Instrumental Variables

Econometric Modelling and Forecasting II
The SEACEN Centre

Tanisa Tawichsri

Puey Ungphakorn Institute for Economic Research (PIER)
Bank of Thailand

May 14, 2026

Overview

Introduction

Motivation: Why Causality is Hard

- A central challenge in policy analysis: distinguishing **causal effects** from spurious correlations
- **Education and wages:** Workers with more education earn higher wages. Does education *cause* higher wages — or do more able, motivated people both acquire more education *and* earn more?
- **Central bank examples:**
  - Does a rate hike *cause* inflation to fall? Or did the central bank raise rates in anticipation of already-declining inflation?
  - Does credit expansion *cause* growth? Or do both respond to the same underlying boom?
- Simple OLS regression cannot resolve these questions without additional structure
- **Instrumental Variables (IV)** is one of the most powerful tools for recovering causal effects when randomisation is impossible

The Endogeneity Problem

- Consider the linear regression model:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

- For OLS to be **consistent**, the regressor must be **exogenous**:

$$\mathbb{E}[x_i \varepsilon_i] = 0$$

  This means x_i is uncorrelated with all unobserved determinants of y_i.
- In applied settings, this fails due to:
  - **Omitted variable bias** — e.g., ability drives both schooling and wages
  - **Simultaneity** — e.g., price and quantity are jointly determined
  - **Measurement error** in x_i attenuates OLS toward zero
- When $\mathbb{E}[x_i \varepsilon_i] \neq 0$: OLS is **biased and inconsistent** — this is the problem of **endogeneity**

OLS Bias: The Wage–Education Example I

- **What we estimate:**

$$\log w_i = \beta_0 + \beta_1 s_i + \varepsilon_i$$

where s_i is years of schooling and ε_i captures all other wage determinants.

- **The true model** includes unobserved ability a_i:

$$\log w_i = \beta_0 + \beta_1 s_i + \underbrace{\beta_2 a_i + u_i}_{=\varepsilon_i}$$

Here a_i is not observed, and u_i collects remaining influences.

OLS Bias: The Wage–Education Example II

- The probability limit of the OLS estimator:

$$\text{plim } \hat{\beta}_1^{\text{OLS}} = \beta_1 + \beta_2 \frac{\text{Cov}(s_i,\ a_i)}{\text{Var}(s_i)}$$

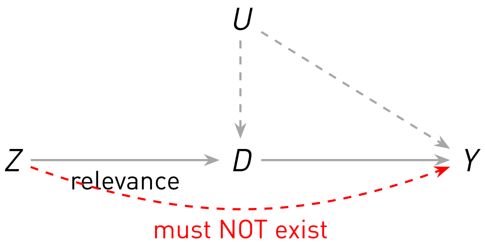
- If ability a_i is positively correlated with schooling *and* wages ($\beta_2 > 0$):
  - $\text{Cov}(s_i, a_i) > 0 \Rightarrow$ OLS **overestimates** the true return to schooling
- **The core challenge:** How do we estimate β_1 when unobserved confounders like ability are present?
  - This is precisely the problem that **instrumental variables (IV)** methods are designed to address

The IV Solution

What is an Instrument?

- Suppose we find an additional variable z_i — an **instrument** — satisfying:
  1. **Relevance:** $\text{Cov}(z_i, x_i) \neq 0$ [testable: check first-stage F-stat]
  2. **Exogeneity:** $\text{Cov}(z_i, \varepsilon_i) = 0$ [untestable: requires theory]
- z_i provides variation in x_i that is "as good as random" and uncontaminated by the confounder
- **Key insight:** **"Good instruments often feel weird"** (Cunningham 2021) — the link from z_i to y_i should make sense *only through* x_i, not through any other channel
- **Examples:**
  - Quarter of birth $\rightarrow$ schooling (Angrist & Krueger 1991): strange, but valid
  - Nearby college $\rightarrow$ schooling (Card 1995): geographic accident, not ability
  - Rainfall $\rightarrow$ agricultural income $\rightarrow$ conflict (Miguel et al. 2004)

The DAG Intuition



- U creates a **backdoor path**: $D \leftarrow U \rightarrow Y$. This is the endogeneity
- Z is connected to D (relevance) but blocked from Y except through D (exclusion)
- The **exclusion restriction** = no direct red arrow from Z to Y. Sometimes called "only through" assumption.

The IV Estimator

- Start from the moment condition $\mathbb{E}[z_i \varepsilon_i] = 0$ and substitute $\varepsilon_i = y_i - \beta_0 - \beta_1 x_i$
- Solving yields the **IV estimator**:

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

- When z_i is binary (e.g., a policy, lottery), this is the **Wald estimator**:

$$\hat{\beta}_1^{\text{Wald}} = \frac{\bar{Y}_{z=1} - \bar{Y}_{z=0}}{\bar{X}_{z=1} - \bar{X}_{z=0}} = \frac{\overbrace{\text{(total effect of Z on Y)}}^{\text{Reduced form}}}{\underbrace{\text{First stage}}_{\text{(effect of Z on D)}}}$$

- **Consistency:**

$$\text{plim}\ \hat{\beta}_1^{\text{IV}} = \frac{\text{Cov}(z_i,\ \beta_1 x_i + \varepsilon_i)}{\text{Cov}(z_i,\ x_i)} = \beta_1 + \frac{\text{Cov}(z_i,\ \varepsilon_i)}{\text{Cov}(z_i,\ x_i)} = \beta_1$$

Two-Stage Least Squares (2SLS)

Two-Stage Least Squares (2SLS)

- With one or more instruments, the practical implementation is **2SLS**
- **Stage 1 — First stage:** Regress the endogenous variable on the instrument(s):

$$x_i = \gamma_0 + \gamma_1 z_i + v_i \quad \Longrightarrow \quad \hat{x}_i = \hat{\gamma}_0 + \hat{\gamma}_1 z_i$$

- **Stage 2 — Second stage:** Regress the outcome on the *fitted* values:

$$y_i = \beta_0 + \beta_1 \hat{x}_i + \eta_i$$

- $\hat{x}_i$ retains only the *exogenous* variation in x_i driven by z_i; the endogenous part is stripped out
- **Important:** Always use a purpose-built `ivregress` / `ivreg` command — manually running two OLS stages gives wrong standard errors

The Precision Trade-off

- 2SLS uses *only the variation in x_i explained by z_i* — typically a small fraction of total variation
- Discarding the endogenous variation reduces precision:

$$\text{Var}\left(\hat{\beta}_1^{\text{2SLS}}\right) \approx \frac{\sigma_\varepsilon^2}{n \cdot \text{Var}(x_i) \cdot R^2_{\text{first stage}}}$$

- Weaker instrument $\Rightarrow$ lower $R^2_{\text{first stage}}$ $\Rightarrow$ larger standard errors
- This is the fundamental trade-off: **consistency for efficiency**
- Wider IV confidence intervals are not a failure — they honestly reflect the limited exogenous variation available
- If you need high precision, you need a strong instrument — there is no shortcut

Application

Rainfall & Civil Conflict — Miguel, Satyanath & Sergenti (2004)

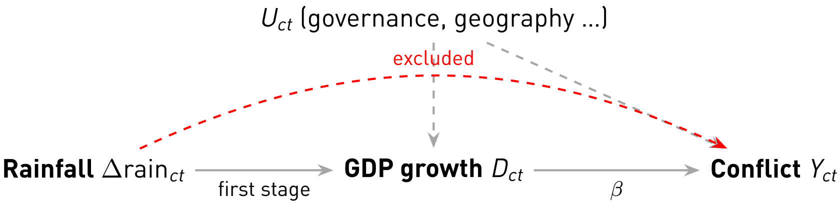
Miguel et al. (2004): Setup

- **Question:** Does a decline in per-capita income *cause* civil conflict in Sub-Saharan Africa?
- **Data:** 41 Sub-Saharan African countries, 1981–1999 (annual panel, 743 country-years)
- **Outcome:** $Y_{ct} = 1$ if civil conflict onset in country *c*, year *t*
- **Treatment:** D_{ct} = per-capita GDP growth (%)
- **Model:**

$$Y_{ct} = \alpha_c + \alpha_t + \beta\, D_{ct} + \varepsilon_{ct}$$

  with country and year fixed effects to absorb time-invariant differences and global shocks.
- **The endogeneity problem:**
  - **Reverse causality:** income decline $\leftrightarrow$ conflict
  - **Omitted variables:** institutions, geography, ethnic fractionalisation, state capacity

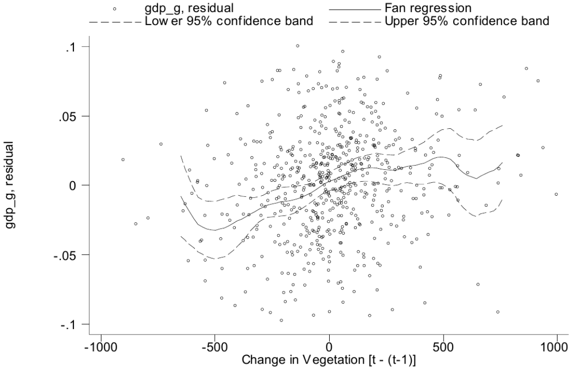
The Rainfall Instrument



- **Relevance:** SSA economies are largely agricultural; rainfall drives crop yields $\Rightarrow$ income
  - First stage: 1pp higher rainfall growth $\approx$ 0.5pp higher GDP growth; **F** $\approx$ **9.4**
- **Exclusion restriction:** Rain affects conflict *only through income* — rain itself does not cause fighting, alter governance, or shift borders
- **Robustness:** Adding rainfall directly to the 2nd-stage regression leaves the IV estimate unchanged

First Stage Visually

Figure 1: Non-Parametric First-Stage Fan Regression (Epanechnikov Kernel), Economic Growth Rate on Changes in Vegetation (conditional on country fixed effects)



- Miguel et al.'s Figure 1: vegetation changes predict GDP growth after removing country fixed effects.
- This is the paper's first-stage evidence: climate shocks move income, giving 2SLS exogenous variation.

OLS vs. 2SLS: Rainfall Results

	OLS	2SLS
GDP growth, $t-1$ ($\hat{\beta}$)	-0.005^{**}	-0.026^{**}
	(0.002)	(0.010)
First stage: rainfall growth, $t-1$	—	2.8^{*}
First-stage F-statistic	—	9.44
Country FE / Year FE	✓	✓
Observations	743	743

Note: LPM. Outcome = civil conflict indicator (0/1).
** $p < 0.05$, * $p < 0.10$. Source: Miguel et al. (2004).

- **Key finding:** A 5pp fall in income growth $\Rightarrow$ $\approx$**12pp higher conflict probability** (2SLS)
- OLS is attenuated toward zero by reverse causality; 2SLS is $5\times$ larger — the true effect is much stronger than OLS suggests

OLS Bias: Two Examples Compared

The direction of bias depends on the source of endogeneity.

	Wages & Education *(omitted variable)*	**Conflict & Income** *(reverse causality)*
Source of bias	Ability $\uparrow$ both s and w	Conflict $\downarrow$ income
OLS vs. truth	Overstates (+bias)	Understates (−bias)
OLS estimate	7.1% per year of schooling	-0.005 per pp growth
2SLS estimate	12.4% per year of schooling	-0.026 per pp growth

- Always ask *before* running IV: which way does the bias go, and by how much? The OLS–2SLS gap is your answer

Weak Instruments & Diagnostics

The Weak Instrument Problem

- When $\text{Cov}(z_i, x_i) \approx 0$, the instrument is **weak**
- The 2SLS bias is approximately (Stock & Yogo 2005):

$$\mathbb{E}\left[\hat{\beta}_1^{\text{2SLS}} - \beta_1\right] \approx \frac{\sigma_{\varepsilon v}}{\sigma_v^2} \cdot \frac{1}{F+1}$$

  - F = first-stage F-statistic; $\sigma_{\varepsilon v}$ = degree of endogeneity
- As $F \to 0$: 2SLS bias $\to$ **OLS bias** — the instrument gives no benefit
- As $F \to \infty$: 2SLS bias $\to 0$
- **Rule of thumb:** First-stage $F > 10$ (Stock & Yogo 2005)
- **Danger:** More instruments $\neq$ stronger identification — each extra weak instrument dilutes valid variation and *increases* bias

Illustration: Graddy (2006) Fulton Fish Market

- **Question:** Price elasticity of demand for fish (price is endogenous — supply and demand move together)
- **Strategy:** Use weather variables to shift supply exogenously

Instrument	First-stage *F*	2SLS Elasticity	Verdict
Wave height (2-day lag)	22.64	-0.96	Strong ✓
Wind speed (3-day lag)	6.58	-1.96	Weak — unreliable
OLS	—	-0.55	(biased)

- Same outcome, same sample — only instrument quality differs
- Weak instrument produces a very different, unreliable estimate
- **Always report and scrutinise the first-stage F-statistic**

Who Does IV Estimate the Effect For?

Local Average Treatment Effect (LATE)

Heterogeneous Treatment Effects

- In reality, treatment effects $\delta_i = Y_i^1 - Y_i^0$ differ across individuals
- With heterogeneous effects, IV does **not** recover the ATE (effect for everyone)
- Individuals can be classified by how the instrument affects their treatment status:

Type	D if $Z = 0$	D if $Z = 1$
Compliers	0	1
Always-takers	1	1
Never-takers	0	0
Defiers	1	0 (ruled out by monotonicity)

- Always-takers and never-takers contribute nothing to the first stage
- IV identifies the effect *only for compliers*

The LATE Theorem (Imbens & Angrist 1994)

- Under relevance, exclusion, and **monotonicity** (no defiers), IV estimates:

$$\delta^{\text{LATE}} = \mathbb{E}[Y_i^1 - Y_i^0 \mid \text{Complier}]$$

- In the binary instrument case, this equals the Wald estimator:

$$\delta^{\text{LATE}} = \frac{\mathbb{E}[Y \mid Z=1] - \mathbb{E}[Y \mid Z=0]}{\mathbb{E}[D \mid Z=1] - \mathbb{E}[D \mid Z=0]}$$

- **Who are the compliers?**
  - **Card (1995):** Students who attended college *because* a college was nearby (credit-constrained, marginal)
  - **Oregon Medicaid lottery:** People who enrolled in Medicaid *because* they won the lottery
  - **Monetary policy:** Firms/households whose spending responded to the policy rate change (not those who were constrained regardless)
- Always ask: **Who are my compliers, and is the LATE the parameter I care about?**

Takeaways: Compliers and Central Bank Relevance

LATE interpretation

- IV identifies the effect of income shocks *transmitted through the agricultural channel*
- Complier intuition: country-years where rainfall/vegetation shocks actually move GDP growth
- "Never-taker" intuition: diversified or industrial economies where rain barely moves GDP
- "Defier" concern: rain can hurt income through floods or tourism disruption
- Always ask: *who responds to my instrument?*

Why this matters for central banks

- Weather shocks $\to$ agricultural output $\to$ food prices $\to$ CPI
- Widely used to instrument for commodity price shocks in emerging market models
- Natural experiment for fiscal multipliers: harvest failures as income shocks
- Pattern repeats across Asia, Africa, Latin America — wherever agriculture is a large share of GDP

Extensions

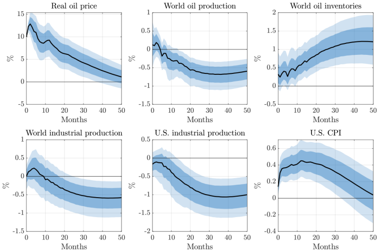
IV logic in macro and applied policy research

Känzig (2021): IV Logic in Macro

OPEC announcement windows provide high-frequency variation in oil supply expectations.

- **Question:** What do oil supply news shocks do to the macroeconomy?
- **Method:** Proxy-SVAR / external-instrument VAR, using the surprise as an instrument inside a structural VAR
- **Instrument:** Oil futures price changes in a narrow window around OPEC production announcements
- **Relevance:** Announcement-day variance is about 3× higher than control days; first-stage **F = 22.67**, robust **F = 10.55**
- **Exclusion challenge:** The event-window surprise should capture oil supply news, not new demand or broader macro news
- **Teaching point:** This is not simple 2SLS, but the IV instinct is the same — isolate cleaner variation in an endogenous macro variable

Känzig (2021): Macroeconomic Effects



First stage regression: F: 22.67, robust F: 10.55, R^2: 4.22%, Adjusted R^2: 4.04%

Figure 3: Impulse responses to an oil supply news shock

Notes: Impulse responses to an oil supply news shock, normalized to increase the real price of oil by 10 percent on impact. The solid line is the point estimate and the dark and light shaded areas are 68 and 90 percent confidence bands, respectively.

Source: Känzig (2021), Figure 3. IRFs to a 10% oil supply news shock. 68%/90% CI.

Pattern: oil prices and CPI rise; U.S. industrial production falls; the dollar depreciates.

Shift-Share (Bartik) Instruments

- **Question:** How does a local labour demand shock affect local wages and employment?
- **Challenge:** Local shocks are correlated with local economic conditions
- **Bartik instrument:**

$$B_\ell = \sum_k \underbrace{z_{\ell k}}_{\text{initial share of industry } k \text{ in area } \ell} \times \underbrace{g_k}_{\text{national growth of industry } k}$$

- **Applications:**
  - Import competition shocks (Autor, Dorn & Hanson 2013)
  - Immigration waves on local labour markets
  - Monetary policy transmission across regions
- **Key question:** Where does identification actually come from — the *shifts* (g_k) or the *shares* ($z_{\ell k}$)?

Goldsmith-Pinkham, Sorkin & Swift (2020)

Same IV question, fancier instrument: what variation is really acting as Z?

- In a Bartik design, the instrument combines local **shares** and national **shifts**
- **GPSS insight:** The identifying variation often comes from the **initial shares**, not just the national growth rates
  - Technical result: this can be written as a GMM/2SLS equivalence
  - Teaching takeaway: defend the part of the instrument that is actually doing the work
- **Connection to today:** This is still relevance + exclusion
  - Relevance: do the shares predict exposure to the shock?
  - Exclusion: are the initial shares unrelated to other outcome trends?

Rotemberg Weights

Which shares are doing the work?

- **Decomposition:** The Bartik estimator is a weighted sum of per-industry IV estimates:
$$\hat{\beta} = \sum_k \hat{\alpha}_k \hat{\beta}_k$$
where $\hat{\alpha}_k$ are **Rotemberg weights** (sum to 1, can be negative)
- Weights reveal *which industry shares* are driving the overall estimate
- **Practical diagnostics:**
  - Examine correlates of high-weight shares — do they suggest alternative channels?
  - Negative weights are a warning sign: no clean LATE interpretation
- **GPSS finding:** Oil & gas gets the largest weight in their labor supply application, but is correlated with many observables $\Rightarrow$ instrument fails
- **Rule:** Defend your top-weight shares, not just the aggregate national shock

Bazot (2024): Financial Deregulation

Does financial deregulation lower the cost of financial intermediation?

Setup

- **Data:** 15 advanced economies, 1970–2014
- **Outcome:** Unit cost of financial intermediation
- **Treatment:** Change in deregulation index
- **Problem:** Reverse causality — governments deregulate when costs are *high*

IV Strategy

- **Instrument:** Lagged *level* of deregulation
- **Relevance:** Low-deregulation countries reform more; **F = 56.18**
- **Exclusion:** Past level should not directly affect future cost *changes*
- **Caution:** Credible if lagged regulation captures convergence, not deeper institutional trends
- **Result:** 10pp $\uparrow$ deregulation $\Rightarrow$ unit cost $\downarrow$ 0.11–0.17 cents

Conclusion

Summary

- **Endogeneity** (omitted variables, simultaneity, measurement error) makes OLS biased and inconsistent
- **IV** solves this by exploiting exogenous variation in the treatment:
  - **Relevance**: instrument moves the treatment (testable: first-stage F)
  - **Exclusion restriction**: instrument does not directly affect the outcome (requires theory)
- **2SLS** is the practical implementation: first stage purges endogeneity, second stage estimates the causal effect
- **Weak instruments** are dangerous — always check F > 10; more instruments $\neq$ better
- **IV estimates the LATE** for compliers; generalising requires additional assumptions
- **Good instruments are rare and valuable** — finding one requires institutional knowledge, creativity, and careful economic reasoning

IV in Central Banking: Where to Apply It

- **Monetary policy transmission:** Romer & Romer narrative shocks isolate policy-rate movements not driven by the current outlook
- **Trade and external shocks:** Shift-share designs use external demand or import shocks to study local exposure
- **Agricultural and commodity shocks:** Rainfall can instrument for agricultural GDP or food-price pressure in emerging markets
- **Oil supply shocks:** Känzig uses OPEC announcement windows to isolate supply-driven oil news from demand-driven oil movements
- **General principle:** Look for policy changes, natural events, or institutional rules that shifted the treatment “accidentally”

Common Pitfalls

- **Exclusion restriction violation:** If z_i has any direct effect on y_i, IV is inconsistent and can be worse than OLS
- **Weak first stage:** $F < 10 \Rightarrow$ 2SLS is biased toward OLS; always report F
- **Over-instrumenting:** Adding many weak instruments dilutes valid variation and reintroduces bias
- **Forgetting LATE:** IV estimates the effect for compliers; don't claim it as the population-average effect
- **Bartik shares:** Shift-share identification rests on initial shares being exogenous — use Rotemberg weights (Goldsmith-Pinkham et al. 2020) to find which shares drive the estimate and test their validity
- **Wrong standard errors:** Never run 2SLS manually as two separate OLS regressions — always use `ivregress 2sls` (Stata) or equivalent

References

- Angrist, J. D., & Krueger, A. B. (1991). Does Compulsory School Attendance Affect Schooling and Earnings? *The Quarterly Journal of Economics*, 106(4), 979–1014.
- Card, D. (1995). Using Geographic Variation in College Proximity to Estimate the Return to Schooling. In L. N. Christofides, E. K. Grant, & R. Swidinsky (Eds.), *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp*. University of Toronto Press.
- Imbens, G. W., & Angrist, J. D. (1994). Identification and Estimation of Local Average Treatment Effects. *Econometrica*, 62(2), 467–475.
- Miguel, E., Satyanath, S., & Sergenti, E. (2004). Economic Shocks and Civil Conflict: An Instrumental Variables Approach. *Journal of Political Economy*, 112(4), 725–753.
- Romer, C. D., & Romer, D. H. (2004). A New Measure of Monetary Shocks: Derivation and Implications. *American Economic Review*, 94(4), 1055–1084.
- Känzig, D. R. (2021). The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements. *American Economic Review*, 111(4), 1092–1125.
- Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik Instruments: What, When, Why, and How. *American Economic Review*, 110(8), 2586–2624.
- Bazot, G. (2024). Deregulation and Financial Intermediation Cost: An International Comparison. *Journal of Money, Credit and Banking*, 56(5), 1129–1171.