_i] = \mathbf{0}$.

Which Exogeneity Condition for Which Property?

Strict exogeneity

$$E[\varepsilon \mid \mathbf{X}] = \mathbf{0}$$

- ▶ The mean of the error does not depend on the regressors.
- ▶ Needed for **unbiasedness**.

Orthogonality (weak exogeneity)

$$E[\mathbf{x}_i \varepsilon_i] = \mathbf{0}$$

- ▶ The error is uncorrelated with the regressors (K conditions).
 - ▶ Needed for **consistency**.
-
- ▶ Strict $\Rightarrow$ weak (law of iterated expectations): $E[\mathbf{x}_i \varepsilon_i] = E[\mathbf{x}_i E[\varepsilon_i \mid \mathbf{x}_i]] = \mathbf{0}$.
 - ▶ Weak $\nRightarrow$ strict: with x_i symmetric around 0 and $\varepsilon_i = x_i^2 - E[x_i^2]$, we have $E[x_i \varepsilon_i] = 0$ but $E[\varepsilon_i \mid x_i] \neq 0$.
 - ▶ Instrumental variables rely on the **weak** form: $E[\mathbf{z}_i \varepsilon_i] = \mathbf{0}$.

Consequences of Endogeneity

- ▶ If $E[\mathbf{x}_i \varepsilon_i] \neq \mathbf{0}$, OLS is **inconsistent** (and, in general, biased too).
- ▶ The problem **does not vanish with more data**: $\hat{\beta}$ converges to $\beta + \mathbf{Q}^{-1}E[\mathbf{x}_i \varepsilon_i] \neq \beta$, a wrong value estimated ever more precisely.
- ▶ If one regressor is endogenous, $E[\mathbf{x}_i \varepsilon_i]$ has one nonzero entry, but $\mathbf{Q}^{-1}$ mixes the entries, and thus:
 - **all** the coefficients may be inconsistent (not only the one attached to the endogenous regressor),
 - the exogenous controls are contaminated too, as soon as they are correlated with the endogenous regressor.
- ▶ Adding more controls does not help if the source of the problem is unobserved.

Three Classic Sources of Endogeneity

- 1 **Omitted variable**: a determinant of y that is correlated with some regressors is left in ε .
 - e.g., ability affects both schooling (x) and wages (y).
- 2 **Simultaneity / reverse causality**: x affects y but y also affects x .
 - e.g., prices and quantities in a market, police and crime.
- 3 **Measurement error** in a regressor: we observe a noisy version of the true one (attenuation bias).
 - e.g., self-reported years of schooling.

Omitted Variable Bias

True model, with $\mathbf{u}$ ($n \times 1$) an **unobserved** variable:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}\gamma + \mathbf{v}, \quad E[\mathbf{x}_i v_i] = \mathbf{0}, \quad \boldsymbol{\varepsilon} = \mathbf{u}\gamma + \mathbf{v}$$

Then $E[\mathbf{x}_i \boldsymbol{\varepsilon}_i] = E[\mathbf{x}_i u_i] \gamma$, and:

$$\text{plim } \hat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \gamma \underbrace{\mathbf{Q}^{-1} E[\mathbf{x}_i u_i]}_{\boldsymbol{\delta}}$$

- ▶ $\boldsymbol{\delta}$: coefficients of the (population) regression of the omitted variable u on **all** the included regressors.
- ▶ Bias $\neq 0$ iff u affects y ($\gamma \neq 0$) **and** is correlated with the regressors ($\boldsymbol{\delta} \neq \mathbf{0}$).
- ▶ Example (wage equation, $u = \text{ability}$): $\gamma > 0$ and $\delta_{\text{schooling}} > 0 \Rightarrow \text{OLS overstates the return to schooling (ability bias)}$.

Omitted Variable Bias: When u Does Not Affect y

Same setting, but the omitted variable u has **no effect on** y once $\mathbf{X}$ is controlled for ($\gamma = 0$):

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{u}\gamma + \mathbf{v} \quad \text{with} \quad \gamma = 0 \quad \implies \quad \mathbf{y} = \mathbf{X}\beta + \mathbf{v}, \quad \varepsilon = \mathbf{v}$$

Then $E[\mathbf{x}_i \varepsilon_i] = E[\mathbf{x}_i v_i] = \mathbf{0}$, and:

$$\text{plim } \hat{\beta} = \beta + \underbrace{\gamma}_{=0} \delta = \beta$$

- ▶ OLS is **consistent**, **even if** u is strongly correlated with the regressors ($\delta \neq \mathbf{0}$): leaving u out is harmless.
 - u is not hidden in the error term. What matters is whether the omitted variable is a **determinant of** y , not whether it is correlated with $\mathbf{X}$.
- ▶ Be careful: $\gamma = 0$ is a statement about the **true model**, holding $\mathbf{X}$ fixed. u may still be correlated with y marginally, through $\mathbf{X}$.
- ▶ Bias requires **both** conditions: $\gamma \neq 0$ **and** $\delta \neq \mathbf{0}$. If either one is zero, there is no omitted variable bias.

2.3. A first look at instrumental variables

Back to the wage equation

- ▶ The model we want to estimate:

$$\ln w_i = \alpha + \beta \text{educ}_i + \varepsilon_i, \quad \varepsilon_i = \gamma u_i + v_i,$$

where u_i is **ability** (unobserved) and v_i is a wage shock.

- ▶ OLS converges to

$$\text{plim } \hat{\beta}_{OLS} = \beta + \frac{\text{Cov}(\text{educ}_i, \varepsilon_i)}{\text{Var}(\text{educ}_i)}.$$

- ▶ If more able people study longer, $\text{Cov}(\text{educ}_i, u_i) > 0$, so $\text{Cov}(\text{educ}_i, \varepsilon_i) \neq 0$: OLS is **inconsistent**, and more data does not help.
- ▶ The same happens with measurement error: the noise in educ_i is also in ε_i .
- ▶ The problem: **part** of the variation in schooling comes from ability. OLS cannot separate the effect of schooling from the effect of ability.

One way out: an instrument

- ▶ The idea consists in finding a variable z_i that **moves schooling for reasons unrelated to ability**, and using only that part of the variation.
- ▶ Example in today's practice: a **scholarship offer drawn by a coin toss**.
- ▶ Two conditions are required for the instrument:
 - **relevance**: $\text{Cov}(z_i, \text{educ}_i) \neq 0$ (the offer changes schooling);
 - **exogeneity**: $\text{Cov}(z_i, \varepsilon_i) = 0$ (the offer is unrelated to ability and affects wages only through schooling).
- ▶ Take the covariance of both sides of the wage equation with z_i :

$$\text{Cov}(z_i, \ln w_i) = \beta \text{Cov}(z_i, \text{educ}_i) + \underbrace{\text{Cov}(z_i, \varepsilon_i)}_{=0} \implies \beta = \frac{\text{Cov}(z_i, \ln w_i)}{\text{Cov}(z_i, \text{educ}_i)}.$$

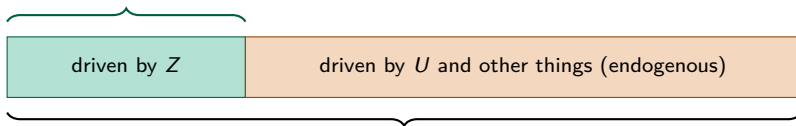
How much the offer moves wages, relative to how much it moves schooling.

Use Only the “Clean” Part of the Variation in X

IV uses

only this part

Variation in X



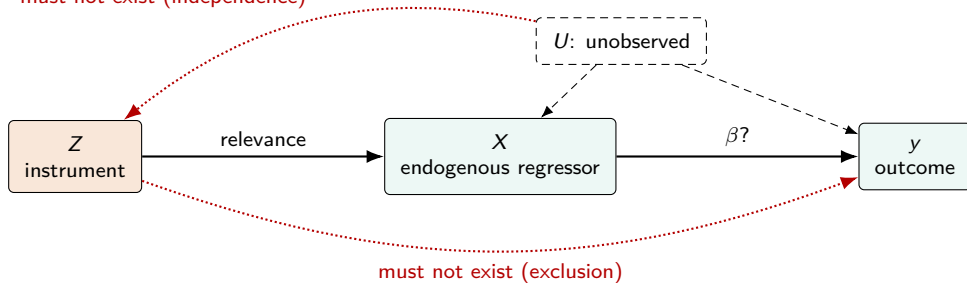
all of it, including the contaminated part

OLS uses

- ▶ An **instrument** Z is a variable that shifts X but has no other route to y .
- ▶ The variation in X induced by Z is **as good as random** with respect to U : we follow it, and only it, to see what happens to y .

What Makes a Valid Instrument?

must not exist (independence)



1. Relevance : Z moves X .

► **Testable**: regress X on Z .

2. Exogeneity : Z is uncorrelated with ε .

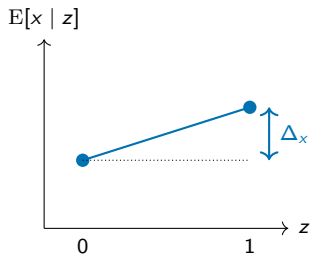
- **Independence**: Z as good as randomly assigned.
- **Exclusion**: Z affects y only through X .
- **Not testable** (ε is unobserved): it is an **argument**.

Binary Instrument: the Wald Estimator

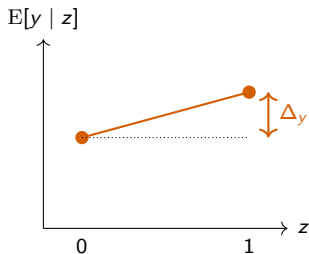
With a binary instrument $z \in \{0, 1\}$, the IV estimator is a ratio of two differences in means:

$$\hat{\beta}_{IV} = \frac{E[y | z = 1] - E[y | z = 0]}{E[x | z = 1] - E[x | z = 0]} = \frac{\text{reduced form}}{\text{first stage}}$$

“By how much does z move y , relative to how much it moves x ?”



first stage



reduced form

3. Practice: does OLS recover the truth? Can an instrument help?

Practice (in class): does OLS recover the truth?

The question

- ▶ We **simulate** wage data, so we know the true return to one year of schooling: $\beta = 0.08$.
- ▶ Does OLS find it? When the sample grows? Can an instrument help?

The data

- ▶ educ (years of schooling, as observed) and lnwage (log wage).
- ▶ Four cases, with hidden names: exog, ovb, irrelevant, measerr.
- ▶ Four sample sizes: $n = 100, 1\,000, 10\,000, 100\,000$.

How to work

- ▶ Download the Zip file that contains the tutorial. Open the R project file, run the chunks one by one.
- ▶ Fill in the blanks (___) and answer the questions.

Timing (≈ 1 hour)

- ▶ 0–15: Setting the environment
- ▶ 15–20: look at the data
- ▶ 20–25: OLS with `lm()` and by hand, $(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$
- ▶ 25–35: the four cases at $n = 1\,000$
- ▶ 35–45: convergence, when n grows
- ▶ 45–55: wrap-up

└ Practice: does OLS recover the truth? Can an instrument help?

Practice debrief (1/4): what is behind the four cases

$$\text{educ}_{\text{true}} = 12 + a u + z + e_x, \quad \ln w = 1.5 + 0.08 \text{educ}_{\text{true}} + \gamma u + v, \quad \text{educ} = \text{educ}_{\text{true}} + \text{err} \times m$$

Case	a (ability $\rightarrow$ educ.)	γ (ability $\rightarrow$ wage)	err	$\text{plim } \hat{\beta}$
exog	0	0.10	0	0.08
ovb	1	0.10	0	$0.08 + 0.10/3.5 \approx 0.109$
irrelevant	1	0	0	0.08
measerr	0	0	1	$0.08 \times 2.5/3.5 \approx 0.057$

- In every case, $\text{plim } \hat{\beta} = \beta + \text{Cov}(\text{educ}, \varepsilon) / \text{Var}(\text{educ})$.
- ovb: bias = $\gamma\delta$ requires **both** $\gamma \neq 0$ and $\delta \neq 0$ (compare with irrelevant).
- measerr: the noise is in educ **and** in $\varepsilon \Rightarrow$ **attenuation bias**, towards zero.

Practice debrief (2/4): more data does not fix endogeneity

- ▶ The standard error shrinks like $1/\sqrt{n}$, in **every** case.
 - ▶ `exog` and `irrelevant` converge to 0.08. `ovb` and `measerr` converge to their plim, not to the truth.
 - ▶ In the `ovb` case:
 - $n = 100$: the 95% CI is about $[0.063, 0.128]$: it contains 0.08, only because it is wide;
 - $n = 100\,000$: the 95% CI is about $[0.108, 0.111]$: it excludes 0.08.
- ⇒ With more data, we become **more and more confident about the wrong number**.
- ▶ A narrow confidence interval measures **precision**, not **closeness to the truth**.

Practice debrief (3/4): an instrument to the rescue

Case ovb, $n = 10\,000$

- ▶ First stage: the offer raises schooling by ≈ 1.02 years; $F \approx 811$ (strong instrument).
- ▶ Reduced form: the offer raises log wages by ≈ 0.082 . Wald ratio: $0.082/1.02 \approx 0.080$.

Estimator	Estimate	Standard error	95% CI
OLS	0.107	0.0017	excludes 0.08
IV	0.080	0.0062	[0.068, 0.092]

Small samples

- ▶ ovb, $n = 100$: $F \approx 0.6$ (weak instrument), $IV = 0.167$, standard error 0.224: useless.
 - ▶ For $n \geq 10\,000$, IV is close to 0.08 in both cases, and the CI shrinks around the truth.
- ⇒ IV trades **precision** for **consistency**: it only uses the part of the variation in schooling driven by the offer.
- ⇒ Relevance can be tested (first stage). Exogeneity and exclusion cannot: here we know they hold because we simulated the data. In real applications, **they must be argued**.

Practice debrief (4/4): Real World Situation

With real data, ability is unobserved and β is unknown. Two ways out:

- 1 use the **structure of the data** (repeated observations, comparison groups): **today**, [Ashenfelter and Card \(1985\)](#);
- 2 find an **instrument**: you just did it, with a coin toss we simulated. **Next session**: the method in detail, and a real instrument ([Angrist and Krueger, 1991](#)).

4. Next session

How we will work with readings

- ▶ Before each session: one main reading, sometimes a shorter second one.
- ▶ You keep the **same five questions** as for today:
 - 1 What is the paper about?
 - 2 Why is this an important question?
 - 3 What is the empirical strategy (or the argument)?
 - 4 What are the main results?
 - 5 Why is this subject of interest to an economist?
- ▶ Each time, we add **one critical question**.
- ▶ In class, you explain the papers **first**; the slides come after.
- ▶ Exercises are done **in class**. At home, you only read (and take notes of what you read).

The course: two questions

A. Does X cause Y ?

Measuring causal impact

- ▶ **S2**: the endogeneity problem; OLS under endogeneity (practice in R).
- ▶ **S3**: instrumental variables: intuition, practice in R, method; application.
- ▶ **S4**: the credibility revolution; randomized trials, natural and quasi-natural experiments, regression discontinuity; panel data and difference-in-differences.

B. How much is X worth?

Assigning value in the absence of a market

- ▶ **S5**: valuation methods (stated, revealed, experimental); climate change and the price of carbon (IAMs, social cost of carbon).
- ▶ **S6**: measuring and valuing climate impacts (agriculture, health, sea-level rise, conflict).
- ▶ **S7**: behavioral economics: beliefs, social norms, nudges; wrap-up.

Reading assignment for Session 3 (October 14, 2026)

Topic: **Instrumental variables with real data**

Everyone reads [Angrist and Krueger \(1991, QJE\)](#), this time **beyond the introduction**:

- ▶ how the instrument works (birth date, school entry, compulsory schooling),
- ▶ the data,
- ▶ the evidence that the quarter of birth moves schooling (the first stage),
- ▶ the main results, compared with OLS.

You do not need to follow every table or every specification.

Keep your notes on [Ashenfelter and Card \(1985\)](#): we come back to it in Session 4.

Reading assignment for Session 3: what to prepare

Part 1. Summary (approximately 1 page; if you already wrote one, improve it)

- 1 What is the paper about?
- 2 Why is this an important question?
- 3 What is the empirical strategy? What data, which source of variation?
- 4 What are the main results?
- 5 Why is this subject of interest to an economist?

Part 2. Critical question

- ▶ What must be true about the quarter of birth for the IV estimate to be causal?
- ▶ Give one reason why it could fail.

Session 3 starts with **you** explaining the paper.

5. Appendix

References I

- Angrist, J. D. and Krueger, A. B. (1991). Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics* 106: 979–1014, doi: 10.2307/2937954.
- Ashenfelter, O. and Card, D. (1985). Using the longitudinal structure of earnings to estimate the effect of training programs. *The Review of Economics and Statistics* 67: 648, doi: 10.2307/1924810.
- Evans, D. (2020). How to write the introduction of your development economics paper. Center for Global Development blog. <https://www.cgdev.org/blog/how-write-introduction-your-development-economics-paper>.
- Head, K. (n.d.). The introduction formula. University of British Columbia.
<https://blogs.ubc.ca/khead/research/research-advice/formula>.
- Mincer, J. A. (1974). Schooling and earnings. In *Schooling, experience, and earnings*. NBER, 41–63.