Econometrics

Practical Session 23

Instrumental Variables: Checking Instrument Validity

Ricardo Gouveia-Mendes
rgouveiamendes@ucp.pt

Spring 2025-26

Católica-Lisbon School of Business and Economics

Theoretical Wrap-up

Weak Instrument Test

A weak instrument creates two problems even in large samples:

1. Bias toward OLS

If $\text{corr}(z, x) \approx 0$, TSLS is essentially estimating an expression of the form 0/0 → the distribution is non-normal and centered near the OLS estimate

2. Invalid inference

t-statistics, confidence intervals, and p-values based on TSLS are unreliable when instruments are weak

Weak Instrument Test

CATÓLICA LISBON BUSINESS & ECONOMICS

KEY CONCEPT 12.5: FIRST-STAGE F-STATISTIC RULE OF THUMB

Compute the F-statistic testing that all instrument coefficients are zero in the first-stage regression

$$H_0 : \pi_1 = \pi_2 = \ldots = \pi_m = 0$$

Rule of thumb: $F > 10$ → instruments are strong. $F \leq 10$ → weak instrument → TSLS is biased and inference is unreliable

Weak Instrument Test

What to do with weak instruments?

Discard weakest IVs

- If overidentified ($m > k$), drop the weakest and use only the strong ones
- SEs may increase, but they were not meaningful anyway with weak IVs

Find stronger instruments

Redesign the study: look for variables more strongly correlated with x while maintaining exogeneity

Use weak-IV robust methods

Anderson–Rubin (AR) test and confidence sets

Test of Overidentifying Restrictions (J-Statistic)

CATÓLICA LISBON BUSINESS & ECONOMICS

When $m > k$ (more instruments than endogenous regressors), we can **partially test** instrument exogeneity

THE HANSEN–SARGAN TEST

$$\hat{u}_i^{TSLS} = \alpha + \sum_{j=1}^{m} \beta_j z_{ji} + \sum_{s=1}^{q} \gamma_s w_{si} + v_i$$

$$H_0 : \beta_j = 0$$

$$J = m \cdot F, \quad J \sim \chi^2_{m-k} \quad \text{under } H_0$$

IMPORTANT LIMITATION

The J-test requires the assumption that **at least k instruments are valid**. It can only detect problems with the extra $(m - k)$ instruments. When $m = k$, $J = 0$ always → no test is possible

Hausman Endogeneity Test (Durbin–Wu–Hausman)

- **Question:** Is x_i actually endogenous?
- **Procedure** (for model $y_i = \beta_0 + \beta_1 x_i + u_i$ with instrument z_i):
  1. Run the first stage: $x_i = \pi_0 + \pi_1 z_i + v_i$. Save residuals $\hat{v}_i$
  2. Augment the structural OLS regression with $\hat{v}_i$:
  $$y_i = \beta_0 + \beta_1 x_i + \delta \hat{v}_i + \varepsilon_i$$
  3. Test $H_0 : \delta = 0$, using robust standard errors
- If δ **significant** (reject H_0), x_i is endogenous → OLS is biased and inconsistent → prefer TSLS
- **Key property:** the coefficient on x_i in the augmented regression is identical to the TSLS estimate.

Practical Checklist for IV Regression

Before trusting any IV estimate, work through this checklist:

1. **Relevance** — is the first-stage F-statistic above 10?
2. **Exogeneity** — provide a convincing economic argument for why the instrument satisfies the exclusion restriction
3. **Overidentification test** — if $m > k$, run the J-test → a rejection signals that at least one instrument may be invalid
4. **Endogeneity test** — run Hausman to confirm that the endogenous regressor is actually endogenous → if not, OLS may suffice

Exercises

Exercise 1 | Demand for Cigarettes (S&W 12.1)

CATÓLICA LISBON BUSINESS & ECONOMICS

The table below reports TSLS estimates of the demand for cigarettes using US state panel data (48 states, 10-year differences 1985–1995). Dependent variable: $\ln\left(Q^{cig}_{1995}\right) - \ln\left(Q^{cig}_{1985}\right)$

Regressor	(1)	(2)	(3)
$\Delta \ln\left(P^{cig}_i\right)$	−0.94 (0.21) [−1.36, −0.52]	−1.34 (0.23) [−1.80, −0.88]	−1.20 (0.20) [−1.60, −0.81]
$\Delta \ln(\text{Inc}_i)$	0.53 (0.34) [−0.16, 1.21]	0.43 (0.30) [−0.16, 1.02]	0.46 (0.31) [−0.16, 1.09]
Intercept	−0.12 (0.07)	−0.02 (0.07)	−0.05 (0.06)
Instrument(s)	Sales tax	Cig.-specific tax	Both taxes
First-stage F	33.7	107.2	88.6
J-stat (p-value)	—	—	4.93 (0.026)

Standard errors in parentheses; 95% CI in brackets. Heteroskedasticity-robust SEs.

Exercise 1 | Demand for Cigarettes (S&W 12.1)

a. A new federal tax is estimated to raise the retail price of cigarettes by \$0.25 per pack. The current price is \$6.75. Use column (1) to predict the change in demand. Construct a 95% confidence interval.

Exercise 1 | Demand for Cigarettes (S&W 12.1)

CATÓLICA LISBON BUSINESS & ECONOMICS

a. A new federal tax is estimated to raise the retail price of cigarettes by $0.25 per pack. The current price is $6.75. Use column (1) to predict the change in demand. Construct a 95% confidence interval.

KEY TAKEAWAYS

- $\Delta \ln(P) \approx \Delta P/P = 0.25/6.75 = 0.037$ (a 3.7% price increase)
- Predicted $\Delta \ln(Q) = -0.94 \times 0.037 = -0.035$ (a **3.5% decrease** in quantity demanded)
- 95% CI for $\Delta \ln(Q)$: use the CI for the price elasticity $[-1.36, -0.52]$:
  - Lower bound: $-1.36 \times 0.037 = -0.05$ (−5%)
  - Upper bound: $-0.52 \times 0.037 = -0.019$ (−1.9%)
- **Conclusion:** the tax is predicted to reduce cigarette consumption by 3.5%, with a 95% CI of [−5%, −1.9%]

Exercise 1 | Demand for Cigarettes (S&W 12.1)

b. Suppose the United States enters a recession and income falls by 5%. Use column (1) to predict the change in demand.

Exercise 1 | Demand for Cigarettes (S&W 12.1)

b. Suppose the United States enters a recession and income falls by 5%. Use column (1) to predict the change in demand.

KEY TAKEAWAYS

- $\Delta \ln(Inc) \approx -0.05$ (a 5% fall in income)
- Predicted $\Delta \ln(Q) = 0.53 \times (-0.05) = -0.027$ (a **2.7% decrease** in quantity demanded)
- Cigarettes appear to be a **normal good** (positive income elasticity), so falling income reduces consumption
- **Note:** the income elasticity (0.53) is not statistically significant at 5% (its 95% CI crosses zero), so this prediction carries substantial uncertainty

Exercise 1 | Demand for Cigarettes (S&W 12.1)

c. Suppose you have additional data from 1993, 1994, 1996, and 1997. How would you expect the estimated price elasticity to change with an 8-year and a 12-year horizon?

Exercise 1 | Demand for Cigarettes (S&W 12.1)

CATÓLICA LISBON BUSINESS & ECONOMICS

c. Suppose you have additional data from 1993, 1994, 1996, and 1997. How would you expect the estimated price elasticity to change with an 8-year and a 12-year horizon?

KEY TAKEAWAYS

- The textbook reports: 5-year elasticity (1985–90) ≈ −0.79; 5-year elasticity (1990–95) ≈ −0.68; 10-year elasticity ≈ −0.94
- **Pattern:** longer time horizons yield **more elastic** demand (larger negative elasticity)
- **Intuition:** cigarettes are addictive; smokers find it harder to adjust in the short run. Over longer horizons, more smokers can quit, switch brands, or reduce consumption in response to higher prices
- **8-year horizon:** expect elasticity between −0.79 and −0.94, roughly around −0.85 to −0.90
- **12-year horizon:** expect elasticity more negative than −0.94, possibly around −1.0 or beyond

Exercise 1 | Demand for Cigarettes (S&W 12.1)

d. Suppose the first-stage F-statistic in column (1) were 63.7 instead of 33.7. Would the regression provide a reliable answer to part (a)?

Exercise 1 | Demand for Cigarettes (S&W 12.1)

CATÓLICA LISBON BUSINESS & ECONOMICS

d. Suppose the first-stage F-statistic in column (1) were 63.7 instead of 33.7. Would the regression provide a reliable answer to part (a)?

KEY TAKEAWAYS

- Yes — the answer to part a) would be essentially unchanged
- The relevant threshold is $F > 10$: both $F = 33.7$ (actual) and $F = 63.7$ (hypothetical) comfortably exceed this threshold → instruments are strong in either case
- A higher F is actually **better**: it means the instrument explains more variation in price, so the TSLS estimate is more precisely identified and the standard errors may be slightly smaller
- **Key point:** $F > 10$ is sufficient to rule out serious weak-instrument bias; having $F = 63.7$ vs $F = 33.7$ does not change the reliability conclusion

Exercise 2 | Property Rights and Income (S&W 12.9)

A researcher collects data from 60 countries (all former colonies) and estimates:

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

where Y_i is GDP per capita and X_i is an index of protection against expropriation (0–10; higher = more secure property rights)

Exercise 2 | Property Rights and Income (S&W 12.9)

a. Explain why the OLS estimates are likely to be unreliable and indicate the likely direction of bias. *(Hint: in which direction does causality run?)*

Exercise 2 | Property Rights and Income (S&W 12.9)

a. Explain why the OLS estimates are likely to be unreliable and indicate the likely direction of bias. *(Hint: in which direction does causality run?)*

KEY TAKEAWAYS

- **Source 1 — Simultaneous causality:** richer countries (high Y) can afford to build and maintain better institutions (high X) → causality runs both ways → $\text{cov}(X_i, u_i) > 0$ → **upward** OLS bias (OLS overstates β_1)
- **Source 2 — Measurement error:** the protection index is survey-based and imperfectly measures true institutional quality → attenuation bias → **downward** OLS bias
- **Net direction:** uncertain — both forces are present. If simultaneity dominates, OLS is upward biased; if measurement error dominates (as in AJR), OLS is downward biased
- **Bottom line:** OLS is inconsistent regardless of the direction, since $E[u|X] \neq 0$

Exercise 2 | Property Rights and Income (S&W 12.9)

b. All countries in the sample were former colonies. Institutions protecting property rights may have originated from European settlements. The decision by Europeans to settle reflected concerns about settler mortality. Explain how settler mortality might be used as an instrument to estimate the effect of more secure property rights on income.

Exercise 2 | Property Rights and Income (S&W 12.9)

KEY TAKEAWAYS

This is exactly the AJR (2001) strategy from Session 21:

- **Instrument:** Z_i = log of 19th-century European settler mortality rate
  - High settler mortality → Europeans did not settle → set up extractive institutions → low X (weak property rights) today
  - **Relevance (**$\text{corr}(Z_i, X_i) \neq 0$**):** AJR find this correlation is strong (first-stage F well above 10)
  - **Exogeneity (**$\text{corr}(Z_i, u_i) = 0$**):** 19th-century disease environment (malaria, yellow fever) affects today's income **only through** the institutional channel
- **Connection:** the 2SLS estimate (0.929) is roughly double the OLS estimate (0.516), suggesting measurement error dominates and OLS substantially understates the true institutional effect on income

Exercise 3 | Omitted Controls in IV (S&W 12.10)

One student specifies the IV regression:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 W_i + u_i \qquad \text{(uses } Z_i \text{ as instrument)}$$

A second student estimates the same model but **omits** W_i:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad \text{where } \varepsilon_i = \beta_2 W_i + u_i$$

Exercise 3 | Omitted Controls in IV (S&W 12.10)

a. The first student says: *"If Z_i and W_i are correlated, then the second student's IV estimator will not be consistent."* Is the first student right?

Exercise 3 | Omitted Controls in IV (S&W 12.10)

a. The first student says: *"If Z_i and W_i are correlated, then the second student's IV estimator will not be consistent."* Is the first student right?

KEY TAKEAWAYS

- **Yes** — the first student is correct
- If W_i is omitted, the error becomes $\varepsilon_i = \beta_2 W_i + u_i$. For the second student's IV to be consistent, we need $\text{cov}(Z_i, \varepsilon_i) = 0$:

$$\text{cov}(Z_i, \varepsilon_i) = \beta_2 \underbrace{\text{cov}(Z_i, W_i)}_{\neq 0 \text{ (by assumption)}} + \underbrace{\text{cov}(Z_i, u_i)}_{=0 \text{ (valid IV)}} = \beta_2 \,\text{cov}(Z_i, W_i) \neq 0$$

  (assuming $\beta_2 \neq 0$). **The exogeneity condition is violated →** the IV estimator is inconsistent
- The second student's instrument is contaminated by its correlation with the omitted control W_i

Exercise 3 | Omitted Controls in IV (S&W 12.10)

b. The second student argues: *"If the true value of β_2 is 0, then my IV estimator will be consistent."* Is the second student correct?

Exercise 3 | Omitted Controls in IV (S&W 12.10)

b. The second student argues: *"If the true value of β_2 is 0, then my IV estimator will be consistent."* Is the second student correct?

KEY TAKEAWAYS

- **Yes** — the second student is correct **in this specific case**
- If $\beta_2 = 0$, then W_i does not belong in the model: the true model is $Y_i = \beta_0 + \beta_1 X_i + u_i$, and the second student's regression is correctly specified
- With the true error being u_i (not $\varepsilon_i = \beta_2 W_i + u_i$): $\text{cov}(Z_i, u_i) = 0$ by assumption → Z is a valid instrument for the true model → the exogeneity condition is satisfied → IV is consistent
- If the omitted variable is genuinely irrelevant ($\beta_2 = 0$), omitting it from the model (and from the first stage) is harmless for consistency