Financial Econometrics | Lecture 6

Joint Hypothesis Tests: the F-statistic for Correlated Regressors

In Class Group Activity

· SJTU, Summer School 2026 · Fatih Kansoy

Lesson of the day. When several regressors are correlated, you must test them together with the F-test, because a string of separate t-tests both inflates the false-alarm rate and can hide a block of factors that each look weak alone yet jointly carry the risk.

Phase 1 – Concept checks

ANSWERS

ConcepTest 1. Answer. Option **(b)**. With two independent 5% tests, the chance of at least one false alarm is $1-(0.95)^2 = 0.0975$, about 9.75%, not 5%. Option (a) is the size-distortion error; (c) applies a correction backwards; (d) adds the rates instead of using $1-0.95^q$.

The false-alarm rate grows with q: about 0.10 at $q=2$, 0.14 at $q=3$, 0.23 at $q=5$. A string of 5% tests is not a 5% test of the group; the F-test fixes the size.

ConcepTest 2. Answer. Option **(b)**. Because the factors are highly correlated, each loading is imprecise on its own (collinearity inflates each standard error), so each t is modest, yet the joint F-test can still reject. Option (a) is the "individual implies joint" error; (c) blames sample size when the issue is correlation; (d) over-generalises with "always".

ConcepTest 3. Answer. Option **(a)**. The equation $b_{\text{coups}} = b_{\text{assass}}$ is one linear restriction, so $q = 1$: test it with a single F (against 3.84) or, equivalently, a t on the difference. Option (b) counts coefficients, not restrictions; (c) uses the wrong statistic (a sum of t's, ignoring correlation); (d) denies the lincom and transform routes.

On the Growth data this gives $F \approx 5.17$, $p \approx 0.027$: the two effects differ.

Phase 2 – The desk problem

WORKED SOLUTIONS

(a) $t_{\text{value}} = \dfrac{0.182}{0.140} = 1.30$ and $t_{\text{qual}} = \dfrac{0.225}{0.150} = 1.50$. Both are below 1.96, so **neither factor is significant** at 5%. A naive analyst would drop both.

(b) The joint $F = 8.9$ exceeds the $q = 2$ critical values (3.00 at 5%, 4.61 at 1%), so **reject the joint null**: the two factors are jointly significant at the 1% level. The apparent contradiction with (a) is resolved by the correlation. The two factors correlate +0.78, so their coefficient estimates are strongly (negatively) correlated: collinearity inflates each individual standard error, making each loading imprecise *alone*, yet together they explain a real share of returns. The F-statistic uses the correlation between the estimates; the two separate t-tests throw that information away.

(c) Difference $= 0.182 - 0.225 = -0.043$; $t = \dfrac{-0.043}{0.27} = -0.16$. Since $|-0.16| < 1.96$, do **not reject** $H_0 : b_{\text{value}} = b_{\text{qual}}$: the data are consistent with the two loadings being equal. But notice this is a weak result, not a strong one: $SE_{\text{diff}} = 0.27$ is large because the two highly correlated factors are hard to tell apart, so the data simply cannot separate the loadings. What the data *can* pin down is the fund's total style exposure (the sum of the two), not how it splits between value and quality.

(d) Homoskedasticity-only F from the R^2 form:

$$F = \frac{(0.118 - 0.010)/2}{(1 - 0.118)/(120 - 2 - 1)} = \frac{0.054}{0.882/117} = \frac{0.054}{0.00754} = 7.16.$$

This differs from the robust $F = 8.9$ because the data are heteroskedastic. Which is larger depends on the data: on the California test-score example the robust F is 5.43 while the homoskedastic one is 8.01. Report the **robust** F, since financial returns cluster in volatility.

Verdict: keep both factors. They are jointly significant though neither is individually significant, and their loadings cannot be told apart.

Phase 3 – The case

REVEAL AND ANSWERS

---

1. $t_{\text{level}} = \dfrac{-0.46}{0.34} = -1.35$ and $t_{\text{slope}} = \dfrac{0.58}{0.39} = 1.49$. Both are below 1.96 in absolute value, so the analyst's individual reading is correct: neither factor is significant alone.

2. The joint $F = 6.40$ exceeds the $q = 2$ critical values (3.00 at 5%, 4.61 at 1%), so the two factors are jointly significant. The proposal to drop both, and to report no interest-rate exposure, is **wrong**.

3. The two factors correlate $+0.85$, so the data cannot pin down LEVEL and SLOPE separately (they move together). Each standard error is large, so each t is small, yet the pair explains real return variation.

4. If the team drops both factors, the reported interest-rate exposure becomes **zero**, which understates the real risk. The fund *is* exposed to rates; the model simply cannot split the exposure cleanly between the two correlated factors. Dropping the block hides a genuine risk from the regulator.

Instructor reveal. The head of risk does not sign off. The joint $F = 6.40$ rejects "no interest-rate exposure" at the 1% level, so the fund clearly has rate risk even though no single factor is individually significant. The fix is not to drop the factors but to report the joint exposure, for example by combining LEVEL and SLOPE into one duration-risk measure, or by reporting the joint F alongside the (wide) individual intervals so the regulator sees that the imprecision is about *attribution*, not about whether the risk exists. Six months later a sharp curve move costs the fund 7% in a quarter, exactly the exposure the "parsimonious" model would have denied. A block of correlated risk factors can each look insignificant and still carry the risk; the joint F is what protects you.

Reserve alternate. Had the joint F come back at 1.8 (below 3.00), not rejecting would have been correct and dropping the block defensible. The difference between the two endings is the one statistic the analyst ignored: let the joint test rule, not a prior belief.

End of worked solutions.