LECTURE 3.3

Functional Form and Model Choice

ECON30130 Econometrics I

Sam Deegan
University College Dublin

Autumn 2027

Where We Are

Block 3 Is Four Lectures on Multiple Regression

Lecture 3.2 tested the coefficients of an equation whose shape was fixed. Today changes the shape and asks which shape to keep. Lecture 3.4 puts categories such as a degree into the equation.

Last Week's Lecture

Lecture 3.2 asked what one sample of 120 can say about a coefficient:

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.

Every Earnings Equation So Far Is a Straight Line

Every equation since Lecture 2.1 adds the same euro per year of schooling. In our population the tenth year and the twentieth both add 1.30 euro.

> **Question**
> Which lecture decided that every year of schooling adds the same euro? Is that decision about the Irish labour market, or about a straight line?

The Data Never Object to a Straight Line

No Irish data chose the straight line; Lecture 2.1 imposed it. Least squares fits a line to curved data too:

- The fit still returns a slope, a standard error and an R^2.
- All three are computed correctly, and none reports that the data curve.

> **Today Changes the Shape and Keeps Least Squares**
> A logarithm changes what a coefficient means. Least squares still fits a log model, because it stays linear in its parameters (Wooldridge, 2020, p. 40).

Lecture Plan

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.

SECTION 1

Logs and Elasticities

One Euro Figure at Every Level of Schooling

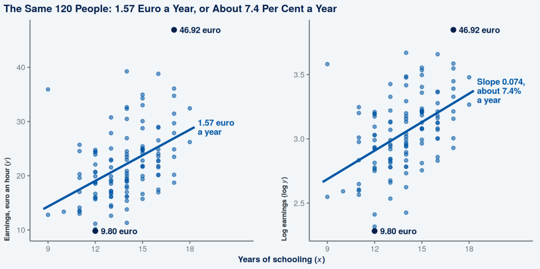

Figure 1: Our 120, all of them in each fit: earnings, then log earnings, on schooling.

> **Each Panel's Line Makes a Different Claim About Ireland**
> The left line adds the same euro at every level of schooling. The right line adds the same proportion, which is more euro at higher earnings. Logs also pull in the long right tail of earnings.

Logs Put the Return in Proportional Terms

Replacing earnings with their natural logarithm turns the euro return into a proportion:

$$\log(y_i) = \beta_0 + \beta_1(x_i - x_0) + u_i \tag{1}$$

- y_i: hourly earnings of person i, in euro
- x_0: the reference level of schooling, thirteen years in our population
- β_1: the change in log earnings per year of schooling
- u_i: everything other than schooling

> **A Semi-Elasticity Is a Hundred Times the Slope**
> β_1 takes a change in years to a proportional change in earnings. Wooldridge calls $100\beta_1$ the semi-elasticity (Wooldridge, 2020, p. 39).

A Hundred Times a Log Change Is Nearly a Percentage

Holding u_i fixed, a year of schooling changes log earnings by β_1. A hundred times a log change is nearly the percentage change:

$$\%\Delta y \approx 100\,\Delta\log(y) = 100\,\beta_1 \Delta x \tag{2}$$

Our $\hat{\beta}_1 = 0.1481$ (standard error 0.0239) means about 14.81 per cent a year.

An Elasticity Has No Units

A slope in euro per year changes whenever euro or years are rescaled. An elasticity (η) compares two proportional changes instead:

$$\eta = \frac{\%\Delta y}{\%\Delta x} = \frac{dy}{dx} \cdot \frac{x}{y} \tag{3}$$

- η: the elasticity of earnings with respect to schooling

Euro over Euro and Years over Years

In dy/dx euro sit over years, and in x/y years sit over euro. Both units divide out (Wooldridge, 2020, eq. A.24 p.676).

Why a Log-Log Slope Is an Elasticity

A log differential is a proportional change, because $d\log(x)/dx = 1/x$ (Wooldridge, 2020, p. 679):

$$\frac{d\log(y)}{d\log(x)} = \frac{dy/y}{dx/x} = \eta \tag{4}$$

An Elasticity of 2.01 from Our Log-Log Fit

Our fit gives $\hat{\eta} = 2.0119$, with a standard error of 0.3185. One per cent more schooling goes with about 2.01 per cent more earnings.

Four Specifications and the Sentence Each Slope Supports

Table 1: The same two variables under four specifications, and the sentence each slope supports.

Specification	Regression	Change in x	Say this out loud
Level–level	y on x	one year	One more year of schooling adds β_1 euro to hourly earnings.
Level–log	y on $\log x$	one per cent	One per cent more schooling adds about $\beta_1/100$ euro to hourly earnings.
Log–level	$\log y$ on x	one year	One more year of schooling raises earnings by about $100\beta_1$ per cent.
Log–log	$\log y$ on $\log x$	one per cent	One per cent more schooling raises earnings by about β_1 per cent.

Only the Level-Level Row Is Exact

In the three log rows, "about" marks the log approximation (Wooldridge, 2020, p. 39). Every β_1 here is a population coefficient, not our estimate.

Our Four Slopes Are in Four Different Units

Fitted to our people, the four slopes are 1.5227, 19.7087, 0.1481 and 2.0119.

> **No Two of the Four Slopes Can Be Ranked**
> Level-level is in euro per year. Log-log is in per cent per one per cent. The other two slopes mix euro and per cent.

A log-log sentence that names euro or a year describes a different row.

A Percentage Change Is Not a Percentage Point

A percentage change is measured against the starting value. A percentage point is the plain difference between two percentages:

- A tax rising from 4 to 6 per cent rises **two percentage points**.
- The same rise is $100 \times (6 - 4)/4 = 50$ **per cent**, because the base is 4.
- Michigan's 1994 rise was reported both ways, both correctly (Wooldridge, 2020, p. 672).

> **Only per Cent Applies to Earnings in Euro**
> Ask what the variable is measured in. Earnings in euro have no percentage to take a difference of. A log-earnings coefficient therefore reads in per cent.

Rules of Thumb for Taking a Logarithm

What a variable is measured in decides whether it is usually logged (Wooldridge, 2020, pp. 187–188):

- **Euro Amounts:** a positive money amount, such as earnings, is often logged.
- **Years:** schooling and other variables in years usually stay in levels.
- **Rates:** a logged rate reads in per cent, an unlogged rate in points.

> **Unemployment from 8 to 9 per Cent**
> The rise is one percentage point. The same rise is a 12.5 per cent increase in the rate (Wooldridge, 2020, p. 188).

Undoing the Logarithm Gives the Exact Percentage Change

Multiplying a log coefficient by a hundred is a shortcut, and only approximate. Exponentiating gives the exact percentage change (Wooldridge, 2020, eq. 6.9 p.186):

$$\%\Delta y = 100\left(e^{\beta_1} - 1\right) \tag{5}$$

Our 0.1481 gives 14.81 per cent by the shortcut, and 15.96 exactly.

The Shortcut's Gap Grows with the Coefficient

Table 2: The shortcut against the exact figure at four values of the coefficient, to four decimals.

β_1	**Shortcut** $100\beta_1$	**Exact**	**Gap**
0.05	5.0000	5.1271	0.1271
0.10	10.0000	10.5171	0.5171
0.25	25.0000	28.4025	3.4025
0.50	50.0000	64.8721	14.8721

Every Gap Is Positive Because the Exponential Bends Upward
At 0.10 the shortcut is half a percentage point low. At 0.50 the shortcut is nearly fifteen points low.

Taking Logs Drops Our Lowest Earner

Our lowest earner is on -0.06 euro an hour, where the logarithm is undefined:

- The level-level slope, 1.5227, uses all 120 people.
- The log-level slope, 0.1481, and the log-log slope, 2.0119, use 119.
- The dropped person is our poorest, so the drop is not random.

> **Adding a Euro Before the Log Changes the Slope**
> Fitting $\log(1 + y_i)$ would keep that person (Wooldridge, 2020, p. 188). The slope would then depend on the added euro, a number we chose.

Reporting a Log Coefficient in Percentage Points

A regression of log earnings on weeks of training returns 0.08. A colleague writes that training raises earnings by eight percentage points.

> **Question**
> The colleague's sentence contains two separate errors. Name both errors, and give the figure the sentence should carry.

Fixing the Unit and the Size in the Training Sentence

The colleague's sentence has one error of unit and one of size:

- **The Unit:** earnings are in euro, so the sentence needs per cent, not points.
- **The Size:** eight is the shortcut; the exact figure is $100(e^{0.08} - 1) = 8.3287$.

> **The Size Error Grows with the Coefficient**
> The unit error does not depend on the coefficient. The size error is 0.33 points at 0.08, and nearly fifteen at 0.50.

SECTION 2

Choosing a Specification

Our Population Was Generated as a Straight Line

We built our 200,000 people with the same euro return at every schooling level:

$$y_i = \beta_0 + \beta_1 (x_i - x_0) + u_i \tag{6}$$

- β_0, β_1: the intercept and the return we set in the generator
- u_i: a normal error, standard deviation 4.50, leaving one of our 120 below zero

> **A Log Specification Here Is a Shape We Imposed**
> The logged equations still fit, and still print a slope and an R^2. None of that output says our population is a straight line.

Why R-Squared Cannot Choose Between Specifications

R^2 never falls when a column is added to the same people (Wooldridge, 2020, p. 197):

- Lecture 2.1's *Adding a Variable That Cannot Matter* added height to schooling.
- The height regression had the higher R^2, whatever height does to earnings.

> **Among Nested Specifications the Largest Always Wins**
> A criterion that never falls always picks the specification with more columns. Choosing needs a criterion that can fall.

Adjusted R-Squared Divides Each Sum by Its Degrees of Freedom

Adjusted R-squared ($\bar{R}^2$) divides each sum of squares by its degrees of freedom (Theil, 1961; Wooldridge, 2020, p. 196):

$$\bar{R}^2 = 1 - \frac{\overbrace{\sum_{i=1}^{n} \hat{u}_i^2}^{\text{SSR}} /(n-k-1)}{\underbrace{\sum_{i=1}^{n}(y_i - \bar{y})^2/(n-1)}_{\text{the sample variance of earnings}}} \tag{7}$$

- $\hat{u}_i$: person i's residual
- $\bar{y}$: the average earnings of our 120

With one slope on our 120, $\bar{R}^2 = 0.3476$ sits 0.0055 below $R^2 = 0.3531$.

Adding a Variable Can Lower Adjusted R-Squared

Dividing by $n - k - 1$ lets $\bar{R}^2$ fall, which R^2 never does (Wooldridge, 2020, p. 197):

- A new variable raises $\bar{R}^2$ if and only if its $|t|$ exceeds one.
- $\bar{R}^2$ can go below zero when R^2 and $n - k - 1$ are both small.

> **Adjusted R-Squared Ranks and Does Not Test**
> A t-statistic of one is nowhere near conventional significance. Lecture 3.2's F-test tests one specification against another.

Predicting Which Specification Has the Highest R-Squared

All four specifications have now been fitted to the same people.

> **Question**
> Before any R^2 is shown: which of the four reports the highest? Does the highest R^2 tell you which shape generated the data?

Why 0.3531 and 0.2476 Are Not a Comparison

Table 3: The four R^2 figures, with the dependent variable and the people behind each.

Specification	Dependent variable	People	R^2
Level-level	earnings in euro	120	0.3531
Level-log	earnings in euro	120	0.3309
Log-level	log earnings	119	0.2476
Log-log	log earnings	119	0.2543

Only Rows Sharing a Dependent Variable Compare

Comparing 0.3531 with 0.2476 is not legitimate (Wooldridge, 2020, p. 188). Dividing by degrees of freedom changes neither the variable nor the sample.

The Highest R-Squared Does Not Identify the Shape

Level-level against level-log is the one legitimate comparison, and 0.3531 beats 0.3309.

> **The Straight Line Is True Only Because We Built It**
> Our population is a straight line. Here the higher figure happens to point at the truth. One sample gives some shape the highest R^2 whatever the truth is.

R^2 ranks shapes by fit in one sample and tests none of them.

Four Fitted Shapes over One Scatter

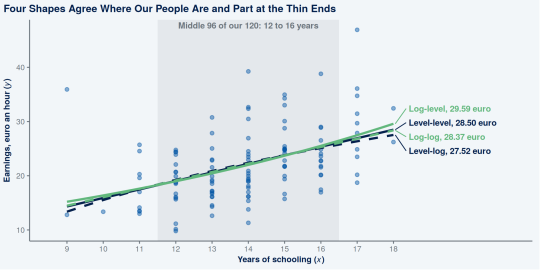

Figure 2: Our 120 and four fitted shapes; the two log-earnings fits are converted back to euro.

> **The Four Shapes Separate Where the Data Thin Out**
> Three of the four bend on the euro scale. In the shaded band, where 96 of our 120 sit, the four stay within 0.44 euro of each other. At 9 and 18 years, two people each, they are 1.83 and 2.07 euro apart.

SECTION 3

Specification Search

The Smallest P-Value Across Twenty Specifications

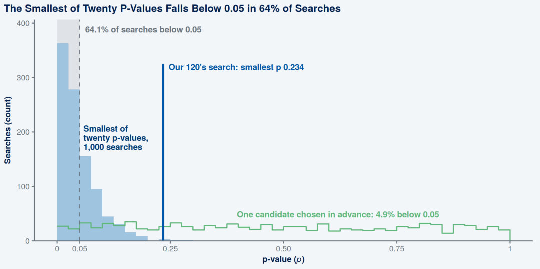

Figure 3: Repeated-sample ghost: smallest of 20 p-values in 1,000 searches; blue is ours.

Each P-Value Is Correct and the Reported One Is Selected

In 64.1 per cent of 1,000 searches on pure noise, the smallest of twenty p-values fell below 0.05. Our 120's own search did not: its smallest was 0.234. Freedman (1983) screened regressors on pure noise, and significant coefficients survived the screen.

Why P-Values Do Not Account for Specification Search

A p-value of 0.04 describes one test chosen before the data were seen.

> **Twenty Specifications Make a Different Procedure**
> Reporting the smallest of twenty p-values is a different procedure. The printed 0.04 looks the same either way.

No line of regression output records how many specifications were fitted.

Published Test Statistics Bunch Just Past the Thresholds

Published test statistics bunch just past the conventional significance thresholds (Brodeur et al., 2016; Brodeur et al., 2020):

- One test chosen in advance in every paper would leave no such bunching.
- Simmons et al. (2011) show how few analysis choices produce a significant result.
- Gelman & Loken (2014) count choices made after seeing the data as a search.

Research Practices That Make the Search Visible

Three practices make a specification search visible to a reader:

- **Pre-Analysis Plans:** fix the specification and outcome in advance (Olken, 2015).
- **Every Specification, Not the Winner:** report the whole set (Simonsohn et al., 2020).
- **Code and Data:** publish both for others to rerun (Christensen & Miguel, 2018).

> **Exploration Is Legitimate When Reported as Exploration**
> Pre-specification forecloses findings reached only by looking (Coffman & Niederle, 2015). A specification found by exploring is reported as exploration.

Keeping a Variable on a Rise in Adjusted R-Squared

An added variable leaves schooling's coefficient almost unchanged and raises $\bar{R}^2$ by 0.004. Its t-statistic is 1.2.

> **Question**
> Does that rise in $\bar{R}^2$ add anything the t-statistic had not said? Now suppose you chose the variable after seeing the output. Does your answer change?

A Rise in Adjusted R-Squared Repeats the t-Statistic

$\bar{R}^2$ rises exactly when the new $|t|$ exceeds one (Wooldridge, 2020, p. 197). The rise and the 1.2 are one fact.

> **The Output Is Identical Before and After Looking**
> A variable chosen in advance supports a pre-specified claim. The same output after looking does not. Nothing in the output shows which case holds.

Counting the rise as a second piece of evidence counts the 1.2 twice.

What Would Have to Be True?

By our log-level fit, a year of schooling raises earnings about 15.96 per cent. The straight line and the logarithm were both our choices.

> **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?

Our Population Has No Single Percentage Return

The 15.96 per cent assumes log earnings rise by the same step each year:

- Our population adds 1.30 euro a year at every level of schooling.
- That 1.30 is 12.0 per cent of the mean 10.80 at thirteen years.
- At eighteen years the same 1.30 is 7.5 per cent of the mean 17.30.

> **No Printout on This Deck Flags the Shape**
> The log fit prints a slope, a standard error and an R^2. Each is computed correctly, and none says the percentage falls with schooling.

Summary: Logs, Specification Choice and Specification Search

1. **Logs and Elasticities.**
   A log slope of 0.1481 means 15.96 per cent a year by the exact formula.

2. **Choosing a Specification.**
   R^2 never falls, and $\bar{R}^2$ rises only when $|t|$ exceeds one.

3. **Specification Search.**
   A p-value of 0.04 describes one test, not the best of twenty.

Before Lecture 3.4

1. **Play:**
   The app, switching specification to watch the shape and the sentence change.
2. **Apply:**
   Problem set 3, one table written in prose.
3. **Review:**
   Wooldridge (2020), sections 6-2a, 6-3a and 6-3b, then 7-1 to 7-4 before Lecture 3.4.

Next Lecture

Lecture 3.4 puts categories, such as holding a degree, into the equation:

1. **A Single Dummy Variable.**
   A 0/1 column shifts the intercept and leaves the two lines parallel.
2. **Multiple Categories.**
   Five qualifications need four dummies, each a gap from the base you chose.
3. **Interactions.**
   A product term lets the slopes differ, and the schooling coefficient describes the base group.
4. **Testing a Difference in Slopes.**
   Equal slopes is a test on one coefficient, and 19 graduates cannot reject it.

Thank you

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

BACKUP

Appendix

Notation Used in This Lecture

Symbol	What it is	First met
i	which person, running over the 120 in our sample	Lecture 1.1
n	how many people are in the sample, here 120	Lecture 1.1
x_i	years of schooling of person i	Lecture 1.1
y_i	hourly earnings of person i, in euro	Logs Put the Return in Proportional Terms
x	a level of schooling, without a subscript	Lecture 1.2
y	earnings, without a subscript	Lecture 1.2
$\log(\cdot)$	the natural logarithm	Logs Put the Return in Proportional Terms

Notation Used in This Lecture

Symbol	What it is	First met
x_0	the reference level of schooling, thirteen years in our population	Lecture 2.1
β_0	the intercept, expected earnings at thirteen years	Lecture 2.1
β_1	the slope on schooling; in the log model, the change in log earnings per year	Logs Put the Return in Proportional Terms
$\hat{\beta}_1$	the fitted β_1, 0.1481 in the log-level fit	Lecture 2.1
u_i	everything other than schooling	Logs Put the Return in Proportional Terms
Δ	the change in what follows	A Hundred Times a Log Change Is Nearly a Percentage
$\%\Delta y$	the percentage change in earnings	A Hundred Times a Log Change Is Nearly a Percentage
η	the elasticity of earnings with respect to schooling	An Elasticity Has No Units

Notation Used in This Lecture

Symbol	What it is	First met
d	an infinitesimally small change in what follows	An Elasticity Has No Units
$\hat{\eta}$	the fitted elasticity, 2.0119	Why a Log-Log Slope Is an Elasticity
e	the base of the natural logarithm	Undoing the Logarithm Gives the Exact Percentage Change
$\hat{y}$	the fitted value of earnings	Lecture 2.1
$\bar{y}$	the average earnings of our 120	Lecture 1.1
$\hat{u}_i$	person i's residual	Lecture 2.1
R^2	the share of variation in the dependent variable a fit accounts for	Lecture 2.1
SSR	the sum of squared residuals	Lecture 2.1
SST	the total sum of squares	Lecture 2.1

Notation Used in This Lecture

Symbol	What it is	First met
k	how many slopes the equation has	Lecture 3.2
$\bar{R}^2$	adjusted R-squared	Adjusted R-Squared Divides Each Sum by Its Degrees of Freedom
t	an estimate divided by its standard error	Lecture 3.2
q	how many coefficients the restriction sets to zero	Lecture 3.2
SSR_r	the sum of squared residuals with the restriction imposed	Lecture 3.2
SSR_u	the sum of squared residuals without the restriction	Lecture 3.2
F	the restricted-against-unrestricted test statistic	Lecture 3.2

Notation Used in This Lecture

Symbol	What it is	First met
G_i	one for a graduate, zero otherwise; taught in Lecture 3.4	Derivation: Why the Centring Constant Cancels From the Slopes
δ_0	the dummy coefficient with schooling centred; Lecture 3.4	Derivation: Why the Centring Constant Cancels From the Slopes
δ_1	the difference in slope between the two groups; Lecture 3.4	Derivation: Why the Centring Constant Cancels From the Slopes
β_0^{raw}	the intercept with schooling uncentred	Derivation: Why the Centring Constant Cancels From the Slopes
δ_0^{raw}	the dummy coefficient with schooling uncentred	Derivation: Why the Centring Constant Cancels From the Slopes

Derivation: Where the Exact Percentage Change Comes From

The body frame stated $\%\Delta y = 100(e^{\beta_1} - 1)$. That formula is the fitted model differenced across one year and exponentiated:

$$\log \hat{y}(x+1) - \log \hat{y}(x) = \beta_1 \quad \Longrightarrow \quad \frac{\hat{y}(x+1)}{\hat{y}(x)} = e^{\beta_1} \quad \Longrightarrow \quad \frac{\Delta \hat{y}}{\hat{y}} = e^{\beta_1} - 1 \tag{8}$$

The Shortcut Is the First Term of a Series

$e^{\beta_1} - 1 = \beta_1 + \beta_1^2/2 + \ldots$, so the shortcut drops everything past the first term. The leading error is $\beta_1^2/2$. That leading error is 0.00125 at $\beta_1 = 0.05$ and 0.125 at $\beta_1 = 0.50$. The gap column on the body frame grows at that same square rate (Wooldridge, 2020, p. 186).

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.]

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: Why the Centring Constant Cancels from the Slopes

Replacing x_i by $x_i - x_0$ rewrites the intercept and the dummy coefficient. The two slopes are unchanged:

$$\beta_0^{\text{raw}} + \beta_1 x_i + \delta_0^{\text{raw}} G_i + \delta_1 x_i G_i = \underbrace{(\beta_0^{\text{raw}} + x_0 \beta_1)}_{\beta_0} + \beta_1 (x_i - x_0) + \underbrace{(\delta_0^{\text{raw}} + x_0 \delta_1)}_{\delta_0} G_i + \delta_1 (x_i - x_0) G_i \quad (9)$$

The Intercept and the Dummy Coefficient Are Relabelled

β_1 and δ_1 appear unchanged on both sides. No slope, fitted value or residual differs. Only the intercept and the dummy coefficient are relabelled. Centring reports those two coefficients at thirteen years of schooling instead of zero (Wooldridge, 2020, p. 193).

Notes and Tools

- **The App:** the specification app, with the four functional forms switchable.
- **The Lecture in One Line:** `lm(log(y) ~ I(x - 13))`, the log-level fit.
- **This Lecture's Lab:** `log()`, `I(x - 13)`, the specification-search script, and Lecture 2.2's hold-out sample.
- **The F-Test in One Line:** `anova(fit_restricted, fit_unrestricted)` in R.
- **Restrictions That Are Not All Zero:** `car::linearHypothesis()`.
- **Adjusted R-Squared and VIFs:** `summary(fit)$adj.r.squared` and `car::vif()`, read alongside O'Brien (2007).
- **Where This Module Sits:** Angrist & Pischke (2017), on what a first econometrics course should teach.

References

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

Brodeur, A., Cook, N., & Heyes, A. (2020). Methods matter: P-hacking and publication bias in causal analysis in economics. *American Economic Review*, *110*(11), 3634–3660. https://doi.org/10.1257/aer.20190687

Brodeur, A., Lé, M., Sangnier, M., & Zylberberg, Y. (2016). Star wars: The empirics strike back. *American Economic Journal: Applied Economics*, *8*(1), 1–32. https://doi.org/10.1257/app.20150044

Christensen, G., & Miguel, E. (2018). Transparency, reproducibility, and the credibility of economics research. *Journal of Economic Literature*, *56*(3), 920–980. https://doi.org/10.1257/jel.20171350

Coffman, L. C., & Niederle, M. (2015). Pre-analysis plans have limited upside, especially where replications are feasible. *Journal of Economic Perspectives*, *29*(3), 81–98. https://doi.org/10.1257/jep.29.3.81

Freedman, D. A. (1983). A note on screening regression equations. *The American Statistician*, *37*(2), 152–155. https://doi.org/10.1080/00031305.1983.10482729

Gelman, A., & Loken, E. (2014). The statistical crisis in science. *American Scientist*, *102*(6), 460–465. https://doi.org/10.1511/2014.111.460

References

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

Olken, B. A. (2015). Promises and perils of pre-analysis plans. *Journal of Economic Perspectives*, *29*(3), 61–80. https://doi.org/10.1257/jep.29.3.61

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Simonsohn, U., Simmons, J. P., & Nelson, L. D. (2020). Specification curve analysis. *Nature Human Behaviour*, *4*(11), 1208–1214. https://doi.org/10.1038/s41562-020-0912-z

Theil, H. (1961). *Economic forecasts and policy* (2nd ed.). North-Holland.

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