FINANCIAL ECONOMETRICS 2026
SHANGHAI JIAO TONG UNIVERSITY
SUMMER SCHOOL 2026
QUIZ 2

Reading time: 10 minutes. **Writing time:** 60 minutes.
Total marks: 100. **Calculators are permitted.**
Instructions: Answer **all** questions. In Section A, circle the letter of the single best answer. In Sections B and C, show your working and write in clear sentences. All standard errors are heteroskedasticity-robust.

Structure:

- **Section A** — Multiple choice: 20 questions, 2.5 marks each (**50 marks**).
- **Section B** — One analytical problem in five parts (**30 marks**).
- **Section C** — One short interpretation (**20 marks**).

Useful facts. $t = \dfrac{\text{estimate} - \text{claimed value}}{SE}$; 95% CI = estimate $\pm 1.96\,SE$ (90%: 1.64; 99%: 2.58); small-sample two-sided 5% t: 2.23 ($df = 10$), 2.09 ($df = 20$), 2.00 ($df = 60$), $\to 1.96$ as df grows; sign of omitted-variable bias = sign$\big[\text{corr}(X, Z) \times \text{effect of } Z \text{ on } Y\big]$; $\bar{R}^2 = 1 - \frac{n-1}{n-k-1}\frac{SSR}{TSS}$; 5% critical values of F: 3.84 ($q = 1$), 3.00 ($q = 2$), 2.60 ($q = 3$); 1%, $q = 2$: 4.61; $F = t^2$ for one restriction.

Section A: Multiple Choice (20 questions × 2.5 marks = 50 marks)

1. The standard error of an estimated regression slope measures:

(A) how much the estimate would vary across repeated samples.

(B) the spread of the outcome Y around its sample mean.

(C) the typical size of the residuals around the fitted line.

(D) the share of the variation in Y the line explains.

2. A regression gives a slope of 1.5 with standard error 0.6. Testing $H_0 : \beta_1 = 0$ at the 5% level:

(A) $t = 0.4$; do not reject.

(B) $t = 2.5$; reject.

(C) $t = 2.5$; do not reject.

(D) $t = 0.9$; reject.

3. A slope is estimated as -2.0 with standard error 0.5. The 95% confidence interval is:

(A) $[-2.50, \ -1.50]$.

(B) $[-2.98, \ -1.02]$.

(C) $[-2.82, \ -1.18]$.

(D) $[-3.29, \ -0.71]$.

4. In the regression $Y_i = \beta_0 + \beta_1 D_i + u_i$, where D_i is 1 for firms with a female CEO and 0 otherwise, β_0 is:

(A) the average of Y among firms with a female CEO.

(B) the gap in average Y between the two groups.

(C) the average of Y among firms without a female CEO.

(D) the sample share of firms with a female CEO.

5. *Heteroskedasticity* means that:

(A) the errors average to zero at every X.

(B) the values of X are all very similar.

(C) the residuals sum to zero by construction.

(D) the spread of the errors changes with X.

6. The rule taught in this course about robust (heteroskedasticity-robust) standard errors is that they:

(A) are always larger than the homoskedasticity-only ones.

(B) are valid whether or not the error spread is constant.

(C) are needed only when the sample size is very small.

(D) are the default output printed by most software.

7. With only $n = 12$ observations, the correct two-sided 5% cutoff for a t-statistic is closest to:

(A) 1.64.

(B) 1.96.

(C) 2.00.

(D) 2.23.

8. Leaving a variable Z out of a regression of Y on X biases the slope on X when:

(A) Z affects Y, whether or not it is related to X.

(B) Z is correlated with X, whether or not it affects Y.

(C) either one of the two conditions holds.

(D) Z affects Y and Z is correlated with X.

9. An omitted variable Z *raises* Y, and Z is *positively* correlated with X. The estimated slope on X in the regression without Z is biased:

(A) downward.

(B) not at all.

(C) upward.

(D) towards exactly zero.

10. If a regression suffers from omitted-variable bias, doubling the sample size:

(A) removes the bias slowly.

(B) removes it once n is large enough.

(C) cuts the bias roughly in half.

(D) does not reduce the bias at all.

11. In the multiple regression $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$, the coefficient β_1 measures the effect on Y of a one-unit change in X_1:

(A) holding X_2 fixed.

(B) and the accompanying change in X_2.

(C) averaged over all values of X_2 and u.

(D) ignoring the presence of X_2.

12. Adding a regressor that is pure noise to a regression will:

(A) lower R^2, while possibly raising $\bar{R}^2$.

(B) lower both R^2 and $\bar{R}^2$ at the same time.

(C) never lower R^2, and possibly lower $\bar{R}^2$.

(D) raise both R^2 and $\bar{R}^2$ in every case.

13. A regression includes an intercept and dummies for *every* sector (technology, financials, other). The standard fix is to:

(A) add one further dummy variable for large firms.

(B) drop one sector dummy, which becomes the baseline.

(C) drop the dependent variable from the equation.

(D) estimate the equation without any sector dummies.

14. Two regressors are highly, but not perfectly, correlated (say 0.95). The consequence is:

(A) unbiased estimates with large standard errors.

(B) biased estimates with small standard errors.

(C) that the software refuses to run the regression.

(D) that the residuals no longer sum to zero.

15. A null hypothesis restricts two coefficients to zero ($q = 2$). If you instead run two separate t-tests at 5% and reject when either is significant, the true false-alarm rate is about:

(A) 5%.

(B) 2.5%.

(C) 10%.

(D) 25%.

16. A robust F-statistic for a joint null with $q = 2$ restrictions comes out at 4.2. The conclusion is:

(A) reject at 1% and at 5%.

(B) reject at neither level.

(C) reject at 1% but not at 5%.

(D) reject at 5% but not at 1%.

17. For a single restriction ($q = 1$), the F-statistic and the t-statistic for the same null are related by:

(A) $F = t$.

(B) $F = 2t$.

(C) $F = t^2$.

(D) $F = t/2$.

18. In a regression with two correlated regressors, each t-statistic is small, yet the joint F-test rejects strongly. The explanation is:

(A) the regressors are correlated and share the credit.

(B) the sample is far too small for reliable t-tests.

(C) a coding error somewhere in the estimation software.

(D) heteroskedasticity in the regression's error terms.

19. A very high R^2 in a regression is, by itself, evidence that:

(A) none of the following.

(B) the included variables cause Y.

(C) there is no omitted-variable bias.

(D) the right variables have been chosen.

20. In a carefully designed regression, the coefficient on a *control* variable (included only to guard the variable of interest):

(A) is the causal effect of the control.

(B) is not given a causal reading.

(C) must exceed 1.96 to keep the control.

(D) equals the bias it removes.

Section B: Analytical Problem (30 marks)

A fund's monthly excess return R (in % per month) is regressed on the market factor MKT and the size factor SMB, using $n = 120$ months of data (robust standard errors in parentheses):

$$\widehat{R} = \underset{(0.10)}{0.20} + \underset{(0.05)}{1.10}\,MKT + \underset{(0.20)}{0.50}\,SMB, \qquad R^2 = 0.62, \quad \bar{R}^2 = 0.61.$$

(a) **[6 marks]** Interpret the coefficient 1.10 on MKT in one sentence, and test whether it differs from zero at the 5% level.

(b) **[6 marks]** The intercept (0.20) is the fund's alpha. Test $H_0 : \alpha = 0$ at the 5% level and give the 95% confidence interval for alpha. What does your conclusion say about the manager?

(c) **[6 marks]** Test whether the fund's market exposure differs from exactly 1, that is $H_0 : \beta_{MKT} = 1$, at the 5% level.

(d) **[6 marks]** A one-factor regression of the same fund on MKT alone gives an alpha of 0.35. Using the omitted-variable-bias sign rule, explain why the one-factor alpha is larger than the three lines above suggest it should be. (The fund's returns load positively on SMB, and assume the fund's SMB exposure is positively related to its market exposure across months.)

(e) **[6 marks]** A sceptic claims that neither factor matters: $H_0 : \beta_{MKT} = 0$ **and** $\beta_{SMB} = 0$. The software reports a robust F-statistic of 8.2 for this joint null. State the conclusion at the 5% and 1% levels, and explain in one sentence why this joint test, rather than the two t-tests, is the right tool.

Section C: Interpretation (20 marks)

Read the following and answer in **no more than five sentences**. Use the concepts from the course; you do not need any calculation.

An analyst regresses the annual stock returns of 300 firms on each firm's ESG score (a measure of environmental and social practices) and finds a positive, statistically significant slope. She concludes that good ESG practices raise returns and recommends investing accordingly. A colleague repeats the exercise with one change: he adds firm size to the regression, since large firms tend to have both high ESG scores and returns that behave differently. In his regression the ESG coefficient falls by more than half and is no longer statistically significant.

Question [20 marks]. Explain what happened to the ESG coefficient and why, name the problem in the analyst's original regression, and state what her one-variable slope was actually measuring. Answer in at most five sentences.