Lecture 12: Asymptotic Consistency of the IV Estimator

Econometrics 2 — *Extending the Two-Step Recipe from OLS to IV*

Instructor: Prof. S. Arvanitis | **Notes:** N. Papavasileiou | **Digitisation & added notes:** T. Kourtalis

Semester: Spring 2026

▷ Amber = Handwritten Notes (professor's words) ◇ Teal = Student's Notes

Recall from Earlier Lectures

Where we are in the course:

- **Lectures 2–3:** Population identification of OLS — $M_n^*(\beta)$ has unique minimum at β_0.
- **Lectures 6–7:** IV estimator introduced; $\hat{\beta}_n^{IV} = (X'ZWZ'X)^{-1}X'ZWZ'Y$; consistency shown informally via $\hat{\beta}_n - \beta_0 = (X'ZWZ'X)^{-1}X'ZWZ'\varepsilon \xrightarrow{p} 0$.
- **Lectures 10–11:** The *two-step consistency recipe* formalised and applied to OLS: **(a)** locally uniform $\widetilde{M}_n \xrightarrow{p} M$, then **(b)** M has a well-separated minimum at β_0.

This lecture applies the exact same two-step recipe to the IV objective. The algebra mirrors Lecture 11; the key change is that Z (instruments) appears everywhere instead of X.

1 Setup: The IV Objective and the Strategy

▷ **Handwritten Notes** (what the professor said)

Let us examine weak consistency for the IV estimator within the framework of the linear model with instrumental variables. The IV objective function is:

$$M_n(\beta) = \frac{1}{n^2}(Y_n - X_n\beta)'\, Z_nWZ_n'\,(Y_n - X_n\beta)$$

Since the model is well-specified, can we design — based on our general theory — sufficient conditions for weak consistency of $\hat{\beta}_n$?

◇ **Student's Notes**

The three-step strategy (identical structure to Lecture 11):

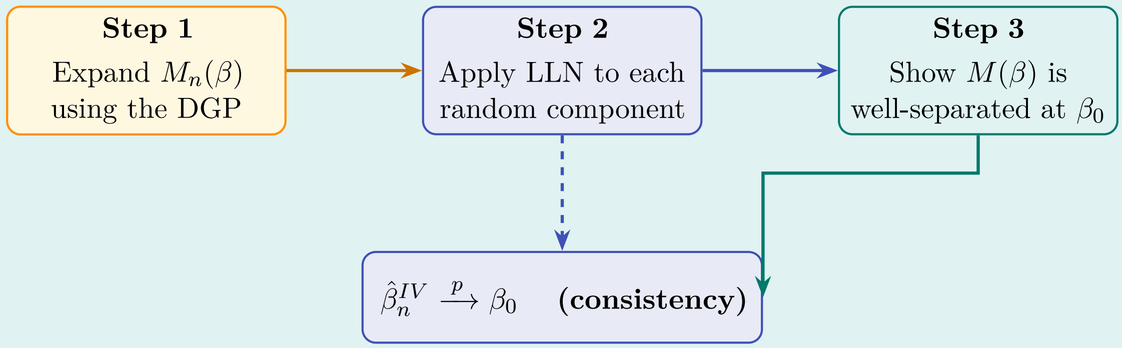

Figure 1: Three-step strategy for IV consistency, identical in structure to the OLS proof in Lecture 11.

High-level comparison: OLS (Lecture 11) vs. IV (Lecture 12)

Aspect	**OLS (Lect. 11)**	**IV (Lect. 12)**
Objective $M_n(\beta)$	$\frac{1}{n}(Y-X\beta)'(Y-X\beta)$	$\frac{1}{n^2}(Y-X\beta)'ZWZ'(Y-X\beta)$
"Inner" matrix	I_n	ZWZ'
Curvature component	$X'X/n$	$\big(X'Z/n\big)\,W\,\big(Z'X/n\big)$
Cross-product component	$X'\varepsilon/n$	$\big(X'Z/n\big)\,W\,\big(Z'\varepsilon/n\big)$
Exogeneity needed for	$\mathbb{E}[X'\varepsilon]=0$	$\mathbb{E}[Z'\varepsilon]=0$
Limit curvature matrix	$Q_{X'X}$	$Q_{Z'X}'\,W\,Q_{Z'X}$

Z simply replaces X in the "projection" role. OLS is the special case $Z = X$, $W = I$.

2 Step 1 — Expanding the IV Objective Function

▷ **Handwritten Notes** (what the professor said)

Since $Y_n = X_n\beta_0 + \varepsilon_n$, substituting:

$$M_n(\beta) = \frac{1}{n^2}\big(\varepsilon_n + X_n(\beta_0-\beta)\big)'Z_nWZ_n'\big(\varepsilon_n + X_n(\beta_0-\beta)\big)$$

Expanding $(a+b)'A(a+b) = a'Aa + b'Ab + 2a'Ab$ with $a = X_n(\beta_0 - \beta)$, $b = \varepsilon_n$:

$$M_n(\beta) = \underbrace{\left(\frac{\varepsilon_n' Z_n}{n}\right) W \left(\frac{Z_n' \varepsilon_n}{n}\right)}_{T_1\text{: noise, constant in } \beta}$$

$$+ \underbrace{(\beta_0 - \beta)' \left(\frac{X_n' Z_n}{n}\right) W \left(\frac{Z_n' X_n}{n}\right) (\beta_0 - \beta)}_{T_2\text{: quadratic (curvature)}}$$

$$+ \underbrace{2(\beta_0 - \beta)' \left(\frac{X_n' Z_n}{n}\right) W \left(\frac{Z_n' \varepsilon_n}{n}\right)}_{T_3\text{: cross term (tilt)}}$$

◇ Student's Notes

The role of each term:

Term	Size	Role
T_1	scalar	Noise floor; does not affect $\arg\min$
T_2	scalar	Creates the bowl; > 0 for $\beta \neq \beta_0$
T_3	scalar	Tilts the bowl; must $\xrightarrow{p} 0$ for correct centring

Dimension ledger — every factor in M_n:

Factor	Size	Limit (next section)
$Z_n' X_n / n$	$q \times p$	$Q_{Z'X}$
$X_n' Z_n / n$	$p \times q$	$Q_{Z'X}'$
$Z_n' \varepsilon_n / n$	$q \times 1$	$0_{q \times 1}$
$\varepsilon_n' Z_n / n$	$1 \times q$	$0_{1 \times q}$
W	$q \times q$	W (fixed, $\succ 0$)

When $q > p$, $Z'X/n$ is *rectangular*—it cannot be inverted directly. That is why the curvature matrix becomes the product $Q_{Z'X}' W Q_{Z'X}$ (a $p \times p$ positive definite matrix).

3 Step 2 — Applying the LLN to Each Component

3.1 Component (a): Instrument–Regressor Covariance

▷ **Handwritten Notes** (what the professor said)

Assumption (a):

$$\frac{Z_n' X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} Z_{(i)}' X_{(i)} \xrightarrow{p} Q_{Z'X}, \qquad \mathrm{rank}(Q_{Z'X}) = p.$$

Sufficient condition: If $(Z_{(i)}, X_{(i)})$ are i.i.d. and $\mathbb{E}[Z_{(1)}' X_{(1)}]$ exists, Kolmogorov's LLN gives $Q_{Z'X} = \mathbb{E}[Z_{(1)}' X_{(1)}]$.

◇ **Student's Notes**

This is the *relevance* condition in asymptotic form.

Concept	Finite-sample (Lect. 6)	Asymptotic (Lect. 12)
Object	$Z_n' X_n \quad (q \times p)$	$Q_{Z'X} = \mathbb{E}[Z'X] \quad (q \times p)$
Condition	$\mathrm{rank}(Z_n' X_n) = p$	$\mathrm{rank}(Q_{Z'X}) = p$
Consequence	$(X'ZWZ'X)$ invertible	$Q_{Z'X}' W Q_{Z'X} \succ 0$
OLS analogue	$\mathrm{rank}(X_n) = p$	$\mathrm{rank}(Q_{X'X}) = p$

By the Continuous Mapping Theorem (transpose is continuous): $X_n' Z_n / n \xrightarrow{p} Q_{Z'X}'$.

3.2 Component (b): Instrument–Error Orthogonality

▷ **Handwritten Notes** (what the professor said)

Assumption (b):

$$\frac{Z_n' \varepsilon_n}{n} = \frac{1}{n}\sum_{i=1}^{n} Z_{(i)}' \varepsilon_{(i)} \xrightarrow{p} 0_{q\times 1}.$$

This holds when $(Z_{(i)}, \varepsilon_{(i)})$ are i.i.d., by Kolmogorov's LLN applied to $\mathbb{E}[Z_{(1)}' \varepsilon_{(1)}] = 0$ (IV exogeneity).

◇ **Student's Notes**

The LIE step (same structure as Lecture 11, with Z replacing X):

$$\mathbb{E}\big[Z'_{(1)}\varepsilon_{(1)}\big] = \mathbb{E}\left[Z'_{(1)}\,\underbrace{\mathbb{E}[\varepsilon_{(1)} \mid Z_{(1)}]}_{=0}\right] = 0_{q\times 1}.$$

Why Z can be exogenous when X is not:

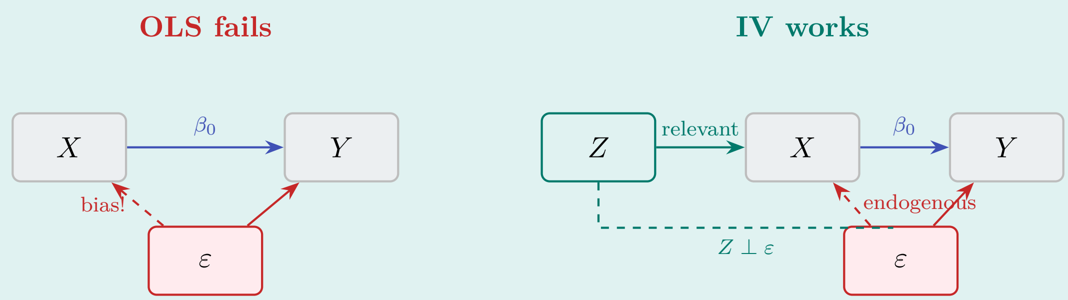

Figure 2: OLS fails when X is correlated with ε. IV routes through Z, which is orthogonal to ε even when X is not.

4 Step 3 — The Limit Function and Positive Definiteness

▷ **Handwritten Notes** (what the professor said)

Given (a) and (b), $M_n(\beta)$ converges locally uniformly to:

$$M(\beta) = \underbrace{0_{1\times q}\,W\,0_{q\times 1}}_{=0} + (\beta_0-\beta)'\,Q'_{Z'X}\,W\,Q_{Z'X}\,(\beta_0-\beta) + \;2(\beta_0-\beta)'\,Q'_{Z'X}\,W\;\underbrace{0_{q\times 1}}_{=0}$$

$$= (\beta_0-\beta)'\;\underbrace{Q'_{Z'X}\,W\,Q_{Z'X}}_{=:\,A}\,(\beta_0-\beta).$$

Because $\operatorname{rank}(Q_{Z'X}) = p$ and $W \succ 0$, the matrix $A = Q'_{Z'X}WQ_{Z'X}$ is **positive definite**, so $M(\beta)$ is well-separated at β_0.
If Θ is convex: $\hat\beta_n \xrightarrow{p} \beta_0$.

■ **Proof / Derivation**

$A = Q'_{Z'X} W Q_{Z'X}$ is positive definite

Write $W = LL'$ (Cholesky; valid since $W \succ 0$). Then for any $v \neq 0_p$:

$$v' A v = v'\, Q'_{Z'X}\, LL'\, Q_{Z'X}\, v = \|L' Q_{Z'X} v\|^2.$$

This is zero iff $L' Q_{Z'X} v = 0$. Since L is invertible, this requires $Q_{Z'X} v = 0$. But $\operatorname{rank}(Q_{Z'X}) = p$ implies the null space of $Q_{Z'X}$ on $\mathbb{R}^p$ is trivial, so $v = 0$. Therefore $v' A v > 0$ for all $v \neq 0$, i.e. $A \succ 0$. □

Well-separation bound: for any β with $\|\beta - \beta_0\| \geq \varepsilon$:

$$M(\beta) \geq \lambda_{\min}(A)\, \|\beta - \beta_0\|^2 \geq \underbrace{\lambda_{\min}(A)\, \varepsilon^2}_{=:\, \delta > 0}.$$

◇ **Student's Notes**

Side-by-side: OLS vs. IV limit bowls

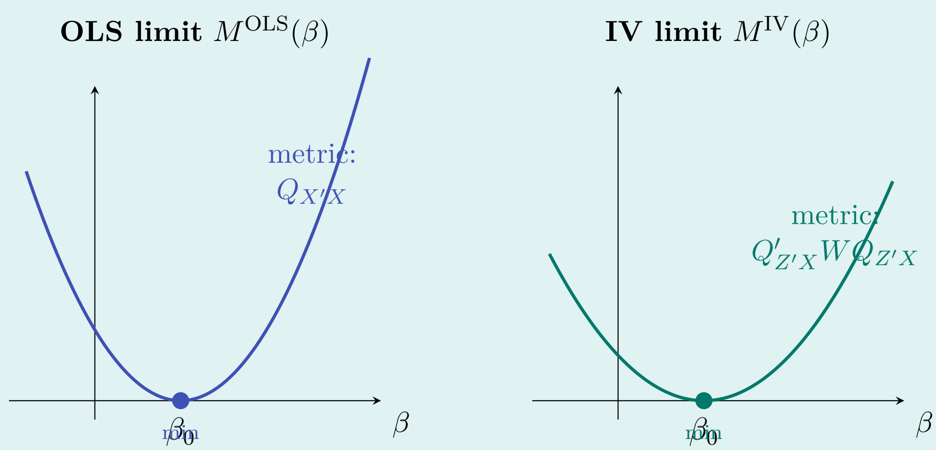

Figure 3: Both limit functions are parabolas minimised at β_0, but with different curvatures. OLS curvature is $Q_{X'X}$; IV curvature is $Q'_{Z'X} W Q_{Z'X}$. Both are positive definite, ensuring unique well-separated minima.

Key Result

Weak Consistency of the IV Estimator

Under:

(i) Well-specified: $Y_n = X_n \beta_0 + \varepsilon_n$, $\beta_0 \in \Theta$

(ii) Relevance: $Z_n' X_n/n \xrightarrow{p} Q_{Z'X}$, $\operatorname{rank}(Q_{Z'X}) = p$

(iii) IV exogeneity: $Z_n' \varepsilon_n/n \xrightarrow{p} 0$

(iv) $W \succ 0$; Θ convex

The IV objective converges to:

$$M(\beta) = (\beta_0 - \beta)' \underbrace{Q_{Z'X}' W Q_{Z'X}}_{\succ 0} (\beta_0 - \beta)$$

which has a unique, well-separated minimum at β_0. Therefore:

$$\boxed{\hat{\beta}_n^{IV} \xrightarrow{p} \beta_0 \qquad \textbf{for any positive definite } W}$$

5 Why W Does Not Affect Consistency

★ **Intuition**

Different choices of W change the *curvature* (steepness) of the bowl $M(\beta)$ but **cannot move its bottom**. The minimum is at β_0 for any $W \succ 0$ because $A = Q_{Z'X}' W Q_{Z'X}$ is always positive definite under the rank condition.

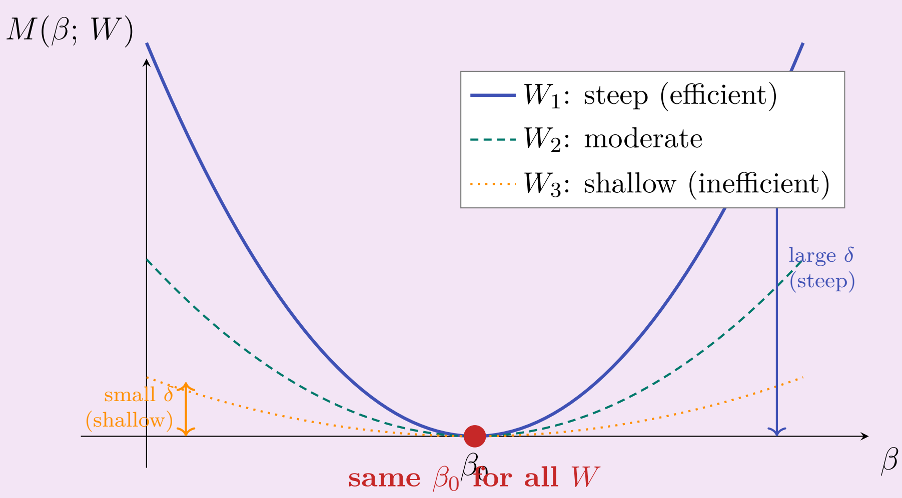

Figure 4: Three choices of W produce bowls with different steepness but the same *minimum at β_0. A steeper bowl gives a larger well-separation gap δ and faster convergence (better efficiency), but all three are consistent.*

Efficient W^*: the choice that minimises the asymptotic variance $V = H^{-1} V H^{-1}$ is

$W^* = [\text{Var}(Z'\varepsilon/\sqrt{n})]^{-1}$, the inverse of the variance of the moment vector. This is **efficient GMM**, derived in later lectures.

6 Complete Logical Flow

◇ Student's Notes

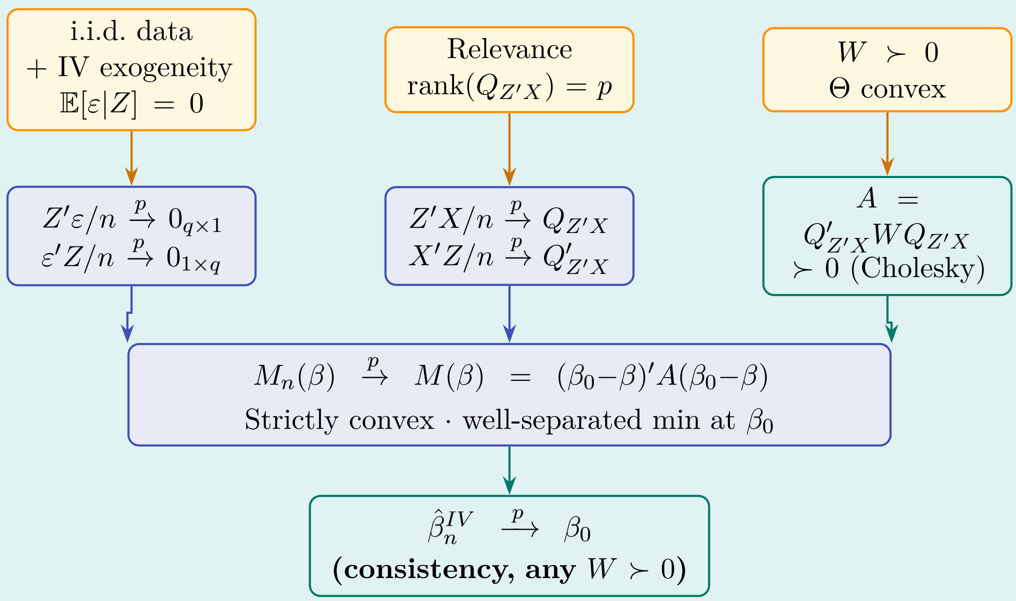

Figure 5: Complete logical flow. Row 1: assumptions. Row 2: LLN + CMT consequences. Row 3: limit function is a p.d. quadratic. Row 4: conclusion.

Origin of each piece:

Piece	Tool	First appeared
Two-step recipe	General M-estimator theory	Lecture 10
LLN for $Z'X/n$	Kolmogorov LLN (i.i.d.)	Lecture 11
LLN+LIE for $Z'\varepsilon/n$	Kolmogorov LLN + LIE	Lecture 11
Cholesky $\Rightarrow A \succ 0$	Linear algebra	Lecture 6
Well-separation	$\delta \geq \lambda_{\min}(A)\varepsilon^2$	Lecture 11

7 Mild Misspecification for IV

◇ Student's Notes

What if $\beta_0 \notin \Theta$?
The expansion of $M_n(\beta)$ never used $\beta_0 \in \Theta$, so $M_n \xrightarrow{p} M$ still holds on all of Θ. If Θ is closed and convex but $\beta_0 \notin \Theta$, the IV estimator converges to the **pseudo-true value**:

$$\beta^* = \arg\min_{\beta\in\Theta} (\beta_0 - \beta)' \, Q'_{Z'X} W Q_{Z'X} \, (\beta_0 - \beta),$$

the $Q'_{Z'X} W Q_{Z'X}$-Mahalanobis projection of β_0 onto Θ.

Case	OLS (Lect. 11)	IV (Lect. 12)
$\beta_0 \in \Theta$	$\hat\beta \xrightarrow{p} \beta_0$	$\hat\beta \xrightarrow{p} \beta_0$
$\beta_0 \notin \Theta$	Proj. w.r.t. $Q_{X'X}$	Proj. w.r.t. $Q'_{Z'X} W Q_{Z'X}$

8 Classroom Q&A

▷ Classroom Q&A

Student question: *Can the parameter space Θ change with n?*

Professor's answer: Yes, Θ can depend on n (writing Θ_n), but then we must explicitly prove the *convergence of the parameter space itself* — e.g. in the sense of Painlevé–Kuratowski set convergence. This is a non-trivial extension not pursued in this course.

◇ Student's Notes

Examples where Θ_n varies:

- **Ridge/Shrinkage:** constraint $\|\beta\|^2 \leq c_n$ with $c_n \to \infty$ (Lecture 5).
- **Sieve estimation:** Θ_n expands to approximate an infinite-dimensional space.
- **Regularised GMM:** penalty parameter shrinks as n grows.

In each case one must show $\Theta_n \to \Theta_\infty$ before applying the two-step recipe.

Quick-Reference Summary

◇ Student's Notes

Lecture 12 at a glance:

Topic	What was accomplished
IV objective expansion	$M_n = T_1(\text{noise}) + T_2(\text{bowl}) + T_3(\text{tilt})$ with Z replacing I
Relevance (a)	$Z'X/n \xrightarrow{p} Q_{Z'X}$, $\text{rank}(Q_{Z'X}) = p$
Exogeneity (b)	$Z'\varepsilon/n \xrightarrow{p} 0$ via LLN + LIE
Positive definiteness	$A = Q'_{Z'X} W Q_{Z'X} \succ 0$ via Cholesky
Limit function	$M(\beta) = (\beta_0 - \beta)' A (\beta_0 - \beta)$; well-separated min at β_0
Main result	$\hat{\beta}_n^{IV} \xrightarrow{p} \beta_0$ for any $W \succ 0$
W and efficiency	W shifts curvature, not location of minimum; optimal $W^* = [\text{Var}(Z'\varepsilon/\sqrt{n})]^{-1}$
Mild misspecification	$\beta_0 \notin \Theta \Rightarrow \hat{\beta}_n \to \beta^*$ ($Q'_{Z'X} W Q_{Z'X}$-projection)

Key formulas:

Object	Expression / Limit
IV objective	$M_n(\beta) = \frac{1}{n^2}(Y - X\beta)' Z W Z' (Y - X\beta)$
Relevance limit	$Z'X/n \xrightarrow{p} Q_{Z'X}, \ \text{rank}(Q_{Z'X}) = p$
Exogeneity limit	$Z'\varepsilon/n \xrightarrow{p} 0_{q\times 1}$
Limit curvature	$A = Q'_{Z'X} W Q_{Z'X} \succ 0$
Limit function	$M(\beta) = (\beta_0 - \beta)' A (\beta_0 - \beta)$
Well-separation	$\delta \geq \lambda_{\min}(A)\, \varepsilon^2 > 0$
Conclusion	$\hat{\beta}_n^{IV} \xrightarrow{p} \beta_0$
Pseudo-true value	$\beta^* = \arg\min_{\beta\in\Theta} (\beta_0 - \beta)' A (\beta_0 - \beta)$ when $\beta_0 \notin \Theta$

Figures:

Fig.	Content
1	Three-step strategy (TikZ flowchart)
2	OLS fails vs. IV works — endogeneity path diagram
3	OLS vs. IV limit bowls side by side
4	Different W: same β_0, different curvature
5	Complete logical flow: assumptions $\to$ LLN $\to$ p.d. limit $\to$ consistency

Lectures 2–12 narrative arc:

Lectures	Theme
2–3	Model setup; population identification
4–5	Analogy principle; extremum estimators; asymptotic properties
6–7	Endogeneity; IV / GMM; just-identified case
8–9	LAD estimator; LP reformulation
10–11	Two-step consistency recipe; OLS proof; mild misspecification
12	**IV consistency via two-step recipe; W-invariance; mild misspecification for IV**