Instrumental Variables

Edicson Luna

University of Pennsylvania

Motivation: the draft lottery

Where we left off

Remember the last thing we said about endogeneity: if $\text{Cov}(x_i, \epsilon_i) \neq 0$, the OLS estimator is **inconsistent**,

$$\hat{\beta} \xrightarrow{p} \beta + \Sigma^{-1}\mathbb{E}[x_i\epsilon_i] \neq \beta.$$

We said this "motivates the need for alternative estimation strategies". Today we study the most important one: *instrumental variables* (IV).

But instead of starting with the algebra, let's start with a famous causal question.

A causal question

Does serving in the military affect your later earnings?

The naive approach: compare the earnings of veterans and non-veterans,

$$\text{wage}_i = \beta_0 + \beta_1 \,\text{veteran}_i + \epsilon_i,$$

and read β_1 as the effect of service.

Why is this a bad idea?

The selection problem

Veterans are not a random group. Who serves is correlated with things we do not observe:

- Health, motivation, family background, schooling, local labor markets . . .

All of these also affect earnings, so they sit inside ϵ_i. Then

$$\mathrm{Cov}(\mathrm{veteran}_i, \epsilon_i) \neq 0,$$

and $\mathrm{veteran}_i$ is endogenous. OLS confounds the effect of *service* with the effect of *who chooses to serve*.

We need a source of variation in military service that is *not* driven by the person's own characteristics.

The Vietnam draft lottery

Angrist (1990) found exactly such a source. During the Vietnam era, who was drafted was decided by a lottery over birth dates.

- Each date of birth got a random lottery number.
- Low numbers were draft-eligible; high numbers were not.
- Eligibility was assigned *by chance*, not by the person.

Call this random assignment z_i (draft-eligible or not). It is going to be our **instrument**.

Why the lottery helps

The lottery number z_i has three features that make it useful:

1. It is random. By construction it is unrelated to the person's health, ability, or background.
2. It moves service. Draft-eligible men were more likely to end up serving. So z_i is correlated with veteran_i.
3. It only touches earnings through service. A birth-date lottery number has no direct effect on wages; its only channel to earnings is via military service.

These are exactly the ingredients of an instrument. Let's see how they identify an effect.

The Wald computation

Compare draft-eligible and draft-ineligible men. Since eligibility is random, any difference in their average earnings must come *through* the difference in their service rates. So

$$\text{effect of service} = \frac{\overbrace{\mathbb{E}[\text{wage} \mid z = 1] - \mathbb{E}[\text{wage} \mid z = 0]}^{\text{reduced form: effect of eligibility on wage}}}{\underbrace{\mathbb{P}(\text{vet} \mid z = 1) - \mathbb{P}(\text{vet} \mid z = 0)}_{\text{first stage: effect of eligibility on service}}}.$$

This ratio is the Wald estimator. Angrist found that serving *lowered* the earnings of white veterans by roughly 15%.

The intuition: scale the (random) effect of eligibility on wages by how much eligibility actually changed service.

Local Average Treatment Effect

Setup and notation

Let's formalize with the potential-outcomes language from before. For each individual:

- $Z_i \in \{0, 1\}$: the instrument (draft-eligible or not).
- $D_i \in \{0, 1\}$: the treatment (veteran or not) — this is endogenous.
- Y_i: the outcome (earnings).

The novelty: the instrument gives treatment its own potential outcomes,

$$D_i(1) = \text{treatment if } Z_i = 1, \qquad D_i(0) = \text{treatment if } Z_i = 0.$$

That is, would this person serve *if* eligible, and *if* not?

Four types of people

Comparing $D_i(0)$ and $D_i(1)$ splits the population into four groups:

- **Compliers:** $D_i(0) = 0, \; D_i(1) = 1$. Serve *because* eligible.
- **Always-takers:** $D_i(0) = D_i(1) = 1$. Serve regardless (e.g. volunteers).
- **Never-takers:** $D_i(0) = D_i(1) = 0$. Never serve regardless.
- **Defiers:** $D_i(0) = 1, \; D_i(1) = 0$. Serve *only if not* eligible.

The instrument only *moves* the compliers. Always-takers and never-takers do the same thing either way, and defiers behave perversely.

The four assumptions

For the instrument to work, we assume:

1. **Independence.** Z_i is as good as random: independent of all potential outcomes. (The lottery.)
2. **Exclusion.** Z_i affects Y_i *only* through D_i. (A lottery number has no direct effect on wages.)
3. **Relevance.** Z_i actually shifts treatment: $\mathbb{P}(D = 1 \mid Z = 1) \neq \mathbb{P}(D = 1 \mid Z = 0)$.
4. **Monotonicity.** No defiers: $D_i(1) \geq D_i(0)$ for everyone.

The LATE theorem

Under those four assumptions, the Wald estimator identifies a very specific object:

$$\frac{\mathbb{E}[Y \mid Z=1] - \mathbb{E}[Y \mid Z=0]}{\mathbb{P}(D=1 \mid Z=1) - \mathbb{P}(D=1 \mid Z=0)} = \mathbb{E}[\, Y_i(1) - Y_i(0) \mid \text{complier}\,].$$

This is the **Local Average Treatment Effect** (LATE): the average causal effect of treatment, *but only for compliers*.

Monotonicity is what removes the always-takers and never-takers from the ratio: their treatment does not respond to Z, so they cancel out.

Why "Local"?

LATE is not the effect for everyone. It is *local* to the compliers — the people whose treatment the instrument actually moved.

- It is not the ATE (average over the whole population).
- We learn nothing about always-takers or never-takers: the instrument never changed their treatment.

A consequence: different instruments recover different LATEs, because they define different sets of compliers. The draft lottery tells us about men whose service was shaped by eligibility — not about career officers.

Instrumental variables in general

Back to the regression

The draft lottery is one instance of a general tool. Return to the endogenous regression

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i, \qquad \mathrm{Cov}(x_i, \epsilon_i) \neq 0.$$

We want a variable z_i — an **instrument** — that plays the

same role the lottery did: it isolates variation in x_i that is unrelated to ϵ_i.

What makes a valid instrument

An instrument z_i must satisfy two conditions:

- **Relevance.** It is correlated with the endogenous regressor:

$$\mathrm{Cov}(z_i, x_i) \neq 0.$$

- **Exogeneity (validity).** It is uncorrelated with the error:

$$\mathrm{Cov}(z_i, \epsilon_i) = 0.$$

Exogeneity bundles two ideas from before: the instrument is as good as random, and it affects y_i *only* through x_i (the exclusion restriction).

Deriving the IV estimator

Take the covariance of the equation with z_i:

$$\mathrm{Cov}(z_i, y_i) = \beta_1 \,\mathrm{Cov}(z_i, x_i) + \underbrace{\mathrm{Cov}(z_i, \epsilon_i)}_{=0}.$$

Since $\mathrm{Cov}(z_i, x_i) \neq 0$ (relevance), we can solve for β_1:

$$\beta_1 = \frac{\mathrm{Cov}(z_i, y_i)}{\mathrm{Cov}(z_i, x_i)}.$$

Replacing population moments by sample analogues gives the IV estimator

$$\hat{\beta}_1^{IV} = \frac{\sum_{i=1}^{n}(z_i - \bar{z})(y_i - \bar{y})}{\sum_{i=1}^{n}(z_i - \bar{z})(x_i - \bar{x})}.$$

Consistency, and a familiar special case

By the LLN, sample covariances converge to population covariances, so

$$\hat{\beta}_1^{IV} \xrightarrow{p} \frac{\text{Cov}(z_i, y_i)}{\text{Cov}(z_i, x_i)} = \beta_1.$$

IV is consistent even though OLS was not. The exogenous instrument breaks the correlation with ϵ_i.

When z_i is a dummy, this formula collapses to

$$\hat{\beta}_1^{IV} = \frac{\bar{y}_1 - \bar{y}_0}{\bar{x}_1 - \bar{x}_0},$$

which is exactly the Wald estimator from the draft lottery. LATE was IV all along.

Two-stage least squares

With many regressors or more instruments than needed, the same idea is implemented as two-stage least squares (2SLS):

- **First stage:** regress the endogenous x_i on the instrument(s) z_i and keep the fitted values $\hat{x}_i$. This is the "clean" part of x_i, driven only by z_i.
- **Second stage:** regress y_i on $\hat{x}_i$.

In matrix form, for the just-identified case, $\hat{\beta}^{IV} = (Z'X)^{-1}Z'Y$. The logic is always the same: throw away the endogenous variation in x, keep only the part the instrument explains.

Using IV in practice

The two conditions are not symmetric

- **Relevance** is *testable*: we can look at the first stage and check that z_i really predicts x_i (e.g. the first-stage F-statistic).
- **Exogeneity** is (mostly) *not testable*: $\text{Cov}(z_i, \epsilon_i) = 0$ involves the unobserved error. It must be defended with economic argument, not a regression.

In the draft example, one must argue that eligibility changed earnings *only* through service — and worry about channels like draft-avoidance schooling that would violate exclusion.

Weak instruments

Relevance is not enough on its own: the instrument must be *strongly* relevant.

If $\mathrm{Cov}(z_i, x_i)$ is close to zero, the denominator of the IV estimator is nearly zero, so

- the estimator is very imprecise (huge standard errors), and
- even a tiny violation of exogeneity gets amplified.

This is the weak instrument problem. A common rule of thumb is to worry when the first-stage F-statistic is below about 10.

Takeaway

- When a regressor is endogenous, OLS is inconsistent; an instrument can restore consistency.
- A valid instrument is relevant ($\text{Cov}(z, x) \neq 0$) and exogenous ($\text{Cov}(z, \epsilon) = 0$, affecting y only through x).
- The draft lottery instruments military service; the Wald ratio (reduced form over first stage) is the simplest IV.
- With heterogeneous effects, IV identifies a LATE: the effect for *compliers*, not the whole population.
- Relevance is testable; exogeneity must be argued. Beware weak instruments.