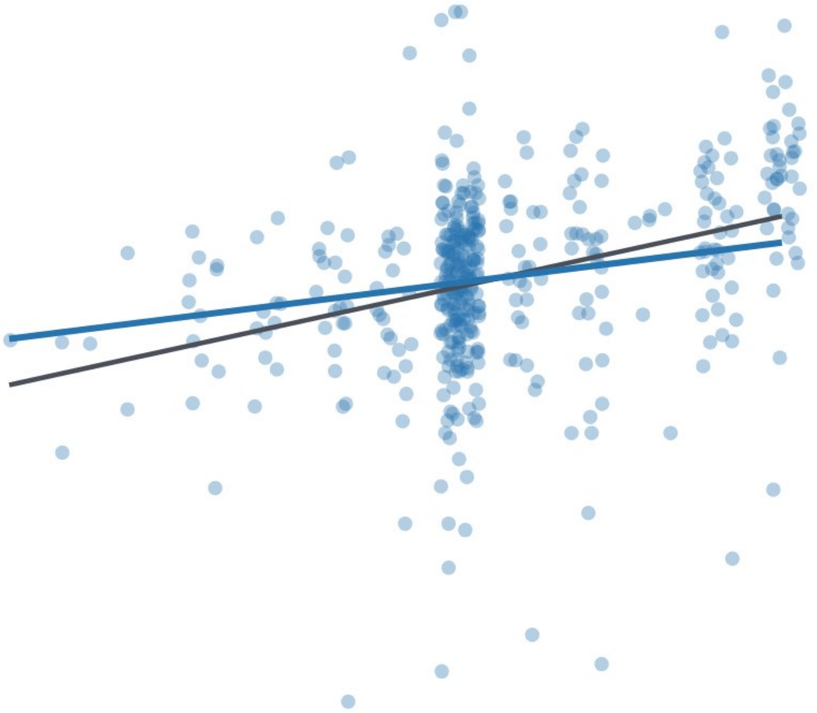
ECON 104 · LECTURE 12 · FALL 2026

Instrumental Variables: Identification, Specification Tests, and Simultaneous Equations



Jake Anderson

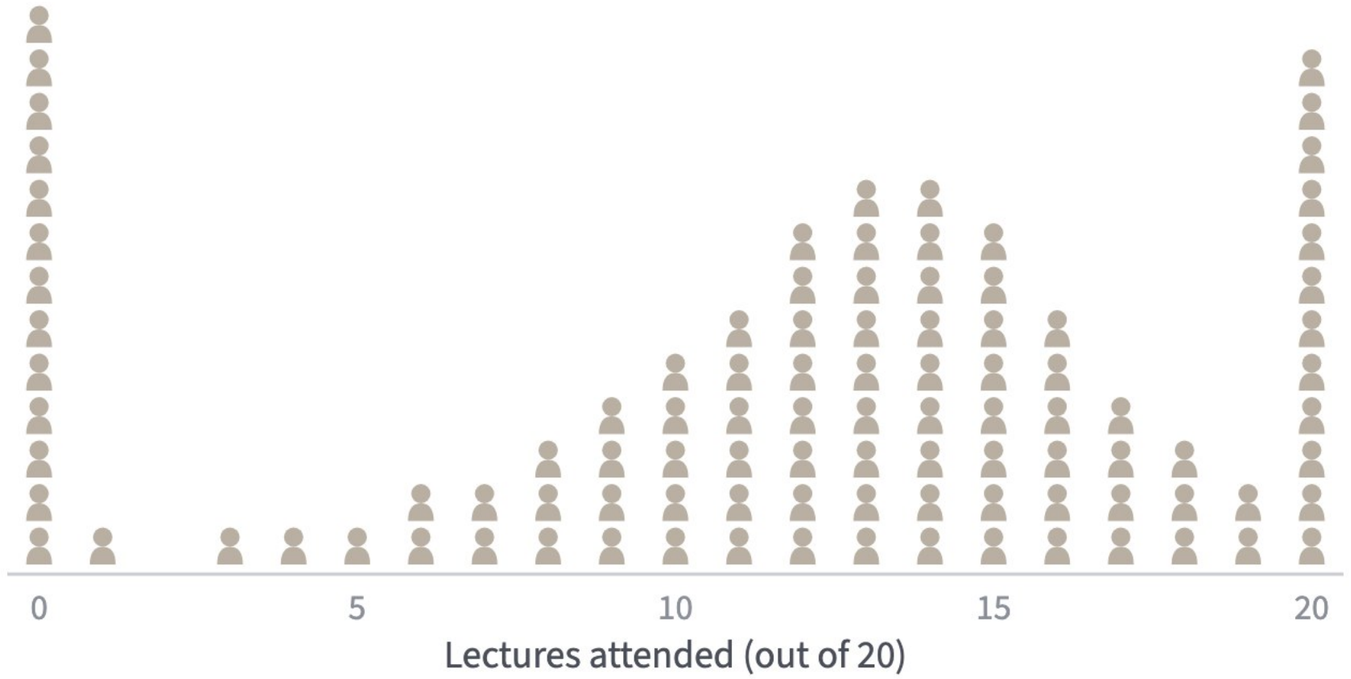
Guest lecture · UCLA

Plan for Today

1. **IV Idea**
2. **In Practice**
3. **2SLS**
4. **Tests**
5. **Simultaneity**

Does Coming to Lecture Raise Your Grade?

$$\underbrace{Y_i}_{\text{final grade}} = \beta_0 + \beta_1 \underbrace{X_i}_{\text{lectures attended}} + \varepsilon_i$$



One quarter of this class: 100 students and 20 lectures. Each figure is one student, stacked by the number of lectures the student attended.

THE REGRESSION

Regress each student's final grade (out of 100) on the number of lectures attended.

MY CLAIM

$\hat{\beta}_1$ is the causal effect of attending one more lecture on the final grade.

My Claim

$$\underbrace{Y_i}_{\text{final grade}} = \beta_0 + \beta_1 \underbrace{X_i}_{\text{lectures attended}} + \varepsilon_i$$

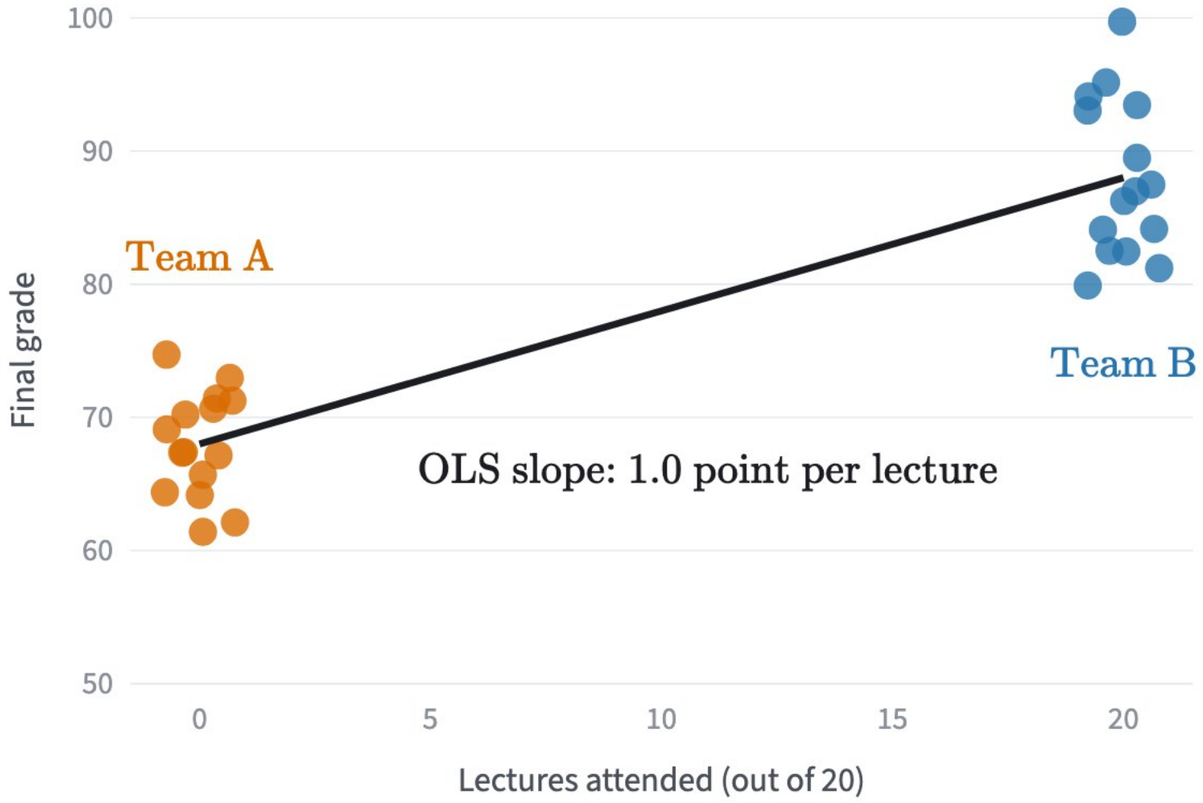
$\hat{\beta}_1$ is the causal effect of attending one more lecture on the final grade.

Is it causal?

Tell me a story that disproves my claim.

A Story: Teaching at Athlete University

$$\underbrace{Y_i}_{\text{final grade}} = \beta_0 + \beta_1 \underbrace{X_i}_{\text{lectures attended}} + \varepsilon_i$$



Every student is on one of two teams.

TEAM A

Practice is at lecture time.

$\Longrightarrow$ Never comes.

TEAM B

The coach requires lecture and studying.

$\Longrightarrow$ Always comes.

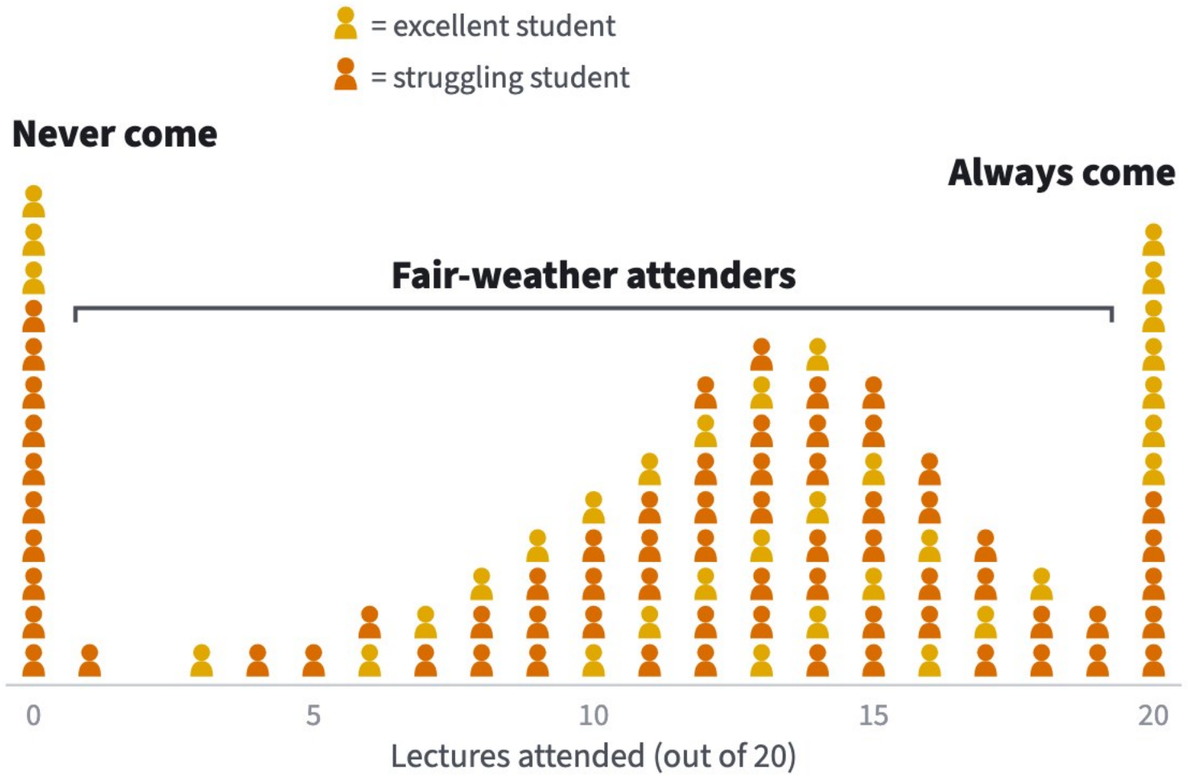
THE REGRESSION

OLS slope: 1.0 point per lecture.

MIXED $\hat{\beta}_1$ captures both the effect of lectures and the differences between the two teams.

Who Comes to Lecture?

$$\underbrace{Y_i}_{\text{final grade}} = \beta_0 + \beta_1 \underbrace{X_i}_{\text{lectures attended}} + \varepsilon_i$$



NEVER COME

Some expect to fail. Others are excellent and do not need to come.

ALWAYS COME

Some struggle and need the help. Others are excellent.

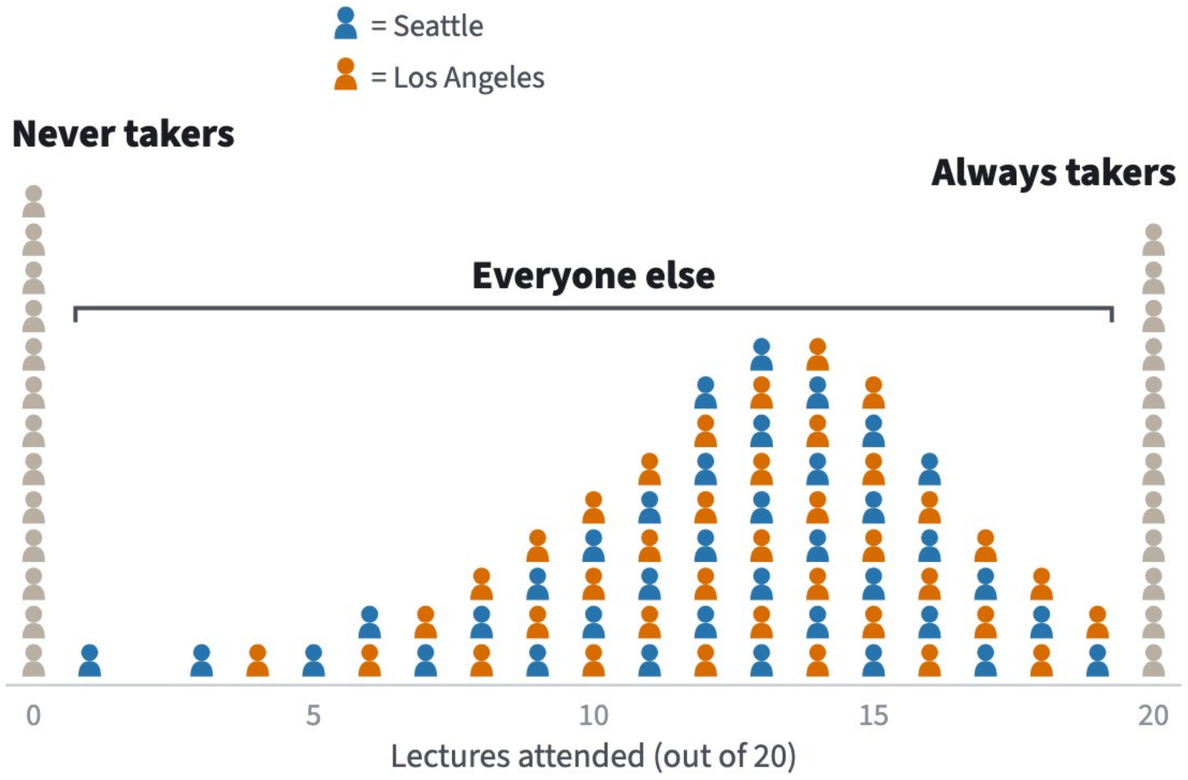
FAIR-WEATHER ATTENDERS

Everyone else: they come when it is easy and skip when it is not.

Attendance depends on ability and motivation, which are in ε_i. So $\text{cov}(X_i, \varepsilon_i) \neq 0$ and OLS is biased. $\implies$ We need a reason to miss lecture that has nothing to do with the student.

Seattle and Los Angeles: Who Is Moved by Rain?

$$\underbrace{Y_i}_{\text{final grade}} = \beta_0 + \beta_1 \underbrace{X_i}_{\text{lectures attended}} + \varepsilon_i$$



Half the class is from Seattle and half from Los Angeles. Both cities send the same mix of ability.

NEVER TAKERS

Do not come, whatever the weather.

ALWAYS TAKERS

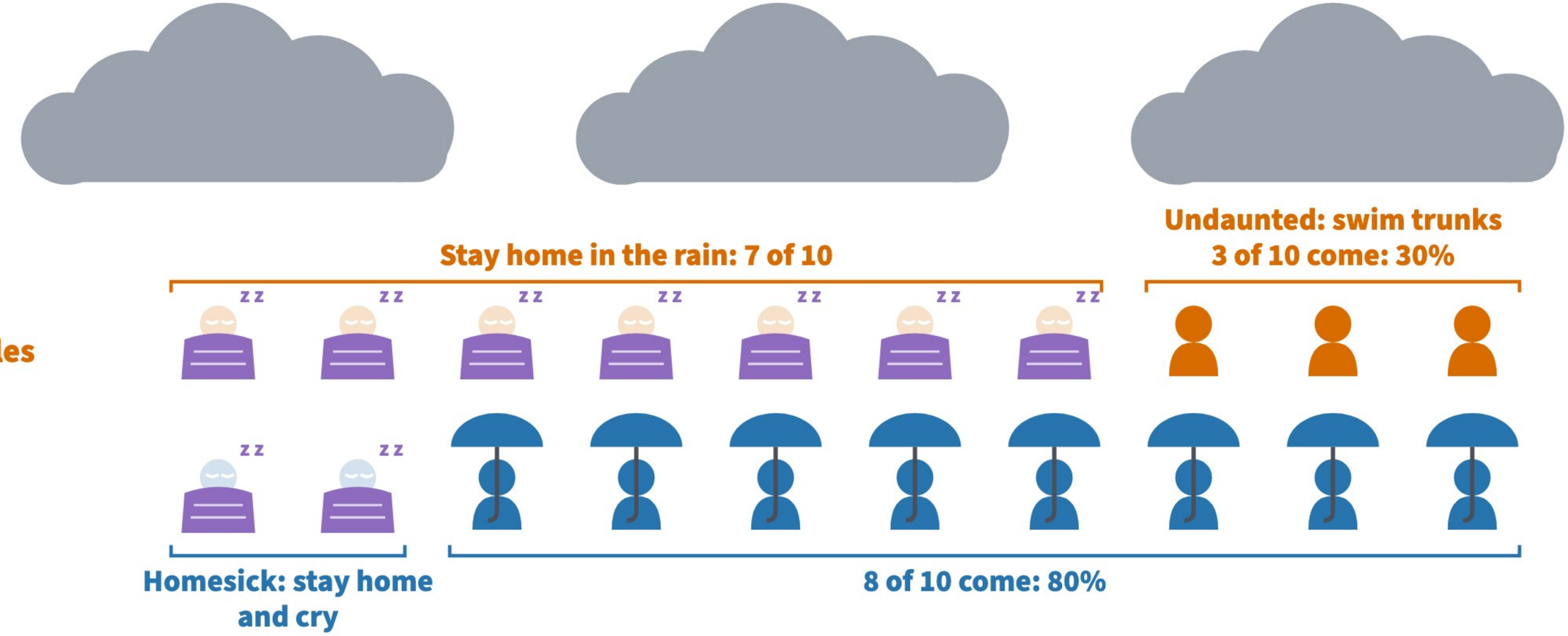
Come, whatever the weather.

EVERYONE ELSE

Seattle students (blue) own umbrellas. LA students (orange) do not.

Where you are from changes attendance in the rain. It does not change ability.

Everyone Else: How Much Rain Keeps You Home?



Rain tolerance is unrelated to academic ability: same mix of high and low performers.

An Instrument Needs Two Conditions

Relevance

The instrument does change the treatment. Rain actually affects who comes to class and who doesn't.

Exogeneity

The instrument affects grades only through X. It can't also be connected directly to Y.

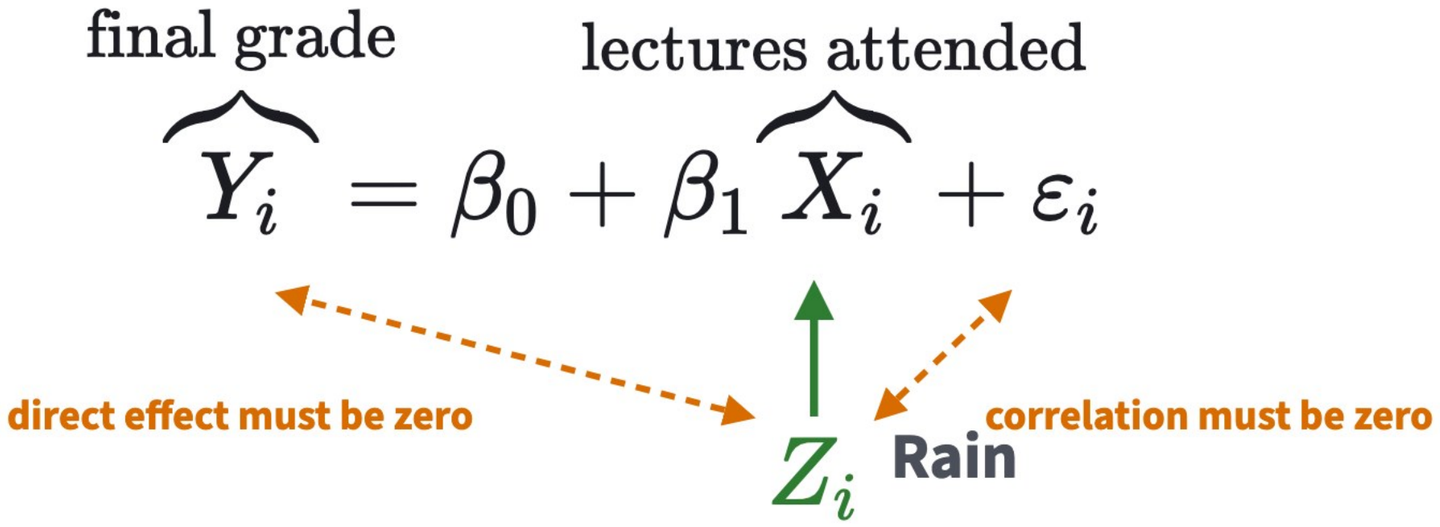
Relevance

$$\overbrace{Y_i}^{\text{final grade}} = \beta_0 + \beta_1 \overbrace{X_i}^{\text{lectures attended}} + \varepsilon_i$$



- The instrument (rain) changes the treatment (class attendance).
- $\text{Cov}(Z_i, X_i) \neq 0$

Exogeneity



- The instrument (rain) affects grades only through attendance.

  Threat: rain keeps students indoors, so they study more at home.

- $\text{Cov}(Z_i, \varepsilon_i) = 0$

  Threat: determination in the rain may go with determination in class and academics.

2SLS: Keep Only the Exogenous Part of X

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

SPLIT X

$$X_i = \overbrace{X_i^{\text{exog}}}^{\text{moved by rain}} + \overbrace{X_i^{\text{endog}}}^{\text{sorting: ability, motivation}}$$

FIRST STAGE

$$X_i = \pi_0 + \pi_1 Z_i + v_i \implies \overbrace{\hat{X}_i}^{\text{exogenous part of } X} = \hat{\pi}_0 + \hat{\pi}_1 Z_i$$

SECOND STAGE

$$Y_i = \beta_0 + \beta_1 \overbrace{\hat{X}_i}^{\text{no sorting}} + \varepsilon_i$$

Our conclusion is only valid for “compliers”: people who are influenced by the rain.
Always takers and never takers do not provide any variation.

First Stage: Does Rain Change Attendance?

$$X_i = \pi_0 + \pi_1 Z_i + v_i$$

X_i = lectures attended · $Z_i = 1$ in the rainy winter, 0 in the dry winter

	Estimate	Std. error	t
Intercept	13.00	0.35	36.7
Rainy winter, Z_i	**−2.00**	0.50	−4.0

$N = 160$ · $R^2 = 0.09$ · $F =$ **16.0**

- $\hat{\pi}_1 = -2.0$: rain lowers attendance by 2.0 lectures.
- F is above 10: rain is a strong instrument.

Reduced Form: Does Rain Change Grades?

$$Y_i = \gamma_0 + \gamma_1 Z_i + u_i$$

Y_i = final grade · $Z_i = 1$ in the rainy winter, 0 in the dry winter

	Estimate	Std. error	t
Intercept	77.00	0.82	93.7
Rainy winter, Z_i	**−1.00**	1.16	−0.9

$N = 160$ · $R^2 = 0.00$

- $\hat{\gamma}_1 = -1.0$: grades are 1.0 point lower in the rainy winter.
- This is the effect of rain on grades, with no attendance in the regression.

OLS: Grade on Attendance

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

	Estimate	Std. error	t
Intercept	55.86	1.39	40.2
Lectures attended, X_i	**1.72**	0.11	15.4

$N = 160$ · $R^2 = 0.60$

- OLS: one more lecture goes with 1.7 more points.
- But ability and motivation are in ε_i, so this slope is biased.
- This model is trained on data that includes people who will never change their attendance, so it is picking up groups that have inherent differences.

The IV Estimate: A Ratio

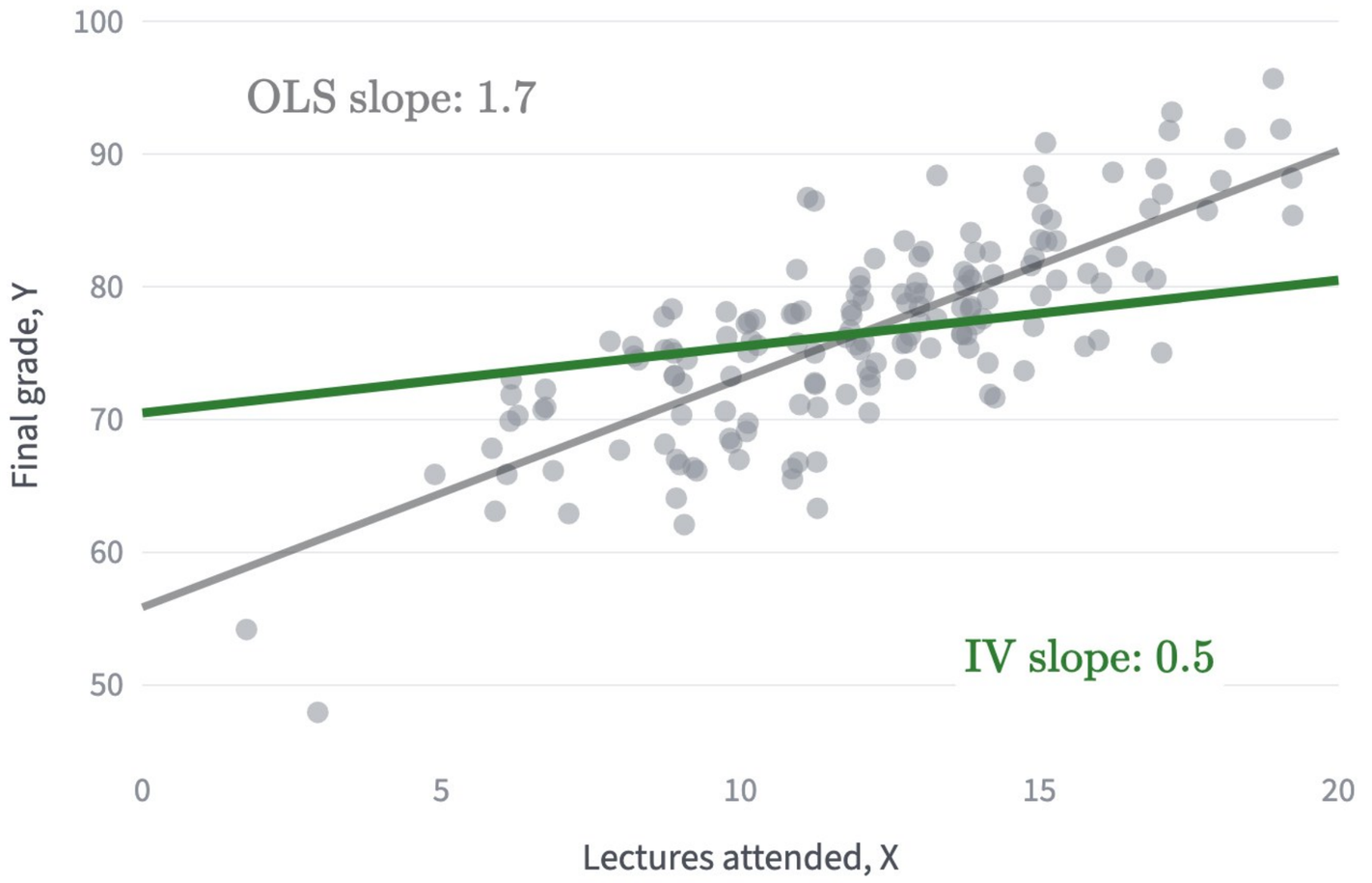
FIRST STAGE Rainy winter: attendance **2.0 lectures lower**.

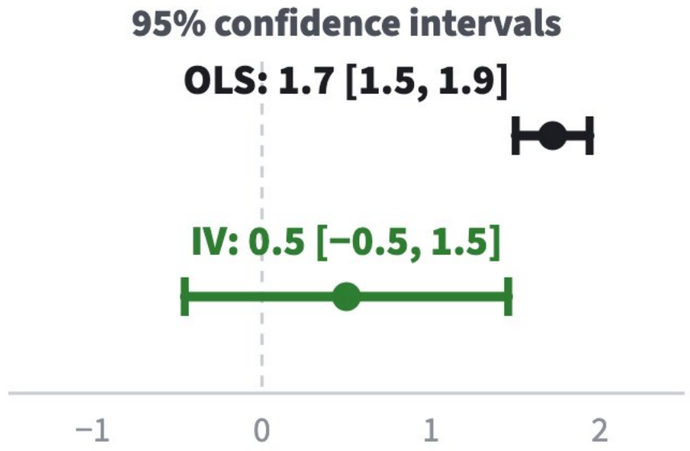
REDUCED FORM Rainy winter: grades **1.0 point lower**.

RATIO $$\hat{\beta}_{IV} = \frac{-1.0}{-2.0} = \mathbf{0.5} \text{ points per lecture}$$

IV VS. OLS OLS gives **1.7**. Too large: it also picks up ability and motivation.

OLS and IV on the Same Data





IV uses less variation.

$\Longrightarrow$ Wider interval.

Immigration Courts: Does Release Change the Outcome?

$$\overbrace{Y_i}^{\text{deported}} = \beta_0 + \beta_1 \overbrace{X_i}^{\text{released}} + \varepsilon_i$$

SETTING After a bond hearing, a judge either releases the person on bond or keeps the person in detention.

TREATMENT $X_i = 1$ if released. $X_i = 0$ if held in detention.

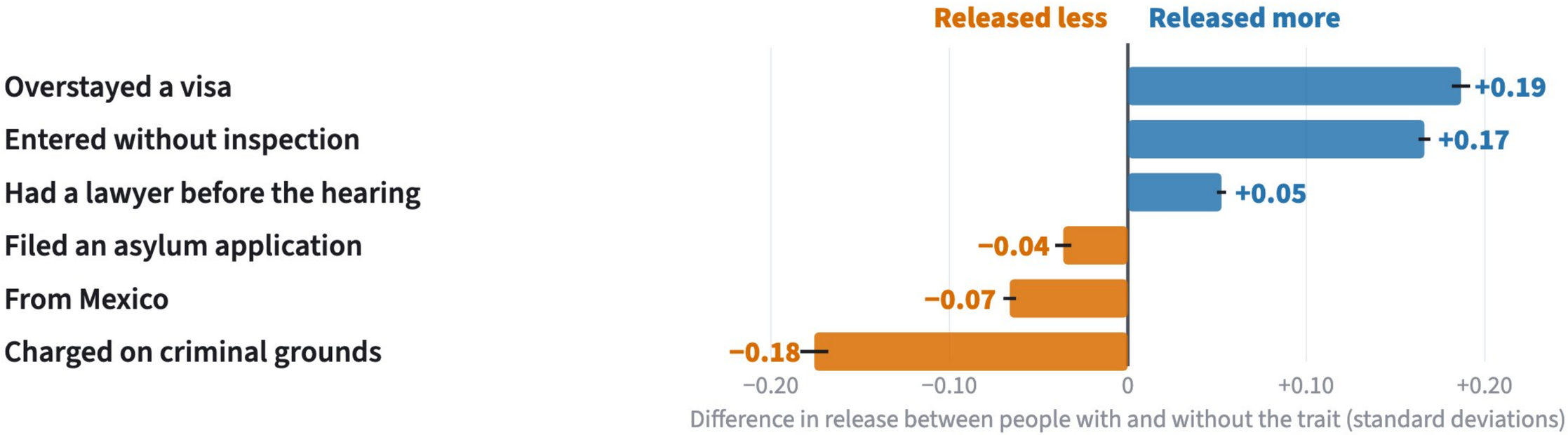
OUTCOME $Y_i = 1$ if deported.

$\hat{\beta}_1$ is the causal effect of release on being deported.

Is it causal?

Released vs. Held: The Groups Differ

$$\overbrace{Y_i}^{\text{deported}} = \beta_0 + \beta_1 \overbrace{X_i}^{\text{released}} + \varepsilon_i$$



EOIR case data, first bond hearings 2010 to 2026, within the same location and week. Released: released within 30 days of the hearing. 95% intervals are narrower than the dots.

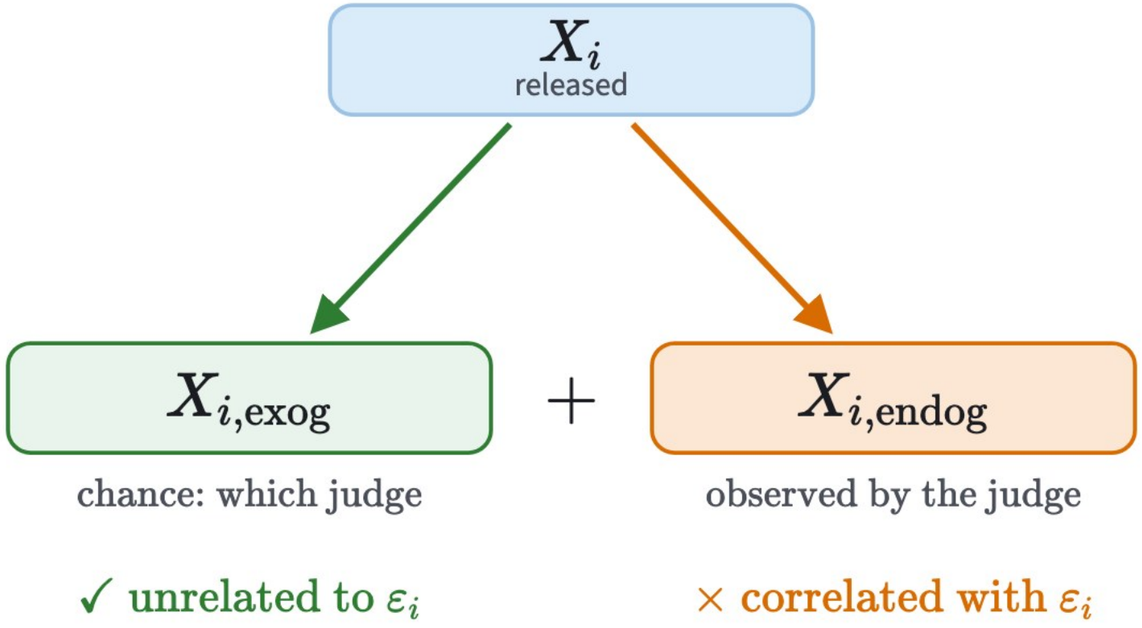
The groups are different $\Longrightarrow$ a naive comparison mixes release with who gets released.

THE IDEAL COMPARISON Two people alike in every way except chance: one released, one held.

The Naive Regression: Release Has Two Parts

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$$

$Y_i = 1$ if person i is deported. $X_i = 1$ if person i is released.



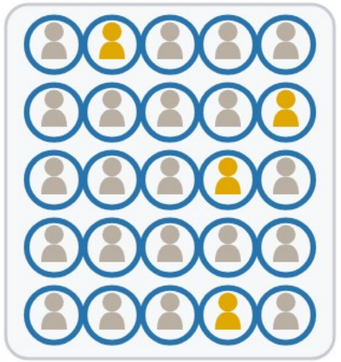
THE PROBLEM

Unseen things, like a strong case or family ties, make release more likely and also change deportation. They sit in ε_i, so $\text{cov}(X_i, \varepsilon_i) \neq 0$ and OLS is biased.

THE IDEA Keep only $X_{i,\text{exog}}$. $\Longrightarrow$ We need something that moves release by chance.

The “Ideal” Experiment: Release by Coin Flip

Heads: released



Record: 16%
Released: 100%

Tails: detained



Record: 16%
Released: 0%

Flip a coin for each person, whatever their case. Heads: released. Tails: detained.

ASSIGNMENT

Chance alone sorts people into two groups: a randomized controlled trial (RCT).

SAME MIX

Both groups have the same share with a criminal record (gold): 16%.

This is a perfect instrument. It is mechanically random.

Thought experiment: the dots illustrate the design and are not data.

= Criminal record = released

Question: Any problems with running this experiment?

Release by Coin Flip: Not Feasible

It is not ethical

Detention takes away a person’s freedom. We cannot detain people for no reason just to run an experiment.

Releasing at random is dangerous

Some people could harm others if released. Detaining them prevents that harm while they are held (incapacitation). A coin would release some of them.

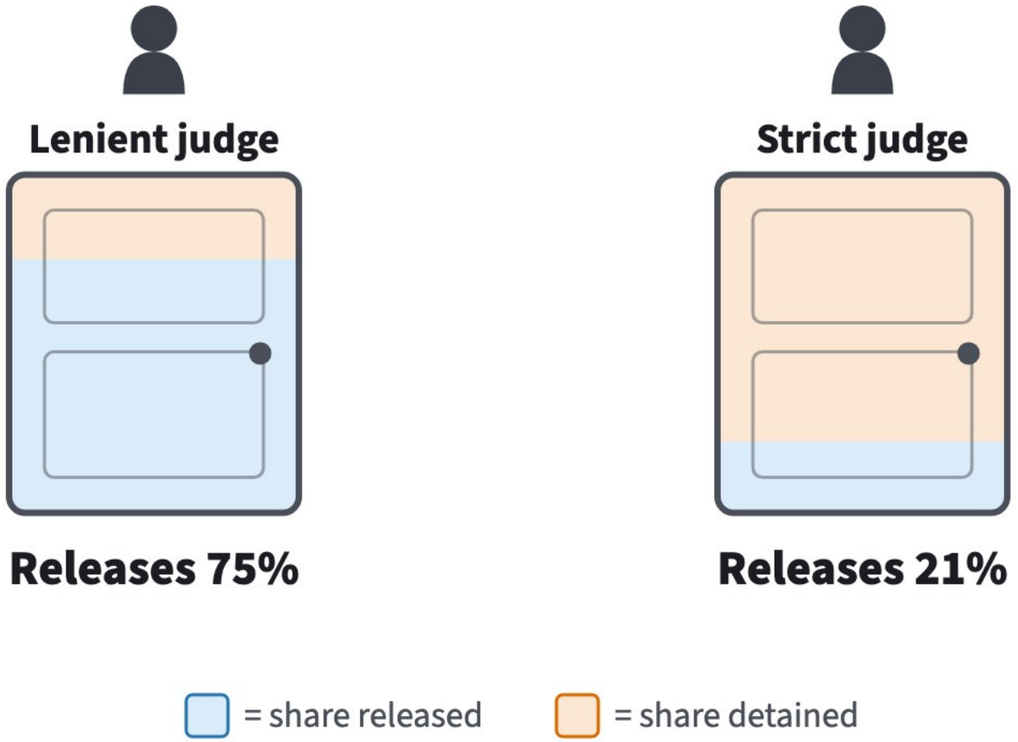
Detaining at random is harmful

Many people should not be detained at all. A coin would detain about half of them.

THE ALTERNATIVE

Something the court already does that works like the coin: it changes who is released, and it has nothing to do with the person.

Put a Judge over Each Door



A JUDGE OVER EACH DOOR

Replace the coin with a judge. Cases go to judges by the court's schedule, not by choice, so the judge works like the coin.

EXTREME JUDGES

If the lenient judge released everyone and the strict judge detained everyone, this would be the coin experiment.

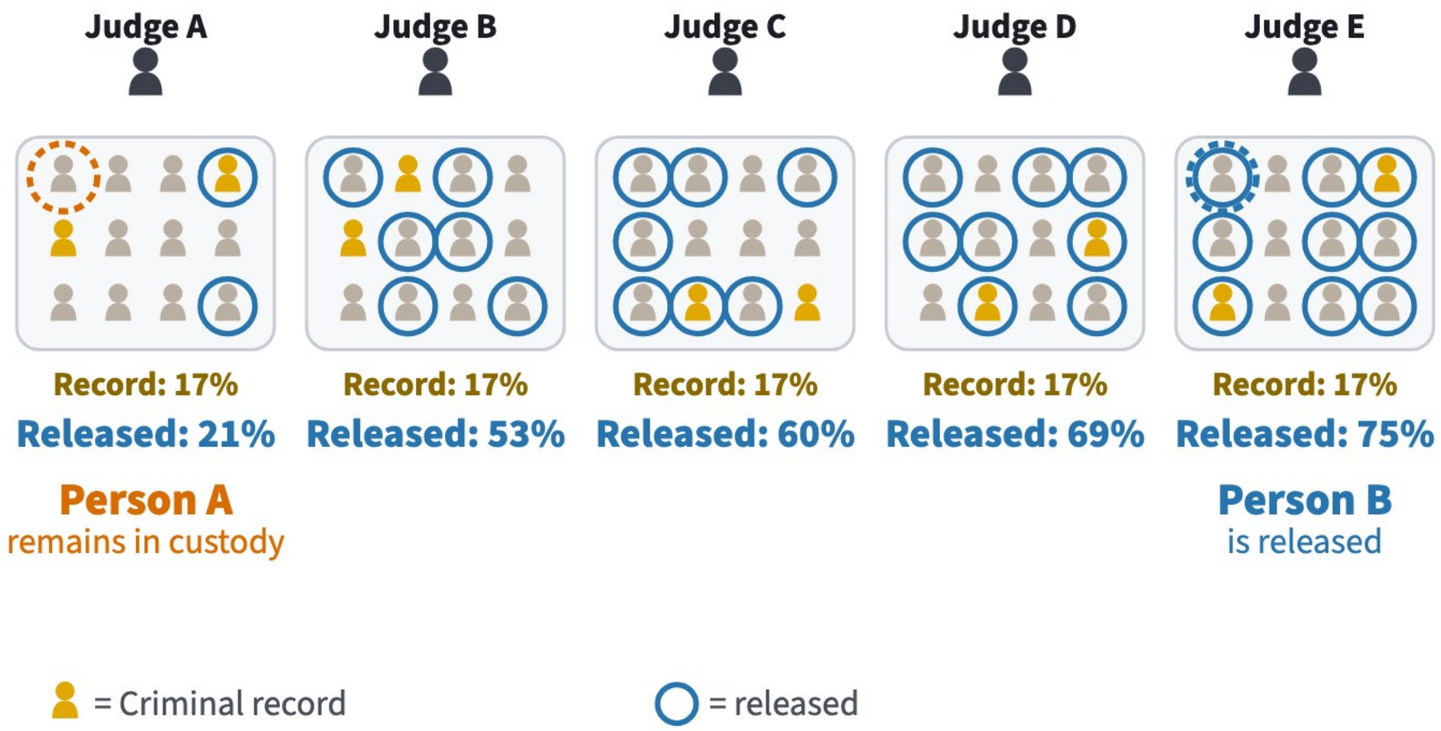
REAL JUDGES

No judge is that lenient or that strict. In one court, they release 75% and 21%.

Which judge you get changes your chance of release: here by 54 percentage points instead of 100.

Release rates: Port Isabel, TX, 2011.

Random Assignment: Who Hears Your Case?



ASSIGNMENT

Each person is assigned to one of five judges by docket, not by choice.

JUDGES GET SIMILAR CASES

Every room has the same mix: the same share with a criminal record (gold).

SOME JUDGES RELEASE MORE

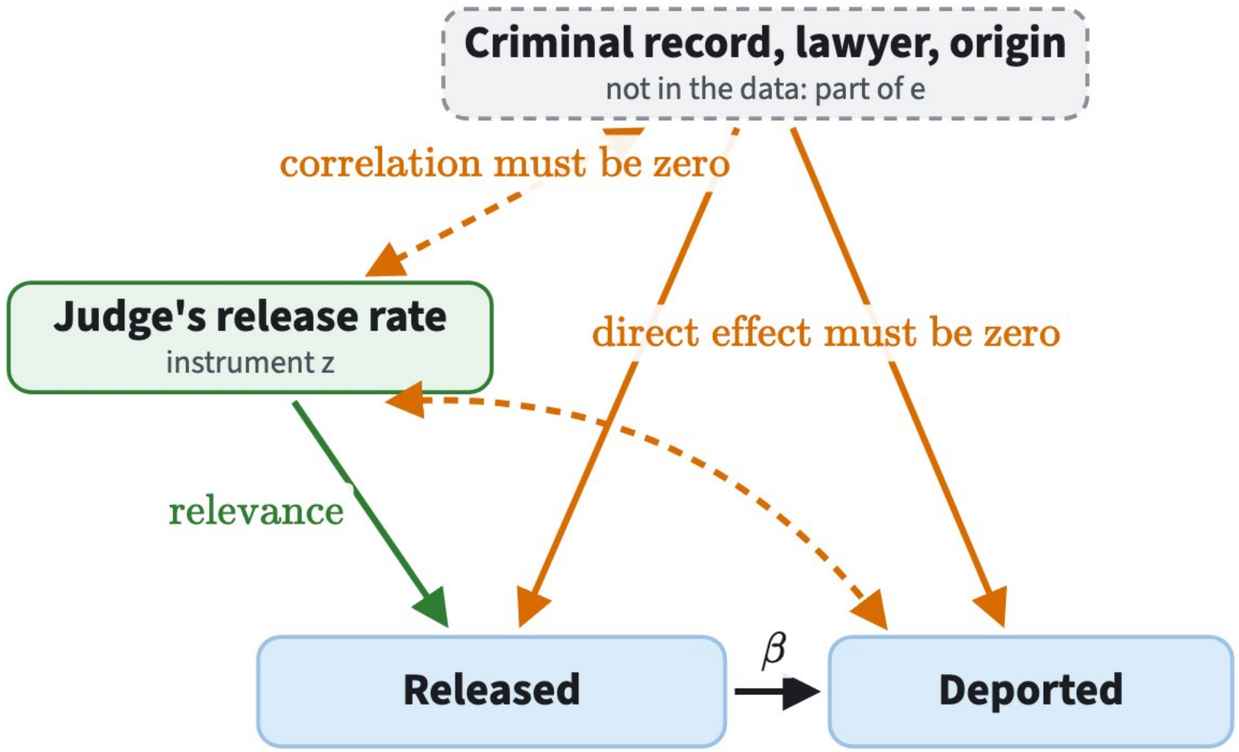
Judges differ: the strict judge releases 21%, the lenient judge 75% (blue ring: released).

TWO ALIKE PEOPLE

Person A gets the strict judge and stays in custody. Person B gets the lenient judge and is released. Chance made the difference.

Release rates are real (Port Isabel, TX, 2011). The dots are illustrative.

The Judge as an Instrument: Two Conditions



Release is not random, but the judge is. The judge's release rate is the instrument z.

Relevance
Some judges release more than others.

Exogeneity
Judges get similar cases. The judge affects deportation only through release.

PASSES EXOGENEITY The balance checks pass.

First Stage: Does the Judge Change Release?

$$\overbrace{X_i}^{\text{released}} = \pi_0 + \pi_1 \overbrace{Z_i}^{\text{judge release rate}} + v_i$$

$X_i = 1$ if released · Z_i = assigned judge's release rate, relative to the court-week average

	Estimate	Std. error	t
Judge release rate, Z_i	**0.772**	0.013	57.2

N = 443,699 · F = **3,274**

- $\hat{\pi}_1 = 0.77$: a judge who releases 10 points more often raises your chance of release by 7.7 points.
- F is far above 10: the judge is a strong instrument.

Reduced Form: Does the Judge Change Deportation?

$$\overbrace{Y_i}^{\text{deported}} = \gamma_0 + \gamma_1 \overbrace{Z_i}^{\text{judge release rate}} + u_i$$

$Y_i = 1$ if deported within 8 years · Z_i = assigned judge's release rate

	Estimate	Std. error	t
Judge release rate, Z_i	**−0.149**	0.013	−11.6

N = 443,699

- $\hat{\gamma}_1 = -0.15$: a judge who releases 10 points more often lowers your chance of being deported by 1.5 points.
- This is the effect of the judge on deportation, with release not in the regression.

The IV Estimate: A Ratio

FIRST STAGE A higher judge release rate raises release: **0.77**.

REDUCED FORM A higher judge release rate lowers deportations: -0.15.

RATIO

$$\hat{\beta}_{IV} = \frac{\text{reduced form}}{\text{first stage}} = \frac{\hat{\gamma}_1}{\hat{\pi}_1}$$

$$\nearrow\ Y_i = \gamma_0 + \gamma_1 Z_i + u_i$$

$$\searrow\ X_i = \pi_0 + \pi_1 Z_i + v_i$$

PLUG IN

$$\hat{\beta}_{IV} = \frac{-0.149}{0.772} = \mathbf{-0.19}$$

ANSWER Release lowers the chance of being deported by 19 points after 8 years (95% interval: -22 to -16). Held people in this group are deported 56% of the time.

IV vs. the Naive Comparison

$$\overbrace{Y_i}^{\text{deported}} = \beta_0 + \beta_1 \overbrace{X_i}^{\text{released}} + \varepsilon_i$$

OLS Released vs. held, no instrument: -0.23 (std. error 0.002).

IV Judge as instrument: -0.19, 95% interval $[-0.22, -0.16]$.

IV VS. OLS $\Longrightarrow$ The raw gap overstates the effect of release: it also reflects who gets released. The IV compares people the judge assignment treats alike.

2SLS: The Two Stages with the Judge

$$\overbrace{Y_i}^{\text{deported}} = \beta_0 + \beta_1 \overbrace{X_i}^{\text{released}} + \varepsilon_i$$

FIRST STAGE

$$X_i = \pi_0 + \pi_1 \overbrace{Z_i}^{\text{judge release rate}} + v_i \quad \Longrightarrow \quad \overbrace{\hat{X}_i}^{\text{release predicted by the judge}} = \hat{\pi}_0 + \underbrace{0.772}_{\hat{\pi}_1} Z_i$$

SECOND STAGE

$$Y_i = \beta_0 + \beta_1 \overbrace{\hat{X}_i}^{\text{no sorting on the case}} + \varepsilon_i \quad \Longrightarrow \quad \hat{\beta}_1 = \mathbf{-0.19}$$

One instrument: the second-stage slope equals the ratio. $\hat{\beta}_{2SLS} = \dfrac{-0.149}{0.772} = -0.19$

Admission Lottery: Does the Charter School Work?

$$\overbrace{\text{score}_i}^{\text{outcome}} = \beta_0 + \beta_1 \overbrace{\text{charter}_i}^{\text{treatment}} + \varepsilon_i$$

MEANING β_1: effect of attending the charter on test scores.

INSTRUMENT Z_i = won the admission lottery.

RELEVANCE Winners attend the charter far more often than losers.

EXOGENEITY Passes: names are drawn at random, so Z_i is unrelated to family motivation.

Naive (OLS): attendees are motivated families, so OLS mixes motivation with the school effect.
IV: about 0.4 standard deviations in math and 0.2 in reading per year of attendance.

Vietnam Draft Lottery: Does Service Change Earnings?

$$\overbrace{\text{earnings}_i}^{\text{outcome}} = \beta_0 + \beta_1 \overbrace{\text{served}_i}^{\text{treatment}} + \varepsilon_i$$

MEANING β_1: effect of military service on later earnings.

INSTRUMENT Z_i = low draft number (called up).

RELEVANCE Low numbers served more often, but the gap is far from 100%.

EXOGENEITY Partly: numbers were random, but men may stay in school to avoid the draft.

Naive (OLS): veterans differ from non-veterans in health, ability, and family.
IV: in the early 1980s, white veterans earned about 15% less than comparable non-veterans.

Quarter of Birth: Returns to Schooling

$$\overbrace{\text{earnings}_i}^{\text{outcome}} = \beta_0 + \beta_1 \overbrace{\text{schooling}_i}^{\text{treatment}} + \varepsilon_i$$

MEANING β_1: return to one more year of schooling.

INSTRUMENT Z_i = born early in the year.

RELEVANCE Weak: early births get about a tenth of a year less schooling.

EXOGENEITY Mostly: birth date looks unrelated to ability, but small differences in health and family appear by season.

Naive (OLS): more able people get more schooling, so OLS may overstate the return.
IV: close to the OLS return, so little sign of ability bias.

Distance to College: Returns to Schooling

$$\overbrace{\text{wage}_i}^{\text{outcome}} = \beta_0 + \beta_1 \overbrace{\text{schooling}_i}^{\text{treatment}} + \varepsilon_i$$

MEANING β_1: return to one more year of schooling.

INSTRUMENT Z_i = grew up near a college.

RELEVANCE Students near a college attend more often.

EXOGENEITY Caveats: families choose where to live. Needs controls for region, race, and parents' education.

Naive (OLS): 0.074 per year of schooling (standard error 0.004).
IV: 0.132 (standard error 0.049): larger than OLS, and far less precise.

Five Instruments, From Fewest to Most Problems

Instrument	Treatment	Relevance	Exogeneity	Main issue
Admission lottery	Charter school	✓	✓	Noncompliance
Draft lottery	Military service	✓	~	Avoiding the draft by staying in school
Random judge	Release	✓	~	Judge also hears the full case
Quarter of birth	Schooling	~	~	Weak first stage
Distance to college	Schooling	✓	~	Families choose where to live

Relevance is checked in the data. Exogeneity needs an argument: the more marks of ~, the more argument the instrument needs.

Returns to Education: The Question and the OLS Estimate

$$\overbrace{\ln(\text{wage}_i)}^{\text{wage}} = \beta_1 + \beta_2 \overbrace{\text{educ}_i}^{\text{years of school}} + \beta_3\, \text{exper}_i + \beta_4\, \text{exper}_i^2 + e_i$$

How much does one more year of school raise wages?

$$\text{OLS: } b_2 = 0.1075 \quad (\text{s.e. } 0.0141,\ N = 428)$$

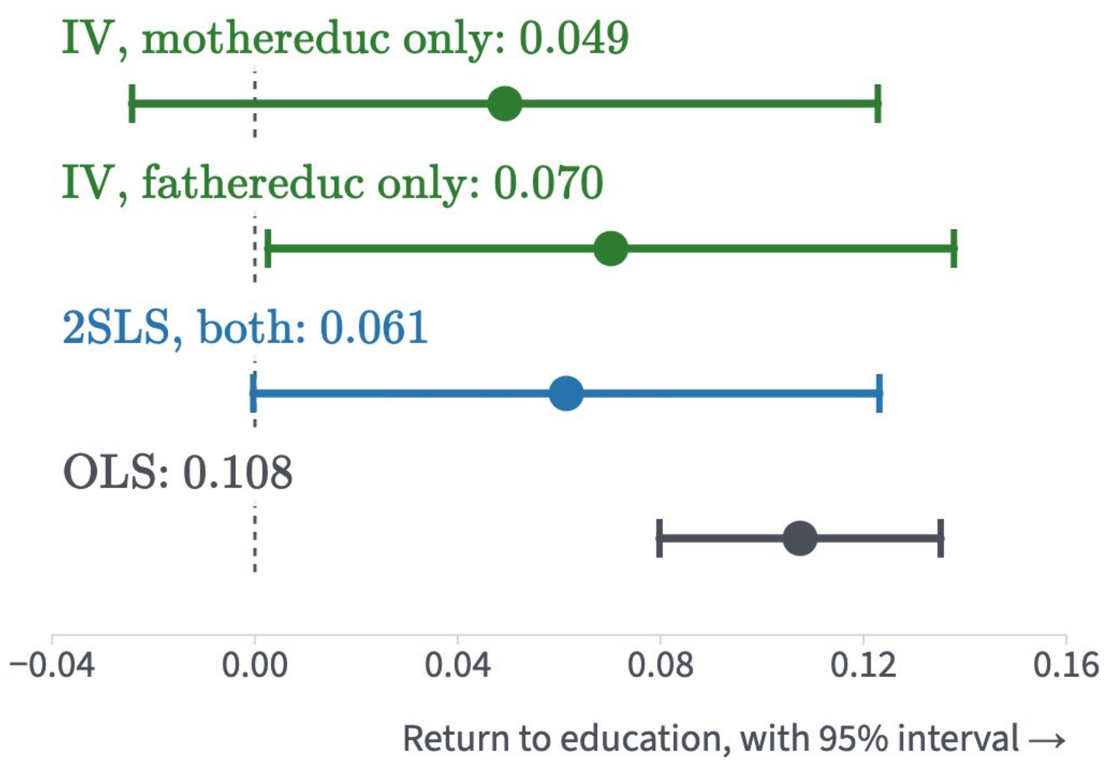
$\Longrightarrow$ one more year of school, wages about 10.8% higher

Wage data: 428 married women in the labor force in 1975 (from the textbook).

Concern: ability, motivation, and family background are in e_i. They raise schooling and wages, so part of the 10.8% may not come from schooling.

$\Longrightarrow$ We need an instrument for education.

IV with Two Instruments: 2SLS Uses Both



OLS

0.108, interval 0.08 to 0.14. Narrow, but biased upward by ability.

ONE INSTRUMENT AT A TIME

mothereduc alone: 0.049. fathereduc alone: 0.070. Each throws away the other instrument ⇒ wide intervals.

2SLS: USE BOTH AT ONCE

Stage 1 combines both instruments into one prediction of educ. Result: 0.061, the narrowest IV interval (std. error 0.031 vs. 0.037 and 0.034).

$\Longrightarrow$ 2SLS gives one estimate that is more precise than either instrument alone. It is still less precise than OLS: the price of removing the bias.

Two-Stage Least Squares: Stage 1

REGRESS the endogenous variable on all exogenous variables

$$\text{educ}_i = \gamma_1 + \gamma_2\,\text{exper}_i + \gamma_3\,\text{exper}_i^2 + \theta_1\,\text{mothereduc}_i + \theta_2\,\text{fathereduc}_i + v_i$$

Stage 1, $N = 428$	**Estimate**	**Std. error**
mothereduc	0.1576	0.0359
fathereduc	0.1895	0.0338
Joint test $\theta_1 = \theta_2 = 0$: $F = 55.40$		

- $\hat{\theta}_1 = 0.158$: one more year of her mother's schooling $\implies$ 0.16 more years of her own.
- $\hat{\theta}_2 = 0.190$: one more year of her father's schooling $\implies$ 0.19 more years of her own.
- $F = 55.40 > 10 \implies$ the instruments strongly predict educ: relevant, not weak.

Two-Stage Least Squares: Stage 2

$$\widehat{\text{educ}}_i = \hat{\gamma}_1 + \hat{\gamma}_2\,\text{exper}_i + \hat{\gamma}_3\,\text{exper}_i^2 + \hat{\theta}_1\,\text{mothereduc}_i + \hat{\theta}_2\,\text{fathereduc}_i$$

SAVE

the fitted values $\widehat{\text{educ}}_i$: the education we predict from the parents' education and experience. It contains no ability, if the instruments are valid.

STAGE 2

regress the outcome on $\widehat{\text{educ}}_i$ in place of educ_i, with the same exogenous variables

$$\ln(\text{wage}_i) = \beta_1 + \beta_2\,\widehat{\text{educ}}_i + \beta_3\,\text{exper}_i + \beta_4\,\text{exper}_i^2 + e_i$$

RESULT

$\hat{\beta}_2 = 0.0614 \implies$ one more year of education raises wages by about 6.1%, compared with 10.8% from OLS.

Standard Errors: Use an IV Command

Estimate of β_2	Coefficient	Std. error
Stage 2 run by hand with `lm()`	0.0614	**0.0330**
IV command, `ivreg()`	0.0614	**0.0313**

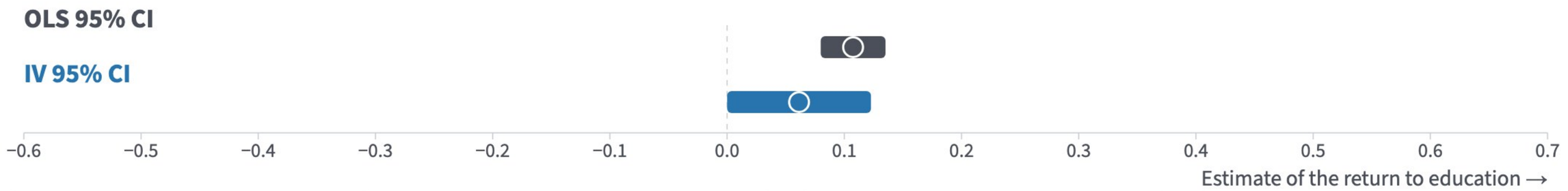
DIFFERENT RESIDUALS

Run by hand, stage 2 computes the residuals with $\widehat{\text{educ}_i}$. The correct residuals use the actual educ_i: $\hat{e}_i = \ln(\text{wage}_i) - \hat{\beta}_1 - \hat{\beta}_2 \, \text{educ}_i - \dots$

```
library(AER)
ivreg(log(wage) ~ educ + exper + I(exper^2) |
        mothereduc + fathereduc + exper + I(exper^2),
      data = mroz)
```

$\Longrightarrow$ Same coefficient, wrong standard error: always use an IV command for inference.

Standard Errors: A Weaker First Stage Widens the Interval



FIRST STAGE	F-STATISTIC	IV STD. ERROR	IV 95% CI
0.268	55.4	0.0313	[0.00, 0.12]

Weak — Strong

$$\mathrm{se}(\hat{\beta}_{IV}) = \overbrace{\frac{\sigma}{\sqrt{n}\,\mathrm{sd}(Z)}}^{\text{reduced-form noise}} \times \overbrace{\frac{1}{|\pi|}}^{\text{first stage}}$$

$$F \approx \left(\frac{\pi}{\mathrm{se}(\pi)}\right)^2, \quad \mathrm{se}(\pi) = \frac{\sigma_v}{\sqrt{n}\,\mathrm{sd}(Z)}$$

$\Longrightarrow$ As $\pi \to 0$, the standard error blows up and the interval becomes useless.

Weak Instruments: The First-Stage F

WEAK INSTRUMENT

z moves x only a little (π near 0). Standard errors blow up, and the 2SLS estimate tends toward the OLS estimate.

FIRST STAGE

$$\text{educ}_i = \gamma_1 + \gamma_2\,\text{exper}_i + \gamma_3\,\text{exper}_i^2 + \overbrace{\theta_1\,\text{mothereduc}_i + \theta_2\,\text{fathereduc}_i}^{\text{excluded instruments}} + v_i$$

THE F-TEST

$$H_0 : \theta_1 = \theta_2 = 0$$

Joint test that every instrument coefficient is zero. Rejecting says some instrument is relevant, not that the set is strong.

RULE OF THUMB $F > 10$ (Staiger and Stock 1997; Stock and Yogo 2005)

ONE ENDOGENOUS REGRESSOR

This F test is valid with one endogenous regressor. With several, software handles it differently.

Weak Instruments: The F-Statistic in the Wage Data

UNRESTRICTED $\text{educ}_i = \gamma_1 + \gamma_2\,\text{exper}_i + \gamma_3\,\text{exper}_i^2 + \theta_1\,\text{mothereduc}_i + \theta_2\,\text{fathereduc}_i + v_i$

RESTRICTED $\text{educ}_i = \gamma_1 + \gamma_2\,\text{exper}_i + \gamma_3\,\text{exper}_i^2 + v_i \qquad (\theta_1 = \theta_2 = 0)$

$$F = \frac{(SSE_R - SSE_U)\,/\,q}{SSE_U\,/\,(n-p)}$$ ($q = 2$ instruments, $n = 428$, $p = 5$ in the unrestricted model)

$$F = \frac{(2219.22 - 1758.58)\,/\,2}{1758.58\,/\,(428-5)} = 55.40$$

STRONG $F = 55.40 > 10$: the two instruments are jointly relevant.

F > 10 IS A MINIMUM BAR

mothereduc alone: $F = 73.95$. Stock and Yogo (2005) set a stricter bar, 16.38 for one instrument.

Identification: Good, Bad, and Lucky Variables

We can estimate the coefficients only if there are enough instruments for the endogenous regressors.

$$y = \overbrace{\beta_1 + \cdots + \beta_G x_G}^{G \text{ exogenous: Good}} + \overbrace{\beta_{G+1} x_{G+1} + \cdots + \beta_K x_K}^{B \text{ endogenous: Bad}} + e$$

G: GOOD

Exogenous regressors. They serve as their own instruments. Wage data: intercept, exper, exper^2 $\implies$ $G = 3$.

B: BAD

Endogenous regressors. Each needs an instrument. Wage data: educ $\implies$ $B = 1$.

L: LUCKY

Instruments: outside the model, correlated with the Bad variables. Wage data: mothereduc, fathereduc $\implies$ $L = 2$. Lucky because good instruments are hard to find.

Identification: Under, Exactly, and Over Identified

UNDER

$$L < B$$

Too few instruments to estimate $\beta_{G+1}, \ldots, \beta_K$. You need more instruments.

EXACTLY

$$L = B$$

Just enough instruments. Wage data with mothereduc only: $L = 1 = B$.

OVER

$$L > B$$

More instruments than needed. Wage data with both parents: $L = 2 > B = 1$.

$\Longrightarrow$ Necessary condition: $L \geq B$, at least one instrument per endogenous regressor. The instruments must also be relevant: a count alone is not enough.

Specification Tests: Two Questions

Hausman **Is x endogenous?**

$H_0 : \text{cov}(x_i, e_i) = 0 \implies$ OLS is fine, and more precise than IV.

Sargan **Are the extra instruments valid?**

Needs more instruments than endogenous variables, $L > B$.

> Wage data: is education endogenous, and are mothereduc and fathereduc valid?

Hausman Test: The Education Example

Data: 428 married women in the labor force in 1975 (from the textbook).

RETURNS TO EDUCATION

$$\ln(\text{wage}_i) = \beta_1 + \beta_2 \overbrace{\text{educ}_i}^{\text{endogenous?}} + \beta_3 \, \text{exper}_i + \beta_4 \, \text{exper}_i^2 + e_i$$

THE WORRY

Ability is in e_i and raises both education and wages $\implies \text{cov}(\text{educ}_i, e_i) > 0$.

THE INSTRUMENTS

mothereduc and fathereduc.

> Question: is education really endogenous, so that we need IV?

Hausman Test, Step 1: The First Stage

REGRESS EDUC ON THE EXOGENOUS VARIABLES AND THE INSTRUMENTS

$$\text{educ}_i = \gamma_1 + \gamma_2 \, \text{exper}_i + \gamma_3 \, \text{exper}_i^2 + \theta_1 \, \text{mothereduc}_i + \theta_2 \, \text{fathereduc}_i + v_i$$

SAVE THE RESIDUAL

$$\hat{v}_i = \text{educ}_i - \widehat{\text{educ}_i}$$

> $\hat{v}_i$ is the part of education that the instruments and experience do not explain.

With several endogenous variables, run one first stage for each and save each residual.

Hausman Test, Step 2: Add $\hat{v}$ to the Wage Equation

STEP 1: FIRST STAGE

$$\text{educ}_i = \gamma_1 + \gamma_2\,\text{exper}_i + \gamma_3\,\text{exper}_i^2 + \theta_1\,\text{mothereduc}_i + \theta_2\,\text{fathereduc}_i + v_i$$



save the residual $\hat{v}_i$

STEP 2: WAGE EQUATION PLUS $\hat{v}$

$$\ln(\text{wage}_i) = \beta_1 + \beta_2\,\text{educ}_i + \beta_3\,\text{exper}_i + \beta_4\,\text{exper}_i^2 + \delta\,\overbrace{\hat{v}_i}^{\text{from step 1}} + e_i$$

> Re-estimate by OLS. The coefficient we look at is δ.

$\hat{v}$ is the part of education unrelated to the instruments. If it still predicts wages, it carries the correlation with e_i.

If δ is significant $\implies$ educ is endogenous: use IV (2SLS).
If δ is not significant $\implies$ no evidence that educ is endogenous: use OLS, which is more precise.

Hausman Test, Step 3: A t-Test on $\hat{v}$

HYPOTHESES

$$H_0 : \delta = 0 \iff \text{educ is exogenous}$$
$$H_1 : \delta \neq 0 \iff \text{educ is endogenous}$$

REJECT H_0 Education is endogenous $\implies$ use IV.

FAIL TO REJECT No evidence of endogeneity $\implies$ use OLS, which is more precise.

If several variables are tested, use an F-test on all the residual coefficients.

Hausman Test: Is Education Endogenous?

Step 2 regression	Estimate	Std. error	t	p
educ	0.0614	0.0310	1.98	0.048
$\hat{v}$	0.0582	0.0348	1.67	0.095

also exper, exper^2, intercept; $N = 428$

REJECT AT 10% $p = 0.095 < 0.10$: some evidence that education is endogenous, though not at 5%.

DECISION

Weak evidence, but enough concern to use IV. The first-stage F of 55.4 says the instruments are strong.

Sargan Test: Requires Extra Instruments

ENDOGENOUS VARIABLES: $B = 1$

educ

INSTRUMENTS: $L = 2$

mothereduc and fathereduc

EACH INSTRUMENT ALONE

gives its own estimate of the return to education

IF BOTH ARE VALID

the two estimates should agree

$\Longrightarrow$ With $L = B$ there is one estimate and nothing to compare. The test needs $L > B$.

Sargan Test: The Idea

MOTHEREDUC ALONE

0.049

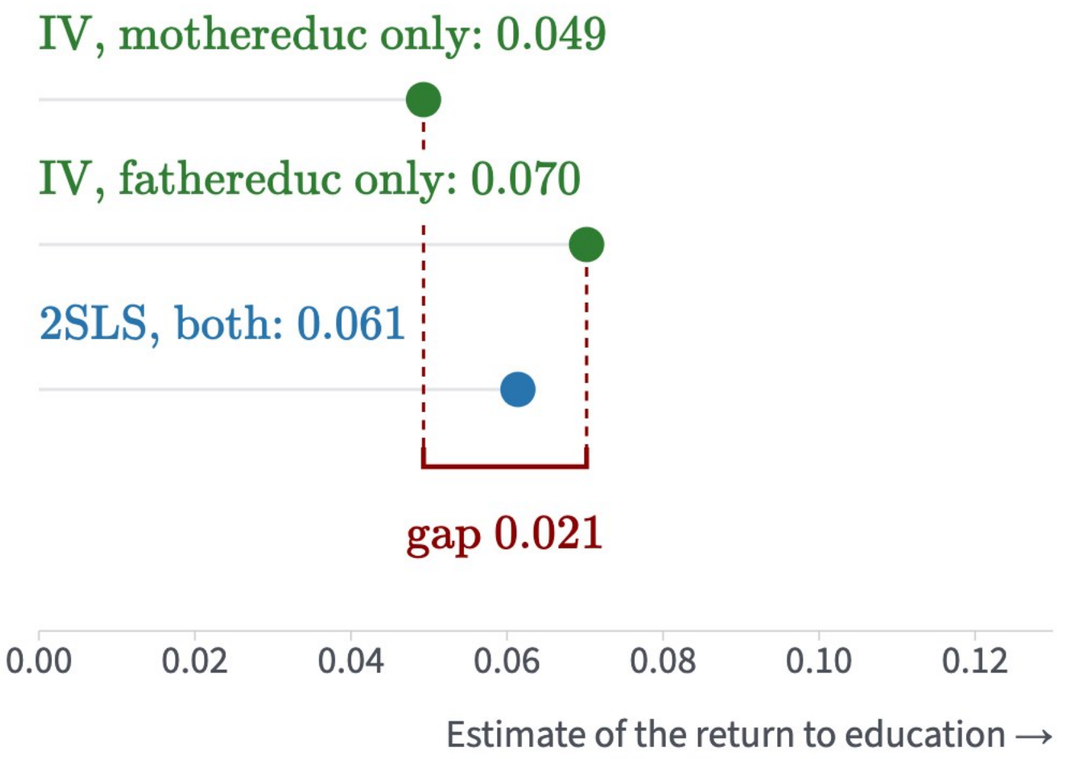
FATHEREDUC ALONE

0.070

GAP: 0.021

Is that more than sampling error?

$\Longrightarrow$ The Sargan test answers this.



Sargan Test: The Hypotheses

NULL

H_0: all instruments are valid $\iff$ $\text{cov}(z_j, e) = 0$ for every instrument

ALTERNATIVE

H_1: at least one instrument is invalid

$\implies$ Reject H_0: at least one instrument is invalid.

Sargan Test: The Three Steps

STEP 1

Estimate the 2SLS model

↓ **save the residuals $\hat{e}$**

STEP 2

Regress $\hat{e}$ on all exogenous variables and all instruments

↓ **save the R^2**

STEP 3

Compute NR^2 and compare with $\chi^2_{(L-B)}$

Sargan Test, Step 1: Save the 2SLS Residuals

THE MODEL

$$\ln(\text{wage}_i) = \beta_1 + \beta_2\,\text{educ}_i + \beta_3\,\text{exper}_i + \beta_4\,\text{exper}_i^2 + e_i$$

ESTIMATE BY 2SLS

instruments: mothereduc and fathereduc

SAVE THE RESIDUALS, USING THE ACTUAL EDUC

$$\hat{e}_i = \ln(\text{wage}_i) - \hat{\beta}_1 - \hat{\beta}_2\,\text{educ}_i - \hat{\beta}_3\,\text{exper}_i - \hat{\beta}_4\,\text{exper}_i^2$$

$\Longrightarrow$ $\hat{e}_i$ estimates the error e_i.

Sargan Test, Step 2: Regress $\hat{e}$ on Everything Exogenous

FROM STEP 1

the residuals $\hat{e}_i$

regress $\hat{e}$ on

$$\hat{e}_i = \overbrace{\alpha_1 + \alpha_2\,\text{exper}_i + \alpha_3\,\text{exper}_i^2}^{\text{exogenous variables}} + \overbrace{\alpha_4\,\text{mothereduc}_i + \alpha_5\,\text{fathereduc}_i}^{\text{all instruments}} + u_i$$

educ is left out: it is endogenous.

Save the R^2 of this regression.

$\Longrightarrow$ Valid instruments cannot explain $\hat{e}$, so R^2 should be near 0.

Sargan Test, Step 3: Compute NR^2 and Compare

UNDER H_0

$$\overbrace{N \times R^2}^{\text{Sargan statistic}} \sim \chi^2_{(L-B)}$$

DEGREES OF FREEDOM

$$\overbrace{L}^{\text{instruments}} - \overbrace{B}^{\text{endogenous variables}}$$

Sargan Test: The Decision Rule

STATISTIC ABOVE THE CRITICAL VALUE

reject H_0: at least one instrument is invalid

STATISTIC BELOW THE CRITICAL VALUE

fail to reject: no evidence against validity

$\Longrightarrow$ Reject means we must find the invalid instrument.

Software runs all three steps, also with many regressors.

Sargan Test: The Education Numbers

STEP 2 REGRESSION: $\hat{e}$ ON EXPER, EXPER^2, MOTHEREDUC, FATHEREDUC

$$R^2 = 0.000883, \quad N = 428$$

STEP 3 STATISTIC

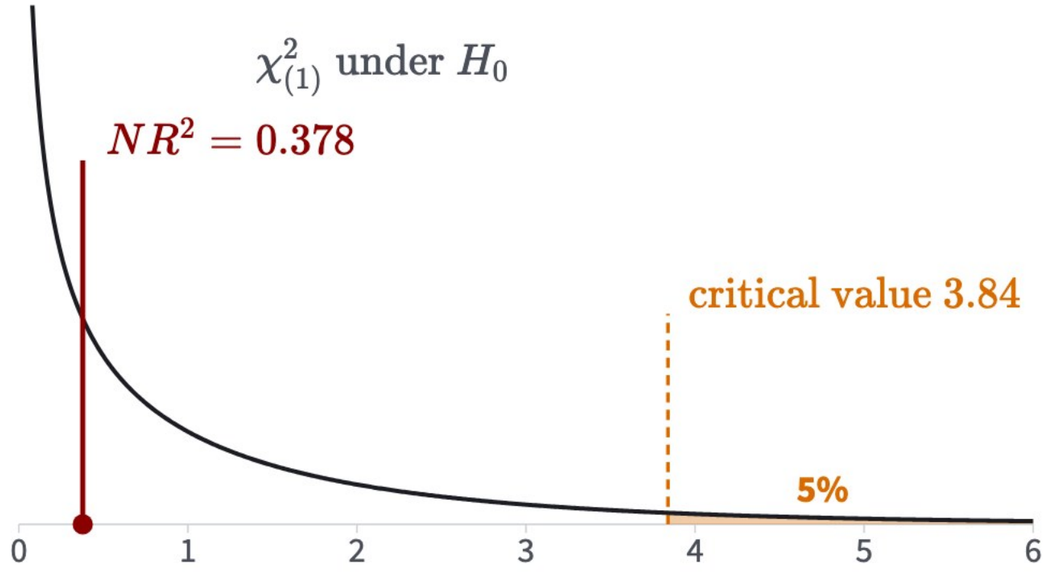
$$N \times R^2 = 428 \times 0.000883 = 0.378$$

DEGREES OF FREEDOM AND CRITICAL VALUE

$$L - B = 2 - 1 = 1, \quad \chi^2_{(0.95,\,1)} = 3.84$$

$$\Longrightarrow 0.378 < 3.84$$

Sargan Test: Are mothereduc and fathereduc Valid?



FAIL TO REJECT $0.378 < 3.84, p = 0.54$

Sargan Test: Conclusion

IN PLAIN LANGUAGE

No evidence that mothereduc or fathereduc is invalid.

THE TWO SEPARATE ESTIMATES

0.049 and 0.070: a gap no larger than chance.

DECISION

Keep the 2SLS estimate: $\hat{\beta}_2 = 0.0614$.

> Caution: passing the test does not prove the instruments are valid. It cannot detect instruments that are all invalid in the same way.

IV Diagnostic Tests: Summary

Test	Question	Wage data
First-stage F	Are the instruments strong?	$F = 55.4$
Hausman	Is x endogenous?	$p = 0.095$
Sargan	Are the instruments valid?	$p = 0.54$

Specification Tests: Decisions After Each Test

Result	Action
First-stage $F < 10$	Find stronger instruments
Hausman rejects	Use IV
Hausman fails to reject	Use OLS
Sargan rejects	Drop the suspect instrument and retest
Sargan fails to reject	Keep the IV estimates

> Passing Sargan does not prove the instruments are valid.

Simultaneous Equations: Two Outcomes Determined Together

SO FAR

One dependent variable. The endogenous regressor came from an omitted variable such as ability.

NOW

A **system** of two or more equations with two or more dependent variables, each a regressor in the other's equation.

CONSEQUENCE

OLS is no longer enough to estimate the system, and the instruments come from the system itself.

Simultaneity in Labor: Wages and Hours

OUTCOMES

Two outcomes for each worker: the wage and the hours worked.

WORKER

$$\text{Hours}_i = \alpha_1 \, \text{Wage}_i + e_{i,h}$$

Hours depend on the wage.

FIRM

$$\text{Wage}_i = \beta_1 \, \text{Hours}_i + e_{i,w}$$

The wage depends on hours.

MECHANISM

A shock to $e_{i,h}$ changes hours, hours change the wage, so Wage_i is correlated with $e_{i,h}$. The same argument applies to Hours_i and $e_{i,w}$.

$\Longrightarrow$ OLS on either equation is biased.

Simultaneity in a Market: Price and Quantity

OUTCOMES

Two outcomes in each market: the price and the quantity sold.

DEMAND

$$Q_i = \alpha_1 P_i + e_{i,d}$$

Buyers respond to price.

SUPPLY

$$Q_i = \beta_1 P_i + e_{i,s}$$

Firms respond to price.

MECHANISM

A demand shock $e_{i,d}$ or a supply shock $e_{i,s}$ moves the price at which the market clears. The price is determined together with Q_i, so P_i is correlated with each error.

$\Longrightarrow$ OLS is biased.

Supply and Demand: A Look Back on ECON 2

DEMAND

$$Q_i = \alpha_1 P_i + \alpha_2 X_i + e_{i,d}$$

SUPPLY

$$Q_i = \beta_1 P_i + e_{i,s}$$

INCOME

X_i is income. **Question:** From ECON 2, which signs should α_1 and β_1 have?

SIGNS

$\alpha_1 < 0$: demand slopes down.
$\beta_1 > 0$: supply slopes up.

Supply and Demand: Endogenous and Exogenous

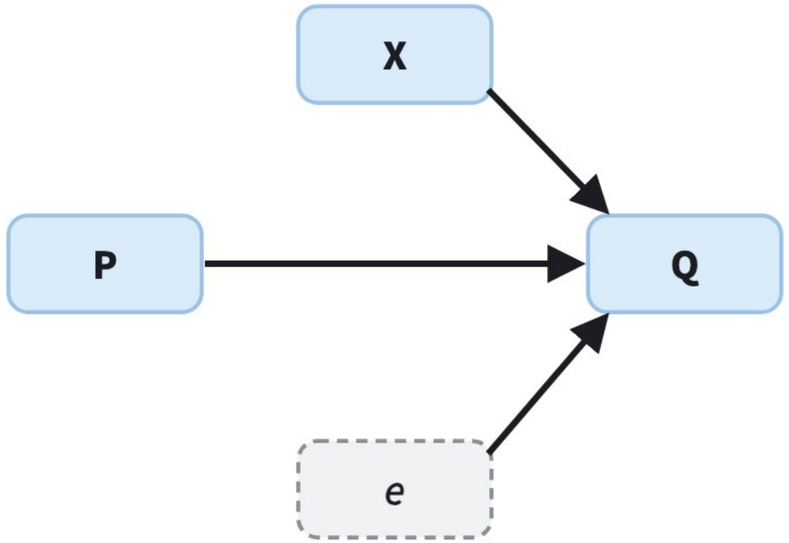
DEMAND $$Q_i = \alpha_1 P_i + \alpha_2 X_i + e_{i,d}$$

SUPPLY $$Q_i = \beta_1 P_i + e_{i,s}$$

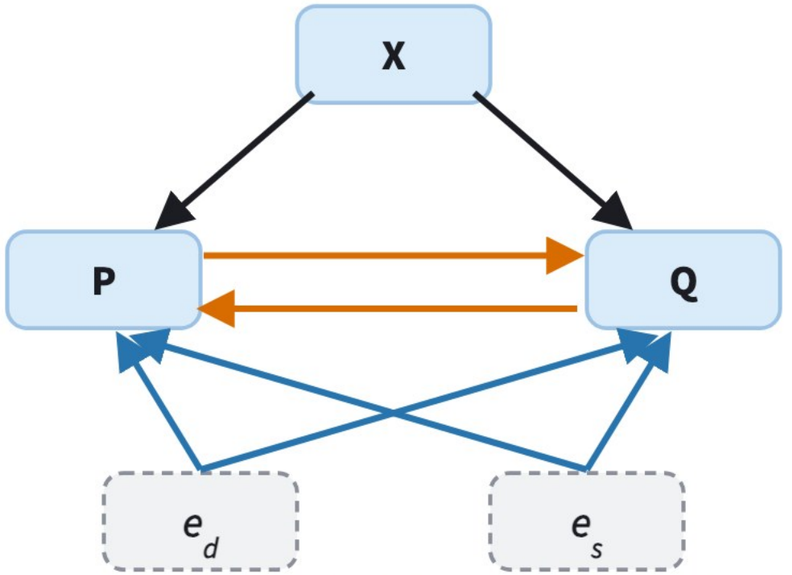
ENDOGENOUS P and Q: their values are determined within the system, so they are random and both are dependent variables.

EXOGENOUS Income X: determined outside the system.

Influence Diagrams: Separate Regressions vs. the System



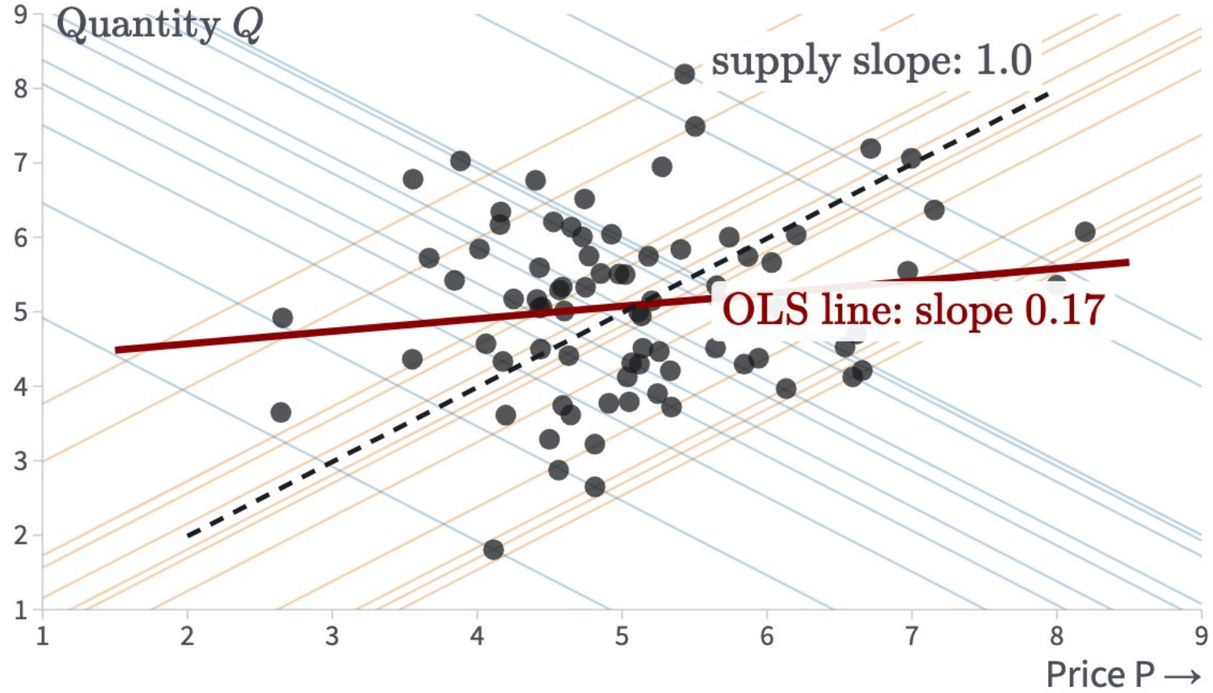
Two separate regressions: P treated as given



The system: P and Q determine each other

P and Q affect each other, and both are affected by e_d and e_s $\Longrightarrow$ P is random and correlated with the errors.

OLS Bias: The Data



Market: demand slope -1, supply slope $+1$, 80 periods with shocks to both curves.

EACH PERIOD Demand and supply both get a shock. We observe only where they cross.

MANY PERIODS The crossing points trace out neither curve.

OLS FAILS Regressing Q on P gives slope 0.17, not the supply slope 1.0.

Reduced Form: Solve for P

REDUCED FORM

Write each endogenous variable in terms of the exogenous variables only.

SET SUPPLY EQUAL TO DEMAND

$$\beta_1 P_i + e_{i,s} = \alpha_1 P_i + \alpha_2 X_i + e_{i,d}$$

SOLVE

$$P_i = \overbrace{\frac{\alpha_2}{\beta_1 - \alpha_1}}^{\pi_1} X_i + \overbrace{\frac{e_{i,d} - e_{i,s}}{\beta_1 - \alpha_1}}^{v_{i,1}}$$

> $\beta_1 - \alpha_1 > 0$, and P contains $-e_{i,s}$: when supply gets a positive shock, the price falls.

Reduced Form: Solve for Q

Substitute P_i into the supply equation (the simpler one):

SUBSTITUTE

$$Q_i = \beta_1 \left(\frac{\alpha_2}{\beta_1 - \alpha_1} X_i + \frac{e_{i,d} - e_{i,s}}{\beta_1 - \alpha_1} \right) + e_{i,s}$$

COMBINE

$$Q_i = \frac{\beta_1 \alpha_2}{\beta_1 - \alpha_1} X_i + \frac{\beta_1 e_{i,d} - \alpha_1 e_{i,s}}{\beta_1 - \alpha_1}$$

RESULT

$$Q_i = \overbrace{\frac{\beta_1 \alpha_2}{\beta_1 - \alpha_1}}^{\pi_2} X_i + \overbrace{\frac{\beta_1 e_{i,d} - \alpha_1 e_{i,s}}{\beta_1 - \alpha_1}}^{v_{i,2}}$$

Reduced Form: Parameters and Errors

$$P_i = \pi_1 X_i + v_{i,1} \qquad Q_i = \pi_2 X_i + v_{i,2}$$

NAMES

π_1, π_2: reduced-form parameters.
$v_{i,1}, v_{i,2}$: reduced-form errors.

FORM

An endogenous variable on the left, exogenous variables on the right: the same form as the first-stage equations of Chapter 10.

OLS WORKS HERE $E(v_{i,1} \mid X_i) = 0$ and $E(v_{i,2} \mid X_i) = 0$, so OLS estimates of π_1 and π_2 are consistent.

Reduced Form: Interpretation

EQUILIBRIUM

The reduced-form equations give the **equilibrium** values of P and Q for a given value of the exogenous variable. Suppose income X rises by one unit.

π_1

The expected change in the equilibrium price, after the market adjusts to the new equilibrium of P and Q.

π_2

The expected change in the equilibrium quantity.

Thank you!

jakeanderson@g.ucla.edu