Econ 103: Introduction to Econometrics

Lecture 17 — F-Tests & Joint Hypothesis Testing

Ryan Longmuir

UCLA

Summer Session C, 2026

Reading: Hill, Griffiths & Lim (5th ed.), §6.1–6.2; Stock & Watson (4th ed.), §7.2.

Where we are

The t-test handles a single restriction — one "equals" sign, even one spanning several coefficients. But many questions are **joint**:

- Does advertising matter *at all* — is $\beta_3 = 0$ and $\beta_4 = 0$ in Big Andy's quadratic model?
- Does a *group* of variables (socioeconomic controls, prices of substitutes) belong?
- Does the model explain *anything* — are all slopes zero?

These have several "equals" signs. A t-test can't do them — and testing one at a time is unreliable. **Today: the F-test.**

- restricted vs. unrestricted models,
- overall significance and the t–F relationship,
- testing economic restrictions (Cobb–Douglas constant returns to scale).

Today's plan

1. Why a new test?
2. The F-statistic: restricted vs. unrestricted
3. Overall significance and the t–F link
4. Testing economic restrictions

Part 1

Why a new test?

Joint hypotheses, and the one-at-a-time trap

A **joint hypothesis** imposes $J \geq 2$ restrictions at once, e.g.

$$H_0: \ \beta_3 = 0 \ \text{ and } \ \beta_4 = 0 \qquad \text{vs.} \qquad H_1: \ \beta_3 \neq 0 \ \text{ or } \ \beta_4 \neq 0.$$

Why not just run two t-tests?

- Error rates compound. Two 5%-tests don't give a 5% joint test — the chance of *some* false rejection is larger.
- It misreads correlated regressors. When two regressors are collinear, *each* t can be insignificant while the pair is jointly decisive. One-at-a-time would wrongly drop both.

We need a test that weighs *all* the restrictions together. That is the F-test.

Part 2

The F-statistic: restricted vs. unrestricted

Two models: with and without the restrictions

The F-test compares the fit of two nested models.

Unrestricted (full)

$$\text{SALES} = \beta_1 + \beta_2 \text{PRICE}$$
$$+\beta_3 \text{ADVERT} + \beta_4 \text{ADVERT}^2 + e$$

Sum of squared errors: SSE_U.

Restricted (impose H_0)

Set $\beta_3 = \beta_4 = 0$:

$$\text{SALES} = \beta_1 + \beta_2 \text{PRICE} + e$$

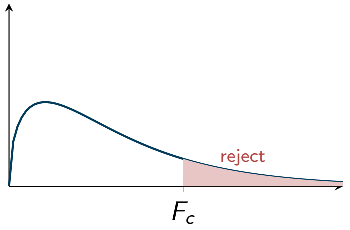
Sum of squared errors: SSE_R.

Dropping variables can only *worsen* fit, so $\mathrm{SSE}_R \geq \mathrm{SSE}_U$ always. The question is whether the increase in SSE from imposing H_0 is *large* (restrictions hurt $\Rightarrow$ reject) or *small* (restrictions harmless $\Rightarrow$ don't reject).

The F-statistic

$$F = \frac{(\mathrm{SSE}_R - \mathrm{SSE}_U)/J}{\mathrm{SSE}_U/(N-K)} \sim F_{(J,\,N-K)} \text{ under } H_0$$

- J = number of restrictions (numerator df).
- $N - K$ = unrestricted df (denominator df).
- Numerator: extra error from imposing H_0, per restriction.
- Denominator: the model's own noise $\hat{\sigma}^2$.
- Large $F \Rightarrow$ restrictions cost a lot $\Rightarrow$ reject H_0 if $F \geq F_c$.



F-distribution; reject in the right tail.

Big Andy's: does advertising matter?

Test $H_0 : \beta_3 = 0,\ \beta_4 = 0$ (advertising, linear *and* quadratic, irrelevant) against "at least one nonzero." Here $J = 2,\ N = 75,\ K = 4$:

$$\mathrm{SSE}_U = 1532.08, \qquad \mathrm{SSE}_R = 1896.39.$$

$$F = \frac{(1896.39 - 1532.08)/2}{1532.08/(75 - 4)} = 8.44.$$

The 5% critical value is $F_{(0.95,2,71)} = 3.13$; $p = 0.0005$.

Since $8.44 > 3.13$, reject H_0 — advertising does affect sales. (We couldn't have learned this cleanly from the two separate t's, since ADVERT and ADVERT^2 are collinear.)

Equivalent R^2 form (S&W): $F = \dfrac{(R^2_U - R^2_R)/J}{(1 - R^2_U)/(N - K)}$ — same number, computed from fit instead of SSE.

Part 3

Overall significance and the t–F link

Testing the whole model at once

The most-reported F-test asks: do the regressors *jointly* explain anything?

$$H_0: \ \beta_2 = \beta_3 = \cdots = \beta_K = 0 \qquad \text{(the model is worthless).}$$

The restricted model is just $y_i = \beta_1 + e_i$, fitted by $\bar{y}$, so $\mathrm{SSE}_R = \mathrm{SST}$. The statistic specializes to

$$F = \frac{(\mathrm{SST} - \mathrm{SSE})/(K-1)}{\mathrm{SSE}/(N-K)} \sim F_{(K-1,\, N-K)}.$$

Example (Big Andy's overall F)

With $\mathrm{SST} = 3115.48$, $\mathrm{SSE} = 1532.08$, $K = 4$*:*

$$F = \frac{(3115.48 - 1532.08)/3}{1532.08/71} = 24.46 \gg F_c = 2.73.$$

Reject — at least one of PRICE, ADVERT, ADVERT^2 *matters. This is the* **overall significance** F *printed on every regression.*

t and F: when are they the same?

For a single restriction ($J = 1$): identical

A two-tailed t-test and the F-test give the same answer, because

$$F = t^2 \qquad \text{and} \qquad F_c = t_c^2.$$

(Big Andy, $H_0 : \beta_2 = 0$: $t = -7.30$, $t^2 = 53.4 = F$.) Same p-value, same conclusion.

But two crucial places where *only* the right tool works:

- One-tailed tests ($H_1 : \beta > c$): use t. Since $F = t^2$ squares away the sign, F *cannot* do one-sided.
- Joint tests ($J \geq 2$): use F. There is no single t.

Custom: test single restrictions with t, joint restrictions with F.

Part 4

Testing economic restrictions

Beyond "$= 0$": restrictions from theory

The real power of the F-test: the restrictions can be *any* linear equalities theory hands us — not just "this coefficient is zero."

Cobb–Douglas and constant returns to scale

A production function $Q = A\,L^{\beta_2} K^{\beta_3}$ in logs is

$$\ln Q = \beta_1 + \beta_2 \ln L + \beta_3 \ln K + e.$$

Constant returns to scale — doubling all inputs doubles output — is exactly

$$H_0:\ \beta_2 + \beta_3 = 1.$$

Impose it (a restricted model with one fewer free parameter), get SSE_R, and form the F with $J = 1$. Increasing returns if the data reject in favor of $\beta_2 + \beta_3 > 1$.

More restriction examples

No money illusion (HGL beer demand)

Log-log demand $\ln Q = \beta_1 + \beta_2 \ln P_B + \beta_3 \ln P_L + \beta_4 \ln P_R + \beta_5 \ln I + e$. Scaling all prices *and* income shouldn't change quantity, so

$$H_0: \ \beta_2 + \beta_3 + \beta_4 + \beta_5 = 0.$$

Is \$1,900 the optimal ad spend?

Andy's optimum satisfies $\beta_3 + 2\beta_4 \text{ADVERT} = 1$. At $\text{ADVERT} = 1.9$,

$$H_0: \ \beta_3 + 3.8\beta_4 = 1.$$

$F = 0.94 < 3.98$ — fail to reject; \$1,900 is compatible with the data.

How it's done in practice

Either rewrite the model to embed the restriction and re-estimate (get SSE_R), or just hand the restriction to software — it computes the F (a **Wald test**) and p-value directly.

Joint restrictions can mix several conjectures

Nothing stops H_0 from bundling *different* economic claims. Andy plans staffing on two assumptions at once: \$1,900 is optimal *and* sales at PRICE $= 6$, ADVERT $= 1.9$ average \$80,000:

$$H_0: \ \beta_3 + 3.8\beta_4 = 1 \quad \text{and} \quad \beta_1 + 6\beta_2 + 1.9\beta_3 + 3.61\beta_4 = 80.$$

Two restrictions ($J = 2$) $\Rightarrow$ an F-test (no t possible). Here $F = 5.74$, $p = 0.005$ — reject: the two plans are *jointly* incompatible with the data, even if each alone might survive.

This is the everyday use of F-tests in research: bundling a model's theoretical restrictions and asking whether the data can live with all of them together.

Recap

The F-test

- joint H_0 ($J \geq 2$ restrictions) $\Rightarrow$ need F, not many t's
- $F = \dfrac{(\text{SSE}_R - \text{SSE}_U)/J}{\text{SSE}_U/(N-K)} \sim F_{(J,N-K)}$
- big SSE jump from H_0 $\Rightarrow$ reject
- Andy ads: $F = 8.44$, reject

Overall significance

- all slopes $= 0$; restricted model $= \bar{y}$
- Andy: $F = 24.46$, model significant

t vs. F

- $J = 1$: $F = t^2$, equivalent (two-tail)
- one-tail $\rightarrow$ t only; joint $\rightarrow$ F only

Economic restrictions

- CRS: $\beta_2 + \beta_3 = 1$
- no money illusion: $\sum \beta = 0$
- optimal ad: $\beta_3 + 3.8\beta_4 = 1$ (not rejected)

Next time (Lecture 18): which variables belong?

F-tests assumed a model. But choosing it is the hard part: omitted-variable bias vs. irrelevant variables, and tools to decide — adjusted R^2, AIC/BIC, the RESET test, and residual diagnostics.

Questions?