ECOM-G315 Econometrics 2
Final Examination
21 December 2017

Name: ______________________________

Student Number: ______________________________

University: ☐ Aalto ☐ Hanken ☐ University of Helsinki

Answer all the question.

The material on page 7 may be useful in answering some of the questions.

Unless otherwise stated, ε_t denotes throughout a white noise process.

1. In the following multiple choice questions, mark the correct alternative by ticking the relevant box (☒). **In each question, only one alternative is correct. If multiple alternatives have been marked, the question yields 0 points.** Each correct choice yields +3 points, and each incorrect choice yields −1 point.

   (a) Consider the following process:

   $$y_t = 1.4 + 0.6y_{t-1} + \varepsilon_t - 1.1\varepsilon_{t-1} + 0.3\varepsilon_{t-2}.$$

   ☐ The process is nonstationary.
   ☐ Using the lag operator L, the process can be written as $(1+0.6L)y_t = (1-0.6L)(1-0.5L)\varepsilon_t$.
   ☐ The process can equivalently be written as an AR(∞) process.
   ☐ The process can equivalently be written as an AR(2) process.

   (b) Consider conditional ordinary least squares (OLS) and maximum likelihood (ML) estimation of the parameters c, θ_1 and θ_2 of the following model:

   $$y_t = c + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \varepsilon_t$$

   from observations $y_1, y_2, \ldots y_T$, when the error term ε_t is normally distributed.

   ☐ The conditional OLS estimator is obtained by minimising $\sum_{t=3}^{T}(y_t - c - \theta_1 y_{t-1} - \theta_2 y_{t-2})^2$ conditional on $y_1 = y_2 = 0$.
   ☐ The conditional OLS estimator is obtained by maximising $\sum_{t=3}^{T}(y_t - c - \theta_1 y_{t-1} - \theta_2 y_{t-2})^2$.
   ☐ The conditional ML estimator is obtained by minimising $\sum_{t=3}^{T}(y_t - c - \theta_1 y_{t-1} - \theta_2 y_{t-2})^2$.
   ☐ The conditional ML estimator is obtained by maximising $\sum_{t=3}^{T}(y_t - c - \theta_1 y_{t-1} - \theta_2 y_{t-2})^2$.

   (c) Suppose ε_{1t}, ε_{2t}, and ε_{3t} are three independent white noise processes, $\theta(L)$ and $\phi(L)$ are first-order and second-order polynomials in the lag operator L, respectively, $\theta(L)x_t = \varepsilon_{1t}$ with $\theta(1) = 0$, $\phi(L)y_t = \varepsilon_{2t}$ with $\phi(2) = \phi(3) = 0$, and $z_t = 2x_t - 0.5y_t + \varepsilon_{3t}$.

   ☐ x_t is weakly stationary.
   ☐ z_t is weakly stationary.
   ☐ $(x_t, z_t)'$ is cointegrated with cointegrating vector $(1, -0.5)'$.
   ☐ $(y_t, z_t)'$ is cointegrated with cointegrating vector $(1, 2)'$.

   (d) Consider the following model:

   $$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \varepsilon_t.$$

   ☐ $y_{T+1|T}$, the optimal one-period forecast of y_{T+1} equals $\phi_1 y_{T-1} + \phi_2 y_{T-2}$.
   ☐ $y_{T+2|T}$, the optimal two-period forecast of y_{T+2} equals $\phi_1 y_{T+1} + \phi_2 y_T$.
   ☐ $y_{T+2|T}$, the optimal two-period forecast of y_{T+2} equals $\phi_1 y_{T+1|T} + \phi_2 y_T$.
   ☐ None of the above alternatives is correct.

(e) Consider the following model:

$$y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}.$$

- ☐ $y_{T+1|T}$, the optimal one-period forecast of y_{T+1} equals $\varepsilon_{T+1} + \theta_1 \varepsilon_T + \theta_2 \varepsilon_{T-1}$.
- ☐ $y_{T+2|T}$, the optimal two-period forecast of y_{T+2} equals $\varepsilon_{T+2} + \theta_1 \varepsilon_{T+1} + \theta_2 \varepsilon_T$.
- ☐ $y_{T+2|T}$, the optimal two-period forecast of y_{T+2} equals zero.
- ☐ None of the above alternatives is correct.

(f) A researcher uses the Bayesian information criterion (BIC) to select the best time series model for the Finnish unemployment rate among the five models (each containing an intercept) that she has estimated by exact maximum likelihood with monthly data from 1996:1–2017:8 (248 observations): AR(1), AR(2), MA(1), MA(2), and ARMA(1,1). The residual variance of each of the models are given in the following table:

Model	AR(1)	AR(2)	MA(1)	MA(2)	ARMA(1,1)
Residual variance	0.46	0.46	0.47	0.46	0.46

Based on this information,

- ☐ the BIC selects the AR(1) model.
- ☐ the BIC selects the MA(1) model.
- ☐ the BIC selects the ARMA(1,1) model.
- ☐ none of the above alternatives is correct.

(g) A researcher is specifying an adequate AR(p) model for the quarterly Finnish inflation rate series (π_t) by sequential testing based on significance tests at the 5% level of significance. He starts out by estimating an AR(6) model,

$$\pi_t = \theta_0 + \theta_1 \pi_{t-1} + \theta_2 \pi_{t-2} + \theta_3 \pi_{t-3} + \theta_4 \pi_{t-4} + \theta_5 \pi_{t-5} + \theta_6 \pi_{t-6} + \varepsilon_t$$

and finds that the p-values of the significance tests of θ_0, θ_1, θ_2, θ_3, θ_4, θ_5, and θ_6 equal 0.201, 0.035, 0.022, 0.063, 0.012, 0.042 and 0.084, respectively. Based on this information,

- ☐ sequential testing selects the AR(4) model.
- ☐ sequential testing selects the AR(5) model.
- ☐ sequential testing selects the AR(6) model.
- ☐ none of the above alternatives is correct.

(h) Consider the following VAR(2) process:

$$\mathbf{y}_t = \Theta_1 \mathbf{y}_{t-1} + \Theta_2 \mathbf{y}_{t-2} + \varepsilon_t.$$

- ☐ The process can equivalently be written as $\Delta \mathbf{y}_t = \Pi \mathbf{y}_{t-1} + \Gamma_1 \Delta \mathbf{y}_{t-1} + \Gamma_2 \Delta \mathbf{y}_{t-2} + \varepsilon_t$.
- ☐ If $I - \Theta_1 - \Theta_2 = 0$, $\mathbf{y}_t$ is a cointegrated $I(1)$ process.
- ☐ If $\mathbf{y}_t = (x_t, z_t)'$, $\Theta_1 = \begin{pmatrix} 0.5 & 2.1 \\ 0.0 & 0.4 \end{pmatrix}$, and $\Theta_2 = \begin{pmatrix} 0.0 & -1.1 \\ 0.0 & 0.0 \end{pmatrix}$, $\mathbf{y}_t$ is weakly stationary.
- ☐ Suppose $\mathbf{y}_t = (x_t, z_t)'$. Then x_t is Granger causal for z_t if the (2,1) elements of Θ_1 and Θ_2 equal zero.

(i) A researcher wants to find out whether two interest rates, the one-month rate (r_{1t}) and the one-year rate (r_{12t}), are cointegrated in accordance with the expectations hypothesis of the term structure of interest rates (i.e., are cointegrated with cointegrating vector $(1, -1)'$). Her data set comprises 812 weekly observations. When she regresses r_{12t} on a constant and r_{1t}, she obtains the following OLS estimates (t-statistics in parentheses):

$$r_{12t} = \underset{(0.08)}{0.95} + \underset{(0.89)}{1.00}\, r_{1t}.$$

She then computes the residuals of this regression model (e_t) and runs the following regression:

$$\Delta e_t = \delta + \rho e_{t-1} + \gamma_1 \Delta e_{t-1} + \gamma_2 \Delta e_{t-2} + \varepsilon_t.$$

Because the value of the ADF test statistic of the residuals based on the latter regression is −3.22,

- ☐ the ADF test rejects the null hypothesis of cointegration at the 5% level of significance.
- ☐ the ADF test rejects the null hypothesis of no cointegration at the 5% level of significance.
- ☐ the results indicate that $(r_{1t}, r_{12t})'$ is cointegrated with cointegrating vector $(1, -1)'$ (at the 5% significance level).
- ☐ none of the above alternatives is correct.

(j) Another researcher tests the implication of the expectations hypothesis in the same data consisting of 812 observations by testing whether $y_t = r_{12t} - r_{1t}$ is a unit root process. Based on the following regression,

$$\Delta y_t = \delta + \rho y_{t-1} + \gamma_1 \Delta y_{t-1} + \varepsilon_t$$

the value of the ADF test statistic is -3.14. Hence,

- ☐ r_{12t} and r_{1t} are not cointegrated at the 5% level of significance.
- ☐ r_{12t} and r_{1t} are cointegrated but not with cointegrating vector $(1, -1)'$ at the 5% level of significance.
- ☐ r_{12t} and r_{1t} are cointegrated with cointegrating vector $(1, -1)'$ at the 5% level of significance.
- ☐ none of the above alternatives is correct.

(k) Yet another researcher tests the implication of the expectations hypothesis in the same data by means of a vector autoregression fitted to the data consisting of 812 weekly observations of r_{1t} and r_{12t}. He obtains the output in Box 1 on page 6. According to the results,

- ☐ r_{1t} and r_{12t} are not cointegrated at the 10% level of significance.
- ☐ r_{1t} and r_{12t} are cointegrated at the 10% level of significance, and assuming they are cointegrated, the cointegrating vector is $(1, -1)'$ at the 10% level of significance.
- ☐ r_{1t} and r_{12t} are not cointegrated at the 10% level of significance, but assuming they are cointegrated, the cointegrating vector is $(-1, 1)'$ at the 10% level of significance.
- ☐ none of the above alternatives is correct.

(l) A researcher is forecasting the Finnish inflation rate (π_t) with an AR(3) model. She estimates the model from monthly data from 1995:1 to 2017:6, and in order to assesits forecast performance, she computes the one-period ahead forecasts ($\pi_{t+1|t}$) over the estimation period. The estimation result of the Mincer-Zarnowitz regression is the following (standard errors in parentheses):

$$\pi_{t+1} = \underset{(0.24)}{0.48} + \underset{(0.50)}{1.10}\, \pi_{t+1|t}$$

The p value of the Wald test of the null hypothesis that the intercept term equals zero and the slope coefficient equals unity is 0.08, while the p value of the Wald test of the null hypothesis that the intercept term equals unity and the slope coefficient equals zero is 0.01. She correctly concludes that

- ☐ the AR(3) model yields unbiased forecasts (at the 5% significance level).
- ☐ the forecast errors of the AR(3) model are significantly forecastable (at the 5% level).
- ☐ the forecast errors of the AR(3) model are autocorreted (at the 5% significance level).
- ☐ None of the above alternatives is correct.

(m) A researcher is surprised because the ADF test does not reject the null hypothesis of a unit root in the Finnish real interest rate $R_t = r_t - \pi_t$ at the 5% level of significance. She was expecting r_t and π_t to be cointegrated in accordance with the Fisher effect hypothesis, and she wonders whether there could be something wrong with her testing setup. Which of the following is **not** a potential reason for not rejecting the unit root hypothesis?

- ☐ Because she had only 112 observations, the power of her test might be low.
- ☐ Because she had only 112 observations, the size of her test might be too high.
- ☐ Although R_t really is stationary, it is so close to being $I(1)$ that the test may have low power.
- ☐ Although r_t and π_t really are cointegrated, they are not cointegrated with cointegrating vector $(1, -1)'$.

(n) A researcher has observations of three time series, x_t, y_t, and z_t. She is unable to reject the null hypothesis of x_t and y_t being unit root processes at the 5% level of significance, but she concludes that z_t is an $I(0)$ process. Moreover, she finds that x_t and y_t are cointegrated with cointegrating vector $(1, -1)'$. She regresses y_t on a constant, x_{t-1}, y_{t-1}, Δx_{t-1}, and z_{t-1} i.e., estimates the following regression model:

$$y_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 y_{t-1} + \beta_3 \Delta x_{t-1} + \beta_4 z_{t-1} + \varepsilon_t. \tag{1}$$

She obtains the following OLS estimates (t-statistics based on a consistent covariance matrix estimator in parentheses):

$$y_t = \underset{(1.14)}{0.12} + \underset{(1.72)}{0.32}\, x_{t-1} - \underset{(-2.46)}{0.98}\, y_{t-1} + \underset{(1.81)}{0.14}\, \Delta x_{t-1} + \underset{(1.38)}{0.22}\, z_{t-1}.$$

The p value (based on the χ^2_2 distribution) of the Wald test of the null hypothesis $\beta_1 = \beta_4 = 0$ equals 0.068, and the p value (based on the χ^2_2 distribution) of the Wald test of the null hypothesis $\beta_3 = \beta_4$ equals 0.048. According to the Breusch-Godfrey test, the residuals are not autocorrelated up to 6 lags (at the 5% significance level). In regression (1),

- ☐ the null hypothesis $\beta_3 = \beta_4$ is rejected at the 5% significance level.
- ☐ the null hypothesis $\beta_3 = \beta_4$ is rejected at the 1% significance level.
- ☐ the critical values of the Wald test of the null hypothesis $\beta_1 = \beta_4$ cannot be taken from the χ^2_2 distribution, so the p value based on it is incorrect.
- ☐ none of the above alternatives is correct.

(o) In regression (1),

- ☐ according to the t test, β_2 is significant at the 5% significance level.
- ☐ according to the t test, β_2 is significantly different from zero at he 1% significance level.
- ☐ β_2 is consistently estimated by OLS, but the significance of β_2 cannot be evaluated by critical values from the standard normal distribution.
- ☐ none of the above alternatives is correct.

(p) In regression (1),

- ☐ according to the t test, β_3 is significant at he 5% significance level.
- ☐ according to the t test, β_3 is significantly different from zero at he 10% significance level.
- ☐ β_3 is consistently estimated by OLS, but the significance of β_3 cannot be evaluated by critical values from the standard normal distribution, so the p value based on it is incorrect.
- ☐ none of the above alternatives is correct.

(q) In regression (1),

- ☐ the null hypothesis $\beta_1 = \beta_4 = 0$ is rejected at the 5% significance level.
- ☐ the null hypothesis $\beta_1 = \beta_4 = 0$ is rejected at the 10% significance level.
- ☐ the critical values of the Wald test of the null hypothesis $\beta_1 = \beta_4 = 0$ cannot be taken from the χ^2_2 distribution.
- ☐ none of the above alternatives is correct.

2. This question yields a maximum of 4 points. Write your answer in the box below each question.

(a) Why is it important to select a consistent estimator of the covariance matrix of the ordinary least squares (OLS) estimator when doing econometric analysis by means of linear regression?

(b) Suppose you are estimating by ordinary least squares the parameters of the following linear regression model:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 z_t + \varepsilon_t,$$

where x_t, y_t and z_t are stationary time series. Describe the steps you need to take to ensure that you select a consistent estimator of the covariance matrix of the OLS estimator.

3. This question yields a maximum of 5 points. Write your answer in the box below each question.

(a) Write down the equations of the estimated restricted vector error correction model in Box 1 on page 6.

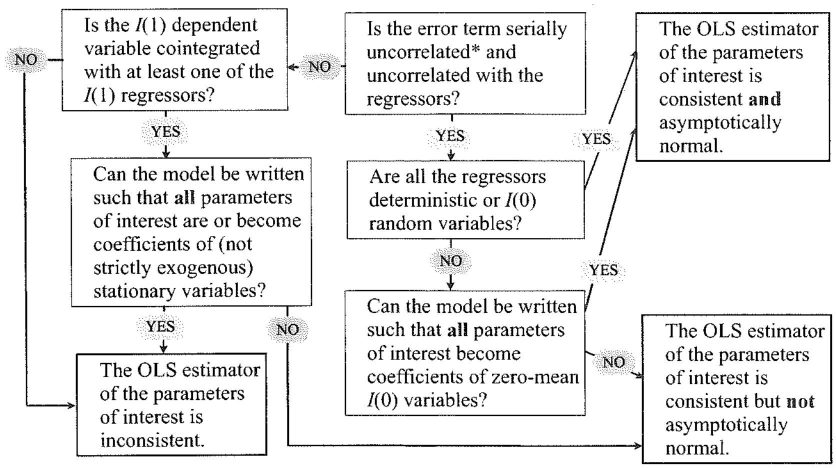
(b) Interpret the estimated restricted vector error correction model in Box 1 on page 6.

```
> y=cbind(r12,r1)
> vecm = ca.jo(y, K=3, ecdet="const", type="trace", spec="transitory")
> summary(vecm)

######################
# Johansen-Procedure #
######################

Test type: trace statistic , without linear trend and constant in cointegration

Values of test statistic and critical values of test:

          test 10pct  5pct  1pct
r <= 1 |  2.92  7.52  9.24 12.97
r = 0  | 17.88 17.85 19.96 24.60

Eigenvectors, normalised to first column:
(These are the cointegration relations)

               r12.l1       r1.l1    constant
r12.l1    1.000000000  1.00000000  1.00000000
r1.l1    -0.964265319 -1.78407362 -0.73397230
constant -0.005025616  0.05145488 -0.08330775

Weights W:
(This is the loading matrix)

         r12.l1      r1.l1       constant
r12.d -1.295431 0.06015652 -4.105492e-16
r1.d  -1.033525 0.07539816 -3.265174e-16

> H1 <- matrix(c(1,-1,0,
+                0,0,1), c(3,2))
> summary(blrtest(vecm, H = H1, r = 1))

######################
# Johansen-Procedure #
######################

Estimation and testing under linear restrictions on beta

The VECM has been estimated subject to:
beta=H*phi and/or alpha=A*psi

     [,1] [,2]
[1,]    1    0
[2,]   -1    0
[3,]    0    1

The value of the likelihood ratio test statistic:
1.95 distributed as chi square with 1 df.
The p-value of the test statistic is: 0.16

Eigenvectors, normalised to first column
of the restricted VAR:

        [,1]    [,2]
[1,]  1.0000  1.0000
[2,] -1.0000 -1.0000
[3,] -0.0025 -0.0618

Weights W of the restricted VAR:

         [,1]    [,2]
r12.d -0.9275  0.0029
r1.d   0.6423 -0.0035
```

Box 1

(Note: `r1` and `r12` denote r_{1t} and r_{12t}, respectively.)

OLS Regression with $I(1)$ Variables Cheat Sheet



- * Not autocorrelated after H-1 lags.

For further details, see Stock, J.H., & Watson M.W. (1988), Variable Trends in Economic Time Series. *Journal of Economic Perspectives* 2, 147 - 174.

Table 8.1 1% and 5% critical values for Dickey–Fuller tests (Fuller, 1976, p. 373)

	Without trend		With trend	
Sample size	1%	5%	1%	5%
$T = 25$	−3.75	−3.00	−4.38	−3.60
$T = 50$	−3.58	−2.93	−4.15	−3.50
$T = 100$	−3.51	−2.89	−4.04	−3.45
$T = 250$	−3.46	−2.88	−3.99	−3.43
$T = 500$	−3.44	−2.87	−3.98	−3.42
$T = \infty$	−3.43	−2.86	−3.96	−3.41

Table 9.2 Asymptotic critical values residual unit root tests for cointegration (with constant term) (Davidson and MacKinnon, 1993)

Number of variables (incl. Y_t)	Significance level 1%	5%	10%
2	−3.90	−3.34	−3.04
3	−4.29	−3.74	−3.45
4	−4.64	−4.10	−3.81
5	−4.96	−4.42	−4.13

Asymptotic critical values of the two-sided t-test from the standard normal distribution

1%	5%	10%
2.58	1.96	1.64