F-Tests and Joint Hypothesis Testing

When Individual t-Tests Are Not Enough

Jake Anderson

March 21, 2026

Outline

Outline

Where We Left Off

Last lecture ended with a puzzle. Consider the wage model:

$$\text{WAGE}_i = \beta_1 + \beta_2\,\text{EDUC}_i + \beta_3\,\text{EXPER}_i + \beta_4\,\text{EXPER}_i^2 + e_i$$

Suppose individual t-tests give these results:

Variable	Coefficient	Std. Error	t-stat
EXPER	0.12	0.08	1.50
EXPER^2	-0.001	0.001	-1.00

Neither is individually significant at 5%. Should we drop both and conclude experience does not affect wages?

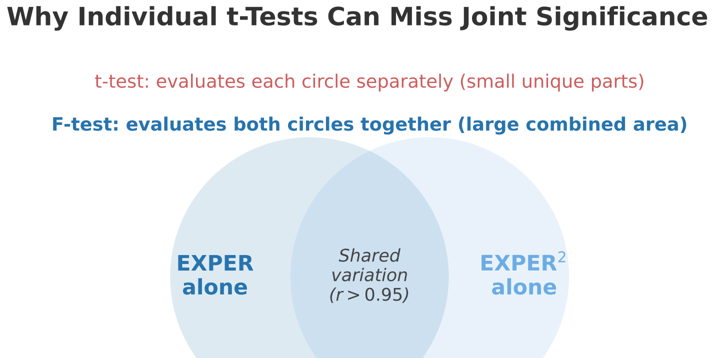
$\Longrightarrow$ **No.** The two variables are highly correlated, inflating both standard errors. We need a test that evaluates them **together**.

The Problem with One-at-a-Time Testing

Each t-test asks: "Does this variable contribute, *given that the other is already in the model*?"

When EXPER and EXPER^2 are correlated ($r > 0.95$ in most datasets):

- Each variable shares so much information with the other that neither looks significant *alone*
- But **together** they capture the curvature of the wage–experience profile



Outline

The Idea: Compare Two Models

The F-test works by comparing two regressions:

Unrestricted model (includes the variables in question):

$$\text{WAGE}_i = \beta_1 + \beta_2\,\text{EDUC}_i + \beta_3\,\text{EXPER}_i + \beta_4\,\text{EXPER}_i^2 + e_i$$

Restricted model (imposes $H_0: \beta_3 = 0,\ \beta_4 = 0$):

$$\text{WAGE}_i = \beta_1 + \beta_2\,\text{EDUC}_i + e_i$$

Each model produces a sum of squared errors: SSE_U and SSE_R.

Because the unrestricted model has more parameters, it always fits at least as well:

$$SSE_R \geq SSE_U \quad \text{always}$$

$\Longrightarrow$ The question is whether $SSE_R - SSE_U$ is large enough to be **statistically meaningful**, or just sampling noise.

The F-Statistic

$$F = \frac{(SSE_R - SSE_U) \ / \ J}{SSE_U \ / \ (N - K)}$$

Symbol	Meaning
SSE_R	Sum of squared errors from the restricted model
SSE_U	Sum of squared errors from the unrestricted model
J	Number of restrictions (hypotheses) in H_0
K	Number of parameters in the unrestricted model
N	Number of observations

Under H_0 (with MR1–MR6): $F \sim F_{(J,\ N-K)}$

Reject H_0 if $F \geq F_c$, where $F_c = F_{(1-\alpha,\ J,\ N-K)}$.

$\Longrightarrow$ The numerator measures how much the restrictions **hurt** the fit (per restriction). The denominator is the baseline noise level (per residual degree of freedom).

Reading the F-Statistic

$$F = \frac{\underbrace{(SSE_R - SSE_U)/J}_{\text{fit lost per restriction}}}{\underbrace{SSE_U/(N-K)}_{\text{baseline noise per df}}}$$

Large F: The restrictions cause a big increase in errors relative to noise.

- The restricted variables were doing real work $\Longrightarrow$ reject H_0

Small F: The restrictions barely change the fit.

- The restricted variables were not contributing much $\Longrightarrow$ do not reject H_0

$\Longrightarrow$ The F-test is always **one-sided** (right tail only), because $SSE_R \geq SSE_U$ guarantees $F \geq 0$.

Outline

Setup: Does Experience Belong in the Model?

Using CPS wage data ($N = 1000$):

Unrestricted model ($K = 4$):

$$\text{WAGE}_i = \beta_1 + \beta_2\,\text{EDUC}_i + \beta_3\,\text{EXPER}_i + \beta_4\,\text{EXPER}_i^2 + e_i$$

$$SSE_U = 97{,}882.50$$

Restricted model ($H_0 : \beta_3 = 0,\ \beta_4 = 0$):

$$\text{WAGE}_i = \beta_1 + \beta_2\,\text{EDUC}_i + e_i$$

$$SSE_R = 100{,}847.00$$

$J = 2$ restrictions. Degrees of freedom: $N - K = 996$.

Computing the F-Statistic

Step 1: Compute F.

$$F = \frac{(SSE_R - SSE_U)/J}{SSE_U/(N-K)}$$

$$= \frac{(100{,}847.00 - 97{,}882.50)\ /\ 2}{97{,}882.50\ /\ 996}$$

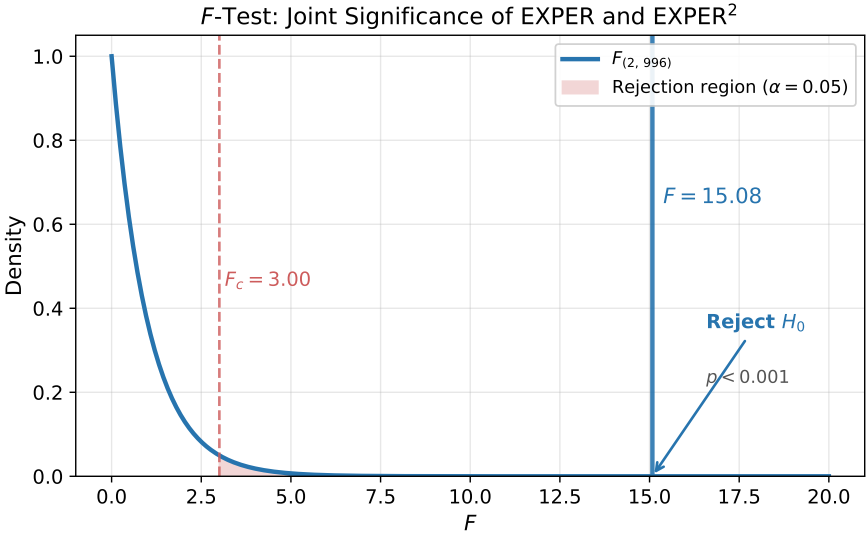
$$= \frac{2{,}964.50\ /\ 2}{98.27}$$

$$= \frac{1{,}482.25}{98.27} = 15.08$$

Step 2: Find the critical value.
At $\alpha = 0.05$: $F_c = F_{(0.95,\ 2,\ 996)} \approx 3.00$

Step 3: Decide.
$15.08 > 3.00 \implies$ **Reject H_0.** The experience terms are jointly significant.

Visualizing the Decision



Outline

Getting J and K Right

The most common computational mistakes involve the degrees of freedom.

J **= number of restrictions in** H_0

- Count the number of "$=$" signs in H_0
- $H_0: \beta_3 = 0,\ \beta_4 = 0 \implies J = 2$
- $H_0: \beta_3 = 0,\ \beta_4 = 0,\ \beta_5 = 0 \implies J = 3$

K **= total parameters in the *unrestricted* model**

- Includes the intercept
- $y = \beta_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + e \implies K = 4$

Denominator df = $N - K$ (from the *unrestricted* model, not the restricted one)

$\implies$ Always use the unrestricted model's parameter count for K.

What the F-Test Can and Cannot Do

Capability	**t-test**	**F-test**
Single restriction ($J = 1$)	Yes	Yes
Multiple restrictions ($J \geq 2$)	No	Yes
One-sided alternative ($>$ or $<$)	Yes	No
Two-sided alternative ($\neq$)	Yes	Yes

When $J = 1$ (two-sided):

$$F = t^2 \qquad \text{and} \qquad F_{(1,\ N-K)} = t^2_{(N-K)}$$

The p-values are identical.

$\implies$ Use t-tests for single restrictions (they handle one-sided alternatives). Use F-tests for joint restrictions ($J \geq 2$).

Outline

Does the Model Explain Anything?

The most common application of the F-test: does the regression model have any explanatory power at all?

$$H_0: \beta_2 = 0,\ \beta_3 = 0,\ \ldots,\ \beta_K = 0$$

$$H_1: \text{at least one } \beta_k \neq 0 \text{ for } k = 2, \ldots, K$$

The restricted model (under H_0) contains only the intercept:

$$y_i = \beta_1 + e_i$$

so $SSE_R = SST$ and $J = K - 1$

The F-statistic simplifies to:

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

$\Longrightarrow$ This is reported automatically by every regression software package as the "overall F-statistic."

The ANOVA Table

Software organizes the components of the overall F-test in an **analysis of variance (ANOVA) table**:

Source	**Sum of Squares**	**df**	**Mean Square**
Regression (explained)	$SST - SSE$	$K - 1$	$\dfrac{SST - SSE}{K - 1}$
Residual (unexplained)	SSE	$N - K$	$\dfrac{SSE}{N - K} = \hat{\sigma}^2$
Total	SST	$N - 1$	

The overall F-statistic is the ratio of the two mean squares:

$$F = \frac{\text{Mean Square Regression}}{\text{Mean Square Residual}} = \frac{(SST - SSE)/(K - 1)}{SSE/(N - K)}$$

$\Longrightarrow$ The overall F-test asks whether the regression mean square is large relative to the residual mean square.

Example: Overall Significance of the Wage Model

For the wage model with $K = 4$, $N = 1000$:

$$SST = 109{,}520.00, \qquad SSE = 97{,}882.50$$

$$\begin{aligned} F &= \frac{(109{,}520.00 - 97{,}882.50)/3}{97{,}882.50/996} \\ &= \frac{11{,}637.50/3}{98.27} \\ &= \frac{3{,}879.17}{98.27} = 39.47 \end{aligned}$$

$F_c = F_{(0.95,\ 3,\ 996)} \approx 2.61$

$39.47 \gg 2.61 \implies$ **Reject** H_0. The model has significant explanatory power ($p \approx 0.000$).

$\implies$ This is the first thing to check in any regression output. If the overall F-test fails to reject, none of the regressors are jointly significant.

Outline

Beyond "Equals Zero" Restrictions

The F-test handles any linear equality hypothesis, not just "coefficient equals zero."

Example: In the Big Andy's Burger Barn model,

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

The marginal effect of advertising is $\beta_3 + 2\beta_4\,\text{ADVERT}$. Optimal advertising satisfies:

$$\text{marginal revenue} = 1 \implies \beta_3 + 2\beta_4\,\text{ADVERT}_0 = 1$$

To test whether \$1,900/month is optimal ($\text{ADVERT}_0 = 1.9$):

$$H_0: \beta_3 + 3.8\beta_4 = 1 \qquad H_1: \beta_3 + 3.8\beta_4 \neq 1$$

$\implies$ This is a single restriction ($J = 1$), but the restriction is **not** "equals zero." The F-test still works; you just need the SSE_R from the model that imposes $\beta_3 + 3.8\beta_4 = 1$.

Deriving the Restricted Model

To get SSE_R, substitute the restriction into the unrestricted model.

From $\beta_3 + 3.8\beta_4 = 1$, solve for β_3:

$$\beta_3 = 1 - 3.8\beta_4$$

Substitute into the original model:

$$\begin{aligned} \text{SALES} &= \beta_1 + \beta_2\,\text{PRICE} + (1 - 3.8\beta_4)\,\text{ADVERT} + \beta_4\,\text{ADVERT}^2 + e \\ &= \beta_1 + \beta_2\,\text{PRICE} + \text{ADVERT} + \beta_4(\text{ADVERT}^2 - 3.8\,\text{ADVERT}) + e \end{aligned}$$

Rearrange to get a standard regression:

$$\underbrace{\text{SALES} - \text{ADVERT}}_{\text{new LHS}} = \beta_1 + \beta_2\,\text{PRICE} + \beta_4\,\underbrace{(\text{ADVERT}^2 - 3.8\,\text{ADVERT})}_{\text{new variable}} + e$$

$\implies$ Run OLS on this transformed equation. Its SSE is SSE_R.

Example: Testing Optimal Advertising

Using the Big Andy data ($N = 75$, $K = 4$):

- $SSE_U = 1{,}532.084$ (unrestricted model)
- $SSE_R = 1{,}552.286$ (restricted model, imposing $\beta_3 + 3.8\beta_4 = 1$)
- $J = 1, \quad N - K = 71$

$$F = \frac{(1{,}552.286 - 1{,}532.084)/1}{1{,}532.084/71} = \frac{20.202}{21.578} = 0.94$$

$F_c = F_{(0.95,\,1,\,71)} = 3.976$

$0.94 < 3.976 \implies$ **Do not reject H_0.** The data are consistent with $1,900/month being the optimal advertising level.

Since $J = 1$, this is equivalent to a two-sided t-test: $t = \sqrt{0.94} = 0.97$, $t_c = 1.994$. Same conclusion.

Outline

Using Restrictions from Economic Theory

Sometimes economic theory provides parameter restrictions *before* looking at the data. If we trust these restrictions, we can **impose them during estimation** rather than test them.

Beer demand example:

$$\ln(Q) = \beta_1 + \beta_2 \ln(\text{PB}) + \beta_3 \ln(\text{PL}) + \beta_4 \ln(\text{PR}) + \beta_5 \ln(I) + e$$

Economic theory (no money illusion): if all prices and income double, quantity demanded should not change:

$$\beta_2 + \beta_3 + \beta_4 + \beta_5 = 0$$

Solve for $\beta_4 = -\beta_2 - \beta_3 - \beta_5$ and substitute:

$$\ln(Q) = \beta_1 + \beta_2 \ln\left(\frac{\text{PB}}{\text{PR}}\right) + \beta_3 \ln\left(\frac{\text{PL}}{\text{PR}}\right) + \beta_5 \ln\left(\frac{I}{\text{PR}}\right) + e$$

$\Longrightarrow$ OLS on this transformed model gives **restricted least squares** estimates.

Properties of Restricted Least Squares

Property	Restriction true	Restriction false
Bias	Unbiased	Biased
Variance	Lower than OLS	Lower than OLS

If the restriction is correct:

- Unbiased *and* lower variance $\implies$ unambiguously better than unrestricted OLS

If the restriction is wrong:

- Lower variance, but biased $\implies$ a tradeoff
- The bias can be substantial if the restriction is far from the truth

$\implies$ Use the F-test to check whether the restriction is supported by the data before relying on restricted estimates.

Outline

The F-Test Procedure: Summary

1. **State hypotheses.** Write H_0 (the restrictions) and H_1 (at least one restriction violated).
2. **Estimate both models.** Run the unrestricted regression (SSE_U) and the restricted regression (SSE_R).
3. **Compute F.**

$$F = \frac{(SSE_R - SSE_U)/J}{SSE_U/(N-K)}$$

4. **Find the critical value.** Look up $F_{(1-\alpha,\ J,\ N-K)}$ or use the p-value.
5. **Decide.** Reject H_0 if $F \geq F_c$ (or if $p < \alpha$).

$\Longrightarrow$ Always verify: $SSE_R \geq SSE_U$. If not, you have the models backward.

Decision Flowchart

When you see a group of variables and wonder whether to keep or drop them:

1. Run the full model (unrestricted). Record SSE_U.
2. Run the model **without** the variables in question (restricted). Record SSE_R.
3. Compute F and compare to $F_{(1-\alpha,\ J,\ N-K)}$.

Result	Action
Reject H_0	Keep the variables
Do not reject H_0	Variables can be dropped

$\Longrightarrow$ Do **not** rely on individual t-tests alone when variables are correlated. The F-test is the correct tool for groups of variables.

Comparison: The Testing Toolkit

Question	J	Test	Statistic
Is $\beta_k = 0$?	1	t-test	$t = \dfrac{b_k}{\text{se}(b_k)}$
Is $\beta_k = c$?	1	t-test	$t = \dfrac{b_k - c}{\text{se}(b_k)}$
Is $c_1\beta_1 + c_2\beta_2 = q$?	1	t-test	$t = \dfrac{\hat{\lambda} - q}{\text{se}(\hat{\lambda})}$
Are $\beta_3 = 0$ and $\beta_4 = 0$?	2	F-test	$F = \dfrac{(SSE_R - SSE_U)/J}{SSE_U/(N-K)}$
Are all slopes $= 0$?	$K-1$	F-test	$F = \dfrac{(SST - SSE)/(K-1)}{SSE/(N-K)}$

$\Longrightarrow$ The t-test is for single restrictions. The F-test is for one or more restrictions. When $J = 1$, $F = t^2$.

Outline

What You Should Take Away

1. **Individual significance $\neq$ joint significance.**
   Variables that are individually insignificant (via t-tests) can be jointly significant (via F-test).

2. **The F-test compares two models.**
   It measures whether imposing the restrictions in H_0 causes a statistically meaningful increase in SSE.

3. **Overall significance.**
   When H_0 sets all slopes to zero, $SSE_R = SST$. This is the overall F-test reported by software.

4. **Restricted least squares.**
   Imposing a restriction always reduces variance but introduces bias unless the restriction is correct.

5. **Practical rule:**
   Before dropping a group of correlated variables, run the F-test. Individual t-tests can mislead when regressors are correlated.

Thank you!

jakeanderson@g.ucla.edu