Эконометрика — МИЭФ, 2023 midterm 2

МИЭФ · Эконометрика · 2023 · midterm 2

QUESTION 1

Multiple-choice test

The following model of determination of the size of dividends is considered:

$$D_t^* = \gamma P_t, \tag{1}$$

$$\Delta D_t = \lambda(D_t^* - D_{t-1}) + \mu \Delta P_t + u_t, \tag{2}$$

where D_t^* is the desirable size of dividends, P_t is current profit, D_t is the actual size of dividends,

$$\Delta D_t = D_t - D_{t-1}, \qquad \Delta P_t = P_t - P_{t-1}.$$

The model is:

1. Adaptive expectations model.
2. Rational expectations model.
3. Partial adjustment model.
4. Error correction model.
5. ADL(1,0) model.

QUESTION 2

Multiple-choice test

The Durbin-Watson d statistic is close to 4. This means that:

1. The correlation coefficient between e_i and e_{i+k} is close to zero for any $k \geq 1$.
2. The correlation coefficient between e_i and e_{i+k} is close to one for any $k \geq 1$.
3. The correlation coefficient between e_i and e_{i+1} is close to one.
4. The correlation coefficient between e_i and e_{i+1} is close to zero.
5. The correlation coefficient between e_i and e_{i+1} is close to minus one.

QUESTION 3

Multiple-choice test

Using 96 observations, a student estimated the production function

$$\ln(Y) = \beta_1 + \beta_2 \ln(K) + \beta_3 \ln(L) + \beta_4 Time,$$

where Y is output, K is capital, and L is labour:

$$\widehat{\ln(Y)} = -6.4 + 0.38 \ln(K) + 0.69 \ln(L) + 0.0057 Time, \qquad R^2 = 0.94,$$

with standard errors

$$(1.3) \qquad (0.14) \qquad (0.12) \qquad (0.0011).$$

Which statement can be made?

1. If all variables $\ln(Y)$, $\ln(K)$, and $\ln(L)$ are $I(1)$, then differencing should be done.
2. If at least one of the variables $\ln(Y)$, $\ln(K)$, and $\ln(L)$ is $I(0)$, this cannot be a spurious regression.
3. This cannot be a spurious regression because the variable $Time$ is included.
4. If differencing is applied to this model, the transformed model will not include a constant.
5. Each variable should be detrended separately instead of including the variable $Time$.

QUESTION 4

Multiple-choice test

Refer to the model

$$Y_t = \alpha_0 + \beta_0 S_t + \beta_1 S_{t-1} + \beta_2 S_{t-2} + \beta_3 S_{t-3} + u_t.$$

The expression

$$\beta_0 + \beta_1 + \beta_2 + \beta_3$$

represents:

1. The short-run change in Y caused by a temporary increase in S.
2. The short-run change in Y caused by a permanent increase in S.
3. The long-run change in Y caused by a permanent increase in S.
4. The long-run change in Y caused by a temporary increase in S.
5. None of the above.

QUESTION 5

Multiple-choice test

A stochastic process $\{x_t : t = 1, 2, \ldots\}$ with a finite second moment, $E(x_t^2) < \infty$, is covariance stationary if:

1. $E(x_t)$ varies, $\mathrm{Var}(x_t)$ varies, and for any t and $h \geq 1$, $\mathrm{Cov}(x_t, x_{t+h})$ depends only on h and not on t.
2. $E(x_t)$ varies, $\mathrm{Var}(x_t)$ varies, and for any t and $h \geq 1$, $\mathrm{Cov}(x_t, x_{t+h})$ depends only on t and not on h.
3. $E(x_t)$ is constant, $\mathrm{Var}(x_t)$ is constant, and for any t and $h \geq 1$, $\mathrm{Cov}(x_t, x_{t+h})$ depends only on h and not on t.
4. $E(x_t)$ is constant, $\mathrm{Var}(x_t)$ is constant, and for any t and $h \geq 1$, $\mathrm{Cov}(x_t, x_{t+h})$ depends only on t and not on h.
5. None of the above.

QUESTION 6

Multiple-choice test

Indicate the incorrect statement:

1. If X_t is a random walk with drift, the first-difference series

$$\Delta X_t = X_t - X_{t-1} = \beta_1 + \varepsilon_t,$$

where ε_t is white noise, is stationary.

2. The time-trend process

$$X_t = \beta_1 + \beta_2 t + \varepsilon_t$$

is non-stationary.

3. The MA(1) process

$$X_t = \varepsilon_t + \alpha_2 \varepsilon_{t-1}$$

is stationary.

4. The AR(1) process

$$X_t = \beta_2 X_{t-1} + \varepsilon_t, \qquad \beta_2 < 1,$$

is asymptotically stationary.

5. A random walk without drift becomes stationary after taking logarithms.

QUESTION 7

Multiple-choice test

Which statement is true?

1. A random walk process is asymptotically stationary.
2. A deterministic time-trend process is asymptotically stationary.
3. The variance of a random walk process increases as a linear function of time.
4. The variance of a random walk with drift decreases as a quadratic function of time.
5. Adding a drift term turns a stationary random walk into a non-stationary process.

QUESTION 8

Multiple-choice test

Which statement about spurious regressions is true?

1. OLS estimates of population parameters are unbiased and the t statistic is valid.
2. OLS estimates of population parameters are unbiased but the t statistic is invalid.
3. Even if the explanatory variables and the dependent variable are independent time-series processes, R^2 can be large.
4. If the explanatory variables and the dependent variable are independent time-series processes, R^2 cannot be large.
5. None of the above.

QUESTION 9

Multiple-choice test

Which statement is true?

1. An error correction model is based on an ADL(1,0) process.
2. An error correction model can be used to study short-run, but not long-run, dynamics between the dependent variable and explanatory variables.
3. An error correction model can be used to study long-run, but not short-run, dynamics between the dependent variable and explanatory variables.
4. An error correction model can be used to study both short-run and long-run dynamics between the dependent variable and explanatory variables.
5. The Dickey-Fuller test can be used to test for autocorrelation in the error terms.

QUESTION 10

Multiple-choice test

Which statement correctly identifies the difference between an autoregressive model and a vector autoregressive model?

1. In an autoregressive model, the dependent variable is a function of its own lag, whereas in a vector autoregressive model, the dependent variable is a function of the lag of an explanatory variable.
2. In an autoregressive model, the dependent variable is a function of the lag of an explanatory variable, whereas in a vector autoregressive model, the dependent variable is a function of its own lag.
3. In an autoregressive model, several series are modelled in terms of their own past, whereas in a vector autoregressive model only one series is modelled in terms of its own past.
4. In an autoregressive model, one series is modelled in terms of its own past, whereas in a vector autoregressive model several series are modelled in terms of their past.
5. None of the above.

QUESTION 11

Multiple-choice test

Which of the following is a reason to use the random-effects approach instead of the fixed-effects approach?

1. It provides unbiased and consistent estimators when disturbance terms are serially correlated.
2. It provides unbiased and consistent estimators when disturbance terms are heteroscedastic.
3. It provides more efficient estimates than the fixed-effects approach.
4. It provides a way to include time-constant explanatory variables in a fixed-effects analysis.
5. It accounts for unobserved heterogeneity, unlike fixed effects.

QUESTION 12

Multiple-choice test

An economist wants to study the effect of income on savings. The economist collects data on 120 identical twins. Which estimation method is most suitable if income is correlated with an unobserved family effect?

1. Random-effects estimation.
2. Fixed-effects estimation.
3. Ordinary least squares estimation.
4. Weighted least squares estimation.
5. OLS after an autoregressive transformation.

Part 2. Free-response questions — one session, 2 hours without a break.

Section A. Answer all questions from this section (original Questions 1-2).

QUESTION 13

Written Question 1 — 25 marks

A student investigates the relationship between the GDP of Australia and the GDP of Fiji, a small island country near Australia whose economy is therefore closely connected to Australia's economy.

Using annual data for 1985-2018, 34 observations, the student regresses the logarithm of Fiji's GDP, F_t, on the logarithm of Australia's GDP, A_t, and obtains:

$$\widehat{F}_t = -4.56 + 0.86A_t, \qquad R^2 = 0.97,$$

with standard errors

$$(0.18) \qquad (0.028). \tag{1}$$

Suspecting that this regression may be spurious, the student investigates F_t and A_t for stationarity and estimates Dickey-Fuller equations for the level series and their first differences,

$$DF_t = \Delta F_t, \qquad DA_t = \Delta A_t.$$

For F_t:

$$\Delta F_t = 0.33 - 0.53F_{t-1} + 0.014time, \qquad R^2 = 0.29, \qquad F = 6.15,$$

with standard errors

$$(0.84) \qquad (0.15) \qquad (0.004), \tag{2a}$$

and

$$\Delta DF_t = -0.90DF_{t-1}, \qquad R^2 = 0.46,$$

with standard error

$$(0.17). \tag{2b}$$

For A_t:

$$\Delta A_t = 1.89 - 0.31A_{t-1} + 0.20\Delta A_{t-1} + 0.01time, \qquad R^2 = 0.20, \qquad F = 2.36,$$

with standard errors

$$(0.78) \qquad (0.13) \qquad (0.18) \qquad (0.004), \tag{3a}$$

and

$$\Delta DA_t = 0.03 - 0.94 DA_{t-1}, \qquad R^2 = 0.47, \qquad F = 26.6,$$

with standard errors

$$(0.007) \qquad (0.18). \tag{3b}$$

(a)

- Briefly explain the theory behind estimating Dickey-Fuller equation (2a). Why is an additional lag included in equation (3a)?
- Use the augmented Dickey-Fuller t test for equations (2a) and (3a) to investigate stationarity of the level series.
- Determine whether DF_t and DA_t are stationary using equations (2b) and (3b).

To test the series for cointegration, the student estimates the following equation for the residuals Z_t from equation (1):

$$\Delta Z_t = -0.48 Z_{t-1}, \qquad R^2 = 0.27,$$

with standard error

$$(0.14). \tag{4}$$

- What is cointegration?
- Why is cointegration important in time-series analysis?
- Conduct the cointegration test.

(b)

The student uses equation (1) as the cointegrating relationship and starts from the ADL(1,1) model

$$F_t = \alpha_1 + \alpha_2 F_{t-1} + \alpha_3 A_t + \alpha_3 A_{t-1} + u_t.$$

The resulting error correction model is

$$DF_t = -0.48Z_{t-1} + 0.93DA_t, \qquad R^2 = 0.39,$$

with standard errors

$$(0.14) \qquad (0.19). \tag{5}$$

- What is the purpose of constructing an error correction model?
- Describe how to derive the error correction model corresponding to the estimated equation in the context of this problem.
- Interpret the coefficients in model (5).
- Explain the advantages and possible disadvantages of the error correction specification compared with model (1) and with a model based only on the differences DF_t and DA_t.

QUESTION 14

Written Question 2 — 25 marks

A student studies how the average monthly wages of employees of a large transportation company, measured in thousands of units of local currency, depend on several factors. The student uses a panel of 550 individuals observed for 7 years. The dependent variable is the natural logarithm of wage, $\ln W_{it}$.

The model is

$$\ln W_{it} = \beta_1 ED_i + \beta_2 EX_{it} + \beta_3 EX_{it}^2 + \beta_4 M_{it} + \beta_5 U_{it} + \sum_{t=1}^{5} \gamma_t T_t + \alpha_i + u_{it}. \qquad (1)$$

Here:

- ED_i is years of full-time education before the start of the work period;
- EX_{it} is work experience;
- EX_{it}^2 is experience squared;
- M_{it} is a dummy for being married;
- U_{it} is a dummy for union membership;
- an individual can change marital and union status in any year;
- T_t is a time dummy;
- α_i is an unobserved-heterogeneity term;
- u_{it} is a disturbance term satisfying the usual regression-model conditions;
- i indexes individuals and $t = 1, \ldots, 7$ indexes time.

The inclusion of only five time dummies prevents multicollinearity and should not be discussed.

(a)

- Briefly characterize the problem of unobserved heterogeneity, represented by α_i, in panel-data analysis.

- What methods can be used to estimate model (1)? Compare their advantages and disadvantages.

- In the context of this model, briefly explain the least-squares dummy-variable, LSDV, method and discuss its advantages and drawbacks.

(b)

The following three models are estimated from the annual panel:

Pooled OLS

$$\widehat{\ln W}_{it} = 1.41 + 0.091ED_i + 0.067EX_{it} - 0.003EX_{it}^2 + 0.108M_{it} + 0.1821U_{it},$$

with standard errors

$$(0.035) \qquad (0.005) \qquad (0.014) \qquad (0.001) \qquad (0.016) \qquad (0.017).$$

Random effects

$$\widehat{\ln W}_{it} = 1.35 + 0.092ED_i + 0.106EX_{it} - 0.005EX_{it}^2 + 0.064M_{it} + 0.1061U_{it},$$

with standard errors

$$(0.042) \qquad (0.011) \qquad (0.015) \qquad (0.001) \qquad (0.017) \qquad (0.018).$$

Fixed effects

$$\widehat{\ln W}_{it} = 1.34 - 0.005EX_{it}^2 + 0.047M_{it} + 0.080U_{it},$$

with standard errors

$$(0.046) \qquad (0.001) \qquad (0.018) \qquad (0.019).$$

Coefficients on the time dummies are not reported; refer to model (1). Numbers in parentheses are standard errors.

- Explain the meaning of the coefficients on ED_i, EX_{it}, EX_{it}^2, M_{it}, and U_{it} in the pooled-OLS equation.

- Give a meaningful interpretation of the unobserved effects in a wage equation of this type.

- Based on that interpretation, explain why the random-effects model might give inconsistent estimates.

- Why do the coefficients on M_{it} and U_{it} differ substantially across the three regressions?

- Explain why the fixed-effects regression does not report estimated coefficients for the first two explanatory variables.

- Which test can justify choosing between random- and fixed-effects regressions? State the statistic and degrees of freedom for this case. What result would you expect in the context of this problem?

Section B. Answer only one question from this section (original Question 3 or Question 4).

QUESTION 15

Written Question 3 — 25 marks

A student is writing a term paper on the factors determining aggregate expenditure on fashionable clothes, F_t. Her original intention was to estimate a linear regression of clothing expenditure on current income:

$$F_t = \beta_1 + \beta_2 I_t + u_t.$$

Her academic adviser explains that expenditure on this category of goods may follow a more advanced relationship. Current income determines only a target or desired level of expenditure:

$$F_t^* = \beta_1 + \beta_2 I_t + u_t,$$

and the target is reached through partial adjustment:

$$F_t - F_{t-1} = \lambda(F_t^* - F_{t-1}), \qquad 0 < \lambda < 1.$$

The student likes this idea but does not understand how to obtain data on the annual target values F_t^*.

(a)

- Explain how this system can be reduced to an estimable model using available data.
- What properties are expected for the resulting estimates?
- Explain the structure and meaning of the partial-adjustment model parameters.
- Explain the dynamic properties of the partial-adjustment model.

(b)

Using data for Brazil, 25 observations from 1998-2022 in local currency at constant prices, the student estimates

$$\widehat{F}_t = -2.09 + 0.013 I_t + 0.89 F_{t-1}, \qquad R^2 = 0.99,$$

with standard errors

$$(1.36) \qquad (0.007) \qquad (0.089).$$

- Interpret the coefficients.

- Compare the short-run and long-run marginal effects of income.
- Explain how to obtain the long-run characteristics of the relationship.

(c)

The academic adviser suggests extending the model with an error correction mechanism:

$$\Delta F_t = \lambda(F_t^* - F_{t-1}) + \delta(I_t - I_{t-1}),$$

where

$$F_t^* = \beta_1 + \beta_2 I_t + u_t.$$

- Show how to obtain an estimable ADL model from this scheme.
- Explain how to recover estimates of all original parameters.
- Why is this scheme called an error correction model?
- How can one test whether it provides a significant improvement over the original partial-adjustment model?

QUESTION 16

Written Question 4 — 25 marks

A student is writing a term paper on the factors determining aggregate expenditure on fashionable clothes, F_t. Her original intention was to estimate a linear regression of clothing expenditure on current income:

$$F_t = \beta_1 + \beta_2 I_t + u_t.$$

Her academic adviser explains that expenditure on fashionable clothes is determined not by current income but by expected income for the next period:

$$F_t = \beta_1 + \beta_2 I^e_{t+1} + u_t, \qquad (*)$$

where I^e_{t+1} denotes expectations of income for period $t+1$. Expectations adapt to current income according to

$$I^e_{t+1} - I^e_t = \lambda(I_t - I^e_t), \qquad 0 < \lambda < 1. \qquad (**)$$

The student likes this idea but does not understand how to obtain data on expected income I^e_{t+1}.

(a)

- Explain how to obtain an estimable Koyck-distribution model:

$$F_t = \beta_1 + \beta_2\lambda I_t + \beta_2\lambda(1-\lambda)I_{t-1} + \beta_2\lambda(1-\lambda)^2 I_{t-2} + \cdots$$
$$+ \beta_2\lambda(1-\lambda)^{s-1} I_{t-s+1} + \beta_2(1-\lambda)^s I^e_{t-s+1} + u_t.$$

- Explain how to drop the unobservable term and use nonlinear estimation to estimate the parameters of the original model.

- What properties are expected for the resulting estimates?

(b)

The Koyck-distribution model can be transformed into the ADL(1,0) model

$$F_t = \lambda\beta_1 + \lambda\beta_2 I_t + (1-\lambda)F_{t-1} + u_t - (1-\lambda)u_{t-1}.$$

Using the Brazil data described in Question 3, the student estimates

$$\widehat{F}_t = -2.09 + 0.013 I_t + 0.89 F_{t-1}, \qquad R^2 = 0.99,$$

with standard errors

$$(1.36) \qquad (0.007) \qquad (0.089).$$

- Interpret the coefficients.
- Compare the short-run and long-run marginal effects of income.
- Explain how to obtain the long-run characteristics of the relationship.

(c)

- Derive the ADL(1,0) equation

$$F_t = \lambda\beta_1 + \lambda\beta_2 I_t + (1-\lambda)F_{t-1} + u_t - (1-\lambda)u_{t-1}$$

from models (*) and (**).

- What are the properties of OLS estimators for this ADL(1,0) equation?