Financial Econometrics 2026
Shanghai Jiao Tong University
Summer School 2026
Quiz 3
With Solutions

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. Standard errors are as stated in each question.

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 $\pm\, 1.96\, SE$; annualised quarterly growth = $400\,\Delta\ln(\text{level})$; first-difference slope = $\dfrac{\sum \Delta X\,\Delta Y}{\sum(\Delta X)^2}$; sign of omitted-variable bias = $\operatorname{sign}\big[\operatorname{corr}(X,Z) \times \text{effect of } Z \text{ on } Y\big]$; AR(1) long-run mean = $\dfrac{\beta_0}{1-\beta_1}$; one-step forecast $\widehat{Y}_{T+1|T} = \widehat{\beta}_0 + \widehat{\beta}_1 Y_T$; forecast error = actual − forecast; 95% forecast interval = forecast $\pm\, 1.96$ RMSFE, with RMSFE $\approx$ SER in large samples; $\text{BIC}(p) = \ln\big(SSR(p)/T\big) + (p+1)\frac{\ln T}{T}$, and AIC replaces $\ln T$ with 2; 5% critical values of F: 3.84 ($q = 1$), 3.00 ($q = 2$).

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

1. In a panel regression, entity fixed effects control for:

(A) omitted factors that vary over time within each entity.

(B) omitted factors that are constant over time within each entity.

(C) omitted shocks that hit every entity in the same period.

(D) measurement error in the regressor of interest.

Solution. (B). An entity fixed effect is one private intercept per entity, so it absorbs everything about the entity that does not move over time (culture, geography, a firm's permanent business model), observed or not. Option (C) describes *time* fixed effects; time-varying confounders (A) survive.

2. A researcher estimates entity fixed effects for a panel of 300 firms by the dummy-variable method, keeping the intercept in the regression. To avoid perfect multicollinearity she must include:

(A) 300 firm dummies.

(B) 1 firm dummy.

(C) 301 firm dummies.

(D) 299 firm dummies.

Solution. (D). With the intercept kept, one firm must be dropped as the baseline: $300 - 1 = 299$ dummies. Including all 300 (A) is the dummy trap, since the dummies would sum to the intercept's constant column.

3. Two banks are observed in 2015 and 2020. Bank North: leverage X goes $10 \to 12$, ROE Y goes $12 \to 10$. Bank South: X goes $16 \to 18$, Y goes $16 \to 13$. The first-difference (within) slope estimate is:

(A) -1.25.

(B) $+1.25$.

(C) -0.63.

(D) -2.50.

Solution. (A). $\Delta X = 2,\ \Delta Y = -2$ (North) and $\Delta X = 2,\ \Delta Y = -3$ (South), so the slope is $\frac{2(-2)+2(-3)}{2^2+2^2} = \frac{-10}{8} = -1.25$. Option (D) sums the two per-bank slopes $(-1) + (-1.5)$; (C) divides by $(\sum \Delta X)^2 = 16$; (B) drops the sign.

4. In the two-way model $Y_{it} = \beta_1 X_{it} + \alpha_i + \lambda_t + u_{it}$, the time effects λ_t absorb:

(A) shocks in period t that are common to all entities, such as a market-wide movement.

(B) each entity's permanent characteristics, such as its home country.

(C) entity-specific factors that change from period to period.

(D) serial correlation in the regression's error terms.

Solution. (A). λ_t is one intercept per period, common to all entities: it soaks up whatever hit everyone that period (in finance, the month's market move). Permanent traits (B) are the job of α_i; entity-specific time-varying factors (C) survive both; serial correlation (D) is handled by clustering, not by intercepts.

5. A fixed-effects slope is -0.60. The ordinary standard error is 0.20; the standard error clustered by entity is 0.40. Using the appropriate standard error, the t-statistic and the 5% verdict are:

(A) $t = -3.0$; reject $H_0 : \beta_1 = 0$.

(B) $t = -1.5$; reject $H_0 : \beta_1 = 0$.

(C) $t = -1.5$; do not reject $H_0 : \beta_1 = 0$.

(D) $t = -3.0$; do not reject $H_0 : \beta_1 = 0$.

Solution. (C). In a panel the valid standard error is the clustered one, because an entity's errors are serially correlated: $t = -0.60/0.40 = -1.5$, and $|-1.5| < 1.96$, so do not reject. The naive $t = -3.0$ (A) overstates precision by treating each year as independent information.

6. Even with entity *and* time fixed effects, a panel regression still fails to control for:

(A) permanent differences in culture across entities.

(B) a common shock that hits all entities in one period.

(C) fixed geographic traits of each entity.

(D) a confounder that varies both across entities and over time.

Solution. (D). Entity effects remove what is fixed within an entity (A, C); time effects remove what is common to all entities in a period (B). A confounder that moves in both dimensions, such as a local business cycle, survives both sets of intercepts.

7. Across all states and years, the slope of the traffic fatality rate on the beer tax is $+0.36$; with state fixed effects it is -0.66. By the omitted-variable-bias sign rule, this pattern is consistent with an omitted *fixed* state trait that:

(A) raises road deaths and is negatively correlated with the beer tax.

(B) raises road deaths and is positively correlated with the beer tax.

(C) raises road deaths but is uncorrelated with the beer tax.

(D) lowers road deaths and is positively correlated with the beer tax.

Solution. (B). The cross-state slope sits *above* the within-state slope, so the bias is upward: $\text{sign}[\text{corr}(X, Z) \times \text{effect of } Z] = (+) \times (+) = +$. A strong drinking culture raises deaths and pushes lawmakers towards higher alcohol taxes. Options (A) and (D) give a negative product (downward bias); (C) gives no bias at all.

8. A firm's regressor X takes the values 4, 6, 8 over three years. After entity demeaning (the within transformation), the firm's transformed values are:

(A) $4,\ 6,\ 8.$

(B) $0,\ 2,\ 4.$

(C) $-2,\ 0,\ 2.$

(D) $-4,\ -2,\ 0.$

Solution. (C). Demeaning subtracts the firm's own time average, $\bar{X}_i = (4+6+8)/3 = 6$: the deviations are $-2,\ 0,\ 2$. Option (B) subtracts the first value and (D) subtracts the last; (A) leaves the data untransformed.

9. The reason standard errors must be clustered by entity in a panel is that:

(A) the slope estimate is biased unless the errors are independent.

(B) the errors are correlated across different entities in the same period.

(C) clustering shrinks the standard errors and so raises power.

(D) an entity's errors are serially correlated, so ordinary standard errors are too small.

Solution. (D). Within one entity the errors repeat themselves over time, so its T observations carry less information than T independent facts; the ordinary formula ignores this and understates the uncertainty. The slope itself stays unbiased (A is wrong), and clustered standard errors are typically *larger*, not smaller (C).

10. In a firm fixed-effects regression, the coefficient on a regressor that never changes within any firm (for example, the firm's country of incorporation):

(A) is estimated more precisely, because the variable is stable.

(B) cannot be estimated, because demeaning removes the variable entirely.

(C) is numerically identical to the pooled OLS coefficient.

(D) is estimated consistently provided the standard errors are clustered.

Solution. (B). Fixed effects learn only from within-entity movement. A variable with no within variation equals its own entity mean, so demeaning turns it into a column of zeros: there is nothing left to estimate. No choice of standard error (D) can revive it.

11. Quarterly real GDP rises from 15,382 to 15,428 (billions). The growth rate at an annual rate, $400\,\Delta\ln(GDP)$, is closest to:

(A) 1.20%.

(B) 0.30%.

(C) 0.075%.

(D) 4.80%.

Solution. (A). $\Delta\ln(GDP) = \ln(15{,}428) - \ln(15{,}382) = 0.0030$, so $400 \times 0.0030 = 1.20\%$; equivalently, the quarterly growth rate $100 \times (15{,}428 - 15{,}382)/15{,}382 = 0.30\%$, annualised $0.30 \times 4 = 1.20\%$. Option (B) forgets to annualise, (C) divides by four instead, and (D) annualises twice.

12. The first autocorrelation of quarterly GDP growth is 0.34. This means that:

(A) 34% of the variance of growth is explained by the previous quarter.

(B) growth this quarter is expected to be 0.34 points higher than last quarter.

(C) above-average growth quarters tend to be followed by above-average growth.

(D) the series is nonstationary, because its memory has not died out at lag one.

Solution. (C). A positive autocorrelation is persistence: high values tend to follow high values, and the link fades at longer lags. The explained variance share would be roughly $\rho_1^2 \approx 0.12$, not 0.34 (A); autocorrelation is a correlation, not an increment (B); and a stationary series can perfectly well have stable, positive memory (D).

13. An AR(1) for GDP growth is $\widehat{Y}_t = 1.991 + 0.344\,Y_{t-1}$. The last observed value is $Y_T = 0.15$, and the next quarter's actual outcome is 1.1. The forecast error is closest to:

(A) $+0.9$.

(B) $+1.1$.

(C) -0.9.

(D) $+2.0$.

Solution. (C). Forecast: $1.991 + 0.344 \times 0.15 \approx 2.0$. Error = actual − forecast = $1.1 - 2.0 = -0.9$: the model was too optimistic. Option (A) subtracts in the wrong order; (B) is the actual value; (D) is the forecast itself.

14. The one-step forecast error has two sources: the future shock u_{T+1} and coefficient-estimation error. As the estimation sample grows very large:

(A) both sources shrink towards zero.

(B) the estimation part shrinks, but the future-shock part remains.

(C) the future-shock part shrinks, but the estimation part remains.

(D) neither source of error is reduced.

Solution. (B). More past data pins the coefficients down ever more precisely, but no amount of history reveals next period's shock: σ_u^2 stays in the MSFE. That is why the RMSFE is approximately the SER in large samples, and why forecasts must be reported with intervals.

15. A one-step forecast of GDP growth is 2.0% with RMSFE = 3.16. The 95% forecast interval is closest to:

(A) $(1.6,\ 2.4)$.

(B) $(-1.2,\ 5.2)$.

(C) $(-3.2,\ 7.2)$.

(D) $(-4.2,\ 8.2)$.

Solution. (D). $2.0 \pm 1.96 \times 3.16 = 2.0 \pm 6.2 = (-4.2,\ 8.2)$. Option (B) uses one RMSFE, (C) uses the 90% multiplier 1.64, and (A) is a much narrower interval for the conditional mean, not for the realisation Y_{T+1} that a forecast interval must cover.

16. A 95% forecast interval for Y_{T+1} is much wider than a 95% confidence interval for the regression line at the same point because:

(A) it must also cover the future shock u_{T+1}, not just the estimated mean.

(B) it uses the small-sample t cutoff in place of 1.96.

(C) it is centred on the long-run mean rather than on the forecast.

(D) it must be computed with clustered standard errors.

Solution. (A). The interval predicts the *realisation* Y_{T+1}, which contains the new shock; the confidence interval only locates the conditional mean. The multiplier is the same 1.96 (B), the centre is the point forecast (C), and clustering (D) is a panel matter, not a forecasting one.

17. A stationary AR(1) is estimated as $Y_t = 1.2 + 0.6\,Y_{t-1} + u_t$. The long-run (unconditional) mean of Y is:

(A) 1.2.

(B) 2.0.

(C) 0.75.

(D) 3.0.

Solution. (D). $\mu = \beta_0/(1-\beta_1) = 1.2/(1-0.6) = 1.2/0.4 = 3.0$. Option (A) reads the intercept as the mean, (B) divides by β_1 itself, and (C) divides by $1+\beta_1$.

18. In an ADL model of GDP growth with two lags of the term spread, the F-statistic on the two spread lags is 4.43 (the 5% critical value for $q = 2$ is 3.00). The correct conclusion is that:

(A) movements in the spread cause movements in GDP growth.

(B) past values of the spread help predict growth, given growth's own lags.

(C) past values of growth also help predict the term spread.

(D) the two spread coefficients are each individually significant at 5%.

Solution. (B). Since $4.43 > 3.00$, the spread *Granger-causes* growth: its lags have predictive content beyond growth's own history. Granger causality is a statement about forecasting, not structural causation (A); the reverse direction (C) was never tested; and a joint rejection does not imply each individual t exceeds 1.96 (D).

19. For an AR(1) with $T = 204$, the fit term is $\ln\bigl(SSR(1)/T\bigr) = 2.289$ and $\frac{\ln T}{T} = 0.026$. The BIC is:

(A) 2.289.

(B) 2.315.

(C) 2.341.

(D) 2.309.

Solution. (C). $\text{BIC}(1) = 2.289 + (1 + 1) \times 0.026 = 2.289 + 0.052 = 2.341$. Option (A) omits the penalty, (B) charges for only one coefficient (forgetting the intercept, $p+1 = 2$), and (D) uses the AIC penalty $2/T = 0.010$ per coefficient.

20. An AR(1) fitted to 516 months of excess stock returns gives a slope of 0.050 with standard error 0.051 and $\bar{R}^2 = 0.0006$. The best reading of this output is that it is:

(A) consistent with efficient markets: past returns carry no usable forecast power.

(B) a sign the model needs more of its own lags before it is informative.

(C) evidence of momentum that a low-cost trader could exploit.

(D) a sign that the sample of 516 months is too short for reliable inference.

Solution. (A). $t = 0.050/0.051 \approx 0.98 < 1.96$ and the $\bar{R}^2$ is essentially zero: exactly what the efficient-markets logic predicts, since a reliably predictable excess return would be traded away. Adding return lags does not change the verdict (B), the point estimate is statistically indistinguishable from zero so there is nothing to exploit (C), and $T = 516$ is a large sample (D).

Section B: Analytical Problem (30 marks)

An analyst at a London asset manager studies the profitability of European banks and then turns to forecasting her own firm's business.

Part 1 (panel). She has a balanced panel of 40 banks observed yearly over 2010 to 2024. She regresses each bank's return on equity ROE (in %) on its leverage ratio LEV (assets/equity) and obtains:

Specification	Slope on LEV	Standard error
Pooled OLS	+0.25	(0.06)
Bank fixed effects, ordinary SE	−0.40	(0.15)
Bank fixed effects, clustered by bank	−0.40	(0.25)

Part 2 (forecasting). She then fits an AR(1) to $T = 120$ quarters of the firm's growth in assets under management, g_t (in % per quarter):

$$\widehat{g}_t \;=\; 0.90 \;+\; 0.55\, g_{t-1}, \qquad \bar{R}^2 = 0.30, \quad \text{SER} = 1.80,$$

and the most recent observation is $g_T = 5.0$.

(a) **[6 marks]** The slope on LEV changes from +0.25 (pooled) to −0.40 (bank fixed effects). Explain what the fixed effects control for, and use the omitted-variable-bias sign rule to describe an omitted bank trait that could produce this pattern.

(b) **[6 marks]** Compute the t-statistic for the fixed-effects slope using the ordinary standard error and again using the clustered standard error. State which inference is valid, why, and the 5% conclusion.

(c) **[6 marks]** Give one example of a confounder that survives *bank* fixed effects, and explain what adding *year* fixed effects would control for and what it still would not.

(d) **[6 marks]** Using the AR(1), compute the one-quarter-ahead forecast $\widehat{g}_{T+1}$ and the long-run mean of g_t. Is the process stationary? Explain in one sentence why the forecast lies between the latest value and the long-run mean.

(e) **[6 marks]** Construct an approximate 95% forecast interval for g_{T+1}, taking the RMSFE to be the SER. Name the two components of the RMSFE, say which one the SER captures here and why the other is small, and explain in one sentence why such a wide interval is the honest way to report this forecast.

Solution. (a) Bank fixed effects give each bank its own intercept, absorbing every **time-invariant** bank trait (business model, home-market institutions, permanent management quality), observed or not; the slope is then estimated only from each bank compared with itself over time. The pooled slope sits *above* the within slope ($+0.25$ vs -0.40), so the omitted trait biases the pooled estimate **upward**: by the sign rule it must *raise ROE* and be *positively* correlated with leverage, $(+)\times(+) = (+)$. A strong, profitable franchise fits: it earns high ROE and can safely operate at higher leverage. Within a bank, years of higher leverage go with *lower ROE*.

(b) Ordinary: $t = -0.40/0.15 = \mathbf{-2.67}$, which would reject $H_0 : \beta_1 = 0$ at 5%. Clustered: $t = -0.40/0.25 = \mathbf{-1.60}$, and $|-1.60| < 1.96$: do not reject. The **clustered** inference is the valid one: a bank's errors are serially correlated (a good run or a local recession persists for several years), so its fifteen yearly observations are not fifteen independent facts, and the ordinary formula understates the uncertainty. The honest 5% conclusion is that the within-bank leverage effect is **not statistically significant**.

(c) Any confounder that varies *within* a bank over time survives: for example, the bank's changing risk appetite, or economic conditions in its home market that shift year by year and move both its leverage and its ROE. Adding year fixed effects controls for shocks **common to all banks in a given year**: the euro-area interest-rate cycle, a market-wide crisis year. It still does not control for bank-specific, time-varying factors, which vary in both dimensions at once.

(d) Forecast: $\widehat{g}_{T+1} = 0.90 + 0.55 \times 5.0 = 0.90 + 2.75 = \mathbf{3.65}\%$. Long-run mean: $\mu = 0.90/(1 - 0.55) = 0.90/0.45 = \mathbf{2.0}\%$. Since $|0.55| < 1$, the AR(1) is **stationary**. The latest value (5.0) is 3.0 points above the mean, and only the fraction 0.55 of that deviation survives one quarter ($0.55 \times 3.0 = 1.65$), so the forecast is pulled part of the way back towards the mean: $2.0 + 1.65 = 3.65$, between 5.0 and 2.0. This is mean reversion.

(e) Interval: $3.65 \pm 1.96 \times 1.80 = 3.65 \pm 3.53 = [\mathbf{0.12},\ \mathbf{7.18}]$, roughly 0.1% to 7.2%. The RMSFE has two components: the **future shock** u_{T+1}, whose spread the SER estimates, and **coefficient-estimation error**, which is small here because $T = 120$ pins the two coefficients down. The interval is wide because most quarterly movement is the unforecastable new shock ($\bar{R}^2 = 0.30$: the lag explains under a third of the variation); reporting the point forecast 3.65% alone would overstate what the model knows.

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 adviser at a wealth manager markets a "return-timing model" for a stock index. Using 20 *years of monthly excess returns, he fitted AR models with one to twelve lags and kept the specification with the highest* $\bar{R}^2$*: an AR(11) with* $\bar{R}^2 = 0.09$ *and three individually significant lag coefficients. His brochure concludes that "the market is predictable and subscribers will beat it." A colleague reruns the exercise and reports two findings: the BIC selects zero lags; and when the AR(11), estimated on the first fifteen years only, is used to forecast the final five years, its RMSFE is larger than the RMSFE of simply forecasting the historical average return every month.*

Question [20 marks]. Explain why the adviser's in-sample evidence does not establish predictability, what each of the colleague's two findings shows, and how the efficient-markets view accounts for all of these results. Answer in at most five sentences.

Solution. Model answer (overfitting exposed by BIC and by forecast performance). The adviser searched across twelve specifications and kept the best-looking one, and with that many tries a high in-sample fit and a few significant t-statistics are close to guaranteed by the search itself rather than by genuine predictability. The BIC charges each extra coefficient a penalty, so its choice of zero lags says that no return lag improves fit by enough to earn its place: the eleven lags are noise dressed as signal. The second finding is the decisive one, because a forecasting model must be judged by its forecast errors on data it was not estimated on, and an RMSFE worse than the historical-mean benchmark means the model's apparent predictability does not survive outside the window it was fitted to. This is exactly what the efficient-markets logic predicts: any reliably forecastable excess return would be traded away as investors act on it, so true predictability should be close to zero and in-sample "discoveries" of it should fail when used for real forecasts. The honest conclusion is that the brochure describes the sample, not the market, and the timing model should not be sold as beating anything.

Marking guide (20 marks).

- **6 marks** — identifies the specification search/overfitting problem: choosing the best-fitting model over many candidates makes good fit and significant coefficients uninformative; the BIC's penalty asks each lag to earn its place, and selecting $p = 0$ means no lag has real predictive content.
- **6 marks** — explains the forecast-evaluation finding: models are judged by forecast errors (RMSFE) on data not used in estimation, and losing to the historical-mean benchmark shows the fitted pattern does not generalise.
- **6 marks** — gives the efficient-markets account: predictable excess returns are traded

away, so near-zero genuine predictability is the expected result and the in-sample fit reflects chance, not exploitable structure.

- **2 marks** — clear writing within the five-sentence limit.

Answers that only say "the model is overfitted" without connecting the BIC choice and the RMSFE comparison to that diagnosis earn partial credit at most: the question asks what each finding shows.