Financial Econometrics 2026
Shanghai Jiao Tong University
Summer School 2026
Midterm Examination
With Solutions

Reading time: 10 minutes. **Writing time:** 120 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. $SE(\bar{Y}) = s_Y/\sqrt{n}$; $t = \dfrac{\text{estimate} - \text{claimed value}}{SE}$; 95% CI = estimate $\pm 1.96\,SE$ (90%: 1.64; 99%: 2.58); $\widehat{\beta}_1 = s_{XY}/s_X^2$; $r = s_{XY}/(s_X s_Y)$; with one regressor $R^2 = r^2$; binary regressor: $\beta_0 = \mathbb{E}[Y \mid D{=}0]$, $\beta_1 = \mathbb{E}[Y \mid D{=}1] - \mathbb{E}[Y \mid D{=}0]$; sign of omitted-variable bias = $\text{sign}\big[\text{corr}(X, Z) \times \text{effect of } Z \text{ on } Y\big]$; $\bar{R}^2 = 1 - \frac{n-1}{n-k-1}\frac{SSR}{TSS}$; 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 $\times$ 2.5 marks = 50 marks)

1. A researcher observes the quarterly return on equity of 35 commercial banks, each bank observed over the same 8 quarters. This data set is:

(A) panel data.

(B) cross-sectional data.

(C) time-series data.

(D) experimental data.

Solution. (A). Many entities, each tracked over several periods: a panel, with $35 \times 8 = 280$ observations. One entity over time is a time series; many entities at one date is a cross-section. The data are observational either way; nothing was randomly assigned (D).

2. A one-year investment returns +8% with probability 0.75 and −4% with probability 0.25. Its expected return is:

(A) 5%.

(B) 6%.

(C) 4%.

(D) 2%.

Solution. (A). $\mathbb{E}[R] = 0.75 \times 8 + 0.25 \times (-4) = 6 - 1 = 5\%$: the probability-weighted average of the outcomes. It is the long-run average per year over many repetitions, not a value that occurs in any single year.

3. Two returns have covariance 6; their standard deviations are 2 and 5. Their correlation is:

(A) 0.30.

(B) 1.20.

(C) 0.24.

(D) 0.60.

Solution. (D). $r = \dfrac{s_{XY}}{s_X s_Y} = \dfrac{6}{2 \times 5} = 0.60$. Dividing by both standard deviations strips the units and confines the number to $[-1, 1]$. Option (B) is impossible for a correlation, which is one quick way to reject it.

4. Daily returns are known to be non-normal, with fat tails. The distribution of the *average* of 250 such returns is nevertheless approximately normal because of:

(A) the law of large numbers.

(B) the central limit theorem.

(C) the assumption of homoskedasticity.

(D) the Gauss-Markov theorem.

Solution. (B). The central limit theorem: whatever the shape of the underlying data, the standardised sample average approaches the standard normal as n grows. The law of large numbers (A) is a different statement, that the average settles near the truth; it says nothing about the shape.

5. A sample of $n = 100$ months has standard deviation $s_Y = 4\%$. The standard error of the sample mean return is:

(A) 0.4%.

(B) 4%.

(C) 0.04%.

(D) 2%.

Solution. (A). $SE(\bar{Y}) = s_Y/\sqrt{n} = 4/\sqrt{100} = 4/10 = 0.4\%$. The data's spread is 4%; the estimate's precision is ten times finer, because averaging cancels noise at the rate $\sqrt{n}$.

6. A test of $H_0 : \mu = 0$ produces a p-value of 0.02. The correct reading is:

(A) there is only a 2% chance that the null hypothesis is true.

(B) under the null, data this extreme occur about 2% of the time.

(C) the estimated mean return equals 2% per month.

(D) the testing procedure has a 2% chance of being wrong.

Solution. (B). The p-value is the probability of the data given the null, not the probability of the null given the data. Since $0.02 < 0.05$, a 5% test rejects. Option (A) is the classic misreading, and (D) confuses the p-value with the significance level chosen in advance.

7. In a sample, the covariance between X and Y is 8 and the variance of X is 5. The OLS slope of Y on X is:

(A) 0.63.

(B) 2.29.

(C) 1.60.

(D) 40.0.

Solution. (C). $\widehat{\beta}_1 = s_{XY}/s_X^2 = 8/5 = 1.60$: how much X and Y move together, per unit of how much X moves on its own. Option (A) inverts the ratio; option (D) multiplies instead of dividing.

8. A single-regressor regression reports $R^2 = 0.09$. This means:

(A) the slope is statistically insignificant.

(B) the regression is misspecified.

(C) 9% of the slope is due to chance.

(D) X explains 9% of the variation in Y.

Solution. (D). R^2 is the share of the sample variation in Y that the fitted line explains, here 9%; the other 91% is everything else. It carries no verdict on significance (A), on specification (B), or on causality; a slope can be precisely estimated and still explain little.

9. A slope is estimated as 0.9 with standard error 0.45. At the 5% level, the test of $H_0 : \beta_1 = 0$:

(A) fails to reject, since $t = 0.5$.

(B) rejects, since $t = 2.0$.

(C) rejects, since $t = 4.5$.

(D) fails to reject, since $t = 2.0$.

Solution. (B). $t = 0.9/0.45 = 2.0 > 1.96$: reject, though only just. A result this close to the boundary is statistically significant by convention but deserves the caution that a slightly different sample could have flipped the verdict; reporting the confidence interval is more informative than the bare verdict.

10. A coefficient is estimated as 3.0 with standard error 1.0. Which claimed value is *rejected* by a two-sided 5% test?

(A) 5.0.

(B) 2.0.

(C) 4.0.

(D) 1.5.

Solution. (A). The 95% confidence interval is $3.0 \pm 1.96 = [1.04,\ 4.96]$: exactly the set of values a 5% test would not reject. Of the four candidates only 5.0 lies outside the interval. Equivalently, $t = (3.0 - 5.0)/1.0 = -2.0$, and $|-2.0| > 1.96$.

11. In $\widehat{Y} = 12 + 3D$, where $D = 1$ for firms paying a dividend and 0 otherwise, the average of Y among dividend payers is:

(A) 12.

(B) 3.

(C) 15.

(D) 9.

Solution. (C). With a binary regressor, $\beta_0 = 12$ is the average for the $D = 0$ group and $\beta_1 = 3$ is the gap, so the $D = 1$ group averages $12 + 3 = 15$. The regression is simply two group means written as one equation.

12. The reason this course always uses heteroskedasticity-robust standard errors is that they are:

(A) smaller than the homoskedasticity-only ones.

(B) exact in small as well as large samples.

(C) required for the slope to be unbiased.

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

Solution. (D). Robust standard errors are correct with or without a constant error spread and agree with the simple formula when the spread is constant, so nothing is lost. They are not systematically smaller (A), and unbiasedness of the slope (C) never depended on the error spread; it is the *standard errors*, hence the tests, that go wrong without them.

13. Omitting a variable Z from a regression of Y on X biases the slope on X when:

(A) Z affects Y, even if unrelated to X.

(B) Z is correlated with X, even if it does not affect Y.

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

(D) Z is measured with error.

Solution. (C). Both conditions must hold at once: the omitted variable must matter for Y (so it hides in the error) and travel with X (so the error moves with the regressor). One condition alone produces noise or nothing; when both hold, the slope on X picks up part of Z's effect, a systematic bias that no sample size can repair.

14. An omitted variable Z *lowers* Y, and Z is *positively* correlated with X. The slope on X estimated without Z is biased:

(A) upward.

(B) not at all.

(C) downward.

(D) towards $+1$.

Solution. (C). Sign of bias $=$ sign[corr$(X, Z) \times$ effect of Z on $Y] = (+) \times (-) = -$: downward. Whenever X is high, the harmful Z tends to be present too, and X is blamed for damage that Z caused. This is exactly the California pattern: districts with large classes also have more English learners, so the lone class-size slope was too negative.

15. In $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + u_i$, the coefficient β_2 is read as the effect of a one-unit change in X_2:

(A) together with the usual change in X_1.

(B) for the average firm only.

(C) ignoring X_1 entirely.

(D) holding X_1 fixed.

Solution. (D). The partial effect: each coefficient in a multiple regression measures its own variable's contribution with the other included variables held fixed. That "holding fixed" is precisely what removes omitted-variable bias from single-regressor slopes: with the other variable in the equation, the coefficient no longer picks up its effect.

16. Compared with R^2, the adjusted $\bar{R}^2$:

(A) is always at least as large as R^2.

(B) charges a penalty for each added regressor.

(C) cannot decrease when a variable is added.

(D) measures the model's out-of-sample fit.

Solution. (B). The factor $(n-1)/(n-k-1)$ is a price per variable: $\bar{R}^2$ rises only if a new regressor improves fit by more than its penalty, can fall, and can even go negative. Neither measure speaks to out-of-sample performance (D); both describe this sample only.

17. A regression includes market capitalisation in millions *and* the same capitalisation in billions as two separate regressors. The software will:

(A) fail to run, or drop one of the two.

(B) report identical coefficients on both.

(C) run normally with no warning.

(D) double the R^2.

Solution. (A). One variable is an exact multiple of the other: perfect multicollinearity. The question "what happens when one moves, holding the other fixed?" has no answer, the formulas divide by zero, and the software refuses or silently drops a column. The cure is to fix the variable list; this is a definition mistake, not a data feature.

18. To test the joint null that three factor loadings are all zero, the correct procedure is:

(A) three t-tests, rejecting if any exceeds 1.96.

(B) one t-test on the sum of the coefficients.

(C) three t-tests at the 1% level each.

(D) one F-test with $q = 3$ restrictions.

Solution. (D). A joint claim is judged jointly. One-at-a-time t-tests have the wrong size (the false-alarm rate compounds well past the promised 5%), and with correlated regressors they can miss a contribution the group makes together. The F-statistic combines the evidence with the correlations handled correctly; at $q = 3$ its 5% cutoff is 2.60.

19. A robust F-statistic of 5.0 is reported for a joint null with $q = 3$ restrictions. At the 5% level:

(A) do not reject, since the cutoff is 5.99.

(B) do not reject, since the cutoff is 7.81.

(C) reject, since the cutoff is 2.60.

(D) reject, since $5.0 > 1.96$.

Solution. (C). The 5% large-sample critical value for $q = 3$ is 2.60, and $5.0 > 2.60$: reject. The cutoff depends on the number of restrictions and *falls* as q rises (3.84, 3.00, 2.60, …), because F averages the evidence across restrictions. The values 5.99 and 7.81 in (A) and (B) come from reading the wrong row or the wrong table of critical values; with the correct cutoff of 2.60, the conclusion is to reject. Comparing an F to 1.96 (D) mixes up the two tests' scales.

20. Across several specifications with different sensible controls, a key coefficient stays close to the same value. This stability is evidence that:

(A) the sample is large enough for reliable inference.

(B) the coefficient is not absorbing what the controls capture.

(C) the regression errors are homoskedastic after all.

(D) the R^2 is computed correctly in every column.

Solution. (B). If a coefficient barely moves as controls are added, it is apparently not absorbing what those controls capture; a coefficient that jumps around as controls change is still picking up the effects of other variables. This is why professionals report a specification table and watch the key coefficient across columns; stability, not fit, is the mark of a trustworthy estimate.

Section B: Analytical Problem (30 marks)

Using a sample of $n = 1{,}000$ workers, hourly wages (in \$) are regressed on a university-degree dummy ($College = 1$ if the worker holds a degree), first alone and then with years of work experience added. Robust standard errors in parentheses:

$$\textbf{Regression 1:}\quad \widehat{Wage} = 11.00 + \underset{(0.50)}{2.00}\,College$$

$$\textbf{Regression 2:}\quad \widehat{Wage} = 8.00 + \underset{(0.40)}{2.50}\,College + \underset{(0.04)}{0.20}\,Exper$$

In this sample, degree holders have *fewer* years of experience on average (they started work later), and experience raises wages.

(a) **[6 marks]** In Regression 1, interpret the two coefficients, test whether the college gap differs from zero at the 5% level, and give its 95% confidence interval.

(b) **[6 marks]** In Regression 2, interpret the coefficient 2.50 precisely, and compute its t-statistic.

(c) **[6 marks]** The college coefficient *rose* from 2.00 to 2.50 when experience was added. Using the omitted-variable-bias sign rule and the facts stated above, explain exactly why Regression 1 understated the college gap.

(d) **[6 marks]** An economist claims each year of experience is worth \$0.25 per hour. Test $H_0 : \beta_{Exper} = 0.25$ at the 5% level using Regression 2.

(e) **[6 marks]** The software reports a robust F-statistic of 24.0 for the joint null $H_0 : \beta_{College} = 0$ and $\beta_{Exper} = 0$. State the conclusion at the 1% level, and explain in one sentence what this joint test asks that the two t-tests do not.

Solution. (a) $\beta_0 = 11.00$: the average wage of workers *without* a degree is \$11.00 per hour. $\beta_1 = 2.00$: degree holders earn on average \$2.00 more per hour, a gap of two group means. Test: $t = 2.00/0.50 = \mathbf{4.0} > 1.96$, reject H_0: the gap is statistically significant. CI: $2.00 \pm 1.96 \times 0.50 = [\mathbf{1.02},\ \mathbf{2.98}]$.

(b) Holding years of experience fixed, degree holders earn on average \$**2.50** more per hour than non-holders: the partial effect of the degree, comparing workers at the *same* experience level. $t = 2.50/0.40 = \mathbf{6.25}$, overwhelmingly significant.

(c) The omitted variable in Regression 1 is experience: it **affects wages** (positively) and is **correlated with** $College$ (negatively, since graduates start work later), so both bias conditions hold. Sign of bias = $\text{sign}[(-) \times (+)] = -$: Regression 1's coefficient is biased *downward*. Intuitively, the raw comparison mixes the degree's positive effect with the experience graduates have not yet accumulated: the coefficient on $College$ picks up part of the negative effect of the missing experience. Holding experience fixed removes this bias, and the gap is larger: 2.50, not 2.00.

(d) $t = \dfrac{0.20 - 0.25}{0.04} = \dfrac{-0.05}{0.04} = \mathbf{-1.25}$. Since $|-1.25| < 1.96$, **do not reject**: the data are consistent with \$0.25, and also with values near \$0.20; failing to reject does not prove the claim, it only leaves it inside the plausible range ($0.20 \pm 1.96 \times 0.04 = [0.122,\ 0.278]$).

(e) For $q = 2$ the 1% critical value is 4.61, and $24.0 > 4.61$: **reject decisively**. The joint test asks whether the two variables *together* explain wages, judged with a single controlled false-alarm rate, which is a different question from whether each variable is individually distinguishable from zero; with correlated regressors the two answers can genuinely differ.

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.

Two research teams study whether holding cash reserves protects firms' stock returns during market crises, using the same data on 2,000 firms. Team A regresses crisis returns on cash holdings alone: the coefficient is large, positive and highly significant, and they report it as the effect of cash, noting their R^2 is high for this kind of data. Team B first lists the factors that plausibly drive crisis returns and also move with cash holdings (firm size, debt levels, industry), includes them as controls, and reports the cash coefficient across several such specifications: it is smaller than Team A's, but similar in every column. The two teams ask you to judge whose estimate of the effect of cash is more credible.

Question [20 marks]. Which team's estimate deserves more trust, and why? Your answer should say what is wrong with Team A's reasoning, what Team B did right, and what role the R^2 and the stability across specifications each play. Answer in at most five sentences.

Solution. Model answer (design and stability beat fit). Team B's estimate deserves more trust. Team A's single-regressor coefficient is exposed to **omitted-variable bias**: size, leverage and industry plausibly affect crisis returns *and* are correlated with cash holdings, so their lone slope picks up part of those variables' effects along with the cash effect, and neither the large t-statistic nor the high R^2 protects against that, since R^2 measures in-sample fit and says nothing about significance, causation, omitted variables, or having the right regressors. Team B did what a credible design requires: they chose controls by *listing the confounders* of the variable of interest, the factors that plausibly affect crisis returns and also move with cash holdings, rather than by chasing fit. The near-constancy of their cash coefficient across specifications is then the key evidence: a coefficient that holds steady as controls are added is apparently not absorbing what the controls capture, while one that jumps around signals possible remaining bias. Team B's smaller, stable coefficient is therefore the more believable measure of what cash itself contributes.

Marking guide (20 marks).

- **8 marks** – diagnoses Team A: omitted-variable bias (confounders that affect returns and correlate with cash), and states that a high R^2/large t does not defend against it.
- **6 marks** – credits Team B's method: controls chosen by listing the confounders of the variable of interest (Key Concept 7.3), not by chasing fit.
- **4 marks** – explains the role of **stability across specifications** as the mark of a trustworthy coefficient.
- **2 marks** – clear writing within the five-sentence limit.

Answers that pick Team A because of the higher R^2 earn very little: that is precisely the fallacy the question tests.