LECTURE 3.2

Multiple Regression: Inference

ECON30130 Econometrics I

Sam Deegan
University College Dublin

Autumn 2027

Where We Are

1. Foundations	2. Simple Regression	3. Multiple Regression	4. Broken Assumptions
1.1 Data and Variation	2.1 Least Squares	3.1 Estimation	4.1 Heteroskedasticity
1.2 Conditional Means	2.2 Units and Fit	**3.2 Inference**	4.2 Specification
1.3 Sampling Variation	2.3 Bias and Precision	3.3 Functional Form	
		3.4 Dummies	

> **Block 3 Builds Multiple Regression in Four Lectures**
> Lecture 3.1 estimated regressions with more than one regressor. Today tests a coefficient, builds its interval, then tests a block of coefficients together.

Last Week's Lecture

Lecture 3.1 put more than one regressor into the regression:

1. **Partialling Out.**
   A coefficient is the slope on the part of its variable the others leave over.

2. **Reading a Partial Coefficient.**
   What controlling for a confounder, a mediator or a collider does.

3. **Omitted Variable Bias.**
   Leaving a variable out shifts the slope by a product of two slopes.

4. **Multicollinearity.**
   Correlated regressors widen standard errors and leave the estimates unbiased.

What One Estimate and Its Standard Error Leave Open

Our 120 people gave a fitted slope of 1.5227 and a standard error of 0.1897. Lecture 1.3 turned a mean and its standard error into a t-statistic, a p-value and a 95 per cent interval.

Question
Our sample gave 1.5227 euro an hour per extra year of schooling. What else would you want printed beside that number before you would be willing to put it in a report?

Lecture Plan

1 **Testing a Coefficient.**
Comparing an estimate with its own standard error, and what that settles.

2 **Confidence Intervals and the Regression Table.**
A range built from the same two numbers, and what a published table prints beside it.

3 **Why the t Survives Non-Normal Errors.**
Our earnings are skewed, and the slope is an average of 120 small errors, so its spread across samples is close to normal.

4 **Testing Restrictions.**
One F-statistic compares the fit with and without a block of coefficients.

SECTION 1

Testing a Coefficient

1. Testing a Coefficient

2. Confidence Intervals and the Regression Table

3. Why the *t* Survives Non-Normal Errors

4. Testing Restrictions

Three Objects Built from an Estimate and Its Standard Error

An estimate and its standard error are enough to build three objects:

- **A t-Statistic:** a t-statistic counts the distance from a claimed value in standard errors.
- **A p-Value:** a p-value describes what the claim would produce, not whether it is true.
- **A Confidence Interval:** an interval reports the estimate and its precision as one object.

> **Built for a Mean in Lecture 1.3**
> Lecture 1.3 built all three for a mean. None of them refers to where the estimate came from.

A Slope Goes in Wherever a Mean Came Out

The slope and a claimed slope ($\beta_{1,0}$) replace the mean and its claimed value (μ_0):

$$\frac{\bar{y} - \mu_0}{\text{se}(\bar{y})} \longrightarrow \frac{\hat{\beta}_1 - \beta_{1,0}}{\text{se}(\hat{\beta}_1)} \tag{1}$$

The interval makes the same swap, with a critical value (c) where the mean used 1.96:

$$\bar{y} \pm 1.96\,\text{se}(\bar{y}) \longrightarrow \hat{\beta}_1 \pm c\,\text{se}(\hat{\beta}_1) \tag{2}$$

A Line Leaves 118 Degrees of Freedom

Dividing the sum of squared residuals by 118 rather than 120 gives Lecture 2.3's $\hat{\sigma}$:

$$\hat{\sigma}^2 = \frac{\text{SSR}}{\underbrace{n-2}_{\text{the degrees of freedom}}} \qquad \Longrightarrow \qquad \hat{\sigma} = \sqrt{\hat{\sigma}^2} \tag{3}$$

- SSR: the sum of the 120 squared residuals, 2692.8 in our sample
- $\hat{\sigma}$: the estimated standard deviation of the population error

> **Two Coefficients Fix Two Residuals**
> The intercept and slope settle two residuals once the other 118 are known; the t counts the same 118.

Worked example: our 120 people

The Standardised Slope Follows a t Distribution

The five assumptions made the slope unbiased and gave its variance, but not its shape. A test needs the shape:

$$\frac{\hat{\beta}_1 - \beta_1}{\text{se}(\hat{\beta}_1)} \sim t_{n-2} \tag{4}$$

Greek for the Population, a Hat for Ours
No sample shows β_1 or σ; our 120 give $\hat{\beta}_1$ and $\hat{\sigma}$. Lecture 2.3 put $\hat{\sigma}$ inside $\text{se}(\hat{\beta}_1)$, and that estimate makes the shape a t.

Worked example: our 120 people

The Exact t Needs Normally Distributed Errors

The t result needs one assumption beyond the five: the population error is normally distributed.

> **Why Normal Errors Give a Normal Slope**
> A weighted sum of normal errors is normal, so $\hat{\beta}_1$ is normal too. Wooldridge states chapter 4 for k regressors, and our k is one (Wooldridge, 2020).

Past empirical evidence suggests normality is a poor assumption for wages (Wooldridge, 2020).

The t-Statistic Compares an Estimate with Its Standard Error

No sample shows β_1, so a test puts a claimed value in its place. An estimate is large or small only against its own standard error (Wooldridge, 2020):

$$t = \frac{\hat{\beta}_1 - \beta_{1,0}}{\text{se}(\hat{\beta}_1)} \qquad \beta_{1,0} = 0 \implies t = \frac{\hat{\beta}_1}{\text{se}(\hat{\beta}_1)} \tag{5}$$

- t: how many standard errors the estimate sits from the claimed value

Our t of 8.03 clears the critical 1.980, so a two-sided five per cent test rejects zero.

> **The Units Cancel in the Ratio**
> Our 1.5227 and its standard error are both euro an hour per year of schooling; 8.03 has no units.

Worked example: our 120 people

A Null Is a Claim About the True Slope

A null hypothesis always states a value for β_1, the slope no sample shows (Wooldridge, 2020).

> **Zero Is a Choice, Not a Default**
> Software tests $\beta_1 = 0$ because zero is often the interesting claim. If theory says the return should be one euro, subtract one. Wooldridge tests an elasticity against one for that reason (Wooldridge, 2020).

We computed 1.5227 ourselves, so nothing about it is left to test.

What Rejecting a Null of Zero Does Not Establish

Rejecting zero says only that a zero return would rarely produce a sample like ours. Rejection does not establish four claims:

- **Not the True Value:** 1.5227 is our estimate, and Lecture 2.3 showed the truth is 1.30.
- **Not a Large Return:** a large *t* can come from a small standard error alone.
- **Not Practical Importance:** whether 1.5227 euro an hour per year of schooling matters is a judgement.
- **Not Causation:** schooling causing earnings is Lecture 4.2's question, and no test here addresses it.

> **A Small p-Value Is Still Only About Zero**
> Our *t* of 8.03 carries a two-sided p-value far below five per cent. The p-value describes a null of zero and nothing else (Wooldridge, 2020).

Why Significance Is About Precision, Not Effect Size

Significance says how precisely a coefficient was measured (Wooldridge, 2020). A star cannot say whether the estimate or its standard error made t large:

- **A Very Large Sample:** even a negligible return is significant, because the standard error is tiny.
- **A Small Sample:** a return that matters is not significant, because its standard error is large.

In Wooldridge's 401(k) example of 1,534 firms, firm size is significant and practically trivial (Wooldridge, 2020).

> **Practical Significance Is a Separate Judgement**
> This arithmetic cannot say whether our 1.5227 matters. That judgement needs facts about Irish wages, not the statistic. Wooldridge says check the statistic first, then discuss the magnitude in units (Wooldridge, 2020).

A Null of 1.30 Rather Than Zero

This module knows the true return, 1.30. Against it, the numerator is our 0.2227 miss, a little over one standard error.

> **Question**
> Would a null of 1.30 be rejected at five per cent? What does your answer tell you about the interval we build next?

A t of 1.17 Does Not Reject 1.30

Against the true slope, the numerator is the miss itself, and our t of 1.17 sits well inside 1.980:

$$\beta_{1,0} = \beta_1 \implies t = \frac{\hat{\beta}_1 - \beta_1}{\text{se}(\hat{\beta}_1)} \tag{6}$$

The sample rejects zero and does not reject 1.30.

> **An Interval Holds Every Value the Test Keeps**
> A test settles one claimed value at a time. The 95 per cent interval holds every value the test keeps, 1.30 included.

Worked example: our 120 people

SECTION 2

Confidence Intervals and the Regression Table

1. Testing a Coefficient

2. Confidence Intervals and the Regression Table

3. Why the *t* Survives Non-Normal Errors

4. Testing Restrictions

Deriving the Interval from the t Distribution

In 95 samples of 100 the standardised slope falls between $-c$ and c:

$$0.95 = \Pr\left(-c < \frac{\hat{\beta}_1 - \beta_1}{\text{se}(\hat{\beta}_1)} < c\right) \tag{7}$$

Multiplying through by the positive standard error keeps both directions:

$$0.95 = \Pr\big(-c\,\text{se}(\hat{\beta}_1) < \hat{\beta}_1 - \beta_1 < c\,\text{se}(\hat{\beta}_1)\big) \tag{8}$$

Solving the Probability Statement for the True Slope

Subtracting $\hat{\beta}_1$ and multiplying by -1 reverses both directions and isolates β_1:

$$0.95 = \Pr\big(\hat{\beta}_1 - c\,\mathrm{se}(\hat{\beta}_1) < \beta_1 < \hat{\beta}_1 + c\,\mathrm{se}(\hat{\beta}_1)\big) \tag{9}$$

On our 120 people the two endpoints are 1.147 and 1.898.

> **Only the Endpoints Differ from Sample to Sample**
> β_1 is one fixed number in every line. $\hat{\beta}_1$ differs from sample to sample, and so do both endpoints. Wooldridge calls the range an interval estimate (Wooldridge, 2020).

Worked example: our 120 people

The Interval Covers the True Slope

Only a population we wrote lets us check an interval against the truth:

- **The Point Estimate:** our 1.5227 misses the true 1.30 by 0.2227.
- **The Interval:** 1.147 to 1.898 contains the true 1.30, well away from either end.

In real work the miss is unknowable and the coverage unverifiable.

> **A Miss of 0.2227 Still Sits Inside the Interval**
> An unbiased rule that misses the truth by 0.2227 has not failed. The interval printed beside 1.5227 reports how far the estimate might have missed. Lecture 2.3's standard error of 0.1897 sets how far the interval reaches each side.

Reading the Interval Out Loud

The wording is examinable. The wrong version gives the truth a 95 per cent chance of sitting in our interval:

"*There is a 95 per cent chance the true return lies between 1.147 and 1.898.*"

That sentence makes β_1 vary. β_1 is one fixed number in a population of 200,000.

> **The Sentence to Use Carries Units**
> Our estimate of the return to a year of schooling is **1.5227 euro an hour**, with a 95 per cent confidence interval from **1.147 to 1.898**. The line compares two people whose schooling differs by one year.

The Ninety-Five per Cent Counts Intervals, Not Truths

A slope leaves unchanged what the 95 per cent counts (Wooldridge, 2020):

- **Before the Sample Is Drawn:** the rule covers β_1 in 95 samples out of 100.
- **After the Endpoints Are Numbers:** 1.30 either sits inside 1.147 to 1.898 or it does not.
- **After, in Real Work:** we can only hope our sample is one of the 95 (Wooldridge, 2020).

> **Coverage Rests on Six Assumptions**
> The 95 per cent needs assumptions one to five and normal errors. If the 120 chose themselves, the rule covers less often. The printed interval will not say so.

The Interval and the Test Are One Statement

The test rejects a value exactly when the interval excludes it (Wooldridge, 2020):

- **A Null of Zero:** zero lies outside 1.147 to 1.898, and $t = 8.03$ said so.
- **A Null of 1.30:** 1.30 lies inside, so the sample does not reject it.

> **An Interval Is Every Test You Could Have Run**
> Every value inside the interval is a null this sample does not reject. Every value outside it is a null the sample rejects. Each endpoint is the value whose t equals the critical 1.980.

The Interval Against the Sampling Distribution

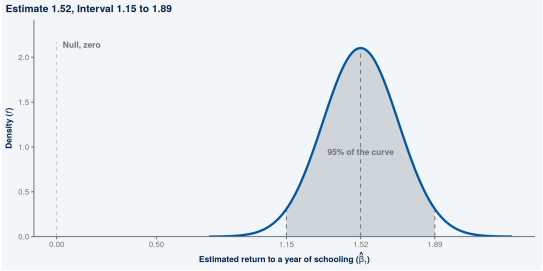

Figure 1: The distribution implied by 1.5227 and 0.1897, with the 95 per cent region shaded.

> **The Shaded Region Excludes Zero and Contains 1.30**
> An interval is built around our estimate, not around the truth. Our interval still covers 1.30.

Comparing a Wide Interval with a Narrow One

Two research teams report the return to a year of schooling:

- **The First Team:** estimates 1.5, with an interval from -0.1 to 3.1.
- **The Second Team:** estimates 0.4, with an interval from 0.3 to 0.5.

> **Question**
> Which study would you rather have been given? Does your answer match which of the two is statistically significant?

The Narrow Interval Is Preferred for Its Width

The second study is the one to want, and its significance is not the reason:

- **The First Team:** −0.1 to 3.1 contains zero, so 1.5 is not significant, and three euro is not ruled out.
- **The Second Team:** 0.3 to 0.5 excludes zero, so 0.4 is significant, and small, and precisely measured.

> **A Wide Interval Is Still a Finding**
> The first team should report −0.1 to 3.1 itself. The range says which returns its data cannot yet rule out.

The Same Estimate at Four Different Sample Sizes

The table is a thought experiment, not four real samples. Our 1.5227 stays frozen while the number of people changes:

People	Slope	Standard error	t	95 per cent interval
30	1.5227	0.3843	3.96	[0.735, 2.310]
120	1.5227	0.1897	8.03	[1.147, 1.898]
480	1.5227	0.0946	16.10	[1.337, 1.709]
1920	1.5227	0.0472	32.23	[1.430, 1.615]

Four times the people halves the standard error and doubles the t.

> **The Last Two Intervals Do Not Cover 1.30**
> Our 1.5227 had already missed by 0.2227. Narrowing the interval around that frozen estimate excludes 1.30 rather than locating it. A real sample of 480 would return its own estimate, usually nearer 1.30.

One Frozen Return Reported at Four Sample Sizes

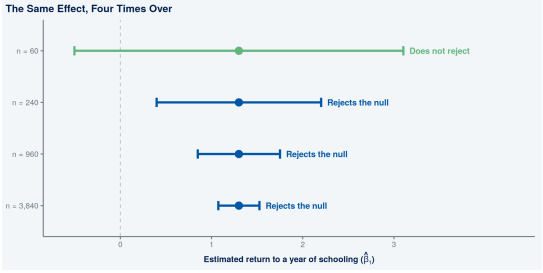

Figure 2: The same estimated return to schooling, reported at four different sample sizes.

> **The Two Narrowest Exclude 1.30**
> The four bars share a centre and differ only in width. The estimate was held fixed, and precision is not accuracy.

What to Read First in a Regression Table

Our regression is laid out below as a published table would set it:

	Hourly earnings
Years of schooling, centred at 13	1.5227
Standard error	(0.1897)
95 per cent interval	[1.147, 1.898]
t against a null of zero	8.03
People	120

Most tables print only the first two lines.

Coefficient First, Stars Last

A reader takes the coefficient in units first, then the standard error. The sample size comes next, with what it is a sample of. The other regressors follow, because Lecture 3.1 showed they can change the coefficient. Stars come last.

What a Table of Stars Leaves Out

Our 8.03 and a *t* just past 1.980 carry the same star. A table of stars costs a reader three things:

- **The Interval:** nobody can rebuild it, so no value can be checked.
- **Precision and Size:** a precise small estimate reads exactly like an imprecise large one.
- **Combining Studies:** pooling results needs standard errors, and this table has none.

> **Report the Interval in Your Own Tables**
> The interval carries the estimate and its precision together. The test sits inside it. A reader recovers the standard error from the two endpoints and 1.980.

SECTION 3

Why the *t* Survives Non-Normal Errors

1. Testing a Coefficient

2. Confidence Intervals and the Regression Table

3. Why the *t* Survives Non-Normal Errors

4. Testing Restrictions

Skewed Earnings Mean the Exact t Does Not Apply

Sections 1 and 2 assumed normal errors, which this module has already denied:

- **Section 1's Assumption:** the exact t needed the population error to be normally distributed.
- **Lecture 1.3's Finding:** earnings here are skewed, so 1.3 called its 1.96 a large-sample value.
- **The Link:** an error is earnings minus the population line, so skewed earnings leave skewed errors.

> **The Sixth Assumption Fails Here**
> Wooldridge reports the same kind of evidence for real wages (Wooldridge, 2020). Section 1's t is therefore not exact.

Which Results Used Normality and Which Did Not

Each of Lecture 2.3's results can be checked against the assumptions it used:

Result	What it needed	Survives?
The slope is unbiased	assumptions one to four	✓
The variance is σ^2/SST_x	assumptions one to five	✓
Least squares is best linear unbiased	assumptions one to five	✓
The standardised slope is exactly t_{118}	those five and normal errors	

Only the Shape Is in Question

Our 1.5227 is still unbiased. The standard error 0.1897 still tells you how far our slope typically sits from the true one. What is not yet established is the shape we compare those two numbers against.

The Slope's Miss Is a Weighted Sum of 120 Small Errors

Our estimate's distance from the truth is one weighted sum of 120 errors:

$$\hat{\beta}_1 - \beta_1 = \sum_{i=1}^{n} w_i \, u_i \tag{10}$$

- u_i: person i's error
- w_i: the weight person i carries

The Errors Supply the Shape
The weights depend on schooling alone, so the errors are what varies here. The shape comes from adding 120 skewed draws together.

The Spread of the Weighted Sum Across Samples

Across samples, the slope's standard deviation is σ over the root of SST_x:

$$\text{sd}(\hat{\beta}_1) = \frac{\sigma}{\sqrt{\text{SST}_x}} \tag{11}$$

- SST_x: the sum of squared deviations of schooling, 633.9 in our sample

Our sample replaces σ with $\hat{\sigma}$ and gets its standard error:

$$\text{se}(\hat{\beta}_1) = \frac{\hat{\sigma}}{\sqrt{\text{SST}_x}} \tag{12}$$

Worked example: our 120 people

Three Steps Take the Sum to a Standard Normal

The central limit theorem applies to the standardised slope in three steps:

- **Adding 120 Pieces:** independent pieces, none dominant, give a nearly normal sum whatever one error's shape.
- **Standardising:** dividing that sum by its own standard deviation gives a standard normal.
- **Estimating σ:** Lecture 2.3 showed $\hat{\sigma}$ closes on σ, so the swap changes nothing in large samples.

> **The Theorem of Lecture 1.3 with Unequal Weights**
> Lecture 1.3 averaged 120 skewed earnings and got a nearly symmetric distribution. Our sum differs only in giving the 120 pieces unequal weights.

The Standardised Slope Is Approximately Standard Normal

Whatever the errors' shape, a large sample leaves the standardised slope nearly standard normal:

$$\frac{\hat{\beta}_1 - \beta_1}{\text{se}(\hat{\beta}_1)} \overset{a}{\sim} \text{Normal}(0, 1) \tag{13}$$

Approximately, Not Exactly

The *a* above the mark reads "is approximately distributed as". That statement holds in large samples and never exactly. Wooldridge's Theorem 5.2 needs only assumptions one to five. He states it for *k* regressors, and our *k* is one (Wooldridge, 2020).

The Exact t Against the Large-Sample Normal

The exact interval takes its critical value c from the t on $n - 2$ degrees of freedom:

$$\hat{\beta}_1 \pm c\,\mathrm{se}(\hat{\beta}_1) \tag{14}$$

The large-sample normal puts 1.96 in place of c, which is 1.980 on our 118:

$$\hat{\beta}_1 \pm 1.96\,\mathrm{se}(\hat{\beta}_1) \tag{15}$$

The Correction Is Not the Assumption

Each endpoint shifts by 0.0038, under half a cent an hour. Estimating σ accounts for that gap. Only a large sample brings the slope's spread across samples close to normal.

Worked example: our 120 people

Skewed Errors and Bunched Schooling Need More People

Three things set how close to normal the standardised slope is at a given n:

- **How Skewed the Errors Are:** a long right tail needs more people than a symmetric error.
- **How the Weights Spread:** bunched schooling puts most of the sum onto a handful of people.
- **How Many People There Are:** more people bring the shape closer to normal, never all the way.

> **Large Enough Depends on the Error Distribution**
> Wooldridge calls nonnormality no serious problem in large samples, and gives no threshold (Wooldridge, 2020). A threshold quoted for one error distribution says nothing about another.

Four Things the Large-Sample Result Does Not Fix

The theorem removes one assumption and leaves the other five standing. Four problems stay exactly as they were:

- **Not Bias:** omitted variable bias at 120 people is still there at 120,000.
- **Not Unequal Spread:** the standard error still assumes assumption five, and Lecture 4.1 breaks it.
- **Not the Shape of the Line:** more data still fits a straight line to a curve.
- **Not a Small Sample:** at thirty people a skewed error leaves the *t* approximate, whatever *c* is used.

> **Assumptions One to Five Still Hold Today**
> This population satisfies all five exactly, because we wrote it. Lecture 3.1 broke the fourth on purpose. The 95 per cent then no longer holds.

Where the Large-Sample Result Leaves the Interval

Our 1.147 to 1.898 came from the exact *t*, and now rests on a different list:

- **Still Needed:** assumptions one to five, and enough people to make the sum near normal.
- **No Longer Needed:** normally distributed errors, which this population does not have.
- **Not Settled:** whether 120 is enough, because that depends on how skewed our errors are.

At 120 people with skewed earnings we are relying on an approximation.

> **Report the Approximation Beside the 95 per Cent**
> The 95 per cent is approximate at this sample size. No line in a table tells a reader how close to 95 per cent the interval really is.

SECTION 4

Testing Restrictions

1. Testing a Coefficient

2. Confidence Intervals and the Regression Table

3. Why the *t* Survives Non-Normal Errors

4. Testing Restrictions

Why Separate t-Tests Cannot Answer a Joint Question

A block of coefficients raises two questions no single *t*-statistic can settle:

1. The four qualification dummies belong as a **block** or not at all. Four dummies give four answers and no verdict.
2. Two coefficients can each look unimpressive on their own. Dropping either one can make the other significant.

> **Four Tests Are Not a Test of the Four**
> Each of the four tests gets its own chance to reject falsely. Those chances accumulate, so the four are not a five per cent test.

Restricted and Unrestricted Fits of the Same Model

Testing a block means fitting the equation twice, once with the restriction imposed:

- **The Unrestricted Fit:** every variable stays in the equation.
- **The Restricted Fit:** the tested coefficients are forced to zero, and the rest are re-estimated.

> **A Joint Hypothesis Restricts Several Coefficients**
> Today's restriction sets the four qualification coefficients to zero. A restriction need not be zeros: it can set two coefficients equal.

The Same 120 People Fitted Twice

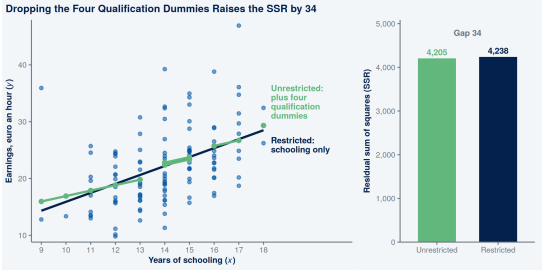

Figure 3: Our 120 with and without four qualification dummies: SSR 4,205 against 4,238.

> **The Gap Between the Bars Is the Fit Given Up**
> Forcing four coefficients to zero can only raise the residual sum of squares. Here it rises by only 34. The question is whether that is more than a different 120 people would give anyway.

The F-Statistic Compares Two Residual Sums of Squares

The F-statistic divides the fit given up per restriction by the estimated error variance:

$$F = \frac{(\text{SSR}_r - \text{SSR}_u)/q}{\underbrace{\text{SSR}_u/(n-k-1)}_{\text{the estimated error variance}}} \tag{16}$$

- SSR_r: residual sum of squares from the restricted fit
- SSR_u: residual sum of squares from the unrestricted fit
- q: how many coefficients the restriction sets to zero

For the qualification block, q is 4 and k is 5, so $n - k - 1$ is 114.

Joint Confidence Regions Versus Separate Intervals

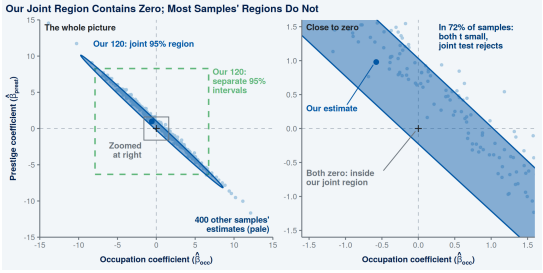

Figure 4: Occupation and prestige: our joint region over 400 repeated samples (pale ghost).

> **Our 120 Do Not Show the Pattern, and Most Samples Do**
>
> Both separate intervals contain zero, and here the joint region does too (p = 0.28). In about 72 per cent of samples, the joint region excludes zero while both intervals contain it.

Whether the Qualification Block Belongs in the Equation

Suppose none of the four qualification dummies has a t-statistic above two.

> **Question**
> Does the block belong in the equation? And write down one hypothesis about this equation that no single row of the table tests.

Four Small t-Statistics Cannot Settle the Block

Four *t*-statistics cannot say whether the four coefficients are zero together:

- **Each *t* Is Conditional:** each asks whether one coefficient is zero, given the other regressors.
- **Separate and Joint Can Disagree:** four small *t*s can be jointly significant (Wooldridge, 2020).
- **What No Row Tests:** two coefficients equal, a block jointly zero, two coefficients summing to one.

> **The Block Needs an F-Test**
> Four zeros are four restrictions, and only the F-statistic tests them together.

What Would Have to Be True?

Our report reads 1.5227 euro an hour, with a 95 per cent interval from 1.147 to 1.898.

Question
What would have to be true for this number to mean what we just said it means? And what is the plausible story where it isn't?

What the Interval from 1.147 to 1.898 Rests On

The 95 per cent rests on a large sample and five assumptions, including these three:

- **Random Sampling:** if earnings decided who answered, the interval is centred in the wrong place.
- **Constant Error Variance:** otherwise 0.1897 no longer tells you how far our slope typically sits from 1.30.
- **The Right Shape:** if the return rises with schooling, one slope averages different returns.

> **Only as Good as Its Assumptions**
> A confidence interval is only as good as the assumptions used to build it (Wooldridge, 2020). Nothing printed in the interval says which one failed.

The App Reports One Estimate Four Ways

Raising the sample from 120 to 480 changes only two of the app's four reported forms:

- **Bare and Starred:** 1.5227 at 480 people reads exactly as 1.5227 at 120 does.
- **Standard Error and Interval:** both halve, so only these two record the extra data.

> **A Star Cannot Record Four Times the Data**
> The bare number and the star are identical at both sample sizes. A reader needs the estimate, the interval and the sample size.

Summary: Testing, Intervals, Large Samples and Restrictions

1. **Testing a Coefficient.**
   Our t of 8.03 clears the critical 1.980 on 118 degrees of freedom.

2. **Confidence Intervals and the Regression Table.**
   The interval runs from 1.147 to 1.898, covers 1.30, and is the line most tables leave out.

3. **Why the t Survives Non-Normal Errors.**
   Our errors are skewed, so the 95 per cent is approximate at 120 people.

4. **Testing Restrictions.**
   A block is tested by the fit it gives up, measured by one F-statistic.

Before Lecture 3.3

1 Play:
The app. Hold the estimate at 1.5227, raise the sample size, and watch the interval narrow around it.

2 Apply:
The problem set rebuilds today's interval by hand from a standard error you compute yourself.

3 Review:
Wooldridge (2020), sections 4-1 to 4-3, 4-5 and 4-6 on today's tests, intervals and tables, and 5-2 on large samples, then 6-2 and 6-3 before Lecture 3.3.

Next Lecture

Lecture 3.3 changes the shape of the regression and asks how to choose among shapes:

1. **Logs and Elasticities.**
   A log reads the return as a proportion, reported in per cent.

2. **Choosing a Specification.**
   R^2 never falls as columns are added; adjusted R^2 can fall but tests nothing.

3. **Specification Search.**
   A p-value describes one test chosen in advance, not the best of twenty.

Thank you

sam.deegan@ucdconnect.ie
sam-deegan.com

BACKUP

Appendix

Notation Carried in from Earlier Lectures

Symbol	What it is	First met
i	which person, running over the 120 in our sample	Lecture 1.1
n	the number of people in the sample, 120 here	Lecture 1.1
$\sum$	a sum over the people in the sample	Lecture 1.1
$\bar{y}$	the average of our 120 hourly earnings, 12.25	Lecture 1.1
μ_0	the value a null hypothesis claims the population mean takes	Lecture 1.3
se$(\cdot)$	an estimate's estimated standard deviation across samples, 0.1897 for our slope	Lecture 1.3
β_1	the population slope, the true return to a year of schooling, 1.30	Lecture 2.1
$\hat{\beta}_1$	the least-squares slope from our 120 people, 1.5227	Lecture 2.1
u_i	person i's error	Lecture 2.1
SSR	the sum of the 120 squared residuals, 2692.8 in our sample	Lecture 2.1

Notation Carried in from Earlier Lectures

Symbol	What it is	First met
w_i	the weight person i carries in the slope	Lecture 2.3
SST_x	the sum of squared deviations of schooling, 633.9 in our sample	Lecture 2.3
σ	the error standard deviation, 4.50 in this population	Lecture 2.3
σ^2	the error variance, 20.25 in this population	Lecture 2.3
$\hat{\sigma}$	the estimated standard deviation of the population error, 4.777	Lecture 2.3
$\hat{\sigma}^2$	the estimated error variance, $\text{SSR}/(n-2) = 22.82$	Lecture 2.3
$\text{sd}(\hat{\beta}_1)$	the slope's standard deviation across samples, 0.1787 here	Lecture 2.3

Notation Introduced in the First Two Sections

Symbol	What it is	First met
$\beta_{1,0}$	a claimed value for the slope	A Slope Goes In Wherever a Mean Came Out
c	leaves 2.5 per cent in each tail, 1.980	A Slope Goes In Wherever a Mean Came Out
$n-2$	120 less two coefficients, 118	A Line Leaves 118 Degrees of Freedom
t_{n-2}	the t distribution on $n-2=118$	The Standardised Slope Follows a t Distribution
k	the number of slopes, one for our line	The Exact t Needs Normally Distributed Errors
t	standard errors from the claim	The t-Statistic Compares an Estimate With Its Standard Error
df	the degrees of freedom, 118 here	Worked Example: The t-Statistic Against Zero
Pr(·)	the probability of the event in brackets	Deriving the Interval From the t Distribution

Notation Introduced in Sections 3, 4 and the Appendix

Symbol	What it is	First met
$\overset{a}{\sim}$	is approximately distributed as	The Standardised Slope Is Approximately Standard Normal
Normal(0, 1)	the standard normal distribution	The Standardised Slope Is Approximately Standard Normal
SSR_r	SSR from the restricted fit	The F-Statistic Compares Two Residual Sums of Squares
SSR_u	SSR from the unrestricted fit	The F-Statistic Compares Two Residual Sums of Squares
q	coefficients set to zero	The F-Statistic Compares Two Residual Sums of Squares
F	the test statistic for the block	The F-Statistic Compares Two Residual Sums of Squares
$F(q,\ n-k-1)$	the null distribution of F	Derivation: The F-Statistic as Restricted Against Unrestricted Fit
$\bar{R}^2$	adjusted R-squared, Lecture 3.3's	Derivation: When Adjusted R-Squared Rises

Worked Example: the Estimated Error Variance

Dividing the sum of squared residuals by 118 rather than 120 gives Lecture 2.3's 4.777:

$$\hat{\sigma}^2 = \frac{\text{SSR}}{\underbrace{n-2}_{\text{the degrees of freedom}}} = \frac{2692.8}{118} = 22.82 \quad \Longrightarrow \quad \hat{\sigma} = 4.777 \tag{17}$$

- SSR: the sum of the 120 squared residuals, 2692.8 in our sample

Back to the slide

Worked Example: Our Degrees of Freedom

Our 120 people, less an intercept and a slope, leave 118 degrees of freedom:

$$\frac{\hat{\beta}_1 - \beta_1}{\text{se}(\hat{\beta}_1)} \sim t_{n-2} \qquad n - 2 = 118 \tag{18}$$

Wooldridge's $n - k - 1$ gives the same 118, because our k is one.

Back to the slide

Worked Example: the t-Statistic Against Zero

Against a null of no return, our estimate sits eight standard errors above zero:

$$t = \frac{\hat{\beta}_1 - \beta_{1,0}}{\mathrm{se}(\hat{\beta}_1)} = \frac{1.5227 - 0}{0.1897} = 8.03 \qquad (19)$$

$$\text{df} = 118 \qquad c = 1.980$$

Beyond c, a t on 118 degrees of freedom leaves 2.5 per cent in each tail.

> **Zero Is Rejected at Five per Cent**
> Our 8.03 clears 1.980, so a two-sided five per cent test rejects zero (Wooldridge, 2020). Angrist & Pischke (2015) treat any $|t|$ past about two as hard to square with a null. Testing 1.30 is separate.

Back to the slide

Worked Example: the t-Statistic Against 1.30

Against 1.30, the same estimate and standard error give a t well inside 1.980:

$$t = \frac{\hat{\beta}_1 - \beta_1}{\text{se}(\hat{\beta}_1)} = \frac{1.5227 - 1.30}{0.1897} = 1.17 \tag{20}$$

The sample rejects zero and does not reject 1.30.

Back to the slide

Worked Example: the Interval on Our 120 People

Our estimate, its standard error and c give two endpoints:

$$\hat{\beta}_1 \pm c\,\mathrm{se}(\hat{\beta}_1)$$

$$1.5227 \pm \underbrace{1.980 \times 0.1897}_{\text{the margin of error}} = 1.5227 \pm 0.3756 = [1.147,\ 1.898] \tag{21}$$

> **1.980 Sits Further Out Than 1.96**
> A t on 118 degrees of freedom has thicker tails than a normal. Wooldridge licenses 1.96 above 120 degrees of freedom, and two standard errors above 50 (Wooldridge, 2020).

Back to the slide

Worked Example: the Spread on Our 120 People

Across samples, the slope's standard deviation is σ over the root of SST_x:

$$\text{sd}(\hat{\beta}_1) = \frac{\sigma}{\sqrt{\text{SST}_x}} = \frac{4.50}{25.18} = 0.1787 \tag{22}$$

- SST_x: the sum of squared deviations of schooling, 633.9 in our sample

Our sample replaces σ with $\hat{\sigma}$ and gets its standard error:

$$\text{se}(\hat{\beta}_1) = \frac{\hat{\sigma}}{\sqrt{\text{SST}_x}} = \frac{4.777}{25.18} = 0.1897 \tag{23}$$

Back to the slide

Worked Example: the Exact t and the Large-Sample Normal on Our 120

Working out both critical values on our 120 shows how much the choice changes:

$$\hat{\beta}_1 \pm c\,\mathrm{se}(\hat{\beta}_1) = 1.5227 \pm 1.980 \times 0.1897 = [\,1.147,\ 1.898\,] \tag{24}$$

The large-sample normal puts 1.960 in place of the exact 1.980:

$$\hat{\beta}_1 \pm 1.960\,\mathrm{se}(\hat{\beta}_1) = 1.5227 \pm 1.960 \times 0.1897 = [\,1.151,\ 1.895\,] \tag{25}$$

Each endpoint shifts by 0.0038, under half a cent an hour.

Back to the slide

Derivation: Why the Distribution Is t and Not Normal

Section 1 stated the exact t and did not derive it. With σ known and the errors normal, the standardised slope would be standard normal exactly:

$$\frac{\hat{\beta}_1 - \beta_1}{\text{sd}(\hat{\beta}_1)} \sim \text{Normal}(0, 1) \qquad \frac{\hat{\beta}_1 - \beta_1}{\text{se}(\hat{\beta}_1)} \sim t_{118} \tag{26}$$

One Estimated Number Separates the Two

Only σ differs: the left uses it and the right uses $\hat{\sigma}$. Lecture 2.3 computed $\hat{\sigma}$ from the 120 residuals. That second source of variation thickens the tails, so 1.980 sits further out than 1.96 (Wooldridge, 2020).

Derivation: the Interval as the Set of Nulls That Survive

Section 1's test runs at every possible null value $\beta_{1,0}$ in turn. The values the sample does not reject are these:

$$\left| \frac{\hat{\beta}_1 - \beta_{1,0}}{\mathrm{se}(\hat{\beta}_1)} \right| \leq c \qquad \Longleftrightarrow \qquad \hat{\beta}_1 - c\,\mathrm{se}(\hat{\beta}_1) \;\leq\; \beta_{1,0} \;\leq\; \hat{\beta}_1 + c\,\mathrm{se}(\hat{\beta}_1) \tag{27}$$

The Same Two Endpoints by a Second Route

The values that survive are exactly 1.147 to 1.898. A null is rejected at five per cent precisely when the interval excludes it. The two frames in section 2 were therefore making one claim (Wooldridge, 2020).

Derivation: the F-Statistic as Restricted Against Unrestricted Fit

Fit the equation twice, with the restriction imposed and without it. Write the restriction's cost as a ratio of two variance estimates. Four steps:

1. Write the restricted equation as the unrestricted one with q zeros.
2. Show that SSR_r is at least as large as SSR_u.
3. Divide the difference by q, and SSR_u by $n - k - 1$.
4. Show that under the null both estimate the same error variance.

> **Step Four Is Where the Assumptions Enter**
> Step four needs normal errors on top of Lecture 2.3's five assumptions. The ratio then follows an $F(q,\ n - k - 1)$ distribution.

[Full derivation. Set as homework; worked solution published with the set.]

Derivation: When Adjusted R-Squared Rises

Show when adding one variable raises $\bar{R}^2$. The threshold is a t-statistic of one in absolute value. Two steps, using only the residual sums of squares:

1. Write $\bar{R}^2$ for both equations, using their residual sums of squares.
2. Take the difference, and what is left compares a squared t-statistic with one.

> **Adjusted R-Squared Rises at a T-Statistic of One**
> Section two's caveat rests on this threshold. A t-statistic of one leaves a coefficient nowhere near conventional significance.

[Full derivation. Set as homework; worked solution published with the set.]

Notes and Tools

- **The App:** the reporting app, with the sample size raised while the estimate is held fixed.
- **The Lecture in One Line:** `confint(lm(y ~ I(x - 13)), level = 0.95)`.
- **This Lecture's Lab:** `summary()` beside `confint()`, then a table that prints standard errors rather than stars.
- **Joint Tests in R:** `anova(fit_restricted, fit_unrestricted)`, or `car::linearHypothesis()` when the restriction is not all zeros.
- **Adjusted R-Squared and VIFs:** `summary(fit)$adj.r.squared` and `car::vif()`, read alongside O'Brien (2007).
- **Today's Reading:** Wooldridge (2020), sections 4-1 to 4-3, 4-5 and 5-2; Angrist & Pischke (2017) on what a first course should teach.

References

Angrist, J. D., & Pischke, J.-S. (2015). *Mastering 'metrics: The path from cause to effect*. Princeton University Press.

Angrist, J. D., & Pischke, J.-S. (2017). Undergraduate econometrics instruction: Through our classes, darkly. *Journal of Economic Perspectives*, *31*(2), 125–144. https://doi.org/10.1257/jep.31.2.125

O'Brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity*, *41*(5), 673–690. https://doi.org/10.1007/s11135-006-9018-6

Wooldridge, J. M. (2020). *Introductory econometrics: A modern approach* (7th ed.). Cengage Learning.