

# Exam 3 Review

Chapters 8, 9, 10, 12


ECON3500: Econometrics and Applications

Spring 2026

# Roadmap

---

Chapter	Topic	Key Skills
8	Nonlinear Relationships	Marginal effects, log interpretation, interactions
9	Assessing Studies	Threats to internal validity
10	Panel Data & DiD	Fixed effects, difference-in-differences
12	Instrumental Variables	2SLS, instrument validity, LATE

 Exam 3: April 23

Closed book. One-page formula sheet (double-sided) + calculator allowed. No Stata.

# Chapter 8: Nonlinear Relationships

# Ch 8 Big Picture

---

We relax the assumption that the effect of  $X$  on  $Y$  is **constant**.

Three tools — all still **linear in parameters**, so OLS still works:

**Polynomial regression** — model curvature with  $X^2$ ,  $X^3$

**Logarithmic transformations** — model percentage changes

**Interaction terms** — allow the effect of  $X_1$  to depend on  $X_2$

# Polynomial Regression: Example

---

$$\widehat{TestScore} = 607.3 + 3.85 \times Income - 0.0423 \times Income^2$$

What is the marginal effect of income at  $Income = 10$  (thousand)?

$$\frac{\Delta TestScore}{\Delta Income} \approx 3.85 + 2(-0.0423)(10) = 3.85 - 0.846 = 3.00$$

At  $Income = 40$ ?

$$\frac{\Delta TestScore}{\Delta Income} \approx 3.85 + 2(-0.0423)(40) = 3.85 - 3.384 = 0.47$$

**Takeaway:** You **cannot** interpret  $\beta_1$  or  $\beta_2$  in isolation. Always compute the **marginal effect at a specific value of  $X$** . You should also know how to find the **turning point**:  $X^* = -\beta_1/2\beta_2$ .

# The Log Interpretation Table

Specification	Model	Interpretation of
Level-level	$Y = \beta_0 + \beta_1 X$	A 1-unit $\uparrow$ in $X \rightarrow \beta_1$ unit $\uparrow$ in $Y$
Log-level	$\ln(Y) = \beta_0 + \beta_1 X$	A 1-unit $\uparrow$ in $X \rightarrow (100 \times \beta_1)\% \uparrow$ in $Y$
Level-log	$Y = \beta_0 + \beta_1 \ln(X)$	A 1% $\uparrow$ in $X \rightarrow (\beta_1/100)$ unit $\uparrow$ in $Y$
Log-log	$\ln(Y) = \beta_0 + \beta_1 \ln(X)$	A 1% $\uparrow$ in $X \rightarrow \beta_1\% \uparrow$ in $Y$ (elasticity)

**Must Know:** This table is one of the **highest-value things** to have on your formula sheet. Know how to apply it.

# Log Interpretation: Practice

---

$$\ln(\widehat{Wage}) = 1.5 + 0.08 \times Educ$$

**This is log-level.** One additional year of education is associated with an approximately **8% increase** in wages.

$$\widehat{Wage} = 12.5 + 3.2 \times \ln(Experience)$$

**This is level-log.** A 1% increase in experience is associated with a  $3.2/100 = \$0.032$  increase in wages.

# Interaction Terms: The Idea

---

The effect of  $X_1$  on  $Y$  may **depend on the value of another variable**  $X_2$ .

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 (X_{1i} \times X_{2i}) + u_i$$

**Marginal effect of  $X_1$ :**

$$\frac{\Delta Y}{\Delta X_1} = \beta_1 + \beta_3 X_2$$

The effect of  $X_1$  **varies with**  $X_2$ .

# Three Types of Interactions

---

## 1. Binary $\times$ Binary ( $D_1 \times D_2$ )

$$Y_i = \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 (D_{1i} \times D_{2i}) + u_i$$

$\beta_3$  = **additional** effect of having *both*  $D_1 = 1$  and  $D_2 = 1$ , beyond their separate effects

## 2. Binary $\times$ Continuous ( $D \times X$ )

$$Y_i = \beta_0 + \beta_1 D_i + \beta_2 X_i + \beta_3 (D_i \times X_i) + u_i$$

Effect of  $X$  when  $D = 0$ :  $\beta_2$

Effect of  $X$  when  $D = 1$ :  $\beta_2 + \beta_3$

$\beta_3$  = how much the **slope** changes when  $D = 1$

## 3. Continuous $\times$ Continuous ( $X_1 \times X_2$ )

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 (X_{1i} \times X_{2i}) + u_i$$

Effect of  $X_1$ :  $\beta_1 + \beta_3 X_2$  (depends on  $X_2$ )

# Interaction Terms: Practice

---

$$\widehat{TestScore} = 700 - 1.0 \times STR - 0.5 \times PctEL - 0.02 \times (STR \times PctEL)$$

where  $STR$  = student-teacher ratio (students per teacher),  $PctEL$  = percent English learners (percentage, 0 to 100), and  $TestScore$  = points on a standardized test.

**Q: What is the effect of reducing STR by 1 in a district where PctEL = 20%?**

$$\frac{\Delta TestScore}{\Delta STR} = -1.0 + (-0.02)(20) = -1.0 - 0.4 = -1.4$$

Reducing STR by 1 raises test scores by **1.4 points** in a district with 20% English learners.

# Interaction Terms: Practice (cont.)

---

Q: What about in a district where PctEL = 0?

$$\frac{\Delta TestScore}{\Delta STR} = -1.0 + (-0.02)(0) = -1.0$$

Reducing STR by 1 raises test scores by **1.0 point**.

**Knowledge Check:** Does the interaction term suggest a **larger** or **smaller** effect of reducing class size in districts with many English learners?

**Larger** — the negative effect of STR is amplified when PctEL is high.

# Ch 8 Summary

---

Tool	Use When...	Key Skill
Polynomial ( $X^2$ )	Effect of $X$ changes as $X$ increases	Marginal effect at specific $X$
Logs	Think in percentages; skewed data	Log interpretation table
Interactions	Effect of $X_1$ depends on $X_2$	Marginal effect at specific $X_2$

## Common Mistakes:

Interpreting  $\beta_1$  alone in a polynomial or interaction model

Forgetting to multiply by 100 (or divide by 100) for log models

Including the interaction but forgetting to include the **constituent terms**

# Chapter 9: Assessing Studies

---

# Ch 9 Big Picture

---

Chapter 9 gives a **framework** for evaluating regression studies.

**Internal validity:** Are the statistical inferences correct *for the population studied*? (Getting the right answer)

**External validity:** Can the results be *generalized* to other populations and settings? (Applying it elsewhere)

Internal validity is what we can **test and address** with econometric tools. External validity requires **judgment** — no mechanical test.

# The Five Threats to Internal Validity

---

When we estimate  $Y_i = \beta_0 + \beta_1 X_i + u_i$ , internal validity requires:

$E(u_i|X_i) = 0$  — no bias

Correct standard errors — valid inference

**Five threats that can violate these:**

**Omitted variable bias**

**Wrong functional form**

**Measurement error** (errors-in-variables)

**Sample selection bias**

**Simultaneous causality**

# Threat 1: Omitted Variable Bias

---

An omitted variable  $Z$  causes bias if:

$Z$  is a determinant of  $Y$  (i.e.,  $Z$  is in the error term), **and**

$Z$  is correlated with  $X$

**ECON 3500 solutions:** Include the omitted variable, use proxy variables, use panel data — difference-in-differences and fixed effects methods (Ch 10) — or use instrumental variables (Ch 12).

**Watch Out:** Adding “kitchen sink” controls can introduce its own problems — only control for variables that satisfy the OVB conditions. Bad controls can **create** bias rather than remove it.

# Threat 3: Measurement Error

---

The variable you observe ( $X_{\tilde{}}$ ) differs from the true variable ( $X$ ):

$$X_{\tilde{i}} = X_i + w_i$$

where  $w_i$  is the measurement error.

Type	Where	Bias Effect
Classical ME in $X$	Independent variable	Attenuation bias (toward zero)
ME in $Y$	Dependent variable	No bias in $\hat{\beta}_1$ ; larger SEs
Non-classical ME in $X$	Correlated with $X$ or $u$	Bias direction uncertain

# Attenuation Bias (Classical ME in $X$ )

---

$$\hat{\beta}_1 \xrightarrow{p} \beta_1 \times \frac{\sigma_X^2}{\sigma_X^2 + \sigma_w^2}$$

The ratio  $\frac{\sigma_X^2}{\sigma_X^2 + \sigma_w^2} < 1$ , so  $|\hat{\beta}_1|$  is too **small**.

OLS is biased **toward zero** — you underestimate the true effect.

**Solution:** Get better data, use IV (Ch 12).


# Other Threats at a Glance

---

Threat	Problem	Example	Solution
<b>Wrong functional form</b>	True relationship is nonlinear	Income $\rightarrow$ test scores is quadratic, you fit a line	Ch 8 tools (polynomials, logs, interactions)
<b>Sample selection</b>	Sample not representative due to selection on $Y$	Studying wages for employed only (survivorship bias)	Random sampling, Heckman correction

---

# DAGs: Three Types of Paths



Path Type	Structure	Default	Example
<b>Causal</b> (front-door)	$X \rightarrow M \rightarrow Y$	Open	Education $\rightarrow$ Job $\rightarrow$ Earnings
<b>Backdoor</b>	$X \leftarrow Z \rightarrow Y$	Open	Class Size $\leftarrow$ Wealth $\rightarrow$ Scores
<b>Collider</b>	$X \rightarrow Z \leftarrow Y$	Closed	Talent $\rightarrow$ Hiring $\leftarrow$ Connections

**Causal paths** = what we want to estimate. Leave them open.

**Backdoor paths** = confounding. Must close them.

**Collider paths** = safely closed. Don't open them!

# What “Controlling For” Does

**Key identity:** Association =  $\underbrace{\text{Causal Effect}}_{\text{front-door}} + \underbrace{\text{Bias}}_{\text{open backdoor}}$

Confounding = OVB = an **open backdoor path**. Controlling for a variable **blocks all paths** through it — but this can be good or bad:

Path Type	Default	Controlling...	Result
Backdoor ( $X \leftarrow Z \rightarrow Y$ )	Open	<b>Closes</b> it	Removes bias ✓
Collider ( $X \rightarrow Z \leftarrow Y$ )	Closed	<b>Opens</b> it	Creates bias ✗
Mediator ( $X \rightarrow Z \rightarrow Y$ )	Open	<b>Closes</b> it	Blocks causal effect ✗

## Bad Controls:

**Don't control for mediators** — you'll block the effect you're trying to measure

**Don't control for colliders** — you'll introduce bias where there was none

**Do control for confounders** — they create open backdoor paths (= OVB)

# DAG Practice

---

**You want to estimate the causal effect of Fertilizer on Crop Yield.**

You believe that Soil Quality affects both Fertilizer use (farmers apply more fertilizer to poor soil) and Crop Yield (better soil → higher yields). Rainfall also independently affects Crop Yield but does *not* affect how much fertilizer farmers apply.

The DAG:

Fertilizer → Crop Yield

Soil Quality → Fertilizer

Soil Quality → Crop Yield

Rainfall → Crop Yield

## **Knowledge Check:**

List all paths from Fertilizer to Crop Yield

Classify each as causal or backdoor

What should you control for?

## **Paths:**

Fertilizer  $\rightarrow$  Crop Yield (causal — leave open)

Fertilizer  $\leftarrow$  Soil Quality  $\rightarrow$  Crop Yield (backdoor — **control for Soil Quality**)

Rainfall has no path connecting Fertilizer and Crop Yield, so it does not need to be controlled for (though including it may reduce noise).

After controlling for Soil Quality, the only open path is causal. Effect is identified.

# Collider Bias: Practice

---

You want to know whether **Academic Ability** affects **Athletic Ability**.

Both cause College Admission:

Academic Ability → College Admission

Athletic Ability → College Admission

There is **no direct causal link** between Academic Ability and Athletic Ability.

**Knowledge Check:** Should you study this relationship **among admitted students only**?

**No!** Conditioning on College Admission (the collider) **opens** a spurious path between Academic Ability and Athletic Ability.

Among admitted students, low Academic Ability becomes correlated with high Athletic Ability — because students who got in despite weak academics likely got in on athletic merit, and vice versa.

This is **collider bias** (a.k.a. selection bias). The two abilities are independent in the full population, but conditioning on the collider creates a spurious negative association.

# Ch 9: What to Focus On

---

Be able to **name and define** each of the five threats

Given a research scenario, **identify which threat(s)** apply

Know the **direction of bias** when possible (especially OVB and attenuation bias)

Know at least one **solution** for each threat

**Draw and interpret** simple DAGs — identify causal vs. backdoor vs. collider paths

Know what to **control for** and what **not to control for** (mediators, colliders)

Explain why confounding = OVB = open backdoor path

# Chapter 10: Panel Data & DiD

---

# Ch 10 Big Picture

---

**Panel data** = same units observed over multiple time periods.

Two powerful tools:

**Fixed effects** — control for all time-invariant unobserved characteristics

**Difference-in-differences** — estimate causal effects by comparing changes over time between treatment and control groups

# Why Panel Data?

---

**Panel data** = the **same units** observed over multiple time periods (not pooled cross-sections, which draw different samples each period).

The key advantage: we can control for **unobserved, time-invariant** characteristics.

$$Y_{it} = \beta_0 + \beta_1 X_{it} + \alpha_i + u_{it}$$

$\alpha_i$  = **entity fixed effect** — captures everything about unit  $i$  that doesn't change over time

Examples: ability, geography, institutional culture, baseline preferences

**Key Insight:** Fixed effects eliminate OVB from **any** time-invariant omitted variable — even ones you can't observe or measure.

# Entity Fixed Effects: Mechanics

---

## Approach 1: First-difference estimator

Take the difference between two periods:

$$Y_{it} - Y_{i,t-1} = \beta_1(X_{it} - X_{i,t-1}) + (u_{it} - u_{i,t-1})$$

The fixed effect  $\alpha_i$  **drops out** because  $\alpha_i - \alpha_i = 0$ .

## Approach 2: Entity dummies

Include a dummy variable for each entity:

$$Y_{it} = \beta_1 X_{it} + \gamma_1 D_{1i} + \gamma_2 D_{2i} + \dots + u_{it}$$

## Approach 3: Demeaning

Subtract the entity mean:  $\tilde{Y}_{it} = Y_{it} - \bar{Y}_i$ , then regress  $\tilde{Y}$  on  $\tilde{X}$ .

All three approaches give the **same**  $\hat{\beta}_1$ .

# Time Fixed Effects

---

$$Y_{it} = \beta_1 X_{it} + \alpha_i + \lambda_t + u_{it}$$

$\lambda_t$  = **time fixed effect** — captures anything affecting all units in period  $t$

Examples: macroeconomic shocks, federal policy changes, technological trends

**Entity FE + Time FE together** control for:

All time-invariant unit characteristics ( $\alpha_i$ )

All unit-invariant time shocks ( $\lambda_t$ )

**Limitation:** Fixed effects **cannot** control for omitted variables that change over time *and* vary across units. For that, you may need other methods such as IV (Ch 12) or DiD.

# Difference-in-Differences: Setup

DiD estimates a causal effect by comparing **changes over time** between a **treatment group** and a **control group**.

	Before	After	Change
<b>Treatment</b>	$\bar{Y}_{T,before}$	$\bar{Y}_{T,after}$	$\Delta\bar{Y}_T$
<b>Control</b>	$\bar{Y}_{C,before}$	$\bar{Y}_{C,after}$	$\Delta\bar{Y}_C$

DiD estimate:

$$\begin{aligned}\hat{\delta}_{DiD} &= (\bar{Y}_{T,after} - \bar{Y}_{T,before}) - (\bar{Y}_{C,after} - \bar{Y}_{C,before}) \\ &= \Delta\bar{Y}_T - \Delta\bar{Y}_C\end{aligned}$$

# DiD as a Regression

---

$$Y_{it} = \beta_0 + \beta_1 \text{Treat}_i + \beta_2 \text{After}_t + \beta_3 (\text{Treat}_i \times \text{After}_t) + u_{it}$$

Coefficient	Meaning
$\beta_0$	Control group mean, before
$\beta_1$	Difference between treatment and control, before
$\beta_2$	Change over time for control group
$\beta_3$	<b>DiD estimate</b> = causal effect of treatment

**The DiD Coefficient:**  $\beta_3$  is the coefficient on the **interaction term**  $\text{Treat} \times \text{After}$ . This is the causal effect of interest.

# DiD: Calculating Group Means

---

You should be able to recover each cell of the 2x2 table from the regression coefficients:

	Before ( )	After ( )
Control (Treat = 0)	$\beta_0$	$\beta_0 + \beta_2$
Treatment (Treat = 1)	$\beta_0 + \beta_1$	$\beta_0 + \beta_1 + \beta_2 + \beta_3$

**Practice:** If  $\hat{\beta}_0 = 53$ ,  $\hat{\beta}_1 = -3$ ,  $\hat{\beta}_2 = -1$ ,  $\hat{\beta}_3 = -2$ :

Treatment group, after period =  $53 + (-3) + (-1) + (-2) = 47$  ✓

# The Parallel Trends Assumption

## Parallel Trends

In the **absence of treatment**, the treatment and control groups would have followed the **same trend** over time.

This is the **key identifying assumption** of DiD.

It is **not directly testable** — we never observe what would have happened to the treatment group without treatment.

We can check **pre-treatment trends** — if they are parallel before treatment, it is more plausible they would have remained parallel.

**If Parallel Trends Fails:** The DiD estimate is **biased**. Any pre-existing divergence between groups gets picked up as a “treatment effect.”

# DiD Calculation: Practice

---

A state raises its minimum wage in 2020. You have average teen employment rates:

	2019	2021
Treatment state	50%	47%
Control state	53%	52%

Calculate the DiD estimate of the minimum wage effect on teen employment.

**Step 1:** Change in treatment:  $47 - 50 = -3$  pp

**Step 2:** Change in control:  $52 - 53 = -1$  pp

**Step 3:** DiD =  $(-3) - (-1) = -2$  pp

The minimum wage increase is associated with a **2 percentage point decrease** in teen employment.

# Ch 10: What to Focus On

---

**Distinguish** cross-sectional, time-series, and panel data

Explain what **entity fixed effects** control for and what they **cannot** control for

Know how the **first-difference estimator** eliminates fixed effects (the algebra)

**Calculate a DiD estimate** from a 2x2 table

State the **parallel trends assumption** and explain why it matters

Know which coefficient in the regression is the DiD estimate ( $\beta_3$ , the interaction)

**Standard Errors in Panel Data:** With panel data, observations within the same entity are likely correlated over time. Use **clustered standard errors** (cluster at the entity level) — otherwise SEs are too small and you over-reject.

# Chapter 12: Instrumental Variables

# Ch 12 Big Picture

---

When  $E(u_i | X_i) \neq 0$  (endogeneity), OLS is **biased and inconsistent**.

**Instrumental variables (IV)** solve this by finding a variable  $Z$  that:

Affects  $X$  (so we can use it to predict  $X$ )

But has no direct relationship with  $Y$  except through  $X$

Think of IV as **isolating** the “good” variation in  $X$  — the part unrelated to the error term.

# Why Do We Need IV?

---

Three threats produce endogeneity ( $E(u_i | X_i) \neq 0$ ):

Threat	Example
Omitted variable bias	Ability biases returns to schooling
Simultaneous causality	Police $\leftrightarrow$ crime
Measurement error	Self-reported income

When we cannot fix these with controls, panel data, or experiments, we turn to **instrumental variables**.

# The Three Conditions for a Valid Instrument

---

An instrument  $Z$  must satisfy:

## **i** 1. Relevance

$Z$  is correlated with  $X$ :  $\text{corr}(Z_i, X_i) \neq 0$

*$Z$  must actually affect  $X$ . Testable with the first-stage regression.*

# Instrument Validity (cont.)

---

## **i** 2. Exogeneity (Independence)

$Z$  is uncorrelated with the error term:  $\text{corr}(Z_i, u_i) = 0$

$Z$  must be “as good as randomly assigned.” **Not directly testable** — requires reasoning.

## **i** 3. Exclusion Restriction

$Z$  affects  $Y$  **only through**  $X$ , not directly.

There is no direct causal path from  $Z$  to  $Y$  that bypasses  $X$ . **Not directly testable.**

# Classic Example: Returns to Education

---

**Problem:** Estimate the causal effect of schooling on wages.

$$\ln(\text{wage}_i) = \beta_0 + \beta_1 \text{schooling}_i + u_i$$

OVB: Ability is in  $u_i$  and correlated with schooling  $\rightarrow \hat{\beta}_1$  is biased upward.

**Proposed instrument:** Quarter of birth ( $Z$ )

**Relevance:** Compulsory schooling laws mean people born earlier in the year can drop out sooner  $\rightarrow$  less schooling. ✓

**Exogeneity:** Quarter of birth is essentially random. ✓ (plausibly)

**Exclusion:** Quarter of birth doesn't directly affect wages — only through schooling. ✓ (debatable)

**Caution:** Exclusion restriction is always **arguable**, not provable. This is where most debates about IV studies focus.

# Two-Stage Least Squares (2SLS)

## i Stage 1: First-Stage Regression

Regress  $X$  on  $Z$  (and any controls  $W$ ):

$$X_i = \pi_0 + \pi_1 Z_i + \pi_2 W_i + u_i$$

Save the predicted values  $\hat{X}_i$ .

## i Stage 2: Second-Stage Regression

Regress  $Y$  on  $\hat{X}$  (and any controls  $W$ ):

$$Y_i = \beta_0 + \beta_1 \hat{X}_i + \beta_2 W_i + u_i$$

**Why this works:**  $\hat{X}_i$  contains only the variation in  $X$  that comes from  $Z$ . Since  $Z$  is exogenous,  $\hat{X}_i$  is uncorrelated with  $u_i$ .

**Important:** Use a proper IV/2SLS command (e.g., `ivregress 2sls` in Stata). Running two separate regressions gives **wrong standard errors**.

# The Wald Estimator (Binary Instrument)

---

When the instrument  $Z$  is **binary** (0 or 1), the IV estimate simplifies:

$$\beta_1^{\text{Wald}} = \frac{\text{Reduced form}}{\text{First stage}}$$

**Reduced form:** regress  $Y$  on  $Z$  — total effect of  $Z$  on  $Y$

**First stage:** regress  $X$  on  $Z$  — effect of  $Z$  on  $X$

**Wald estimate:** the ratio — causal effect of  $X$  on  $Y$ , scaled by how much  $Z$  moves  $X$

**Intuition:** The reduced form tells you how much  $Z$  moves  $Y$ . The first stage tells you how much  $Z$  moves  $X$ . Dividing gives you how much  $X$  moves  $Y$ .

# Weak Instruments

---

If  $Z$  is only **weakly correlated** with  $X$ , the IV estimator has serious problems:

Large standard errors

Biased (toward OLS) in finite samples

Unreliable hypothesis tests

**The  $F > 10$  Rule of Thumb:** Test instrument relevance with the **first-stage F-statistic**.

$F > 10$ : Instrument is sufficiently strong ✓

$F < 10$ : Weak instrument — IV estimates are unreliable ✗

This is a **necessary condition** you should always check.

**How to get the F-statistic:** Run the first-stage regression and look at the F-statistic on the excluded instrument(s).

# LATE: What Does IV Actually Estimate?

---

## Local Average Treatment Effect (LATE)

IV estimates the causal effect for **compliers** — the subpopulation whose treatment status is changed by the instrument.

IV estimates the effect **only for compliers** — not for the whole population.

This is a **local** estimate, not a global one.

The IV estimate may differ from the ATE if compliers are not representative — think carefully about **who the estimate applies to**.

# Evaluating Instruments: Practice

---

A researcher wants to estimate the effect of **years of education** on earnings. She proposes using **distance from the nearest college** (where the person grew up) as an instrument.

Evaluate each condition:

**1. Relevance:** Living closer to a college lowers the cost of attending → more schooling. ✓ (testable)

**2. Exogeneity:** Is distance from college correlated with unobserved determinants of earnings? Growing up near a college often means growing up in an urban area — urban/rural differences in family income, school quality, and job markets could violate this. ⚠️

**3. Exclusion:** Does distance affect earnings **only through** education? Possibly not — proximity to a college town may independently affect local labor markets and earnings. ⚠️

**Verdict:** Relevance is plausible, but exogeneity and exclusion are **debatable** — the researcher would need to add controls for urban/rural, family background, and local labor market conditions.

# Ch 12 Summary

---

Concept	Key Point
Endogeneity	$E(u_i   X_i) \neq 0 \rightarrow$ OLS is biased
Valid instrument	Relevant + exogenous + exclusion restriction
2SLS	First stage predicts $X$ ; second stage uses $\hat{X}$
Weak instruments	$F < 10$ in first stage $\rightarrow$ unreliable
LATE	IV estimates effect for compliers only

## Common Mistakes:

Claiming an instrument is valid without arguing exclusion restriction

Running 2SLS as two separate regressions (wrong SEs)

Ignoring the weak instrument problem

Interpreting IV as the ATE instead of the LATE

# Putting It All Together

---

# Connections Across Chapters

---

These chapters are deeply linked:

**Ch 8 → Ch 9:** Wrong functional form (Ch 9) is solved by nonlinear models (Ch 8)

**Ch 9 → Ch 10:** Omitted variable bias (Ch 9) can be addressed by fixed effects (Ch 10) when the omitted variable is time-invariant

**Ch 9 → Ch 12:** OVB, simultaneous causality, and measurement error (Ch 9) are all forms of endogeneity addressed by IV (Ch 12)

**Ch 10 + Ch 12:** IV can be combined with panel data/fixed effects when endogeneity remains after controlling for fixed effects

# Decision Tree: Which Tool?

---

You suspect your OLS estimate is biased. What do you do?

If the problem is...	Try...
Wrong functional form	Polynomials, logs, interactions (Ch 8)
Time-invariant omitted variable + panel data available	Entity fixed effects (Ch 10)
Common time shock	Time fixed effects (Ch 10)
Clear before/after treatment with a control group	DiD (Ch 10)
Endogeneity remains (OVB, simultaneity, meas. error)	Instrumental variables (Ch 12)
Non-random sample	Address selection, or be cautious

# Formula Sheet Suggestions

---

Your double-sided formula sheet should include:

**Log interpretation table** (Ch 8) — you will almost certainly need this

Marginal effect formulas for polynomials and interactions (Ch 8)

Five threats to internal validity (Ch 9)

DiD formula and regression setup (Ch 10)

Three conditions for a valid instrument (Ch 12)

2SLS steps (Ch 12)

$F > 10$  rule (Ch 12)