This paper is not to be removed from the Examination Halls

UNIVERSITY OF LONDON **EC2020 ZA**

BSc degrees and Diplomas for Graduates in Economics, Management, Finance and the Social Sciences, the Diplomas in Economics and Social Sciences and Access Route

Elements of Econometrics

Monday, 11 May 2015 : 14:30 to 17:30

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Extracts from statistical tables are given after the final question on this paper.

Graph paper is provided at the end of this question paper. If used, it must be detached and fastened securely inside the answer book.

A calculator may be used when answering questions on this paper and it must comply in all respects with the specification given with your Admission Notice. The make and type of machine must be clearly stated on the front cover of the answer book.

If more questions are answered than requested, only the first answers attempted will be counted.

PLEASE TURN OVER

© University of London 2015

SECTION A

Answer all questions from this section.

1. Explain the concept of consistency of an estimator. Show that in a simple regression model of Y_i on X_i, the ordinary least squares estimate of the slope is consistent. **(8 marks)**

2. In the model

$$y_t = \alpha x_t + u_t \; ; \quad t = 1, 2, \ldots, T$$

x_t is an explanatory variable which can be regarded as fixed in repeated samples.

u_t is an unobserved disturbance for which it is assumed that

$$E(u_t) = 0$$

$$E(u_s u_t) = \sigma^2 \quad \text{if } s = t$$

$$\qquad\quad = 0 \quad \text{if } s \neq t$$

An estimator of α is $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{y_t}{x_t}\right)$.

Under the assumptions above show that the estimator is unbiased and consistent. Comment briefly on the efficiency of the estimator. **(8 marks)**

3. (a) If a random variable X has a distribution with probability density function $f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, show that the maximum likelihood (ML) estimator of the mean (μ) of the random variable X is the sample mean. **(4 marks)**

(b) State the statistical properties of the ML estimators. **(4 marks)**

© University of London 2015

4. Suppose that business expenditure for a new plant (Y_t) is explained by the relation

$$\ln(Y_t) = \alpha + \beta \ln(X_t^*) + u_t,$$

where u_t is a random variable, ln is the natural logarithm and X_t^* is the level of expected sales (which is unobserved) and is formed by

$$\ln(X_t^*) - \ln(X_{t-1}^*) = \gamma\left(\ln(X_{t-1}) - \ln(X_{t-1}^*)\right).$$

X_t is the level of actual sales. Derive a linear relationship that can be used to estimate α and β, using the observable variables Y_t and X_t. **(8 marks)**

5. Discuss how dummy variables can be used to test

(a) change in intercept, **(3 marks)**

(b) change in slope and, **(3 marks)**

(c) changes in both intercept and slope. **(2 marks)**

© University of London 2015

SECTION B

Answer three questions from this section.

6. The Cobb-Douglas production function can be written as follows:

$$\ln Y_t = \alpha_0 + \alpha_1 \ln L_t + \alpha_2 \ln K_t + u_t \quad ; \quad t = 1,2,\ldots,T \qquad \text{(i)}$$

where Y_t is real output, L_t is a measure of labour input, K_t is a measure of real capital input, and u_t is an unobserved random disturbance with $E(u_t) = 0$.

The following estimates of (i) were obtained by ordinary least squares (OLS) using 15 annual observations from the Taiwanese agricultural sector.

$$\ln Y_t = \underset{(2.44)}{-3.329} + \underset{(0.54)}{1.498} \ln L_t + \underset{(0.10)}{0.489} \ln K_t + e_t \qquad \text{(ii)}$$

where e_t are OLS residuals, standard errors are in parentheses and $R^2 = 0.89$.

(a) Give an economic interpretation of the estimated coefficients. Are the estimated slope parameters of the expected sign? Explain. **(4 marks)**

(b) Test the slope parameters for significance, and explain what assumptions your tests require in order to be valid. **(6 marks)**

(c) If the Taiwanese agricultural sector has constant returns to scale then $\alpha_1 + \alpha_2 = 1$. **Discuss** whether the estimates in (ii) support this restriction. **(3 marks)**

(d) The equation was also estimated in the following restricted form

$$[\ln Y_t - \ln L_t] = \underset{(0.42)}{1.712} + \underset{(0.09)}{0.612} [\ln K_t - \ln L_t] + \upsilon_t \qquad \text{(iii)}$$

where υ_t are OLS residuals, standard errors are in parentheses and $R^2 = 0.77$.

Test the restriction(s) in (iii), and show that (iii) incorporates the restriction of constant returns to scale. **(7 marks)**

© University of London 2015

7 (a) Explain the problem of identification in the context of simultaneous equation models. **(3 marks)**

(b) In the model

$$y_{1t} = \alpha y_{2t} + u_{1t}$$
$$y_{2t} = \beta_1 x_t + \beta_2 y_{1t} + u_{2t} \quad t = 1, 2, \ldots, T$$

where x_t is an exogenous variable.

i. Examine the identification of both equations. **(4 marks)**

ii. Obtain the ordinary least squares estimator of α, and examine its consistency. **(7 marks)**

iii. Derive the two-stage least squares of α and also prove its consistency stating carefully any assumptions you need. **(3 marks)**

(c) What is meant by indirect least squares? Explain. **(3 marks)**

8. A study of applications for home mortgages used the linear probability model

$$MORT_i = \beta_0 + \beta_1 INC_i + \beta_2 AGE_i + \beta_3 PROP_i + u_i \; ; \quad i = 1, 2, \ldots, 700$$

where

$MORT_i$ = 1 if a mortgage is granted to the i-th applicant: 0 otherwise
INC_i = income of the i-th applicant in thousands of pounds
AGE_i = age of the i-th applicant in years
$PROP_i$ = age of the property for which the mortgage is being applied.

(a) The estimated coefficient for INC_i was 1.02 with standard error 0.51. What is the interpretation of this coefficient? **(4 marks)**

(b) Why is R^2 meaningless in probit and logit models? What measures of 'goodness of fit' are applicable to probit and logit models? **(7 marks)**

(c) Using a two variable linear model, show that the ordinary least squares estimator will be heteroscedastic if the dependent variable takes only values 0 and 1. **(9 marks)**

© University of London 2015

9. (a) Explain the meaning of spurious regression. **(4 marks)**

(b) The following equations were estimated by ordinary least squares.

$$Y_t = \underset{(0.1305)}{3.0920} + \underset{(0.0103)}{0.6959}\, X_t + \hat{u}_t \qquad (1)$$

$R^2 = 0.99$, $F = 4523.25$, $s = 0.0236$, $DW = 0.557$, $T = 740$.

$$\Delta\hat{u}_t = -\underset{(0.0845)}{0.2161}\,\hat{u}_{t-1} + \underset{(0.1592)}{0.2349}\,\Delta\hat{u}_{t-1} + \underset{(0.1631)}{0.2029}\,\Delta\hat{u}_{t-2} + \hat{\varepsilon}_t \qquad (2)$$

$R^2 = 0.1799$; $s = 0.0115$; $T = 737$.

Where s is the standard error of the residuals, T is the number of observations, $\hat{u}_t$ and $\hat{\varepsilon}_t$ are OLS residuals, and standard errors are in parentheses.

Do the results above indicate that Y_t and X_t are cointegrated? Specify clearly all the assumptions you have made. **(6 marks)**

[Note: Critical value at 5% significant level from MacKinnon table is -3.3377].

(c) Consider a model

$$y_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}\,; \quad t = 1, 2, \ldots, T$$

$E(u_t) = 0$; $E(u_t^2) = \sigma^2$; and $E(u_s u_t) = 0$ for all $s, t = 1, 2, \ldots, T$.

i. Is y_t stationary? Explain. **(5 marks)**

ii. Calculate the autocorrelation function of y_t. **(5 marks)**

© University of London 2015

10. Let the regression equation be

$$Y_t = \beta_1 + \beta_2 X_t + u_t \; ; \quad t = 1, 2, \ldots, T$$

where

$$u_t = \rho u_{t-1} + \varepsilon_t \quad \text{for all } t \quad ; \quad |\rho| < 1$$
$$E(\varepsilon_t) = 0$$
$$E(\varepsilon_s \varepsilon_t) = \sigma_\varepsilon^2 \quad \text{if } s = t$$
$$= 0 \quad \text{if } s \neq t$$

(a) Derive

i. the variance of u_t ; and

ii. $E(u_t u_{t-1})$. **(7 marks)**

(b) Explain the consequences of this model specification on ordinary least squares estimators for β_1 and β_2. **(3 marks)**

(c) Explain how would you test the null hypothesis $H_0 : \rho = 0$ against the alternative $H_1 : \rho \neq 0$. Specify all the assumptions needed for this test. **(5 marks)**

(d) Discuss in detail a method of estimation which gives best linear unbiased estimates of β_1 and β_2 . **(5 marks)**

END OF PAPER

© University of London 2015

This paper is not to be removed from the Examination Halls

UNIVERSITY OF LONDON **EC2020 ZB**

BSc degrees and Diplomas for Graduates in Economics, Management, Finance and the Social Sciences, the Diplomas in Economics and Social Sciences and Access Route

Elements of Econometrics

Monday, 11 May 2015 : 14:30 to 17:30

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Extracts from statistical tables are given after the final question on this paper.

Graph paper is provided at the end of this question paper. If used, it must be detached and fastened securely inside the answer book.

A calculator may be used when answering questions on this paper and it must comply in all respects with the specification given with your Admission Notice. The make and type of machine must be clearly stated on the front cover of the answer book.

If more questions are answered than requested, only the first answers attempted will be counted.

PLEASE TURN OVER

© University of London 2015

SECTION A

Answer all questions from this section.

1. Let the regression equation be:

$$y_t = \beta x_t + u_t \quad ; \quad t = 1, 2, \ldots, T.$$

Where $E(u_t) = 0$; $E(u_t^2) = \sigma^2$ and $E(u_s u_t) = 0$ if $s \neq t$. x values are fixed in repeated samples.

Obtain the ordinary least squares estimator (OLS) of β. Show that the OLS estimator of β is linear and unbiased. **(8 marks)**

2. Suppose that business expenditure for a new plant (Y_t) is explained by the relation

$$\ln(Y_t) = \alpha + \beta \ln(X_t^*) + u_t,$$

where u_t is a random variable, ln is the natural logarithm and X_t^* is the level of expected sales (which is unobserved) and is formed by

$$\ln(X_t^*) - \ln(X_{t-1}^*) = \gamma\left(\ln(X_{t-1}) - \ln(X_{t-1}^*)\right).$$

X_t is the level of actual sales. Derive a linear relationship that can be used to estimate α and β, using the observable variables Y_t and X_t. **(8 marks)**

3. Let the probability density function of a random variable X be $f(x;\theta)$. Explain the procedure to use likelihood ratio test for testing the null hypothesis $H_0 : \theta = \theta_0$. **(8 marks)**

4. What is meant by a common factor test in the context of a linear model with an autocorrelated error? How would you perform a common factor test and what hypothesis would you be testing? **(8 marks)**

5. Consider the model

$$y_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2} \quad ; t = 1, 2, \ldots, T$$

$E(u_t) = 0$; $E(u_t^2) = \sigma^2$; and $E(u_s u_t) = 0$ if $s \neq t$ for all $s, t = 1, 2, \ldots, T$.

Is y_t stationary? Explain in detail. **(8 marks)**

© University of London 2015

SECTION B

Answer three questions from this section.

6. Let the model be:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + u_t \quad ; t = 1, 2, \ldots, T$$

$E(u_t) = 0$ for all t. A researcher suspects that the variance of the disturbance term is $Var(u_t) = \sigma^2 X_{t1}$.

(a) Explain how to test the null hypothesis $H_0 : Var(u_t) = \sigma^2$ against the alternative hypothesis $H_1 : Var(u_t) = \sigma^2 X_{t1}$, for all $t = 1, 2, \ldots, T$. **(6 marks)**

(b) Suggest in detail an estimation procedure which gives best linear unbiased estimates of the parameters. **(6 marks)**

(c) Consider the model given by

$$y_t = \alpha x_t + u_t \; ; \quad t = 1, 2, \ldots, T$$

where $E(u_t) = 0$; $E(u_t^2) = \sigma^2 x_t^2$; $E(u_s u_t) = 0$ if $s \neq t$, for all s and t. x_t is an observed non-random variable.

The density function of u_t is

$$f(u_t; x, t) = \left(2\pi\sigma^2 x_t^2\right)^{-1/2} \exp\left[-\frac{1}{2}\left(\frac{u_t}{\sigma x_t}\right)^2\right].$$

Derive the maximum likelihood estimators of α and σ^2. **(8 marks)**

© University of London 2015

7. (a) Let the model be

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + u_t \; ; \quad t = 1, 2, \ldots, T .$$

Suppose the relevant variable Z_t was omitted from the model. Discuss the effects of this omission on unbiasedness and consistency of the ordinary least squares estimators. **(5 marks)**

(b) The following regression was estimated by ordinary least squares (OLS) on 745 observations of annual data from the UK.

$$q_t = \underset{(1.035)}{18.210} - \underset{(0.087)}{0.718}\, y_t + e_{1t} \quad ; t = 1, 2, \ldots, 745 \quad \text{(i)}$$

$$R^2 = 0.613, \quad F = 67.97, \quad dw = 1.68$$

where q_t is the log of the quantity of cigarettes purchased, y_t is the log of real disposable income. Standard errors are in brackets. e_{1t} is an OLS residual. dw is the Durbin-Watson statistic.

Test the null hypothesis that the coefficient of y_t is -1. Comment on the result of your test. Do you think the estimates in (i) are plausible? Explain. **(5 marks)**

(c) A second regression was estimated using the same data sample

$$q_t = \underset{(2.214)}{0.399} - \underset{(0.041)}{0.343}\, p_t + \underset{(0.181)}{0.737}\, y_t + e_{2t} \; ; \quad t = 1, 2 \ldots, 745 \quad \text{(ii)}$$

$$R^2 = 0.855, \quad F = 124.28, \quad dw = 2.01$$

where the variables are defined as before, p_t is the log of the relative prices of cigarettes, e_{2t} is an OLS residual. Standard errors in brackets. dw is the Durbin-Watson statistic.

Test the joint hypothesis that both slope coefficients in (ii) are zero. **(2 marks)**

(d) Compare the results in (ii) with those in (i). Provide an explanation in terms of omitted variable bias of the difference between the estimates of the coefficient of y_t in the two equations. **(8 marks)**

© University of London 2015

8. Consider the simple linear regression model

$$Y_t = \beta_0 + \beta_1 X_t + u_t \quad ; t = 1, 2, \ldots, T$$

where β_0 and β_1 are unknown coefficients. A "goodness of fit" measure for ordinary least squares estimation of the above model is defined by

$$R^2 = \frac{\text{Explained Sum of Squares (ESS)}}{\text{Total Sum of Squares (TSS)}}$$

(a) Prove that

$$R^2 = 1 - \frac{RSS}{TSS} \ ,$$

where the residual sum of squares $RSS = \sum_{t=1}^{T} \hat{u}_t^2$, $\hat{u}_t = Y_t - \hat{\beta}_0 - \hat{\beta}_1 X_t$;

$TSS = \sum_{t=1}^{T} (Y_t - \bar{Y})^2$; and $ESS = \sum_{t=1}^{T} (\hat{Y}_t - \bar{Y}_t)^2$ for all t. **(8 marks)**

(b) Hence, or otherwise, show that

$0 \le R^2 \le 1$. **(4 marks)**

(c) Briefly explain what happens to R^2 if an extra explanatory variable is added to the regression model. **(4 marks)**

(d) Has R^2 any drawbacks? Explain. **(4 marks)**

© University of London 2015

9. A study of applications for home mortgages used the linear probability model

$$MORT_i = \beta_0 + \beta_1 INC_i + \beta_2 AGE_i + \beta_3 PROP_i + u_i \; ; \quad i = 1, 2, \ldots, 700$$

where

$MORT_i$ = 1 if a mortgage is granted to the i-th applicant: 0 otherwise
INC_i = income of the i-th applicant in thousands of pounds
AGE_i = age of the i-th applicant in years
$PROP_i$ = age of the property for which the mortgage is being applied.

(a) The estimated coefficient for INC_i was 1.02 with standard error 0.51. What is the interpretation of this coefficient? **(4 marks)**

(b) Why is R^2 meaningless in probit and logit models? What measures of 'goodness of fit' are applicable to probit and logit models? **(7 marks)**

(c) Using a two variable linear model, show that the ordinary least squares estimator will be heteroscedastic if the dependent variable takes only values 0 and 1. **(9 marks)**

10. Consider a two equation linear model

$q_t = \beta_1 p_t + \beta_2 y_t + u_t$ (i) demand equation
$q_t = \alpha p_t + e_t$ (ii) supply equation

$t = 1, 2, \cdots, T$; $E(u_t) = E(e_t) = 0; E(u_t^2) = \sigma_u^2; E(e_t^2) = \sigma_e^2; E(u_t e_t) = \sigma_{ue}$;
$E(u_s e_t) = 0$ if $s \neq t$, for all $s, t = 1, 2, \cdots, T$ and variables are defined as:

q_t = demand for good
p_t = price for good
y_t = personal disposable income.
Quantity demanded is equal to the quantity supplied
u_t and e_t are disturbance terms.

(a) Examine the identifiability of the above two equations. **(6 marks)**

(b) Derive the two-stage least squares estimator of α and also examine its consistency. **(10 marks)**

(c) Discuss, without derivation, what is the indirect least squares estimator of α. **(4 marks)**

END OF PAPER

© University of London 2015

Examiners' commentaries 2015

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2014–15. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2014). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

General remarks

Learning outcomes

At the end of this course, and having completed the Essential reading and activities, you should be able to:

- describe and apply the classical regression model and its application to cross-section data
- describe and apply the:
  - Gauss–Markov conditions and other assumptions required in the application of the classical regression model
  - reasons for expecting violations of these assumptions in certain circumstances
  - tests for violations
  - potential remedial measures, including, where appropriate, the use of instrumental variables
- recognise and apply the advantages of logit, probit and similar models over regression analysis when fitting binary choice models
- competently use regression, logit and probit analysis to quantify economic relationships using standard regression programmes (Stata and EViews) in simple applications
- describe and explain the principles underlying the use of maximum likelihood estimation
- apply regression analysis to fit time-series models using stationary time series, with awareness of some of the econometric problems specific to time series applications (for example, autocorrelation) and remedial measures
- recognise the difficulties that arise in the application of regression analysis to nonstationary time series, know how to test for unit roots, and know what is meant by cointegration.

Common mistakes committed by candidates

- A large number of candidates were not able to distinguish between sample variance and covariance, and population variance and covariance (this is happening year after year). They treat them as the same. This results in incorrect analysis and candidates lose significant marks.

  Consider an example: Suppose data is deviation from respective sample means and the regression model is:

  $$y_t = \beta x_t + u_t, \quad t = 1, 2, \ldots, T.$$

  The ordinary least squares estimator of β is:

  $$\widehat{\beta} = \frac{\sum_{t=1}^{T} x_t y_t}{\sum_{t=1}^{T} x_t^2} = \beta + \frac{\sum_{t=1}^{T} x_t u_t}{\sum_{t=1}^{T} x_t^2}.$$

  In terms of variances and covariances (a large number of candidates prefer this terminology) this can be written as:

  $$\widehat{\beta} = \beta + \frac{\mathrm{Cov}(x, u)}{\mathrm{Var}(x)}.$$

  Here $\mathrm{Cov}(x, u)$ and $\mathrm{Var}(x)$ are sample[$\mathrm{Cov}(x, u)$] and sample[$\mathrm{Var}(x)$].

  Candidates should realise that $\sum_{t=1}^{T} u_t$, $\sum_{t=1}^{T} x_t u_t$, $\mathrm{Cov}(x, u)$ and $\mathrm{Var}(x)$ given above are sample moments and as such $\sum_{t=1}^{T} u_t \neq 0$, $\sum_{t=1}^{T} x_t u_t \neq 0$ and $\mathrm{Cov}(x, u) \neq 0$. But, if we take expectation, then:

  $$\mathrm{E}[u_t] = 0$$

  by assumption. Then:

  $$\mathrm{E}\left[\sum_{t=1}^{T} x_t u_t\right] = \sum_{t=1}^{T} x_t [\mathrm{E}(u_t)] = 0$$

  as the x_t are fixed they can be taken out of the expectation, and so:

  $$\mathrm{E}[\mathrm{Cov}(x, u)] = \mathrm{E}\left[\frac{1}{T}\sum_{t=1}^{T} x_t u_t\right] = 0$$

  as previously argued. This makes $\mathrm{E}(\widehat{\beta}) = \beta$, i.e. $\widehat{\beta}$ is an unbiased estimator for β.

  To prove consistency take plim to get:

  $$\begin{aligned}
  \mathrm{plim}\left(\widehat{\beta}\right) &= \beta + \mathrm{plim}\left(\frac{\frac{1}{T}\sum_{t=1}^{T} x_t u_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2}\right) \\
  &= \beta + \frac{\mathrm{plim}\left(\frac{1}{T}\sum_{t=1}^{T} x_t u_t\right)}{\mathrm{plim}\left(\frac{1}{T}\sum_{t=1}^{T} x_t^2\right)} \\
  &= \beta + \frac{\mathrm{plim}(\text{sample } \mathrm{Cov}(x, u))}{\mathrm{plim}(\text{sample } \mathrm{Var}(x))} \\
  &= \beta + \frac{\text{population } \mathrm{Cov}(x, u)}{\text{population } \mathrm{Var}(x)}.
  \end{aligned}$$

By assumption, population $\text{Cov}(x, u) = 0$ and population $\text{Var}(x) > 0$, hence $\text{plim}\left(\widehat{\beta}\right) = \beta$, in other words $\widehat{\beta}$ is a consistent estimator of β.

Remember that in general:

$$\begin{aligned} \text{plim(sample variance)} &= \text{population variance,} \\ \text{plim(sample covariance)} &= \text{population covariance.} \end{aligned}$$

This concept has been used in many questions. This simple mistake of not distinguishing between sample variance and covariance and population variance and covariance results in a significant loss of marks which might result in the loss of a degree class or even be the difference between pass and fail.

- Candidates struggled to give competent answers to the interpretation of empirical results. When interpreting an empirical result you should discuss the significance of the coefficients, magnitude and sign of the coefficients. Also, you should make sure that the GM conditions hold.
- Just as last year, many candidates did not appear to read the questions carefully enough and often omitted to give answers to parts of questions which asked for details of such things as the assumptions necessary for a particular result to be true.

Key steps to improvement

Essential reading for this course includes the subject guide and the following.

> Dougherty, C. *Introduction to econometrics.* (Oxford: Oxford University Press, 2011) 4th edition [ISBN 9780199567089];
> http://global.oup.com/uk/orc/busecon/economics/dougherty4e/

Apart from Essential readings you should do some supplementary readings. One very good book of the same level is:

> Gujarati, D.N. and D.C. Porter *Basic econometrics.* (McGraw–Hill, 2009, International edition) 5th edition [ISBN 9780071276252].

To understand the subject clearly it is important to supplement C. Dougherty, *Introduction to econometrics* (fourth edition) with the subject guide **EC2020 Elements of econometrics** (2014), especially Chapter 10 which covers maximum likelihood.

It is very important to carefully go through the subject guide. The subject guide contains solutions to the questions given in the main book and also some additional questions and solutions. Working through these will improve your understanding of the subject.

The chapter in the subject guide on maximum likelihood (Chapter 10) includes some additional theory which has not been covered in the main book. It is important to read the additional theory given in the subject guide to have a better understanding of the principles of maximum likelihood and tests based on the likelihood function.

Please check the VLE course page for resources for this subject such as a downloadable copy of the subject guide **EC2020 Elements of econometrics** (2014), PowerPoint slideshows that provide graphical treatment of the topics covered in the text, datasets and statistical tables. Candidates should utilise data sets using standard regression programmes (STATA or EViews). This will help in the understanding of the subject.

Examination revision strategy

Many candidates are disappointed to find that their examination performance is poorer than they expected. This may be due to a number of reasons. The *Examiners' commentaries* suggest ways of addressing common problems and improving your performance. One particular failing is '**question spotting**', that is, confining your examination preparation to a few questions and/or topics which have come up in past papers for the course. This can have serious consequences.

We recognise that candidates may not cover all topics in the syllabus in the same depth, but you need to be aware that examiners are free to set questions on **any aspect** of the syllabus. This means that you need to study enough of the syllabus to enable you to answer the required number of examination questions.

The syllabus can be found in the Course information sheet in the section of the VLE dedicated to each course. You should read the syllabus carefully and ensure that you cover sufficient material in preparation for the examination. Examiners will vary the topics and questions from year to year and may well set questions that have not appeared in past papers. Examination papers may legitimately include questions on any topic in the syllabus. So, although past papers can be helpful during your revision, you cannot assume that topics or specific questions that have come up in past examinations will occur again.

If you rely on a question-spotting strategy, it is likely you will find yourself in difficulties when you sit the examination. We strongly advise you not to adopt this strategy.

Examiners' commentaries 2015

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2014–15. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2014). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone A

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Section A

Answer all questions from this section.

Question 1

Explain the concept of consistency of an estimator. Show that in a simple regression model of Y_i on X_i, the ordinary least squares estimate of the slope is consistent.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (4th edition) Chapters R.14 (Probability limits and consistency) and 8.3 (Asymptotic properties of OLS regression estimators).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 3A.7 (Consistency of least-squares estimators).

Approaching the question

Definition of consistency is required. Sufficient condition of consistency should also be given. Probability limit or sufficient condition of consistency should be used to show the consistency of

the OLS estimator of the slope. Answer is:

Definition:

$\widehat{\beta}$ is a consistent estimator of β if:

$$\lim_{T\to\infty} P(|\widehat{\beta}-\beta|>\varepsilon)\to 0$$

in short plim $\widehat{\beta}=\beta$. ε is an arbitrarily small positive number.

Sufficient condition for consistency:

i. $\mathrm{E}(\widehat{\beta})=\beta$ or Asy. $\mathrm{E}(\widehat{\beta})\to\beta$.

ii. $\mathrm{Var}(\widehat{\beta})\to 0$ as $T\to\infty$, where T is the sample size.

If the sufficient condition holds then the definition holds.

Consistency of the estimator of the slope parameter:

Let the model be:

$$Y_t=\beta_1+\beta_2X_t+u_t;\quad t=1,2,\ldots,T.$$

OLS estimator of β_2 is:

$$\widehat{\beta}_2=\frac{\sum_{t=1}^{T}(X_t-\bar{X})(Y_t-\bar{Y})}{\sum_{t=1}^{T}(X_t-\bar{X})^2}=\beta_2+\frac{\sum_{t=1}^{T}(X_t-\bar{X})(u_t-\bar{u})}{\sum_{t=1}^{T}(X_t-\bar{X})^2}.$$

$$\text{plim}\ \widehat{\beta}_2=\beta_2+\frac{\text{plim}\frac{1}{T}\sum_{t=1}^{T}(X_t-\bar{X})(u_t-\bar{u})}{\text{plim}\frac{1}{T}\sum_{t=1}^{T}(X_t-\bar{X})^2}=\beta_2+\frac{\sigma_{Xu}}{\sigma_X^2}=\beta_2+\frac{0}{\sigma_X^2}=\beta_2\ \Rightarrow\ \text{consistent.}$$

σ_{Xu} is the population covariance between X and u. By assumption it is zero. σ_X^2 is the population variance of X and it is >0.

Question 2

In the model

$$\boldsymbol{y_t=\alpha x_t+u_t;\quad t=1,2,\ldots,T}$$

x_t is an explanatory variable which can be regarded as fixed in repeated samples.

u_t is an unobserved disturbance for which it is assumed that

$$\begin{aligned}\mathbf{E}(u_t)&=0\\ \mathbf{E}(u_su_t)&=\sigma^2\ \text{if}\ s=t\\ &=0\ \text{if}\ s\neq t\end{aligned}$$

An estimator of α is $\frac{1}{T}\sum_{t=1}^{T}\left(\frac{y_t}{x_t}\right)$.

Under the assumptions above show that the estimator is unbiased and consistent. Comment briefly on the efficiency of the estimator.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapters R.6 (Unbiasedness and efficiency), R.14 (Probability limits and consistency) and 8.3 (Asymptotic properties of OLS regression estimators).

Gujarati, D.N. and D.C. Porter. *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 3 A.7 (Consistency of least-squares estimators).

Approaching the question

Sufficient condition of the consistency should be used to show the consistency of the given estimator. Answer is:

$$\widehat{\alpha} = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{y_t}{x_t}\right) = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{\alpha x_t + u_t}{x_t}\right) = \alpha + \frac{1}{T}\sum_{t=1}^{T}\left(\frac{u_t}{x_t}\right)$$

hence:

$$\mathrm{E}(\widehat{\alpha}) = \alpha + \frac{1}{T}\sum_{t=1}^{T}\frac{\mathrm{E}(u_t)}{x_t} = \alpha \;\Rightarrow\; \text{unbiased.}$$

To show consistency, the sufficient condition of consistency will be used.

$$\mathrm{Var}(\widehat{\alpha}) = \mathrm{E}\left((\alpha - \widehat{\alpha})^2\right) = \mathrm{E}\left(\frac{1}{T}\sum_{t=1}^{T}\left(\frac{u_t}{x_t}\right)\right)^2 = \sigma^2\frac{1}{T^2}\sum_{t=1}^{T}\left(\frac{1}{x_t^2}\right)$$

which will tend to zero as $T \to \infty$.

As the estimator is unbiased and also the variance of the estimator tends to zero as $T \to \infty$, the sufficient condition of consistency holds, hence the estimator is consistent. The given estimator is not efficient because under the assumptions above, the ordinary least squares estimator is the most efficient estimator.

Question 3

(a) If a random variable X has a distribution with probability density function $f(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$, show that the maximum likelihood (ML) estimator of the mean (μ) of the random variable X is the sample mean.

(4 marks)

(b) State the statistical properties of the ML estimators.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapter 10.6 (An introduction to maximum likelihood estimation).

Dougherty, C. Subject guide, Chapter 10 (Binary choice and limited dependent variable models, and maximum likelihood estimation).

Approaching the question

(a) Log likelihood function should be derived. It should be differentiated with respect to and equated to zero to obtain the ML estimator. Detailed answer is:

The likelihood function is defined by:

$$L = \prod_{i=1}^{n} \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(X_i-\mu)^2}{2\sigma^2}\right) = (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{\sum(X_i-\mu)^2}{2\sigma^2}\right).$$

Hence, taking logs we get:

$$\ln L = -\frac{n}{2}\ln(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2.$$

To maximise differentiate with respect to μ, and set the derivative equal to zero.

$$\frac{\partial \ln L}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu) = 0$$

which gives:

$$\tilde{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}.$$

(b) Properties of ML estimators should be discussed. Answer is:

- MLE is consistent.
- MLE is invariant to the transformation of parameters; for example, if $\widehat{\theta}$ is the MLE of θ, then $\widehat{\theta}^2$ is the MLE of θ^2. If $\widehat{\theta}$ is the MLE of θ, then $\exp(\widehat{\theta})$ is the MLE of $\exp(\theta)$.
- MLE is efficient in large samples in the sense that variance of the MLE reaches the Cramer–Rao lower bound (CRLB) in large samples.
- MLE is asymptotically normally distributed.
- If a sufficient estimator exists, then MLE is a function of the sufficient estimator.

Question 4

Suppose that business expenditure for a new plant (Y_t) is explained by the relation

$$\ln(Y_t) = \alpha + \beta\ln(X_t^*) + u_t,$$

where u_t is a random variable, ln is the natural logarithm and X_t^* is the level of expected sales (which is unobserved) and is formed by

$$\ln(X_t^*) - \ln(X_{t-1}^*) = \gamma(\ln(X_{t-1}) - \ln(X_{t-1}^*)).$$

X_t is the level of actual sales. Derive a linear relationship that can be used to estimate α and β, using the observable variables Y_t and X_t.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapter 11.4 (Models with lagged dependent variable).

Gujarati, D.N. and D.C. Porter. *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 17.5 (Rationalization of the Koyck model: The adaptive expectations model).

Approaching the question

To get an estimable equation we need to eliminate the expected values from the equation. Multiply through by $(1-\gamma)$ and lag to get a new equation. Subtract the new equation from the original equation to get the result. Answer is:

Expectations are formed by:

$$\ln(X_t^*) - (1-\gamma)\ln(X_{t-1}^*) = \gamma \ln(X_{t-1}).$$

To get an estimable equation we need to eliminate the expected values from the equation. Multiply through by $(1-\gamma)$ and lag to give:

$$(1-\gamma)\ln(Y_{t-1}) = (1-\gamma)\alpha + (1-\gamma)\beta \ln(X_{t-1}^*) + (1-\gamma)u_{t-1}.$$

Now subtract from the original equation:

$$\begin{aligned}
\ln(Y_t) - (1-\gamma)\ln(Y_{t-1}) &= \alpha\gamma + \beta(\ln(X_t^*) - (1-\gamma)\ln(X_{t-1}^*)) + (u_t - (1-\gamma)u_{t-1}) \\
\ln(Y_t) - (1-\gamma)\ln(Y_{t-1}) &= \alpha\gamma + \beta(\gamma \ln(X_{t-1})) + (u_t - (1-\gamma)u_{t-1})
\end{aligned}$$

or:

$$\ln(Y_t) = \alpha\gamma + \beta\gamma\ln(X_{t-1}) + (1-\gamma)\ln(Y_{t-1}) + (u_t - (1-\gamma)_{t-1}).$$

Parameters are estimated by non-linear techniques. If these procedures are not available then a grid search can be used where γ is given values between 0 and 1 in steps of 0.1 and the remaining parameters estimated by OLS.

Question 5

Discuss how dummy variables can be used to test

(a) change in intercept,

(3 marks)

(b) change in slope and,

(3 marks)

(c) changes in both intercept and slope.

(2 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapters 5.1 (Illustration of the use of a dummy variable) and 5.3 (Slope dummy variables).

Dougherty, C. Subject guide, Chapter 5 (Dummy variables).

Gujarati, D.N. and D.C. Porter. *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 9 (Dummy variable regression models).

Approaching the question

(a) **Only intercept has changed.**

Specify the model as:

$$Y_t = \beta_0 + \beta_1 X_t + \alpha Z_t + u_t; \quad t = 1, 2, \ldots, T \qquad \text{(i)}$$

where Z_t is a dummy variable which takes values:

$$Z_t = \begin{cases} 1 & \text{for war period} \\ 0 & \text{for peace period.} \end{cases}$$

Estimating (i) by OLS we get:

$$Y_t = \widehat{\beta}_0 + \widehat{\beta}_1 X_t + \widehat{\alpha} Z_t + \widehat{u}_t; \quad t = 1, 2, \ldots, T \qquad \text{(ii)}$$

From (ii), we can write two separate regressions for two different periods as:

$$\begin{aligned} Y_t &= (\widehat{\beta}_0 + \widehat{\alpha}) + \widehat{\beta}_1 X_t + \widehat{u}_t \qquad \text{(war period)} \qquad \text{(iii)} \\ &= \widehat{\beta}_0 + \widehat{\beta}_1 X_t + \widehat{u}_t \qquad \text{(peace period)} \qquad \text{(iv)} \end{aligned}$$

To test whether the intercept has changed or not, the hypotheses are:

$\mathrm{H}_0 : \alpha = 0$ (intercept has not changed).

$\mathrm{H}_1 : \alpha \neq 0$ (intercept has changed).

It can be tested by a t test.

(b) **Only slope has changed.**

Specify the model as:

$$Y_t = \beta_0 + \beta_1 X_t + \alpha X_t Z_t + u_t; \quad t = 1, 2, \ldots, T \qquad \text{(v)}$$

where Z_t is a dummy variable which takes values:

$$Z_t = \begin{cases} 1 & \text{for war period} \\ 0 & \text{for peace period.} \end{cases}$$

Estimating (v) by OLS we get:

$$Y_t = \widehat{\beta}_0 + \widehat{\beta}_1 X_t + \widehat{\alpha} X_t Z_t + \widehat{u}_t; \quad t = 1, 2, \ldots, T \qquad \text{(vi)}$$

From (vi), we can write two separate regressions for two different periods as:

$$\begin{aligned} Y_t &= \widehat{\beta}_0 + (\widehat{\beta}_1 + \widehat{\alpha}) X_t + \widehat{u}_t \qquad \text{(war period)} \qquad \text{(vii)} \\ &= \widehat{\beta}_0 + \widehat{\beta}_1 X_t + \widehat{u}_t \qquad \text{(peace period)} \qquad \text{(viii)} \end{aligned}$$

To test whether the slope has changed or not, the hypotheses are:

$\mathrm{H}_0 : \alpha = 0$ (slope has not changed).

$\mathrm{H}_1 : \alpha \neq 0$ (slope has changed).

This can be tested by a t test. If we do not reject H_0 then we can apply OLS to (v) and get the estimated parameters. If H_0 is rejected then our estimated equations for two different periods are given by (vii) and (viii).

(c) **Intercept and slope both have changed.**

Specify the model as:

$$Y_t = \beta_0 + \beta_1 X_t + \alpha_1 Z_t + \alpha_2 X_t Z_t + u_t; \quad t = 1, 2, \ldots, T. \qquad \text{(ix)}$$

where Z_t is a dummy variable which takes values:

$$Z_t = \begin{cases} 1 & \text{for war period} \\ 0 & \text{for peace period.} \end{cases}$$

Estimating (ix) by OLS we get:

$$Y_t = \widehat{\beta}_0 + \widehat{\beta}_1 X_t + \widehat{\alpha}_1 Z_t + \widehat{\alpha}_2 X_t Z_t + \widehat{u}_t; \quad t = 1, 2, \ldots, T \qquad \text{(x)}$$

From (x), we can write two separate regressions for two different periods as:

$$\begin{aligned} Y_t &= (\widehat{\beta}_0 + \widehat{\alpha_0}) + (\widehat{\beta}_1 + \widehat{\alpha}_2) X_t + \widehat{u}_t \qquad \text{(war period)} \qquad \text{(xi)} \\ &= \widehat{\beta}_0 + \widehat{\beta}_1 X_t + \widehat{u}_t \qquad \text{(peace period)} \qquad \text{(xii)} \end{aligned}$$

To test jointly whether both intercept and slope has changed or not hypotheses are:

$\mathrm{H}_0 : \alpha_1, \alpha_2 = 0$ (both intercept and slope have not changed).

$\mathrm{H}_1 : \alpha_1, \alpha_2 \neq 0$ (both intercept and slope have changed).

This can be tested by an F test. If we do not reject H_0 then we can apply OLS to (ix) and get the estimated parameters. If H_0 is rejected then our estimated equations for two different periods are given by (xi) and (xii).

Section B

Answer three questions from this section.

Question 6

The Cobb–Douglas production function can be written as follows:

$$\ln Y_t = \alpha_0 + \alpha_1 \ln L_t + \alpha_2\, nK_t + u_t; \quad t = 1, 2, \ldots, T \qquad \text{(i)}$$

where Y_t is real output, L_t is a measure of labour input, K_t is a measure of real capital input, and u_t is an unobserved random disturbance with $\mathrm{E}(u_t) = 0$.

The following estimates of (i) were obtained by ordinary least squares (OLS) using 15 annual observations from the Taiwanese agricultural sector.

$$\begin{aligned} \ln Y_t &= \underset{(2.44)}{-3.329} + \underset{(0.54)}{1.498} \ln L_t + \underset{(0.10)}{0.489} \ln K_t + e_t \qquad \text{(ii)} \end{aligned}$$

where e_t are OLS residuals, standard errors are in parentheses and $R^2 = 0.89$.

(a) **Give an economic interpretation of the estimated coefficients. Are the estimated slope parameters of the expected sign? Explain.**

(4 marks)

(b) **Test the slope parameters for significance, and explain what assumptions your tests require in order to be valid.**

(6 marks)

(c) **If the Taiwanese agricultural sector has constant returns to scale then $\alpha_1 + \alpha_2 = 1$. Discuss whether the estimates in (ii) support this restriction.**

(3 marks)

(d) **The equation was also estimated in the following restricted form**

$$\begin{aligned} [\ln Y_t - \ln L_t] &= \underset{(0.42)}{1.712} + \underset{(0.09)}{0.612}[\ln K_t - \ln L_t] + \nu_t \end{aligned}$$

where ν_t are OLS residuals, standard errors are in parentheses and $R^2 = 0.77$.

Test the restriction(s) in (iii), and show that (iii) incorporates the restriction of constant returns to scale.

(7 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapters 2.6 (Testing hypotheses relating to the regression coefficients) and 6.5 (Testing linear restriction).

Dougherty, C. Subject guide, Additional exercise sections A6.9 in Chapter 6 (Specification of regression variables).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapters 5.8 (Hypothesis testing: Some practical aspects) and 7.9 (Cobb–Douglas production function: More on functional form).

Approaching the question

(a) Interpretation of α_1, α_2 and $\alpha_1 + \alpha_2$ should be given. Answer is:

The properties of the Cobb–Douglas production function are well known.

α_1 is the (partial) elasticity of output with respect to the labour input. Hence, a 1 per cent rise in the labour input, holding capital constant, will increase output by 1.5 per cent on average.

Likewise, α_2 is the (partial) elasticity of output with respect to the capital input, holding the labour input constant. Hence, a 1 per cent rise in the capital input, holding labour constant, will increase output by 0.49 per cent on average.

The sum $\alpha_1 + \alpha_2$ gives information about the returns to scale; that is, the response of output to a proportionate change in the inputs. The estimated slope parameters are of the expected sign as both are positive – more input should produce more output.

(b) Significance of both the slope parameters should be tested. Assumptions should be explicitly provided. Answer is:

The t statistics are:

$$t_{\alpha_1} = \frac{1.498}{0.54} = 2.77 \qquad \text{and} \qquad t_{\alpha_2} = \frac{0.489}{0.10} = 4.89.$$

The 5 per cent critical value for the two-sided t distribution with 12 degrees of freedom is 2.18. Hence, reject the null hypothesis in both cases. The test requires the following assumptions:

- The model is linear in parameters and correctly specified.
- There is some variation in the regressor in the sample.
- The disturbance term has zero expectation: $\mathrm{E}(u_t) = 0$ for all t.
- The disturbance term is homoscedastic: $\mathrm{E}(u_t^2) = \sigma_u^2$ for all t.
- The values of the disturbance term have independent distribution: $\mathrm{E}(u_i u_j) = 0$ for $i \neq j$.
- The disturbance term has a normal distribution.

(c) It should be mentioned that enough information is not given to test it. Answer is:

The estimated coefficients sum to 1.987, which gives the value of the return to scale. The results suggest that, over the period of estimation, the Taiwanese agricultural sector was characterised by increasing returns to scale. However, we do not know whether 1.987 is significantly different from 1 without a formal statistical test, in which case we would need a measure of the standard error of the sum of the coefficients, or use an F test.

(d) The F test for linear restriction should be used. Answer is:

The F test is given by:

$$F = \frac{(R_U^2 - R_R^2)/q}{(1 - R_U^2)/(n-k-1)} = \frac{0.89 - 0.77}{(1 - 0.89)/12} = 13.04.$$

The 5 per cent critical value for $F(1, 12)$ is 4.75, thus reject the null hypothesis of constant returns to scale as 1.987 is significantly different from 1.

Equation (iii) incorporates the restriction of constant returns to scale as:

$$\begin{aligned} \ln Y_t &= \alpha_0 + \alpha_1 \ln L_t + \alpha_2 \ln K_t \\ \ln Y_t &= \alpha_0 + (1 - \alpha_2) \ln L_t + \alpha_2 K_t. \end{aligned}$$

Re-arranging gives:

$$(\ln Y_t - \ln L_t) = \alpha_0 + \alpha_2 (\ln K_t - \ln L_t)$$

which is the same as (iii).

Question 7

(a) Explain the problem of identification in the context of simultaneous equation models.

(3 marks)

(b) In the model

$$\begin{aligned} y_{1t} &= \alpha y_{2t} + u_{1t} \\ y_{2t} &= \beta_1 x_t + \beta_2 y_{1t} + u_{2t} \quad t = 1, 2, \ldots, T \end{aligned}$$

where x_t is an exogenous variable.

i. Examine the identification of both equations.

(4 marks)

ii. Obtain the ordinary least squares estimator of α, and examine its consistency.

(7 marks)

iii. Derive the two-stage least squares of and also prove its consistency stating carefully any assumptions you need.

(3 marks)

(c) What is meant by indirect least squares? Explain.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapters 9.2 (Simultaneous equations bias) and 9.3 (Instrumental variable estimation).

Dougherty, C. Subject guide, Chapter 9 (Simultaneous equation estimation).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 19.3 (Rules of identification) and 20.4 (Estimation of an overidentified equation: The method of two-stage least squares (2SLS)).

Approaching the question

(a) To answer this question, order condition of identification should be used. Answer is:

Order Condition of Identification (Necessary condition of identification):

$$R \geq G - 1$$

where:

$$\begin{aligned} R &= \text{Number of restrictions imposed on the equation under consideration} \\ &= \text{In our case, number of variables excluded from the equation.} \\ G &= \text{Number of jointly dependent variables in the model} \\ &= \text{Number of equations in the model.} \end{aligned}$$

If:

$R = G - 1$	Equation under consideration is exactly identified.
$R > G - 1$	Equation under consideration is over identified.
$R < G - 1$	Equation under consideration cannot be identified.

(b) i. The concept of the order of identification given in part (a) should be used. Detailed answer is:

In the first equation $R = 0$ and $G - 1 = 1$. That is $R < G - 1$. Equation is under identified.

In the second equation $R = 1$ and $G - 1 = 1$. That is $R = G - 1$. Equation is exactly identified.

ii. It is required to derive the OLS estimator. Probability limit should be used to examine consistency. Answer is:
The OLS estimator of α is given by:

$$\widehat{\alpha} = \frac{\sum y_{1t} y_{2t}}{\sum y_{2t}^2}.$$

To show consistency we consider:

$$\text{plim } (\widehat{\alpha}) = \frac{\text{plim } \left(\frac{1}{T}\sum y_{1t}y_{2t}\right)}{\text{plim } \left(\frac{1}{T}\sum y_{2t}^2\right)} = \alpha + \frac{\text{plim } \left(\frac{1}{T}\sum y_{2t}u_{1t}\right)}{\text{plim } \frac{1}{T}\sum y_{2t}^2} \neq \alpha$$

since:

$$\text{plim } \left(\frac{1}{T}\sum y_{2t}u_{1t}\right) = \text{plim } \left(\frac{1}{T}\sum(\beta_1 x_t + \beta_2 y_{1t} + u_{2t})\right) u_{1t} \neq 0$$

and:

$$\text{plim } \left(\frac{1}{T}\sum y_{2t}^2\right) \neq 0.$$

This implies $\widehat{\beta}_{1,OLS}$ is an inconsistent estimator of β_1.

iii. It is required to derive the 2SLS estimator. Probability limit should be used to examine consistency. Answer is:
The two-stage estimator of α is:

$$\tilde{\alpha} = \frac{\sum y_{1t} z_t}{\sum y_{2t} z_t}$$

where z_t is the linear combination of instruments (in this case x_t only). To show consistency consider:

$$\text{plim } (\tilde{\alpha}) = \text{plim } \frac{\sum y_{1t} z_t}{\sum y_{2t} z_t} = \text{plim } \frac{\sum(\alpha y_{2t} + u_t) z_t}{\sum y_{2t} z_t} = \alpha + \text{plim } \left(\frac{\frac{1}{T}\sum z_t u_t}{\frac{1}{T}\sum y_{2t} z_t}\right) = \alpha + \frac{\text{Cov}(z_t, u_t)}{\text{Cov}(y_{2t}, z_t)} = \alpha$$

since z_t is correlated with y_{2t} but uncorrelated with u_t. Covariances given are population covariances.

(d) A brief discussion of ILS estimator is required. Answer is:

- Obtain the reduced form, from the given simultaneous equation model. There is a relationship between the reduced form (RF) parameters and the structural parameters.
- Estimate the RF parameters by the OLS. The estimates will be consistent as in the RF all explanatory variables are exogenous.
- As the RF parameters and the structural parameters are related, once the RF form parameters have been estimated, the estimates of the structural parameters can be obtained. These estimates will be consistent.

Question 8

A study of applications for home mortgages used the linear probability model

$$\boldsymbol{MORT_i = \beta_0 + \beta_1 INC_i + \beta_2 AGE_i + \beta_3 PROP_i + u_i; \quad i = 1, 2, \ldots, 700}$$

where

$\boldsymbol{MORT_i}$ **= 1 if a mortgage is granted to the i-th applicant: 0 otherwise**

$\boldsymbol{INC_i}$ **= income of the i-th applicant in thousands of pounds**

$\boldsymbol{AGE_i}$ **= age of the i-th applicant in years**

$\boldsymbol{PROP_i}$ **= age of the property for which the mortgage is being applied.**

(a) The estimated coefficient for INC_i was 1.02 with standard error 0.51. What is the interpretation of this coefficient?

(4 marks)

(b) Why is R^2 meaningless in probit and logit models? What measures of 'goodness of fit' are applicable to probit and logit models?

(7 marks)

(c) Using a two variable linear model, show that the ordinary least squares estimator will be heteroscedastic if the dependent variable takes only values 0 and 1.

(9 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapters 10.1 (Linear probability model) and 10.2 (Logit analysis).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 15.2 (The linear probability model (LPM)).

Approaching the question

(a) It is important to discuss that in LPM the estimated probability of an event occurring may be greater than one or less than zero. Answer is:

$t = 1.02/0.51 = 2$ which is significantly different from 0. As INC increases by 1000, $MORT$ increases by 1.48 but since the estimated $MORT$ can be interpreted as a probability the prediction is likely to lie outside [0, 1].

(b) It should be discussed that as the dependent variable takes only two values R^2 is meaningless.

A brief discussion of the likelihood ratio test and pseudo-R^2 should be given. Answer is given below:

The definition of R^2 is:

$$R^2 = \frac{\text{ESS}}{\text{TSS}} = \frac{\text{TSS} - \text{RSS}}{\text{TSS}} = 1 - \frac{\text{RSS}}{\text{TSS}}$$

where TSS is the total sum of squares, ESS is the explained sum of squares and RSS is the residual sum of squares. Under logit and probit the dependent variable only takes two states, 0 and 1, hence TSS will take different values dependent on the coding of 'success' or 'failure' even though the independent variables are the same.

The possibilities for measuring goodness of fit are (i) the pseudo-R^2 defined by $1 - (\ln L / \ln L_0)$ where $\ln L$ is the unrestricted log likelihood and $\ln L_0$ is the log likelihood that would have been obtained with only the intercept in the regression. This has a minimum of 0 but the maximum will be less than 1 and, unlike R^2, it does not have a natural interpretation.

The alternative is (ii) the likelihood ratio statistic defined by $2\ln(L/L_0) = 2[\ln L - \ln L_0]$ which has in large samples, a chi-squared distribution with q degrees of freedom where q is the number of restrictions imposed by the null hypothesis. The null hypothesis is that the coefficients of the variables are all jointly zero.

The parameters of $\ln L$ are estimated by the maximum likelihood under the alternative hypothesis and $\ln L_0$ is estimated by the maximum likelihood under the null.

(c) It is required to derive the variance of the disturbance term. Detailed answer is:

Let the model be:

$$Y_i = \beta_1 + \beta_1 X_i + u_i; \quad i = 1, 2, \ldots, n \qquad \text{(i)}$$

where:

$$Y_i = \begin{cases} 1 & \text{if event occurs} \\ 0 & \text{if not.} \end{cases}$$

As Y_i takes only two values 1 or 0, therefore u_i can take only two values $1 - \beta_0 - \beta_1 X_i$ when $Y_i = 1$ and $-\beta_0 - \beta_1 X_i$ when $Y_i = 0$. Based on this we can write the probability distribution of u_i as:

Y_i	u_i	$f(u_i)$
1	$1 - \beta_0 - \beta_1 X_i$	$\beta_0 + \beta_1 X_i$
0	$-\beta_0 - \beta_1 X_i$	$1 - \beta_0 - \beta_1 X_i$

This probability distribution also satisfies the assumption that:

$$\mathrm{E}(u_i) = (1 - \beta_0 - \beta_1 X_i)(\beta_0 + \beta_1 X_i) + (-\beta_0 - \beta_1 X_i)(1 - \beta_0 - \beta_1 X_i) = 0.$$

We can write $\mathrm{Var}(u_i)$ as:

$$\begin{aligned} \mathrm{Var}(u_i) = \mathrm{E}(u_i^2) &= (1 - \beta_0 - \beta_1 X_i)^2(\beta_0 + \beta_1 X_i) + (-\beta_0 - \beta_1 X_i)^2(1 - \beta_0 - \beta_1 X_i) \\ &= (1 - \beta_0 - \beta_1 X_i)(\beta_0 + \beta_1 X_i)[(1 - \beta_0 - \beta_1 X_i) + (\beta_0 + \beta_1 X_i)] \\ &= (\beta_0 + \beta_1 X_i)(1 - \beta_0 - \beta_1 X_i) \\ &= \mathrm{E}(Y_i)[1 - \mathrm{E}(Y_i)] \\ &= P_i(1 - P_i), \quad \text{for all } i = 1, 2, \ldots, n. \end{aligned}$$

Hence the disturbance term is heteroscedastic. This will make the OLS estimators inefficient.

Question 9

(a) Explain the meaning of spurious regression.

(4 marks)

(b) The following equations were estimated by ordinary least squares.

$$\begin{aligned} \boldsymbol{Y_t} &= \mathbf{3.0920} + \mathbf{0.6959}\boldsymbol{X_t} + \widehat{u}_t \qquad (1) \\ &\quad\ \ (\mathbf{0.1305}) \quad (\mathbf{0.0103}) \end{aligned}$$

$$\boldsymbol{R^2} = \mathbf{0.99},\ \boldsymbol{F} = \mathbf{4523.25},\ \boldsymbol{s} = \mathbf{0.0236},\ \mathbf{DW} = \mathbf{0.557},\ \boldsymbol{T} = \mathbf{740}.$$

$$\begin{aligned} \boldsymbol{\Delta\widehat{u}_t} &= -\mathbf{0.2161}\widehat{u}_{t-1} + \mathbf{0.2349}\boldsymbol{\Delta}\widehat{u}_{t-1} + \mathbf{0.2029}\boldsymbol{\Delta}\widehat{u}_{t-2} + \widehat{\varepsilon}_t \\ &\quad\ (\mathbf{0.0845}) \qquad (\mathbf{0.1592}) \qquad (\mathbf{0.1631}) \end{aligned}$$

$$\boldsymbol{R^2} = \mathbf{0.1799};\ \boldsymbol{s} = \mathbf{0.0115};\ \boldsymbol{T} = \mathbf{737}.$$

Where s is the standard error of the residuals, T is the number of observations, $\widehat{u}_t$ and $\widehat{\varepsilon}_t$ are OLS residuals, and standard errors are in parentheses.

Do the results above indicate that Y_t and X_t are cointegrated? Specify clearly all the assumptions you have made.

(6 marks)

[Note: Critical value at 5% significant level from MacKinnon table is -3.3377].

(c) Consider a model

$$\boldsymbol{y_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}; \quad t = 1, 2, \ldots, T}$$

$$\mathbf{E}(\boldsymbol{u_t}) = \mathbf{0};\ \mathbf{E}(\boldsymbol{u_t^2}) = \boldsymbol{\sigma^2};\ \text{and } \mathbf{E}(\boldsymbol{u_s u_t}) = \mathbf{0} \text{ for all } \boldsymbol{s, t} = \mathbf{1, 2, \ldots, T}.$$

i. Is y_t stationary? Explain.

(5 marks)

ii. Calculate the autocorrelation function of y_t.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapters 13.1 (Stationarity and nonstationarity), 13.2.(Spurious regressions) 13.3 (Graphical techniques for detecting nonstationarity) and 13.5 (Cointegration).

Dougherty, C. Subject guide, Chapter 13 (Introduction to nonstationary time series).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 21.3 (Stochastic Processes), 21.8 (Tests for stationarity) and 21.11 (Cointegration: Regression of a unit root time series on another unit root time series).

Approaching the question

(a) Concept of spurious regression should be explained with a simple model: Answer is:

Spurious regression was first demonstrated by Granger and Newbold who showed, using Monte Carlo techniques, that a regression involving 2 non-stationary series could give rise to spurious results in that the t statistics over-rejected the null hypothesis of a zero coefficient for 2 independent random walk series.

If Y_t and X_t are non-stationary and we regress Y_t and X_t, that is:

$$Y_t = \pi_0 + \pi_1 X_t + v_t$$

then even if there is no relationship between Y and X, the regression will produce a t ratio **which will reject the null hypothesis** $\mathrm{H}_0 : \pi_1 = 0$.

The reason for this result is that if $\mathrm{H}_0 : \pi_1 = 0$ then:

$$Y_t = \pi_0 + v_t.$$

Suppose Y_t is $I(1)$. Since Y_t is $I(1)$ and π_0 is constant it follows that v_t must be $I(1)$. This violates the standard distributional theory based on the assumption that v_t is stationary, i.e. v_t is $I(0)$. Hence the misleading result.

(b) Clear concept of cointegration is required. Assumptions should be stated explicitly. Answer is:

H_0 : No cointegration.

H_1: Cointegration.

Cointegration test statistic is $-0.2161/0.0845 = -2.557$.

5 per cent critical value given in the Mackinnon table is -3.3377. Thus, cannot reject the null of no cointegration.

Main assumption is that the error terms in both equations have constant variances and no serial correlation. Also need to assume that the specifications are correct (for example no structural breaks).

(c) It should be shown that mean, variance and covariances are independent of time. Part (ii) is based on part (i). Answer is:

i. We have:

$$\begin{aligned}
\mathrm{E}(y_t) &= 0 \\
\mathrm{Var}(y_t) &= \mathrm{E}(y_t^2) \\
&= \mathrm{E}(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2})^2 \\
&= \mathrm{E}(u_t^2) + \theta_1^2 \mathrm{E}(u_{t-1}^2) + \theta_2^2 \mathrm{E}(u_{t-2}^2); \quad \text{as } \mathrm{E}(u_s u_t) = 0 \text{ if } s \neq t \\
&= (1 + \theta_1^2 + \theta_2^2)\sigma^2 \\
\mathrm{Cov}(y_1, y_{t-1}) &= \mathrm{E}(y_t y_{t-1}) \\
&= \mathrm{E}(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2})(u_{t-1} + \theta_1 u_{t-2} + \theta_2 u_{t-3}) \\
&= (\theta_1 + \theta_1\theta_2)\sigma^2
\end{aligned}$$

$$\begin{aligned}
\mathrm{Cov}(y_1, y_{t-2}) &= \mathrm{E}(y_t y_{t-2}) \\
&= \mathrm{E}(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2})(u_{t-2} + \theta_1 u_{t-3} + \theta_2 u_{t-4}) \\
&= \theta_2 \sigma^2 \\
\mathrm{Cov}(y_t, y_{t-s}) &= \mathrm{E}(y_t y_{t-s}) = 0 \ \text{ for all } s > 2.
\end{aligned}$$

Hence as mean, variance and covariances are constant over time, y_t is weakly stationary. If u's are normally distributed then this also implies strong stationarity.

(ii) Autocorrelation function is defined as:

$$\rho_s = \frac{\mathrm{Cov}(y_t, y_{t-s})}{\sqrt{\mathrm{Var}(y_t)}\sqrt{\mathrm{Var}(y_{t-s})}} = \frac{\mathrm{Cov}(y_t, y_{t-s})}{\mathrm{Var}(y_t)}; \quad \text{as } \mathrm{Var}(y_t) = \mathrm{Var}(y_{t-s}).$$

Hence:

$$\rho_s = \begin{cases} 1 & \text{if } s = 0 \\ \frac{\theta_1 + \theta_1\theta_2}{1+\theta_1^2+\theta_2^2} & \text{if } s = 1 \\ \frac{\theta_2}{1+\theta_1^2+\theta_2^2} & \text{if } s = 2 \\ 0 & \text{if } s > 2. \end{cases}$$

Question 10

Let the regression equation be

$$\boldsymbol{Y_t = \beta_1 + \beta_2 X_t + u_t; \quad t = 1, 2, \ldots, T}$$

where

$$\begin{aligned}
\boldsymbol{u_t} &= \boldsymbol{\rho u_{t-1} + \varepsilon_t} \quad \textbf{for all } \boldsymbol{t}; \quad \boldsymbol{|\rho| < 1} \\
\mathbf{E}(\boldsymbol{\varepsilon_t}) &= \mathbf{0} \\
\mathbf{E}(\boldsymbol{\varepsilon_s \varepsilon_t}) &= \boldsymbol{\sigma_s^2} \ \textbf{ if } \boldsymbol{s = t} \\
&= \mathbf{0} \ \textbf{ if } \boldsymbol{s \neq t}
\end{aligned}$$

(a) Derive

i. the variance of u_t; and

ii. $\mathbf{E}(u_t u_{t-1})$

(7 marks)

(b) Explain the consequences of this model specification on ordinary least squares estimators for β_1 and β_2.

(3 marks)

(c) Explain how would you test the null hypothesis $\mathrm{H}_0 : \rho = 0$ against the alternative $\mathrm{H}_1 : \rho \neq 0$. Specify all the assumptions needed for this test.

(5 marks)

(d) Discuss in detail a method of estimation which gives best linear unbiased estimates of β_1 and β_2.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapters 12.1 (Definition and consequences of autocorrelation), 12.2 (Detection of autocorrelation) and 12.3 (Fitting a model subject to AR(1) autocorrelation).

Dougherty, C. Subject guide, Chapter 12 (Properties of regression models with time series data).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapters 12.4 (Consequences of using OLS in the presence of autocorrelation) and 12.6 (Detecting autocorrelation).

Approaching the question

(a) i. It is required to derive variance and covariance of u_t. For finding these it has to be shown that $\mathrm{E}(u_t) = 0$.

$$u_t = \rho u_{t-1} + \varepsilon_t$$

can be written in lag operator form as:

$$(1 - \rho L)u_t = \varepsilon_t$$

or:

$$u_t = (1 - \rho L)^{-1}\varepsilon_t = (1 + \rho L + \rho^2 L^2 + \cdots)\varepsilon_t = \varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \cdots$$

Therefore, the variance of u_t is:

$$\mathrm{Var}(u_t) = (1 + \rho^2 + \rho^4 + \cdots)\sigma_\varepsilon^2 = \frac{\sigma_\varepsilon^2}{1 - \rho^2}.$$

ii. We have:

$$\begin{aligned} u_t u_{t-1} &= (\varepsilon_t + \rho\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \cdots)(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \rho^2\varepsilon_{t-3} + \cdots) \\ &= [\varepsilon_t + \rho(\varepsilon_{t-1} + \rho^2\varepsilon_{t-2} + \cdots)][\varepsilon_{t-1} + \rho(\varepsilon_{t-2} + \rho^2\varepsilon_{t-3} + \cdots)] \\ &= \varepsilon_t(\varepsilon_{t-1} + \rho\varepsilon_{t-2} + \rho^2\varepsilon_{t-3} + \cdots) + \rho(\varepsilon_t + \rho\varepsilon_{t-2} + \rho^2\varepsilon_{t-3} + \cdots)^2 \end{aligned}$$

Hence:

$$\mathrm{E}(u_t u_{t-1}) = \rho(1 + \rho^2 + \rho^4 + \cdots)\sigma_\varepsilon^2 = \frac{\rho\sigma_\varepsilon^2}{1 - \rho^2}.$$

(b) It should be explained for which properties OLS holds and also the properties for which OLS does not hold. Answer is:

The effect of serially correlated errors is to produce unbiased and consistent but inefficient parameter estimates. The standard errors are incorrect leading to invalid t tests. Hence, the estimates of β_1 and β_2 will be unbiased and consistent but inefficient.

(c) This question is based on the Durbin–Watson test. Assumptions should be stated clearly. Answer is:

Durbin–Watson (DW) Statistic

DW statistic is defined as:

$$\mathrm{DW} = \frac{\sum_{t=2}^{T}(\widehat{u}_t - \widehat{u}_{t-1})^2}{\sum_{t=1}^{T}\widehat{u}_t^2}.$$

The Durbin–Watson statistic can be applied only if:

- the disturbance term follows an AR(1) process
- the model has an intercept term, and
- there is no lagged dependent variable as an explanatory variable.

In the Durbin–Watson table (which can be found in the Appendix of any econometrics book) there are two values:

$$\begin{aligned} d_L &= \text{lower limit} \\ d_U &= \text{upper limit.} \end{aligned}$$

We have:

$$\begin{aligned}
\mathrm{DW} < d_L &\Rightarrow \text{positive autocorrelation} \\
\mathrm{DW} > 4 - d_L &\Rightarrow \text{negative autocorrelation} \\
d_L \leq \mathrm{DW} \leq d_U &\Rightarrow \text{no conclusion} \\
4 - d_U \leq \mathrm{DW} \leq 4 - d_L &\Rightarrow \text{no conclusion.}
\end{aligned}$$

(d) As ρ is unknown, the Cochrane–Orcutt method of estimation or Prais–Winstein method of estimation should be used. Detailed answer is:

Cochrane–Orcutt method of estimation

For simplicity assume the model is:

$$Y_t = \beta_0 + \beta_1 X_t + u_t; \quad t = 1, 2, \ldots, T \qquad \text{(i)}$$

and:

$$u_t = \rho u_{t-1} + \varepsilon_t$$

with $\mathrm{E}(\varepsilon_t) = 0$, $\mathrm{E}(\varepsilon_t^2) = \sigma_\varepsilon^2$ and $\mathrm{E}(\varepsilon_s \varepsilon_t) = 0$ for $s \neq t$.

This means that the disturbance term u_t follows an AR(1) process.

Lag (i) by one period and multiply by ρ to get:

$$\rho Y_{t-1} = \rho\beta_0 + \rho\beta_1 X_{t-1} + \rho u_{t-1}. \qquad \text{(ii)}$$

Subtract (ii) from (i), to get:

$$Y_t - \rho Y_{t-1} = (\rho - 1)\beta_0 + \beta_1(X_t - \rho X_{t-1}) + u_t - \rho u_{t-1} \qquad \text{(iii)}$$

The disturbance term in (iii) is $u_t - \rho u_{t-1} = \varepsilon_t$ which is a well-behaved disturbance term. Hence if ρ is *known*, OLS can be applied to (iii) to obtain best linear unbiased estimators of β_0 and β_1.

If ρ is *not known*, (iii) cannot be estimated as such. Estimation of the parameters requires the following steps:

- Apply OLS to (i) to obtain $\widehat{u}_t$.
- Apply OLS to $\widehat{u}_t = \rho\widehat{u}_{t-1} + \varepsilon_t$ to obtain the OLS estimator of ρ:

$$\widehat{\rho} = \frac{\sum_{t=2}^{T} \widehat{u}_t \widehat{u}_{t-1}}{\sum_{t=2}^{T} \widehat{u}_{t-1}^2}.$$

  Replace ρ in (iii) by $\widehat{\rho}$ and apply OLS to get $\widehat{\beta}_0$ and $\widehat{\beta}_1$.
- Obtain a new set of residuals by replacing β_0 and β_1 in (i) by $\widehat{\beta}_0$ and $\widehat{\beta}_1$ as:

$$\tilde{u} = Y_t - \widehat{\beta}_0 - \widehat{\beta}_1 X_t.$$

- Repeat step (b) to (d). Keep on doing it until the estimate of ρ converges. That will be the final estimate of ρ. Let it be $\widehat{\rho}_F$.
- Replace ρ in (iii) by the final estimate of ρ i.e. $\widehat{\rho}_F$. Apply OLS to (iii) to obtain final estimates of β_0 and β_1.

Examiners' commentaries 2015

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2014–15. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2014). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone B

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Section A

Answer all questions from this section.

Question 1

Let the regression equation be:

$$y_t = \beta x_t + u_t; \quad t = 1, 2, \ldots, T.$$

Where $\mathrm{E}(u_t) = 0$; $\mathrm{E}(u_t^2) = \sigma^2$ and $\mathrm{E}(u_s u_t) = 0$ if $s \neq t$. x values are fixed in repeated samples.

Obtain the ordinary least squares estimator (OLS) of β. Show that the OLS estimator of β is linear and unbiased.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapters R.6 (Unbiasedness and efficiency), 1.3 (Derivation of the regression coefficients) and 2.3 (The random components and Unbiasedness of the OLS regression coefficients).

Dougherty, C. Subject guide, Chapter 2 (Properties of the regression coefficients and hypothesis testing).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 3A.2 (Linearity and unbiasedness properties of least-squares estimators).

Approaching the question

It is required to obtain the residual sum of squares (RSS) and minimise the RSS with respect to $\widehat{\beta}$ to obtain the OLS estimator. To show unbiasedness, it has to be shown that $\mathrm{E}(\widehat{\beta}) = \beta$.

Answer is:

$$y_t = \beta x_t + u_t; \quad t = 1, 2, \ldots, T.$$

The residual sum of squares (RSS) is:

$$\sum \widehat{u}_t^2 = \sum (y_t - \widehat{\beta} x_t)^2$$

and minimising RSS with respect to $\widehat{\beta}$ we get:

$$\frac{\mathrm{d} \sum \widehat{u}_t^2}{\mathrm{d}\widehat{\beta}} = 2 \sum (y_t - \widehat{\beta} x_t)(-x_t) = 0.$$

Solving we get:

$$\widehat{\beta}_{OLS} = \frac{\sum x_t y_t}{\sum x_t^2}.$$

Linear

Define w_t as:

$$w_t = \frac{x_t}{\sum x_t^2}$$

then:

$$\widehat{\beta}_2 = \frac{\sum x_t y_t}{\sum x_t^2} = \sum w_t y_t.$$

This shows that $\widehat{\beta}$ is linear in Y's.

Unbiasedness

We have:

$$\widehat{\beta} = \sum w_t y_t = \sum w_t (\beta x_t + u_t) = \beta \sum w_t x_t + \sum w_t u_t = \beta + \sum w_t u_t$$

as $\sum w_t x_t = 1$. Hence:

$$\mathrm{E}(\widehat{\beta}) = \beta + \sum w_t \mathrm{E}(u_t) = \beta \;\Rightarrow\; \text{unbiased}.$$

Question 2

Suppose that business expenditure for a new plant (Y_t) is explained by the relation

$$\ln(Y_t) = \alpha + \beta \ln(X_t^*) + u_t,$$

where u_t is a random variable, ln is the natural logarithm and X_t^* is the level of expected sales (which is unobserved) and is formed by

$$\ln(X_t^*) - \ln(X_{t-1}^*) = \gamma(\ln(X_{t-1}) - \ln(X_{t-1}^*)).$$

X_t is the level of actual sales. Derive a linear relationship that can be used to estimate α and β, using the observable variables Y_t and X_t.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapter 11.4 (Models with lagged dependent variable).

Gujarati, D.N. and D.C. Porter. *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 17.5 (Rationalization of the Koyck model: The adaptive expectations model).

Approaching the question

To get an estimable equation we need to eliminate the expected values from the equation. Multiply through by $(1-\gamma)$ and lag to get a new equation. Subtract the new equation from the original equation to get the result. Answer is:

Expectations are formed by:

$$\ln(X_t^*) - (1-\gamma)\ln(X_{t-1}^*) = \gamma \ln(X_{t-1}).$$

To get an estimable equation we need to eliminate the expected values from the equation. Multiply through by $(1-\gamma)$ and lag to give:

$$(1-\gamma)\ln(Y_{t-1}) = (1-\gamma)\alpha + (1-\gamma)\beta \ln(X_{t-1}^*) + (1-\gamma)u_{t-1}.$$

Now subtract from the original equation:

$$\begin{aligned}
\ln(Y_t) - (1-\gamma)\ln(Y_{t-1}) &= \alpha\gamma + \beta(\ln(X_t^*) - (1-\gamma)\ln(X_{t-1}^*)) + (u_t - (1-\gamma)u_{t-1}) \\
\ln(Y_t) - (1-\gamma)\ln(Y_{t-1}) &= \alpha\gamma + \beta(\gamma \ln(X_{t-1})) + (u_t - (1-\gamma)u_{t-1})
\end{aligned}$$

or:

$$\ln(Y_t) = \alpha\gamma + \beta\gamma\ln(X_{t-1}) + (1-\gamma)\ln(Y_{t-1}) + (u_t - (1-\gamma)_{t-1}).$$

Parameters are estimated by non-linear techniques. If these procedures are not available then a grid search can be used where γ is given values between 0 and 1 in steps of 0.1 and the remaining parameters estimated by OLS.

Question 3

Let the probability density function of a random variable X be $f(x;\theta)$. Explain the procedure to use likelihood ratio test for testing the null hypothesis $\mathrm{H_0} : \theta = \theta_0$.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapter 10.6 (An introduction to maximum likelihood estimation).

Dougherty, C. Subject guide, Chapter 10 (Binary choice and limited dependent variable models, and maximum likelihood estimation).

Approaching the question

Likelihood Ratio test should be discussed. Detailed answer is:

Suppose we have to test a simple hypothesis $\mathrm{H}_0 : \theta = \theta_0$ against all possible alternatives. Given a simple random sample $X_1, X_2, \ldots, X_T$, a natural way of judging the acceptability, or otherwise of the hypothesis, would be to compare the likelihood functions.

Let:

L_R = Restricted likelihood (likelihood based on the null hypothesis)

L_U = Unrestricted likelihood (likelihood based on the alternative hypothesis). $\widehat{\theta}_{MLE}$ is used.

If the likelihood ratio (LR):

$$\lambda = \frac{L_R}{L_U}$$

is close to unity, then in light of the given sample, H_0 would seem highly plausible. On the other hand if this ratio is close to zero, H_0 would seem to have little validity. Since λ is a random variable, its distribution may be derived, and hence, we can make probability statements about how close the LR is to unity. A test for H_0 is thus provided by a critical region defined by $\lambda < \lambda_0$, where λ_0 is such that $P(\lambda < \lambda_0 \,|\, \mathrm{H}_0) = \alpha$. α is the probability of a Type 1 error.

For large N, $-2\ln\lambda$ has approximately a χ^2 distribution with degrees of freedom equal to the number of *restrictions* imposed by the null hypothesis.

Question 4

What is meant by a common factor test in the context of a linear model with an autocorrelated error? How would you perform a common factor test and what hypothesis would you be testing?

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapter 12.3 (Fitting a model subject to AR (1) autocorrelation).

Dougherty, C. Subject guide, Additional Exercises sections A12.3–A12.5 in Chapter 12 (Properties of regression models with time series data).

Approaching the question

The common factor test is used to test whether autoregressive model or autoregressive distributed lag model is the correct specification. The test should be specified. Detailed answer is:

Given the equations:

$$\begin{aligned} Y_t &= \beta_1 + \beta_2 X_t + u_t \\ u_t &= \rho u_{t-1} + v_t; \quad t = 1, 2, \ldots, T \end{aligned}$$

where v has zero mean, constant variance and zero autocovariance. Combine the two equations to give a new equation, which has the general form:

$$Y_t = \beta_1(1-\rho) + \rho Y_{t-1} + \beta_2 X_t - \beta_2 \rho X_{t-1} + v_t$$

It is a restricted version of the ADL(1,1) model:

$$Y_t = \lambda_1 + \lambda_2 Y_{t-1} + \lambda_3 X_t + \lambda_4 X_{t-1} + \varepsilon_t$$

and is subject to the restriction $\lambda_4 = -\lambda_2\lambda_3$. The test of this restriction is the common factor test.

$\mathrm{H}_0 : \lambda_4 = -\lambda_2\lambda_3$.

$\mathrm{H}_1 : \lambda_4 \neq -\lambda_2\lambda_3$.

Note that the usual F test of the restriction is not appropriate because the restriction is non-linear, so we have to use the test statistic $T\ln(\mathrm{RSS}_R/\mathrm{RSS}_U)$ where RSS_R and RSS_U are the residual sum of squares from the restricted and unrestricted models, respectively. T is the sample size and the test statistic is asymptotically chi-squared with 1 degree of freedom. If the restriction is rejected, we conclude that AR(1) specification is correct. If the restriction is not rejected then unrestricted ADL(1,1) model is the correct specification.

Question 5

Consider the model

$$\boldsymbol{y_t = u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2}; \quad t = 1, 2, \ldots, T.}$$

$\mathrm{E}(u_t) = 0$; $\mathrm{E}(u_t^2) = \sigma^2$; and $\mathrm{E}(u_s u_t) = 0$ if $s \neq t$ for all $s, t = 1, 2, \ldots, T$.

Is y_t stationary? Explain in detail.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapter 13.1 (Stationarity and nonstationarity).

Dougherty, C. Subject guide, Chapter 13 (Introduction to nonstationary time series).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 21.3 (Stochastic Processes).

Approaching the question

It should be shown that mean, variance and covariances are independent of time. Answer is:

We have:

$$\begin{aligned}
\mathrm{E}(y_t) &= 0 \\
\mathrm{Var}(y_t) &= \mathrm{E}(y_t^2) \\
&= \mathrm{E}(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2})^2 \\
&= \mathrm{E}(u_t^2) + \theta_1^2 \mathrm{E}(u_{t-1}^2) + \theta_2^2 \mathrm{E}(u_{t-2}^2); \quad \text{as } \mathrm{E}(u_s u_t) = 0 \text{ if } s \neq t \\
&= (1 + \theta_1^2 + \theta_2^2)\sigma^2 \\
\mathrm{Cov}(y_1, y_{t-1}) &= \mathrm{E}(y_t y_{t-1}) \\
&= \mathrm{E}(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2})(u_{t-1} + \theta_1 u_{t-2} + \theta_2 u_{t-3}) \\
&= (\theta_1 + \theta_1\theta_2)\sigma^2 \\
\mathrm{Cov}(y_1, y_{t-2}) &= \mathrm{E}(y_t y_{t-2}) \\
&= \mathrm{E}(u_t + \theta_1 u_{t-1} + \theta_2 u_{t-2})(u_{t-2} + \theta_1 u_{t-3} + \theta_2 u_{t-4}) \\
&= \theta_2 \sigma^2 \\
\mathrm{Cov}(y_t, y_{t-s}) &= \mathrm{E}(y_t y_{t-s}) = 0 \ \text{ for all } s > 2.
\end{aligned}$$

Hence as mean, variance and covariances are constant over time, y_t is weakly stationary. If u's are normally distributed then this also implies strong stationarity.

Section B

Answer three questions from this section.

Question 6

Let the model be:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + u_t; \quad t = 1, 2, \ldots, T$$

$\mathrm{E}(u_t) = 0$ for all t. A researcher suspects that the variance of the disturbance term is $\mathrm{Var}(u_t) = \sigma^2 X_{t1}$.

(a) **Explain how to test the null hypothesis $\mathrm{H}_0 : \mathrm{Var}(u_t) = \sigma^2$ against the alternative hypothesis $\mathrm{H}_1 : \mathrm{Var}(u_t) = \sigma^2 X_{t1}$, for all $t = 1, 2, \ldots, T$.**

(6 marks)

(b) **Suggest in detail an estimation procedure which gives best linear unbiased estimates of the parameters.**

(6 marks)

(c) **Consider the model given by**

$$y_t = \alpha x_t + u_t; \quad t = 1, 2, \ldots, T$$

where $\mathrm{E}(u_t) = 0$; $\mathrm{E}(u_t^2) = \sigma^2 x_t^2$; $\mathrm{E}(u_s u_t) = 0$ if $s \neq t$, for all s and t. x_t is an observed non-random variable.

The density function of u_t is

$$f(u_t; x, t) = \left(2\pi\sigma^2 x_t^2\right)^{-1/2} \exp\left[-\frac{1}{2}\left(\frac{u_t}{\sigma x_t}\right)^2\right].$$

Derive the maximum likelihood estimators of α and σ^2.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapters 7.2 (Detection of heteroscedasticity) and 7.3 (Remedies for heteroscedasticity).

Dougherty, C. Subject guide, Chapter 7 (Heteroscedasticity).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapters 4 A.1 (Maximum likelihood estimation of two-variable regression model), 11.5 (Detection of heteroscedasticity) and 11.6 (Remedial measures).

Approaching the question

(a) In Dougherty, the Goldfeldt–Quandt test and White test are given. Candidates are required to describe any one of them. Both tests are explained below:

Goldfeld–Quandt test

This test assumes that $\mathrm{Var}(u_t)$ is proportional to the size of one of the RHS variables (say X_1). The observations are ranked by X_1. Run separate regression for the first n_1 ($< n/2$) and the last n_1 observations – the middle $(n - 2n_1)$ observations are not used. If heteroscedasticity is present the RSS from the two regressions will differ. Form the test

$\text{RSS}_2/\text{RSS}_1$ where RSS_1 is the residual sum of squares from the first n_1 observations and RSS_2 is the residual sum of squares from the last n_1 observations. The test statistic will have an F distribution with $(n_1 - k,\, n_1 - k)$ degrees of freedom where k is the number of parameters in the equation. Null hypothesis is that the variances are homoscedastic.

The power of the test depends upon the choice of n_1 in relation to n. Assumptions are that of no autocorrelation and the disturbance term is normally distributed.

White test

The White test looks for more general evidence of association between the variance of the error term and the regressors. Regress the squared residual from the original regression on the regressors from that model, together with the squares and the cross-products of those variables. The test statistic is nR^2 where n is the sample size and R^2 is the R^2 from the White regression. This test statistic has a chi-squared distribution with degrees of freedom equal to the number of regressors. The test assumes the sample size is large. Normality of the disturbance term is not required.

(b) Weighted least squares should be used: Answer is:

The model is:

$$Y_t = \beta_0 + \beta_1 X_{t1} + \beta_2 X_{t2} + u_t; \quad t = 1, 2, \ldots, T. \qquad \text{(i)}$$

Divide (i) by $\sqrt{X_{t1}}$ to get:

$$\frac{Y_t}{\sqrt{X_{t1}}} = \frac{\beta_0}{\sqrt{X_{t1}}} + \beta_1\sqrt{X_{t1}} + \frac{\beta_3 X_{t2}}{\sqrt{X_{t1}}} + \frac{u_t}{\sqrt{X_{t1}}}; \quad t = 1, 2, \ldots, T \qquad \text{(ii)}$$

Equation (ii) is the transformed model, where:

$$\text{E}\left(\frac{u_t}{\sqrt{X_{t1}}}\right) = 0; \quad \text{as } \text{E}(u_t) = 0$$

and:

$$\text{Var}\left(\frac{u_t}{\sqrt{X_{t1}}}\right) = \frac{\text{Var}(u_t)}{X_{t1}} = \frac{\sigma^2 X_{t1}}{X_{t1}} = \sigma^2.$$

Also:

$$\text{E}\left(\frac{u_s}{\sqrt{X_{s1}}}\frac{u_t}{\sqrt{X_{t1}}}\right) = 0; \quad \text{as } \text{E}(u_s u_t) = 0.$$

All the assumptions required to apply OLS are satisfied. Hence, OLS can be applied to (ii) to obtain the best linear unbiased estimators of β_0, β_1 and β_2.

(c) Taking into account that $\text{Var}(u_t) = \sigma^2 X_{t1}^2$, the likelihood function should be derived. To get the answer, the likelihood function has to be differentiated with respect to α and σ^2 and equated to zero. Detailed answer is:

The log likelihood function is:

$$\ln L = -\frac{T}{2}\ln 2\pi - \frac{T}{2}\ln\sigma^2 - \frac{1}{2}\sum \ln x_t^2 - \frac{1}{2\sigma^2}\sum\left(\frac{y_t}{x_t} - \alpha\right)^2.$$

The first order conditions are:

$$\frac{\partial \ln L}{\partial \alpha} = -\frac{1}{\sigma^2}\sum\left(\frac{y_t}{x_t} - \alpha\right) = 0 \qquad \text{(iii)}$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{T}{2\sigma^2} + \frac{1}{2\sigma^4}\sum\left(\frac{y_t}{x_t} - \alpha\right)^2 = 0 \qquad \text{(iv)}.$$

Solving (iii) and (iv), the ML estimators of α and σ^2 are obtained as:

$$\widehat{\alpha}_{MLE} = \frac{1}{T}\sum\left(\frac{y_t}{x_t}\right)$$

$$\widehat{\sigma}^2_{MLE} = \frac{1}{T}\sum\left(\frac{y_t}{x_t} - \widehat{\alpha}_{MLE}\right)^2.$$

Question 7

(a) **Let the model be**

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + u_t; \quad t = 1, 2, \ldots, T.$$

Suppose the relevant variable Z_t was omitted from the model. Discuss the effects of this omission on unbiasedness and consistency of the ordinary least squares estimators.

(5 marks)

(b) **The following regression was estimated by ordinary least squares (OLS) on 745 observations of annual data from the UK.**

$$\begin{aligned} q_t &= 18.210 - 0.718 y_t + e_{1t}; \quad t = 1, 2, \ldots, 745 \qquad \text{(i)} \\ &\ (1.035) \quad (0.087) \end{aligned}$$

$R^2 = 0.613$, $F = 67.97$, dw $= 1.68$

where q_t is the log of the quantity of cigarettes purchased, y_t is the log of real disposable income. Standard errors are in brackets. e_{1t} is an OLS residual. dw is the Durbin–Watson statistic.

Test the null hypothesis that the coefficient of y_t is -1. Comment on the result of your test. Do you think the estimates in (i) are plausible? Explain.

(5 marks)

(c) **A second regression was estimated using the same data sample**

$$\begin{aligned} q_t &= 0.399 - 0.343 p_t + 0.737 y_t + e_{2t}; \quad t = 1, 2, \ldots, 745 \quad \text{(ii)} \\ &\ (2.214) \quad (0.041) \quad (0.181) \end{aligned}$$

$R^2 = 0.855$, $F = 124.28$, dw $= 2.01$

where the variables are defined as before, p_t is the log of the relative prices of cigarettes, e_{2t} is an OLS residual. Standard errors in brackets. dw is the Durbin–Watson statistic.

Test the joint hypothesis that both slope coefficients in (ii) are zero.

(2 marks)

(d) **Compare the results in (ii) with those in (i). Provide an explanation in terms of omitted variable bias of the difference between the estimates of the coefficient of y_t in the two equations.**

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapters 2.6 (Testing hypotheses relating to the regression coefficients), 2.7 (The F test of goodness of fit) and 6.2 (The effect of omitting a variable that ought to be included).

Dougherty, C. Subject guide, Chapter 6.4 and 6.7 (Specification of regression variables).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapters 5.8 (Hypothesis testing: Some practical aspects), 8.4 (Testing the overall significance of the sample regression) and 13.3 (Consequences of model specification errors).

Approaching the question

(a) It is required to derive the omitted variable bias. The sign and size of the bias should be discussed. Answer is:

If the true model is:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Z_t + u_t$$

but you estimate the regression equation:

$$Y_t = \beta_1 + \beta_1 X_t + v_t$$

then the ordinary least squares (OLS) estimator of β_1 is:

$$\begin{aligned}
\widehat{\beta}_1 &= \frac{\sum(X_t - \bar{X})(Y_t - \bar{Y})}{\sum(X_t - \bar{X})^2} \\
&= \frac{\sum(X_t - \bar{X})[(\beta_0 + \beta_1 X_t + \beta_2 Z_t + u_t) - (\beta_0 + \beta_1 \bar{X} + \beta_2 \bar{Z} + \bar{u})]}{\sum(X_t - \bar{X})^2} \\
&= \beta_2 + \frac{\sum(X_t - \bar{X})(Z_t - \bar{Z})}{\sum(X_t - \bar{X})^2} + \frac{\sum(X_t - \bar{X})(u_t - \bar{u})}{\sum(X_t - \bar{X})^2}
\end{aligned}$$

and:

$$\mathrm{E}(\widehat{\beta}_1) = \beta_1 + \beta_2 \frac{\sum(X_t - \bar{X})(Z_t - \bar{Z})}{\sum(X_t - \bar{X})^2}$$

since $\mathrm{E}(u_t) = 0$ and X_t is non-stochastic.

The bias will be:

$$\beta_2 \frac{\sum(X_t - \bar{X})(Z_t - \bar{Z})}{\sum(X_t - \bar{X})^2}$$

which will be positive if β_2 is positive and X_t and Z_t are positively correlated (or β_2 is negative and X_t and Z_t negatively correlated). This bias is known as the omitted variable bias and occurs because β_1 takes up the effect of the missing Z_t.

$$\mathrm{plim}(\widehat{\beta}_1) = \beta_1 + \beta_2 \frac{\mathrm{plim}\ \frac{1}{T}\sum(X_t - \bar{X})(Z_t - \bar{Z})}{\mathrm{plim}\ \frac{1}{T}\sum(X_t - \bar{X})^2} + \frac{\mathrm{plim}\ \frac{1}{T}\sum(X_t - \bar{X})(u_t - \bar{u})}{\mathrm{plim}\ \frac{1}{T}\sum(X_t - \bar{X})^2}$$

If X and Z are uncorrelated in the limit then:

$$\mathrm{plim}\ \left(\frac{1}{T}\sum(X_t - \bar{X})(Z_t - \bar{Z})\right) = 0$$

and:

$$\mathrm{plim}\ \left(\frac{1}{T}\sum(X_t - \bar{X})^2\right) = \sigma_X^2.$$

Thus $\mathrm{plim}(\widehat{\beta}_1) = \beta_1$ and $\widehat{\beta}_1$ is a consistent estimator of β_1.

Due to the omitted variable, standard error of coefficients and the test statistics are in general invalidated.

If X and Z are orthogonal then $\widehat{\beta}_1$ is both unbiased and consistent.

(b) Required to use a t test. Null and alternative hypotheses should be clearly specified. Answer is:

$t = (-0.718 + 1)/0.087 = 3.24$. 5 per cent critical value (two sided) $= \pm 1.96$ (approximately), thus reject the null hypothesis. The null is that the income elasticity of demand for cigarettes is minus one. This is rejected but the point estimate of the income elasticity is significantly negative. Since cigarettes are known to be bad for health this might be a plausible estimate.

(c) An F test should be used to test joint significance of both slope coefficients. Answer is:

Use F statistic. 5 per cent critical value for $F(2, 742) = 2.96$ (approximately) and F given is 124.28. Reject the null.

(d) Sign of the bias and correlation between y_t and p_t should be discussed. Answer is:

The estimated coefficient of y_t has gone from -0.718 to 0.737 as a result of the inclusion of p_t in (ii). p_t has a highly significant coefficient in (ii); the estimates in (i) suffer from omitted variable bias. The bias on the coefficient of y_t is negative. Since the coefficient of p_t in (ii) is negative, this shows that the correlation between p_t and y_t is positive. This is clearly true for recent UK data. Incomes have risen and so has the relative price of cigarettes as the tax on cigarettes has risen.

Question 8

Consider the simple linear regression model

$$\boldsymbol{Y_t = \beta_0 + \beta_1 X_t + u_t; \quad t = 1, 2, \ldots, T}$$

where β_0 and β_1 are unknown coefficients. A "goodness of fit" measure for ordinary least squares estimation of the above model is defined by

$$\boldsymbol{R^2 = \frac{\text{Explained Sum of Squares (ESS)}}{\text{Total Sum of Squares (TSS)}}}$$

(a) Prove that

$$\boldsymbol{R^2 = 1 - \frac{\text{RSS}}{\text{TSS}},}$$

where the residual sum of squares RSS $= \sum_{t=1}^{T} \widehat{u}_t^2$, $\widehat{u} = Y_t - \widehat{\beta}_0 - \widehat{\beta}_1 X_t$;

TSS $= \sum_{t=1}^{T} (Y_t - \bar{Y})^2$; and ESS $= \sum_{t=1}^{T} (\widehat{Y}_t - \bar{Y}_t)^2$ for all t.

(8 marks)

(b) Hence, or otherwise, show that

$$\boldsymbol{0 \leq R^2 \leq 1.}$$

(4 marks)

(c) Briefly explain what happens to R^2 if an extra explanatory variable is added to the regression model.

(4 marks)

(d) Has R^2 any drawbacks? Explain.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapter 1.6 (Goodness of fit: R^2).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 3.5 (The coefficient of determination R^2: A measure of 'goodness of fit').

Approaching the question

(a) To answer Part (a), it is important to discuss that $\sum \widehat{u}_t X_t = 0$. Detailed answer is:

After running a regression we know:

$$Y_t = \widehat{Y}_t + e_t$$

hence:

$$\text{Var}(Y) = \text{Var}(\widehat{Y}) + \text{Var}(e) + 2\text{Cov}(\widehat{Y}, e).$$

At this stage we need to show that the covariance term vanishes hence:

$$\text{Cov}(\widehat{Y}, e) = \text{Cov}(\widehat{\beta}_1 + \widehat{\beta}_2 X, e) = \text{Cov}(\widehat{\beta}_1, e) + \text{Cov}(\widehat{\beta}_2 X, e) = 0 + \widehat{\beta}_2 \text{Cov}(X, e).$$

But $\text{Cov}(X, e) = 0$ from the minimisation of the residual sum of squares hence:

$$\text{Var}(Y) = \text{Var}(\widehat{Y}) + \text{Var}(e)$$

from which it follows that $\text{TSS} = \text{ESS} + \text{RSS}$.

The result $R^2 = 1 - \text{RSS}/\text{TSS}$ follows.

(b) It should be discussed that as both RSS and TSS are the sum of squares, they are positive. Answer is:

$$R^2 = 1 - \frac{\text{RSS}}{\text{TSS}}$$

and $\text{RSS} \leq \text{TSS}$.

RSS and TSS are sum of squares, therefore they are ≥ 0 .

If all the variations in the dependent variable are explained by the explanatory variables, then $\text{RSS} = 0$ and:

$$R^2 = 1 - \frac{0}{\text{TSS}} = 1.$$

If the explanatory variables do not explain any variation in the dependent variable, then $\text{RSS} = \text{TSS}$ and:

$$R^2 = 1 - \frac{\text{RSS}}{\text{TSS}} = 1 - \frac{\text{TSS}}{\text{TSS}} = 0.$$

Hence, $0 \leq R^2 \leq 1$.

(c) It should be discussed that TSS does not change with the addition of extra explanatory variables. Answer is:

As extra variables are added to an equation the RSS must reduce if the estimated coefficient for that variable is not zero. Hence R^2 must increase since the TSS will remain the same.

(d) Discussion of the situations under which R^2 should not be used is required. Detailed answer is:

The main drawback to R^2 is explained in (c) but it also is not a good measure of the fit of a regression if there is no constant term in the regression (assuming that the true intercept is not zero), and it cannot be used to compare regressions where the dependent variable is not the same; that is, one cannot say that one regression is better than another on the basis of the R^2 value, if the regressions have different dependent variables.

Question 9

A study of applications for home mortgages used the linear probability model

$$\boldsymbol{MORT_i = \beta_0 + \beta_1 INC_i + \beta_2 AGE_i + \beta_3 PROP_i + u_i; \quad i = 1, 2, \ldots, 700}$$

where

$\boldsymbol{MORT_i}$ **= 1 if a mortgage is granted to the** $\boldsymbol{i}$**-th applicant: 0 otherwise**

$\boldsymbol{INC_i}$ **= income of the** $\boldsymbol{i}$**-th applicant in thousands of pounds**

$\boldsymbol{AGE_i}$ **= age of the** $\boldsymbol{i}$**-th applicant in years**

$\boldsymbol{PROP_i}$ **= age of the property for which the mortgage is being applied.**

(a) The estimated coefficient for $\boldsymbol{INC_i}$ **was 1.02 with standard error 0.51. What is the interpretation of this coefficient?**

(4 marks)

(b) Why is R^2 meaningless in probit and logit models? What measures of 'goodness of fit' are applicable to probit and logit models?

(7 marks)

(c) Using a two variable linear model, show that the ordinary least squares estimator will be heteroscedastic if the dependent variable takes only values 0 and 1.

(9 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapters 10.1 (Linear probability model) and 10.2 (Logit analysis).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 15.2 (The linear probability model (LPM)).

Approaching the question

(a) It is important to discuss that in LPM the estimated probability of an event occurring may be greater than one or less than zero. Answer is:

$t = 1.02/0.51 = 2$ which is significantly different from 0. As INC increases by 1000, $MORT$ increases by 1.48 but since the estimated $MORT$ can be interpreted as a probability the prediction is likely to lie outside [0, 1].

(b) It should be discussed that as the dependent variable takes only two values R^2 is meaningless.

A brief discussion of the likelihood ratio test and pseudo-R^2 should be given. Answer is given below:

The definition of R^2 is:

$$R^2 = \frac{\text{ESS}}{\text{TSS}} = \frac{\text{TSS} - \text{RSS}}{\text{TSS}} = 1 - \frac{\text{RSS}}{\text{TSS}}$$

where TSS is the total sum of squares, ESS is the explained sum of squares and RSS is the residual sum of squares. Under logit and probit the dependent variable only takes two states, 0 and 1, hence TSS will take different values dependent on the coding of 'success' or 'failure' even though the independent variables are the same.

The possibilities for measuring goodness of fit are (i) the pseudo-R^2 defined by $1 - (\ln L / \ln L_0)$ where $\ln L$ is the unrestricted log likelihood and $\ln L_0$ is the log likelihood that would have been obtained with only the intercept in the regression. This has a minimum of 0 but the maximum will be less than 1 and, unlike R^2, it does not have a natural interpretation.

The alternative is (ii) the likelihood ratio statistic defined by $2\ln(L/L_0) = 2[\ln L - \ln L_0]$ which has in large samples, a chi-squared distribution with q degrees of freedom where q is the number of restrictions imposed by the null hypothesis. The null hypothesis is that the coefficients of the variables are all jointly zero.

The parameters of $\ln L$ are estimated by the maximum likelihood under the alternative hypothesis and $\ln L_0$ is estimated by the maximum likelihood under the null.

(c) It is required to derive the variance of the disturbance term. Detailed answer is:

Let the model be:

$$Y_i = \beta_1 + \beta_1 X_i + u_i; \quad i = 1, 2, \ldots, n \qquad \text{(i)}$$

where:

$$Y_i = \begin{cases} 1 & \text{if event occurs} \\ 0 & \text{if not.} \end{cases}$$

As Y_i takes only two values 1 or 0, therefore u_i can take only two values $1 - \beta_0 - \beta_1 X_i$ when $Y_i = 1$ and $-\beta_0 - \beta_1 X_i$ when $Y_i = 0$. Based on this we can write the probability distribution of u_i as:

Y_i	u_i	$f(u_i)$
1	$1-\beta_0-\beta_1 X_i$	$\beta_0+\beta_1 X_i$
0	$-\beta_0-\beta_1 X_i$	$1-\beta_0-\beta_1 X_i$

This probability distribution also satisfies the assumption that:

$$\mathrm{E}(u_i) = (1-\beta_0-\beta_1 X_i)(\beta_0+\beta_1 X_i) + (-\beta_0-\beta_1 X_i)(1-\beta_0-\beta_1 X_i) = 0.$$

We can write $\mathrm{Var}(u_i)$ as:

$$\begin{aligned}
\mathrm{Var}(u_i) = \mathrm{E}(u_i^2) &= (1-\beta_0-\beta_1 X_i)^2(\beta_0+\beta_1 X_i) + (-\beta_0-\beta_1 X_i)^2(1-\beta_0-\beta_1 X_i) \\
&= (1-\beta_0-\beta_1 X_i)(\beta_0+\beta_1 X_i)[(1-\beta_0-\beta_1 X_i) + (\beta_0+\beta_1 X_i)] \\
&= (\beta_0+\beta_1 X_i)(1-\beta_0-\beta_1 X_i) \\
&= \mathrm{E}(Y_i)[1-\mathrm{E}(Y_i)] \\
&= P_i(1-P_i), \quad \text{for all } i = 1, 2, \ldots, n.
\end{aligned}$$

Hence the disturbance term is heteroscedastic. This will make the OLS estimators inefficient.

Question 10

Consider a two equation linear model

$$\begin{aligned}
q_t &= \beta_1 p_t + \beta_2 y_t + u_t \qquad \text{(i)} \qquad \text{demand equation} \\
q_t &= \alpha p_t + e_t \qquad \text{(ii)} \qquad \text{supply equation}
\end{aligned}$$

$t = 1, 2, \ldots, T$; $\mathrm{E}(u_t) = \mathrm{E}(e_t) = 0$; $\mathrm{E}(u_t^2) = \sigma_u^2$; $\mathrm{E}(e_t^2) = \sigma_e^2$; $\mathrm{E}(u_t e_t) = \sigma_{ue}$; $\mathrm{E}(u_s e_t) = 0$ if $s \neq t$, for all $s, t = 1, 2, \ldots, T$ and variables are defined as:

$$\begin{aligned}
q_t &= \text{demand for good} \\
p_t &= \text{price for good} \\
y_t &= \text{personal disposable income}.
\end{aligned}$$

Quantity demanded is equal to the quantity supplied.

u_t and e_t are disturbance terms.

(a) Examine the identifiability of the above two equations.

(6 marks)

(b) Derive the two-stage least squares estimator of α and also examine its consistency.

(10 marks)

(c) Discuss, without derivation, what is the indirect least squares estimator of α.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics.* (4th edition) Chapter 9.3 (Instrumental variable estimation).

Dougherty, C. Subject guide, Chapter 9 (Simultaneous equation estimation).

Gujarati, D.N. and D.C. Porter *Basic econometrics.* (5th edition) [ISBN 9780071276252], Chapter 19.3 (Rules of identification) and 20.4 (Estimation of an overidentified equation: The method of two-stage least squares (2SLS)).

Approaching the question

(a) To answer this question order condition of identification should be used. Answer is:

Only Order condition of identification is required.

Order Condition: $R \geq G - 1$, where R is the number of restrictions imposed on the equation under consideration and G is the number of jointly dependent variables or total number of equations in the complete model.

1st equation:

$R = 0$ and $G - 1 = 2 - 1 = 1 \;\Rightarrow\; R < G - 1$. Equation is underidentified.

2nd equation:

$R = 1$ and $G - 1 = 2 - 1 = 1 \;\Rightarrow\; R = G - 1$. Equation is exactly identified.

(b) The two-stage least squares estimator should be derived and to prove consistency, it should be shown that $\text{plim}(\widehat{\alpha}) = \alpha$. Detailed answer is:

Model is:

$$\begin{aligned} q_t &= \beta_1 p_t + \beta_2 y_t + u_t \qquad \text{(i)} \qquad \text{demand equation} \\ q_t &= \alpha p_t + e_t \qquad \text{(ii)} \qquad \text{supply equation} \end{aligned}$$

$t = 1, 2, \ldots, T$. T is the sample size. u and e are disturbance terms.

$\text{E}(u_t) = \text{E}(e_t) = 0$; $\text{E}(u_t^2) = \sigma_u^2$; $\text{E}(e_t^2) = \sigma_e^2$; $\text{E}(u_t e_t) = \sigma_{ue}$; $\text{E}(u_t e_s) = 0$; $t = 1, 2, \ldots, T$.

q and p are endogenous variables.

In 2SLS, the RHS endogenous variable (in this case p), is replaced by its estimated value (in this case $\widehat{p}$). After that we apply OLS to the equation.

We want to estimate α_1. Estimation involves two stages.

First stage:

Write down the RF corresponding to the RHS endogenous variable, in this case p. Let it be:

$$p_t = \pi y_t + v_t. \qquad \text{(iii)}$$

As RF parameters can always be consistently estimated by the OLS, apply OLS to (iii) to get:

$$\widehat{p}_t = \widehat{\pi} y_t; \qquad \widehat{\pi} = \frac{\sum_{t=1}^{T} p_t y_t}{\sum_{t=1}^{T} y_t^2} \qquad \text{(iv)}$$

We can also write p_t as:

$$p_t = \widehat{p}_t + \widehat{v}_t \qquad \text{(v)}$$

where $\widehat{v}_t$ is the OLS estimate of v_t.

Second stage:

In the second stage instead of p_t write $\widehat{p}_t + \widehat{v}_t$ in (ii) to get:

$$q_t = \alpha \widehat{p}_t + (\alpha_1 \widehat{v}_t + e_t) \qquad \text{(vi)}$$

Apply OLS to (vi) to get 2SLS estimator of α as:

$$\begin{aligned}
\widehat{\alpha}_{2SLS} &= \frac{\sum_{t=1}^{T} \widehat{p}_t q_t}{\sum_{t=1}^{T} \widehat{p}_t^2} \\
&= \frac{\widehat{\pi} \sum_{t=1}^{T} q_t y_t}{\widehat{\pi}^2 \sum_{t=1}^{T} y_t^2}; \qquad \widehat{p}_t = \widehat{\pi} y_t \text{ from (iv)} \\
&= \frac{\sum_{t=1}^{T} q_t y_t}{\sum_{t=1}^{T} y_t^2} \frac{\sum_{t=1}^{T} y_t^2}{\sum_{t=1}^{T} p_t y_t}; \qquad \widehat{\pi} = \frac{\sum_{t=1}^{T} p_t y_t}{\sum_{t=1}^{T} y_t^2} \text{ from (iv)} \\
&= \frac{\sum_{t=1}^{T} q_t y_t}{\sum_{t=1}^{T} p_t y_t} \qquad \text{2SLS estimator of } \alpha.
\end{aligned}$$

Consistency of the 2SLS estimator:

We have:

$$\begin{aligned}
\widehat{\alpha}_{2SLS} &= \frac{\sum_{t=1}^{T} q_t y_t}{\sum_{t=1}^{T} p_t y_t} \\
&= \frac{\sum_{t=1}^{T} y_t(\alpha p_t + e_t)}{\sum_{t=1}^{T} p_t y_t} \\
&= \alpha \frac{\sum_{t=1}^{T} p_t y_t}{\sum_{t=1}^{T} p_t y_t} + \frac{\sum_{t=1}^{T} y_t e_t}{\sum_{t=1}^{T} p_t y_t} \\
&= \alpha + \frac{\sum_{t=1}^{T} y_t e_t}{\sum_{t=1}^{T} p_t y_t}.
\end{aligned}$$

Also:

$$\text{plim}(\widehat{\alpha}_{2SLS}) = \alpha + \frac{\text{plim}\left(\frac{1}{T}\sum_{t=1}^{T} y_t e_t\right)}{\text{plim}\left(\frac{1}{T}\sum_{t=1}^{T} p_t y_t\right)} = \alpha + \frac{\text{Cov}(y_t, e_t)}{\text{Cov}(p_t, y_t)} = \alpha$$

since $\text{Cov}(y_t, e_t) = 0$ by definition, as y_t is exogenous.

(c) This question is based upon the fact that the given equation is exactly identified. Answer is: ILS and 2SLS estimators are same as the equation under consideration is exactly identified.

This paper is not to be removed from the Examination Halls

UNIVERSITY OF LONDON **EC2020 ZA**

BSc degrees and Diplomas for Graduates in Economics, Management, Finance and the Social Sciences, the Diplomas in Economics and Social Sciences

Elements of Econometrics

Thursday, 5 May 2016 : 14:30 to 17:30

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Extracts from statistical tables are given after the final question on this paper.

Graph paper is provided at the end of this question paper. If used, it must be detached and fastened securely inside the answer book.

A calculator may be used when answering questions on this paper and it must comply in all respects with the specification given with your Admission Notice. The make and type of machine must be clearly stated on the front cover of the answer book.

If more questions are answered than requested, only the first answers attempted will be counted.

PLEASE TURN OVER

© University of London 2016

SECTION A

Answer all questions from this section.

1. A random variable Y has unknown population mean μ_Y and population variance σ_Y^2. A sample of n observations $\{Y_1,...,Y_n\}$ has been generated. Consider the following estimator:

$$\tilde{Y} = \frac{1}{n}\left(\frac{1}{2}Y_1 + \frac{3}{2}Y_2 + \frac{1}{2}Y_3 + \frac{3}{2}Y_4 + ... + \frac{1}{2}Y_{n-1} + \frac{3}{2}Y_n\right)$$

where the number of observations n is assumed to be even for convenience. Show that $\tilde{Y}$ is a consistent estimator of μ_Y. **(8 marks)**

2. Consider a two variable linear model:

$$Y_i = \beta_0 + \beta_1 X_i + u_i \qquad i = 1,2,...,n$$

where $E(u_i \mid X) = 0$ for all i; $E(u_i^2) = \sigma^2$; and $E(u_i u_j) = 0$ if $i \neq j$. Show that the ordinary least squares (OLS) estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased estimators of β_0 and β_1, respectively. **(8 marks)**

3. Show that the R^2 in the regression of Y on X (with an intercept) is the squared value of the sample correlation between X and Y (i.e. $R^2 = r_{XY}^2$). **(8 marks)**

4. Explain how you would estimate the parameters α and β for the model $(M_t / P)_t = \alpha r_t^{\beta}$ given data on money (M_t), prices (P_t) and the nominal interest rate (r_t). Interpret the results. **(8 marks)**

5. What do you understand by an instrumental variable (IV)? How you would estimate β in the model $y_t = \beta x_t + u_t$, where x_t and u_t are correlated, using z_t, $t = 1,2,...,T$ as an IV? Examine the consistency of the IV estimator. **(8 marks)**

© University of London 2016

SECTION B

Answer three questions from this section.

6. A researcher wants to test the hypothesis that drug companies practice price discrimination against less developed countries. To do this she estimates two demand for pharmaceuticals equations using data from 32 countries.

Equation A: $P_i = \underset{(6.27)}{32.60} + \underset{(0.24)}{1.66}\,GDPN_i - \underset{(0.25)}{0.88}\,CD_i + e_{1i} \qquad R^2 = 0.70 \;\; i = 1,\ldots,32$

Equation B: $\ln(P_i) = \underset{(0.24)}{2.26} + \underset{(0.13)}{0.95}\ln(GDPN_i) - \underset{(0.09)}{0.45}\ln(CD_i) + e_{2i} \qquad R^2 = 0.71 \;\; i = 1,\ldots,32$

where P_i is the price index for pharmaceuticals in each country, $GDPN_i$ is the gross domestic product (GDP) per capita for each country, and CD_i is the per capita consumption of pharmaceutical drugs. $\ln(P_i)$, $\ln(GDPN_i)$, $\ln(CD_i)$ are the logarithms of these quantities. Figures in parentheses are the estimated standard errors.

(a) For each of the estimated equations, are the estimated partial regression coefficients individually statistically significant at the 5 percent level of significance? Specify and justify the alternative hypothesis in each test. **(3 marks)**

(b) Test the hypothesis that the slope coefficients are all simultaneously zero in each equation. What do you conclude from your results? **(3 marks)**

(c) Show that in a log-log model, such as (B), the slope parameters are elasticities. **(4 marks)**

(d) The researcher claims that: (i) as the population of a country becomes richer the cost of drugs decreases, and hence the demand becomes less elastic; and (ii) as people consume more drugs the market becomes more developed, and hence more competitive. To what extent, if at all, do the results in (A) and (B) support these claims? **(6 marks)**

(e) Are serial correlation and/or heteroscedasticity likely to be a problem in the linear and log-linear regression models (A) and (B)? Explain. **(4 marks)**

© University of London 2016

7. The following estimates were calculated from a sample of 7,634 women respondents from the General Household Survey 1995. The dependent variable takes the value 1 if the woman was in paid employment, and 0 otherwise.

	OLS	Logit	Probit
high	0.093 (0.015)	0.423 (0.071)	0.259 (0.043)
noqual	–0.210 (0.013)	–0.898 (0.056)	–0.554 (0.035)
age	0.038 (0.003)	0.173 (0.124)	0.107 (0.008)
age2	–0.051 (0.003)	–0.230 (0.069)	–0.142 (0.009)
mar	0.024 (0.009)	0.103 (0.057)	0.063 (0.035)
Constant	–0.068 (0.049)	–2.587 (0.225)	–1.593 (0.137)

Where high is one if the respondent has a higher educational qualification, zero otherwise; noqual is one if the respondent has no qualifications, zero otherwise; age is age in years; age2 is $(\text{age} \times \text{age})/100$; mar is one if married, zero otherwise. Conventionally calculated standard errors are given in brackets for the ordinary least squares (OLS) results and asymptotic standard errors in brackets elsewhere.

(a) Explain briefly how Probit estimates are calculated when the model has no intercept and only one explanatory variable. **(6 marks)**

(b) For all three sets of estimates, test the null hypothesis that the coefficient of mar is zero. Which test statistics would you consider more reliable? Explain. **(6 marks)**

(c) Using the OLS and Probit estimates, calculate the estimated probabilities of being in employment for a married woman aged 40 with a higher educational qualification. Comment on your results. **(5 marks)**

(d) Test the null hypothesis that all the slope coefficients of the **probit** model are jointly equal to zero, given that:

$$\ln L_R = -416.01$$
$$\ln L_U = -321.25$$

where $\ln L_R$ and $\ln L_U$ are the logs of the likelihood from the restricted and the unrestricted **probit** models, respectively. **(3 marks)**

© University of London 2016

8. The relationship between a dependent variable Y and an explanatory variable X is given by the linear model

$$Y_i = \alpha + \beta X_i + u_i; \qquad i = 1, 2, \ldots, n$$

for an i.i.d. sample $\{Y_i, X_i\}$, but where the errors u_i are correlated with the explanatory variable, $E(X_i u_i) = \sigma_{xu} \neq 0$.

(a) Assuming that $\operatorname{plim}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2\right) = q_{xx}$, where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, show that the OLS estimator of β is inconsistent. **(4 marks)**

(b) Three possible instrumental variables z_1, z_2 and z_3 are available in the data set. What properties do these instruments need to have in order to be able to use them to estimate β consistently by the Instrumental Variable, or the Two Stage Least Squares (2SLS), estimator? **(4 marks)**

(c) Using all three instruments, describe in detail the 2SLS estimator. **(4 marks)**

(d) Assuming conditional homoskedasticity of the errors, $E(u_i^2 \mid z_i) = \sigma^2$, explain in detail how you would test whether $E(z_i u_i) = 0$. What are the consequences for the 2SLS estimator if the test result implies a rejection of this hypothesis? **(4 marks)**

(e) Explain in detail how would you test whether $E(X_i u_i) = 0$ (Durbin-Wu-Hausman test). Which estimator of β would you use if this hypothesis does not get rejected? **(4 marks)**

© University of London 2016

9. Answer the following questions:

(a) Explain what a trend stationary series and what a difference stationary series are. What is an important difference between the two types of stationarity? **(5 marks)**

(b) Let the model be:

$$Y_t = \varphi Y_{t-1} + u_t\,; \quad t = 1, 2, ..., T$$

where u_t are independently and identically distributed as $N(0, \sigma^2)$. Explain how would you test $H_0 : |\varphi| = 1$ against the alternative of a zero mean, covariance stationary AR(1) process. Give the assumptions this test requires. **(5 marks)**

(c) Consider an ADL(1,1) model:

$$Y_t = \alpha_1 + \alpha_2 Y_{t-1} + \alpha_3 X_t + \alpha_4 X_{t-1} + u_t$$

where both Y_t and X_t are I(1). Express the ADL(1,1) model in an error correction form and interpret the coefficients of the error correction model. Discuss the advantages of the error correction form. **(10 marks)**

10. (a) Consider a model:

$$Y_t = \beta X_t + u_t\,; \qquad t = 1, 2, ..., T$$

where $E(u_t) = 0; E(u_t^2) = \sigma^2$ and $E(u_s u_t) = 0$ if $s \neq t$ for all $s, t = 1, 2, ..., T$. X_t is fixed in repeated samples. The density function of u_t is given by

$$f(u_t) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{u_t^2}{2\sigma^2}\right)$$

i. Derive the likelihood function. **(6 marks)**

ii. Obtain the maximum likelihood estimators of β and σ^2. **(8 marks)**

(b) Consider a model:

$$Y_t = \alpha + \beta X_t + u_t\,; \qquad t = 1, 2, ..., T$$

where $E(u_t) = 0;\ E(u_t^2) = \sigma^2$ and $E(u_s u_t) = 0$ if $s \neq t$ for all $s, t = 1, 2, ..., T$. u is normally distributed. The parameters α and β have been estimated by maximum likelihood.

Explain how the hypothesis that the coefficients are jointly equal to zero can be tested. **(6 marks)**

END OF PAPER

© University of London 2016

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	**0·40**	0·6554	**0·80**	0·7881	**1·20**	0·8849	**1·60**	0·9452	**2·00**	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	**0·45**	0·6736	**0·85**	0·8023	**1·25**	0·8944	**1·65**	0·9505	**2·05**	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	**0·50**	0·6915	**0·90**	0·8159	**1·30**	0·9032	**1·70**	0·9554	**2·10**	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	**0·55**	0·7088	**0·95**	0·8289	**1·35**	0·9115	**1·75**	0·9599	**2·15**	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	**0·60**	0·7257	**1·00**	0·8413	**1·40**	0·9192	**1·80**	0·9641	**2·20**	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	**0·65**	0·7422	**1·05**	0·8531	**1·45**	0·9265	**1·85**	0·9678	**2·25**	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	**0·70**	0·7580	**1·10**	0·8643	**1·50**	0·9332	**1·90**	0·9713	**2·30**	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	**0·75**	0·7734	**1·15**	0·8749	**1·55**	0·9394	**1·95**	0·9744	**2·35**	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	**0·80**	0·7881	**1·20**	0·8849	**1·60**	0·9452	**2·00**	0·9772	**2·40**	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

When $x > 3·3$ the formula $1-\Phi(x) \doteqdot \frac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\frac{1}{x^2}+\frac{3}{x^4}-\frac{15}{x^6}+\frac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	5·0	1·6449	3·0	1·8808	2·0	2·0537	1·0	2·3263	0·10	3·0902
45	0·1257	4·8	1·6646	2·9	1·8957	1·9	2·0749	0·9	2·3656	0·09	3·1214
40	0·2533	4·6	1·6849	2·8	1·9110	1·8	2·0969	0·8	2·4089	0·08	3·1559
35	0·3853	4·4	1·7060	2·7	1·9268	1·7	2·1201	0·7	2·4573	0·07	3·1947
30	0·5244	4·2	1·7279	2·6	1·9431	1·6	2·1444	0·6	2·5121	0·06	3·2389
25	0·6745	4·0	1·7507	2·5	1·9600	1·5	2·1701	0·5	2·5758	0·05	3·2905
20	0·8416	3·8	1·7744	2·4	1·9774	1·4	2·1973	0·4	2·6521	0·01	3·7190
15	1·0364	3·6	1·7991	2·3	1·9954	1·3	2·2262	0·3	2·7478	0·005	3·8906
10	1·2816	3·4	1·8250	2·2	2·0141	1·2	2·2571	0·2	2·8782	0·001	4·2649
5	1·6449	3·2	1·8522	2·1	2·0335	1·1	2·2904	0·1	3·0902	0·0005	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

This paper is not to be removed from the Examination Halls

UNIVERSITY OF LONDON **EC2020 ZB**

BSc degrees and Diplomas for Graduates in Economics, Management, Finance and the Social Sciences, the Diplomas in Economics and Social Sciences

Elements of Econometrics

Thursday, 5 May 2016 : 14:30 to 17:30

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Extracts from statistical tables are given after the final question on this paper.

Graph paper is provided at the end of this question paper. If used, it must be detached and fastened securely inside the answer book.

A calculator may be used when answering questions on this paper and it must comply in all respects with the specification given with your Admission Notice. The make and type of machine must be clearly stated on the front cover of the answer book.

If more questions are answered than requested, only the first answers attempted will be counted.

PLEASE TURN OVER

© University of London 2016

SECTION A

Answer all questions from this section.

1. What do you understand by an instrumental variable (IV)? How you would estimate a model $y_t = \beta x_t + u_t$ using an IV z_t; $t = 1, 2, ..., T$? Examine the consistency of the IV estimator.

(8 marks)

2. Derive the order of integration of x_t in the following models. Assume in each case that u_t is stationary, x_0 is fixed and $E(u_s u_t) = 0$ if $s \neq t$.

(a) $x_t = \alpha_0 + u_t + u_{t-1}$; $t = 1, 2, \ldots, T$. **(3 marks)**

(b) $x_t = \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 t + u_t$; $|\alpha| < 1$; $t = 1, 2, \ldots, T$. **(5 marks)**

3. Consider a model:

$$Y_i = \alpha + \beta X_i + u_i \quad ; \quad i = 1, 2, \ldots, 6$$

where $E(u_i) = 0$; $E(u_i^2) = \sigma^2$ and $E(u_i u_j) = 0$ if $i \neq j$.

The observations on X_i are

X_1	X_2	X_3	X_4	X_5	X_6
1	2	3	4	5	6

The OLS estimator of β is $\hat{\beta}$ and $V(\hat{\beta}) = \dfrac{\sigma^2}{17.5}$.

An alternative estimator of β is $\tilde{\beta} = \dfrac{1}{8}[Y_6 + Y_5 - Y_2 - Y_1]$.

Compare the sampling variance of $\tilde{\beta}$ with that of $\hat{\beta}$. **(8 marks)**

© University of London 2016

4. Consider a model:

$$y_t = \alpha y_{t-1} + u_t$$
where
$$u_t = \theta\varepsilon_{t-1} + \varepsilon_t \ ; \ t = 1, 2, \ldots, T$$

and $E(\varepsilon_t) = 0, E(\varepsilon_t^2) = \sigma^2$ and $E(\varepsilon_s \varepsilon_t) = 0$ if $s \neq t$ for all $s, t = 1, 2, \ldots, T$.

Show that the OLS estimator of α is inconsistent if $\theta \neq 0$. **(8 marks)**

5. Consider an ADL(1, 1) model:

$$Y_t = \alpha_1 + \alpha_2 Y_{t-1} + \alpha_3 X_t + \alpha_4 X_{t-1} + u_t \ ; \ t = 1, 2, \ldots, T.$$

where both Y_t and X_t are I(1). u_t is the disturbance term where $E(u_t) = 0; E(u_t^2) = \sigma^2$ and $E(u_s u_t) = 0$ for all $s \neq t$. Express the ADL(1, 1) model in an error correction form and interpret the coefficients of the error correction model. **(8 marks)**

SECTION B

Answer three questions from this section.

6. (a) Describe an adaptive expectations model. **(6 marks)**

(b) Koyck investigated the relationship between investment in railcars and the volume of freight carried on the U.S. railroads using annual data for the period 1884–1939. Assuming that the desired stock of railcars in year t depended on the volume of the freight in year $t-1$ and $t-2$ and a time trend, and assuming that investment in railcars was subject to a partial adjustment process, he fitted the following regression using ordinary least squares:

$$\hat{I}_t = 0.077F_{t-1} + 0.017F_{t-2} - 0.003t - 0.110K_{t-1} \ ; R^2 = 0.85 \ ; t = 1, 2. \ldots, T$$

where $I_t = K_t - K_{t-1}$ is investment in railcars in year t (thousands), K_t is the stock of railcars at the end of year t (thousands), and F_t is the volume of freight handled in year t (ton-miles).

i. Explain how Koyck's model can be derived from an adjustment equation $K_t - K_{t-1} = \lambda(K_t^* - K_{t-1})$ and a behavioural equation with dependent variable K_t^*, using variables F (lagged) and t (time). **(6 marks)**

ii. Using Koyck's estimated equation estimate the parameters of your behavioural equation. What are the implications for the behaviour in the long run? **(8 marks)**

© University of London 2016

7. The following estimates were calculated from a sample of 7,634 women respondents from the General Household Survey 1995. The dependent variable takes the value 1 if the woman was in paid employment, and 0 otherwise.

	OLS	Logit	Probit
high	0.093 (0.015)	0.423 (0.071)	0.259 (0.043)
noqual	–0.210 (0.013)	–0.898 (0.056)	–0.554 (0.035)
age	0.038 (0.003)	0.173 (0.124)	0.107 (0.008)
age2	–0.051 (0.003)	–0.230 (0.069)	–0.142 (0.009)
mar	0.024 (0.009)	0.103 (0.057)	0.063 (0.035)
Constant	–0.068 (0.049)	–2.587 (0.225)	–1.593 (0.137)

Where high is one if the respondent has a higher educational qualification, zero otherwise; noqual is one if the respondent has no qualifications, zero otherwise; age is age in years; age2 is $(\text{age} \times \text{age})/100$; mar is one if married, zero otherwise. Conventionally calculated standard errors are given in brackets for the ordinary least squares (OLS) results and asymptotic standard errors in brackets elsewhere.

(a) Explain briefly how Probit estimates are calculated when the model has no intercept and only one explanatory variable. **(6 marks)**

(b) For all three sets of estimates, test the null hypothesis that the coefficient of mar is zero. Which test statistics would you consider more reliable? Explain. **(6 marks)**

(c) Using the OLS and Probit estimates, calculate the estimated probabilities of being in employment for a married woman aged 40 with a higher educational qualification. Comment on your results. **(5 marks)**

(d) Test the null hypothesis that all the slope coefficients of the **probit** model are jointly equal to zero, given that:

$$\ln L_R = -416.01$$
$$\ln L_U = -321.25$$

where $\ln L_R$ and $\ln L_U$ are the logs of the likelihood from the restricted and the unrestricted **probit** models, respectively. **(3 marks)**

© University of London 2016

8. Economists have tried to examine the catch-up hypothesis where it is predicted that poorer countries will grow faster than richer countries to 'catch-up' to the richer countries. To assess this hypothesis an economist regresses the growth rate of GDP (gross domestic product), gr_i, on log of GDP per capita, $\ln(gdp_i)$, as well as a regression including other variables which might influence growth rates. The following regression results come from regressions using data from 96 countries:

$$gr_i = \underset{(0.023)}{-0.058} - \underset{(0.002)}{0.009} \ln(gdp_i) + e_{1i} \qquad \text{(A)}$$

$n = 96$, $R^2 = 0.122$ and $s = 0.025$

where estimated standard errors are given in brackets, gr_i is the growth rate of the ith country over the period 1985-1990, $\ln(gdp_i)$ is the logarithm of GDP for country i, e_{1i} is the residual and s is the estimated standard error of the residuals.

A second regression yields

$$gr_i = -0.019 + \underset{(0.023)}{0.001} \ln(gdp_i) + \underset{(0.002)}{0.12} inv_i + e_{2i} \qquad \text{(B)}$$

$n = 96$, $R^2 = 0.196$ and $s = 0.024$

where e_{2i} is the estimated residual and inv_i is the ratio of investment to GDP in 1985.

(a) Test, for each regression, the hypothesis that the coefficient on $\ln(gdp_i)$ is zero. What does the result of your tests say about the 'catch-up' hypothesis? **(7 marks)**

(b) The R^2 measure is higher for equation (B) than for (A). Is this what you would expect? Explain. **(5 marks)**

(c) The regression results may be affected by heteroskedasticity. Explain what you understand about heteroskedasticity, why the estimation results might be affected by heteroskedasticity and what the effects are if heteroskedasticity is present. **(8 marks)**

9. (a) Explain the concept of the likelihood function and the maximum likelihood estimator. **(6 marks)**

(b) Suppose that an event occurs with probability p. In a simple random sample of size n the event occurs m times.

(i) Show that the maximum likelihood estimator (MLE) of p is $\frac{m}{n}$. Verify that the MLE maximises the likelihood function. **(8 marks)**

(ii) Derive the likelihood ratio statistic for the null hypothesis $p = p_0$. If $m = 40$ and $n = 100$, test the null hypothesis $p = 0.5$. **(6 marks)**

© University of London 2016

10. (a) There are six missing values (denoted by x's) in the given Stata output

i. Model SS
ii. Model df
iii. R-squared
iv. avetemp t value
v. harvrain 95% Conf. Interval.

Give the formulae for them and obtain their values based on the given output.

```
      Source |       SS       df       MS              Number of obs =      27
-------------+------------------------------           F(  3,    23) =   21.90
       Model |  xxxxxxxxxx    xx   2.58519261          Prob > F      =  0.0000
    Residual |  2.71468685    23  .118029863           R-squared     =  xxxxxx
-------------+------------------------------           Adj R-squared =  0.7069
       Total |  10.4702647    26  .402702488           Root MSE      =  .34355
------------------------------------------------------------------------------
     lnprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    wintrain |    .001282   .0005765     2.22   0.036     .0000894    .0024747
     avetemp |   .7123178   .1087674     xxxx   0.000     .4873154    .9373202
    harvrain |  -.0036242   .0009646    -3.76   0.001    xxxxxxxxx   xxxxxxxxx
       _cons |  -13.44433   1.969396    -6.83   0.000    -17.51834   -9.370326
------------------------------------------------------------------------------
```

where the explanatory variables are

wintrain: the level of winter rain (October - March) in millimetres
avetemp: average temperature in the growing season (April-September)
harvrain: level of harvest rain (August-September) in millimetres

and the dependent variable (lnprice) is the price of mature red wine from the Bordeaux region of France at harvest time. The data relate to wines of different vintages (ages) in 1987. **(10 marks)**

(b) Interpret the regression results in (a). **(3 marks)**

(c) If the age of the wine is included in the equation the estimates become:

```
      Source |       SS       df       MS              Number of obs =      27
-------------+------------------------------           F(  4,    22) =   26.39
       Model |  8.66443586     4  2.16610897           Prob > F      =  0.0000
    Residual |  1.80582883    22  .082083129           R-squared     =  0.8275
-------------+------------------------------           Adj R-squared =  0.7962
       Total |  10.4702647    26  .402702488           Root MSE      =   .2865
------------------------------------------------------------------------------
     lnprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    wintrain |   .0011668    .000482     2.42   0.024     .0001671    .0021665
     avetemp |   .6163926   .0951755     6.48   0.000     .4190107    .8137745
    harvrain |  -.0038606   .0008075    -4.78   0.000    -.0055353   -.0021858
         age |   .0238474   .0071667     3.33   0.003     .0089846    .0387103
       _cons |  -12.31227   1.677212    -7.34   0.000    -15.79059   -8.833945
------------------------------------------------------------------------------
```

i. What is the interpretation of the coefficient of the variable age? **(3 marks)**

ii. If the true model is given by the specification in (c) do you think that the model in (a) should show evidence of specification error? Does it? Carefully explain your observations. **(4 marks)**

END OF PAPER

© University of London 2016

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	5·0	1·6449	3·0	1·8808	2·0	2·0537	1·0	2·3263	0·10	3·0902
45	0·1257	4·8	1·6646	2·9	1·8957	1·9	2·0749	0·9	2·3656	0·09	3·1214
40	0·2533	4·6	1·6849	2·8	1·9110	1·8	2·0969	0·8	2·4089	0·08	3·1559
35	0·3853	4·4	1·7060	2·7	1·9268	1·7	2·1201	0·7	2·4573	0·07	3·1947
30	0·5244	4·2	1·7279	2·6	1·9431	1·6	2·1444	0·6	2·5121	0·06	3·2389
25	0·6745	4·0	1·7507	2·5	1·9600	1·5	2·1701	0·5	2·5758	0·05	3·2905
20	0·8416	3·8	1·7744	2·4	1·9774	1·4	2·1973	0·4	2·6521	0·01	3·7190
15	1·0364	3·6	1·7991	2·3	1·9954	1·3	2·2262	0·3	2·7478	0·005	3·8906
10	1·2816	3·4	1·8250	2·2	2·0141	1·2	2·2571	0·2	2·8782	0·001	4·2649
5	1·6449	3·2	1·8522	2·1	2·0335	1·1	2·2904	0·1	3·0902	0·0005	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

Examiners' commentaries 2016

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2015–16. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2011). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

General remarks

Learning outcomes

At the end of this course, and having completed the Essential reading and activities, you should be able to:

- describe and apply the classical regression model and its application to cross-section data
- describe and apply the:
  - Gauss–Markov conditions and other assumptions required in the application of the classical regression model
  - reasons for expecting violations of these assumptions in certain circumstances
  - tests for violations
  - potential remedial measures, including, where appropriate, the use of instrumental variables
- recognise and apply the advantages of logit, probit and similar models over regression analysis when fitting binary choice models
- competently use regression, logit and probit analysis to quantify economic relationships using standard regression programmes (Stata and EViews) in simple applications
- describe and explain the principles underlying the use of maximum likelihood estimation
- apply regression analysis to fit time-series models using stationary time series, with awareness of some of the econometric problems specific to time series applications (for example, autocorrelation) and remedial measures
- recognise the difficulties that arise in the application of regression analysis to nonstationary time series, know how to test for unit roots, and know what is meant by cointegration.

Common mistakes committed by candidates

- A large number of candidates were not able to distinguish between sample variance and covariance, and population variance and covariance (this is happening year after year). They treat them as the same. This results in incorrect analysis and candidates lose significant marks.

  Consider an example: Suppose data are deviations from the respective sample means and the regression model is:
  $$y_t = \beta x_t + u_t, \quad t = 1, 2, \ldots, T.$$
  The ordinary least squares estimator of β is:
  $$\widehat{\beta} = \frac{\sum_{t=1}^{T} x_t y_t}{\sum_{t=1}^{T} x_t^2} = \beta + \frac{\sum_{t=1}^{T} x_t u_t}{\sum_{t=1}^{T} x_t^2}.$$
  In terms of variances and covariances (a large number of candidates prefer this terminology), this can be written as:
  $$\widehat{\beta} = \beta + \frac{\mathrm{Cov}(x, u)}{\mathrm{Var}(x)}.$$
  Here $\mathrm{Cov}(x, u)$ and $\mathrm{Var}(x)$ are sample[$\mathrm{Cov}(x, u)$] and sample[$\mathrm{Var}(x)$].

  Candidates should realise that $\sum_{t=1}^{T} u_t$, $\sum_{t=1}^{T} x_t u_t$, $\mathrm{Cov}(x, u)$ and $\mathrm{Var}(x)$ given above are sample moments and as such $\sum_{t=1}^{T} u_t \neq 0$, $\sum_{t=1}^{T} x_t u_t \neq 0$ and $\mathrm{Cov}(x, u) \neq 0$. However, if we take the expectation, then:
  $$\mathrm{E}(u_t) = 0$$
  by assumption. Therefore:
  $$\mathrm{E}\left[\sum_{t=1}^{T} x_t u_t\right] = \sum_{t=1}^{T} x_t [\mathrm{E}(u_t)] = 0$$
  as the x_ts are fixed so they can be taken out of the expectation, and so:
  $$\mathrm{E}[\mathrm{Cov}(x, u)] = \mathrm{E}\left[\frac{1}{T}\sum_{t=1}^{T} x_t u_t\right] = 0$$
  as previously argued. This makes $\mathrm{E}(\widehat{\beta}) = \beta$, i.e. $\widehat{\beta}$ is an unbiased estimator of β.

  To prove consistency take the plim to get:
  $$\begin{aligned}
  \mathrm{plim}\left(\widehat{\beta}\right) &= \beta + \mathrm{plim}\left(\frac{\frac{1}{T}\sum_{t=1}^{T} x_t u_t}{\frac{1}{T}\sum_{t=1}^{T} x_t^2}\right) \\
  &= \beta + \frac{\mathrm{plim}\left(\frac{1}{T}\sum_{t=1}^{T} x_t u_t\right)}{\mathrm{plim}\left(\frac{1}{T}\sum_{t=1}^{T} x_t^2\right)} \\
  &= \beta + \frac{\mathrm{plim}(\text{sample } \mathrm{Cov}(x, u))}{\mathrm{plim}(\text{sample } \mathrm{Var}(x))} \\
  &= \beta + \frac{\text{population } \mathrm{Cov}(x, u)}{\text{population } \mathrm{Var}(x)}.
  \end{aligned}$$

By assumption, population $\text{Cov}(x, u) = 0$ and population $\text{Var}(x) > 0$, hence $\text{plim}\left(\widehat{\beta}\right) = \beta$, in other words $\widehat{\beta}$ is a consistent estimator of β.

Remember that in general:

$$\text{plim(sample variance)} = \text{population variance}$$

and:

$$\text{plim(sample covariance)} = \text{population covariance.}$$

This concept has been used in many questions. This simple mistake of not distinguishing between sample variance and covariance, and population variance and covariance, results in a significant loss of marks which might result in the loss of a degree class or even be the difference between pass and fail.

- Candidates struggled to give competent answers to the interpretation of empirical results. When interpreting an empirical result you should discuss the significance of the coefficients, magnitude and sign of the coefficients. Also, you should make sure that the Gauss–Markov conditions hold. Gauss–Markov conditions have to be explicitly specified. Only writing that the Gauss–Markov conditions hold is not sufficient.
- Just as last year, many candidates did not appear to read the questions carefully enough and often omitted to give answers to parts of questions which asked for details of such things as the assumptions necessary for a particular result to be true.

Key steps to improvement

Essential reading for this course includes the subject guide and the following.

> Dougherty, C. *Introduction to econometrics.* (Oxford: Oxford University Press, 2011) 4th edition [ISBN 9780199567089];
> http://global.oup.com/uk/orc/busecon/economics/dougherty4e/

Apart from Essential readings you should do some supplementary readings. One very good book of the same level is:

> Gujarati, D.N. and D.C. Porter *Basic econometrics.* (McGraw–Hill, 2009, International edition) 5th edition [ISBN 9780071276252].

To understand the subject clearly it is important to supplement C. Dougherty, *Introduction to econometrics* (fourth edition) with the subject guide **EC2020 Elements of econometrics** (2014), especially Chapter 10 which covers maximum likelihood.

It is very important to carefully go through the subject guide. The subject guide contains solutions to the questions given in the main textbook and also some additional questions and solutions. Working through these will improve your understanding of the subject.

The chapter in the subject guide on maximum likelihood (Chapter 10) includes some additional theory which has not been covered in the main textbook. It is important to read the additional theory given in the subject guide to have a better understanding of the principles of maximum likelihood and tests based on the likelihood function.

Please check the VLE course page for resources for this subject such as a downloadable copy of the subject guide **EC2020 Elements of econometrics** (2014), PowerPoint slideshows that provide graphical treatment of the topics covered in the textbook, datasets and statistical tables. Candidates should utilise datasets using standard regression programmes (STATA or EViews). This will help in the understanding of the subject.

Examination revision strategy

Many candidates are disappointed to find that their examination performance is poorer than they expected. This may be due to a number of reasons. The *Examiners' commentaries* suggest ways of addressing common problems and improving your performance. One particular failing is '**question spotting**', that is, confining your examination preparation to a few questions and/or topics which have come up in past papers for the course. This can have serious consequences.

We recognise that candidates may not cover all topics in the syllabus in the same depth, but you need to be aware that the examiners are free to set questions on **any aspect** of the syllabus. This means that you need to study enough of the syllabus to enable you to answer the required number of examination questions.

The syllabus can be found in the Course information sheet in the section of the VLE dedicated to each course. You should read the syllabus carefully and ensure that you cover sufficient material in preparation for the examination. Examiners will vary the topics and questions from year to year and may well set questions that have not appeared in past papers. Examination papers may legitimately include questions on any topic in the syllabus. So, although past papers can be helpful during your revision, you cannot assume that topics or specific questions that have come up in past examinations will occur again.

If you rely on a question-spotting strategy, it is likely you will find yourself in difficulties when you sit the examination. We strongly advise you not to adopt this strategy.

Examiners' commentaries 2016

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2015–16. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone A

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly.

Section A

Answer all questions from this section.

Question 1

A random variable Y has unknown population mean μ_Y and population variance σ_Y^2. A sample of n observations $\{Y_1, \ldots, Y_n\}$ has been generated. Consider the following estimator:

$$\tilde{Y} = \frac{1}{n}\left(\frac{1}{2}Y_1 + \frac{3}{2}Y_2 + \frac{1}{2}Y_3 + \frac{3}{2}Y_4 + \cdots + \frac{1}{2}Y_{n-1} + \frac{3}{2}Y_n\right)$$

where the number of observations n is assumed to be even for convenience. Show that $\tilde{Y}$ is a consistent estimator of μ_Y.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) Chapter R.14 (Probability limits and consistency).

Gujarati, D.N. and D.C. Porter. *Basic Econometrics* (fifth edition) (ISBN 9780071276252) Chapter A.7 (Statistical Inference: Estimation).

Approaching the question

The definition of consistency should be given. To answer this question, the sufficient condition of consistency should be used. The answer is as follows.

$\tilde{Y}$ is a consistent estimator of μ_Y if $\tilde{Y} \xrightarrow{p} \mu_Y$. $\tilde{Y}$ is constructed by applying a weight of 1/2 to the $n/2$ 'odd' observations, and a weight of 3/2 to the remaining $n/2$ observations. We have:

$$\begin{aligned}
\mathrm{E}(\tilde{Y}) &= \frac{1}{n}\left(\frac{1}{2}\mathrm{E}(Y_1) + \frac{3}{2}\mathrm{E}(Y_2) + \frac{1}{2}\mathrm{E}(Y_3) + \frac{3}{2}\mathrm{E}(Y_4) + \cdots + \frac{1}{2}\mathrm{E}(Y_{n-1}) + \frac{3}{2}\mathrm{E}(Y_n)\right) \\
&= \frac{1}{n}\left(\frac{1}{2} \times \frac{n}{2} \times \mu_Y + \frac{3}{2} \times \frac{n}{2} \times \mu_Y\right) \\
&= \mu_Y
\end{aligned}$$

and:

$$\begin{aligned}
\mathrm{var}(\tilde{Y}) &= \frac{1}{n^2}\left(\frac{1}{4}\mathrm{var}(Y_1) + \frac{9}{4}\mathrm{var}(Y_2) + \cdots + \frac{1}{4}\mathrm{var}(Y_{n-1}) + \frac{9}{4}\mathrm{var}(Y_n)\right) \\
&= \frac{1}{n^2}\left(\frac{1}{4} \times \frac{n}{2} \times \sigma_Y^2 + \frac{9}{4} \times \frac{n}{2} \times \sigma_Y^2\right) \\
&= 1.25\,\frac{\sigma_Y^2}{n}.
\end{aligned}$$

Therefore, $\tilde{Y}$ is an unbiased estimator of μ_Y, and because $\mathrm{var}(\tilde{Y}) \to 0$ as $n \to \infty$, $\tilde{Y}$ is consistent.

Question 2

Consider a two variable linear model:

$$\boldsymbol{Y_i = \beta_0 + \beta_1 X_i + u_i \qquad i = 1, 2, \ldots, n}$$

where $\mathrm{E}(u_i \mid X) = 0$ for all i; $\mathrm{E}(u_i^2) = \sigma^2$; and $\mathrm{E}(u_i u_j) = 0$ if $i \neq j$. Show that the ordinary least squares (OLS) estimators $\widehat{\beta}_0$ and $\widehat{\beta}_1$ are unbiased estimators of β_0 and β_1, respectively.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) 2.3 (The random components and unbiasedness of the OLS regression coefficients).

Dougherty, C. Subject guide Chapter 2 (Properties of the regression coefficients and hypothesis testing).

Gujarati, D.N. and D.C. Porter. *Basic Econometrics* (fifth edition) (ISBN 9780071276252) Chapter 3A.2 (Linearity and Unbiasedness Properties of Least-Squares Estimators).

Approaching the question

The derivation of the OLS estimators is not required. To show unbiasedness, candidates should take expectations of the estimators. The answer is as follows.

The OLS estimator of the slope β_1 is:

$$\begin{aligned}\widehat{\beta}_1 &= \frac{\sum(X_i-\bar{X})(Y_i-\bar{Y})}{\sum(X_i-\bar{X})^2}\\ &= \frac{\sum(X_i-\bar{X})([\beta_0+\beta_1X_i+u_i]-[\beta_0+\beta_1\bar{X}+\bar{u}])}{\sum(X_i-\bar{X})^2}\\ &= \beta_1+\frac{\sum(X_i-\bar{X})(u_i-\bar{u})}{\sum(X_i-\bar{X})^2}.\end{aligned}$$

Also:

$$\begin{aligned}\sum(X_i-\bar{X})(u_i-\bar{u}) &= \sum(X_i-\bar{X})u_i-\bar{u}\sum(X_i-\bar{X})\\ &= \sum(X_i-\bar{X})u_i-\bar{u}\left(\sum X_i-n\bar{X}\right)\\ &= \sum(X_i-\bar{X})u_i\end{aligned}$$

since $\sum X_i = n\bar{X}$. Hence:

$$\widehat{\beta}_1=\beta_1+\frac{\sum(X_i-\bar{X})u_i}{\sum(X_i-\bar{X})^2}=\beta_1+\sum a_iu_i$$

where $a_i=(X_i-\bar{X})/\sum(X_i-\bar{X})^2$. Therefore:

$$\mathrm{E}(\widehat{\beta}_1)=\mathrm{E}(\beta_1)+\mathrm{E}\left(\sum a_iu_i\right)=\beta_1+\sum a_i\,\mathrm{E}(u_i)=\beta_1$$

since $\mathrm{E}(u_i)=0$ by assumption.

The OLS estimator of the intercept β_0 is:

$$\widehat{\beta}_0=\bar{Y}-\widehat{\beta}_1\bar{X}$$

where $\widehat{\beta}_1$ is the OLS estimator of the slope β_1. Taking the expectation of $\widehat{\beta}_0$, we have:

$$\begin{aligned}\mathrm{E}(\widehat{\beta}_0)=\mathrm{E}(\bar{Y}-\widehat{\beta}_1\bar{X}) &= \mathrm{E}\left[\left(\beta_0+\beta_1\bar{X}+\frac{1}{n}\sum u_i\right)-\widehat{\beta}_1\bar{X}\right]\\ &= \beta_0+\mathrm{E}(\beta_1-\widehat{\beta}_1)\bar{X}+\frac{1}{n}\sum\mathrm{E}(u_i\,|\,X_i)\\ &= \beta_0\end{aligned}$$

where the third equality in the above equation has used the facts that $\widehat{\beta}_1$ is unbiased so $\mathrm{E}(\beta_1-\widehat{\beta}_1)=0$, and $\mathrm{E}(u_i\,|\,X_i)=0$.

Question 3

Show that the R^2 in the regression of Y on X (with an intercept) is the squared value of the sample correlation between X and Y (i.e. $R^2=r^2_{XY}$).

(8 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) Chapter 1.6 (Goodness of fit: R^2).

Dougherty, C. Subject guide Chapter 1 (Simple regression analysis).

Approaching the question

R^2 is given by:

$$R^2 = \frac{\text{Explained Sum of Squares (ESS)}}{\text{Total Sum of Squares (TSS)}}.$$

Simple algebraic manipulations of ESS and TSS is required to answer this question. The answer is as follows.

The coefficient of determination, R^2, is given by:

$$R^2 = \frac{\text{ESS}}{\text{TSS}} = \frac{\sum(\widehat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2}.$$

Given that:

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1\bar{X} \qquad \text{and} \qquad \widehat{\beta}_1 = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sum(X_i - \bar{X})^2}$$

we have:

$$\begin{aligned}
\text{ESS} = \sum(\widehat{Y}_i - \bar{Y})^2 &= \sum(\widehat{\beta}_0 + \widehat{\beta}_1 X_i - \bar{Y})^2 \\
&= \sum\left[\widehat{\beta}_1(X_i - \bar{X})\right]^2 \\
&= \widehat{\beta}_1^2 \sum(X_i - \bar{X})^2 \\
&= \frac{\left[\sum(X_i - \bar{X})(Y_i - \bar{Y})\right]^2}{\sum(X_i - \bar{X})^2}.
\end{aligned}$$

Hence:

$$\begin{aligned}
R^2 = \frac{\text{ESS}}{\sum(Y_i - \bar{Y})^2} &= \frac{\left[\sum(X_i - \bar{X})(Y_i - \bar{Y})\right]^2}{\sum(X_i - \bar{X})^2 \sum(Y_i - \bar{Y})^2} \\
&= \left[\frac{\frac{1}{n-1}\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\frac{1}{n-1}\sum(X_i - \bar{X})^2}\sqrt{\frac{1}{n-1}\sum(Y_i - \bar{Y})^2}}\right]^2 \\
&= \left[\frac{s_{XY}}{s_X s_Y}\right]^2 \\
&= r_{XY}^2.
\end{aligned}$$

Question 4

Explain how you would estimate the parameters α and β for the model $(M_t/P_t) = \alpha r_t^{\beta}$ given data on money (M_t), prices (P_t) and the nominal interest rate (r_t). Interpret the results.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) Chapter 4.2 (Logarithmic transformations).

Dougherty, C. Subject guide Chapter 4 (Transformation of variables).

Gujarati, D.N. and D.C. Porter. *Basic Econometrics* (fifth edition) (ISBN 9780071276252) Chapter 6.4 (Functional Forms of Regression Models).

Approaching the question

To answer this question, a logarithmic transformation of the given model is required. The answer is as follows.

We transform the equation into an estimable form by taking logs to give:

$$\log\left(\frac{M_t}{P_t}\right) = \log(\alpha) + \beta\log(r_t)$$

that is, a log-log model. Now we can estimate the parameters by using OLS. The estimate of α is given by exp(constant), and the slope is the estimate of β. In this model β is the interest rate elasticity of demand for money (the percentage change in the demand for money for a 1 per cent change in the interest rate). Alternatively, use a non-linear estimation process.

Question 5

What do you understand by an instrumental variable (IV)? How you would estimate $\boldsymbol{\beta}$ in the model $\boldsymbol{y_t = \beta x_t + u_t}$, where $\boldsymbol{x_t}$ and $\boldsymbol{u_t}$ are correlated, using $\boldsymbol{z_t}$, $\boldsymbol{t = 1, 2, \ldots, T}$ as an IV? Examine the consistency of the IV estimator.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fourth edition): Chapter 8.5 (Instrumental variables).

Dougherty, C. Subject guide (2011): Chapter 8.5 (Stochastic regressor and measurement errors).

Approaching the question

Candidates should define an instrumental variable (IV) and when it is used. It is required to prove that the IV estimator is consistent. The answer is as follows.

Consider the model:

$$y_t = \beta x_t + u_t, \qquad t = 1, 2, \ldots, T.$$

If x_t is not independently distributed of u_t then the OLS estimator of β will be inconsistent. Consider a variable z that is correlated with x but not correlated with u. z can be considered as an instrumental variable. An estimator of β based on z is known as an instrumental variable (IV) estimator. It is defined as:

$$\widehat{\beta}_{IV} = \frac{\sum z_t y_t}{\sum z_t x_t}.$$

It can be shown that $\widehat{\beta}_{IV}$ is a consistent estimator of β.

$$\widehat{\beta}_{IV} = \frac{\sum z_t y_t}{\sum z_t x_t} = \frac{\sum z_t(\beta x_t + u_t)}{\sum z_t x_t} = \beta + \frac{\sum z_t u_t}{\sum z_t x_t}$$

and:

$$\text{plim}(\widehat{\beta}_{IV}) = \beta + \frac{\text{plim}\left(\sum z_t u_t / T\right)}{\text{plim}\left(\sum z_t x_t / T\right)} = \beta$$

hence $\widehat{\beta}_{IV}$ is a consistent estimator of β.

Note that:

$$\text{plim}\left(\sum z_t u_t / T\right) = 0 \qquad \text{and} \qquad \text{plim}\left(\sum z_t x_t / T\right) \neq 0.$$

Section B

Answer three questions from this section.

Question 6

A researcher wants to test the hypothesis that drug companies practice price discrimination against less developed countries. To do this she estimates two demand for pharmaceuticals equations using data from 32 countries.

$$\textbf{Equation A:} \quad \underset{(6.27)}{P_i = 32.60} + \underset{(0.24)}{1.66 GDPN_i} - \underset{(0.25)}{0.88 CD_i} + e_{1i} \qquad R^2 = 0.70 \quad i = 1, \ldots, 32$$

$$\textbf{Equation B:} \quad \underset{(0.24)}{\ln(P_i) = 2.26} + \underset{(0.13)}{0.95 \ln(GDPN_i)} - \underset{(0.09)}{0.45 \ln(CD_i)} + e_{2i} \qquad R^2 = 0.71 \quad i = 1 \ldots, 32$$

where P_i is the price index for pharmaceuticals in each country, $GDPN_i$ is the gross domestic product (GDP) per capita for each country, and CD_i is the per capita consumption of pharmaceutical drugs. $\ln(P_i)$, $\ln(GDPN_i)$, $\ln(CD_i)$ are the logarithms of these quantities. Figures in parentheses are the estimated standard errors.

(a) **For each of the estimated equations, are the estimated partial regression coefficients individually statistically significant at the 5 percent level of significance? Specify and justify the alternative hypothesis in each test.**

(3 marks)

(b) **Test the hypothesis that the slope coefficients are all simultaneously zero in each equation. What do you conclude from your results?**

(3 marks)

(c) **Show that in a log-log model, such as (B), the slope parameters are elasticities.**

(4 marks)

(d) **The researcher claims that: (i) as the population of a country becomes richer the cost of drugs decreases, and hence the demand becomes less elastic; and (ii) as people consume more drugs the market becomes more developed, and hence more competitive. To what extent, if at all, do the results in (A) and (B) support these claims?**

(6 marks)

(e) **Are serial correlation and/or heteroscedasticity likely to be a problem in the linear and log-linear regression models (A) and (B)? Explain.**

(4 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) Chapters 2.6 (Testing hypotheses relating to the regression coefficients), 3.5 (Goodness of fit: R^2) and 4.2 (Logarithmic transformations).

Dougherty, C. Subject guide Chapter 2 (Properties of the regression coefficients and hypothesis testing).

Approaching the question

In part (a) individual coefficients are being tested, the t test is appropriate for this. Part (b) involves a test of a joint hypothesis. To conduct the test, an F test should be used. In part (d) the sign and significance of the coefficients should be explained. The answer is as follows.

(a) The critical t value for both equation (A) and equation (B) is 2.045 (degrees of freedom 29, 5% significance level for a 2-tailed test). The estimated t values are:

$$\text{Equation (A)} \qquad t = \frac{32.6}{6.27} = 5.2, \qquad t = \frac{1.66}{0.24} = 6.92 \qquad \text{and} \qquad t = -\frac{0.88}{0.25} = -3.52$$

$$\text{Equation (B)} \qquad t = \frac{2.26}{0.24} = 9.42, \qquad t = \frac{0.95}{0.13} = 7.31 \qquad \text{and} \qquad t = -\frac{0.45}{0.09} = -5.$$

These are all larger in absolute terms than the critical value. Hence we can say that all the coefficients are significantly different from zero. The estimated parameters show that as $GDPN$ increases so does the price of pharmaceuticals, the income elasticity of the price for pharmaceuticals is 0.95 (equation (B)). The estimated parameters also show that as CD increases the price for pharmaceuticals decreases with elasticity -0.45.

(b) To test the hypothesis that all coefficients are simultaneously zero we need to use an F test, which would normally be based on the residual sum of squares (RSS). Here the only information we have is the R^2, which contains the same information. The F test is:

$$F = \frac{R^2/k}{(1-R^2)/(n-k-1)} = \frac{0.70/2}{(1-0.70)/(32-2-1)} = 33.83$$

for equation (A), and the F critical value is 3.33 with $(2, 29)$ degrees of freedom. Hence we reject the null hypothesis of zero coefficients. Equation (B) is similar: replace the value of R^2 with 0.71, and repeat the computation. The answer is similar.

(c) If the equation is $\log(Y_i) = \beta_0 + \beta_1 \log(X_i) + u_i$, and we differentiate it we get:

$$\frac{\partial Y_i}{\partial X_i}\frac{1}{Y_i} = \beta_1 \frac{1}{X_i} \qquad \text{or} \qquad \frac{\partial Y_i}{\partial X_i}\frac{X_i}{Y_i} = \beta_1 = \text{elasticity.}$$

(d) The hypothesis might suggest that as the income per capita increases the price of drugs falls. It also says that as the income of the population increases the slope coefficient on equation (B) falls. We have shown that the first statement is not correct – the coefficient on $GDPN$ is positive and significant. The second statement is not covered by the regressions, which simply model fixed elasticities.

The second hypothesis says that as the consumption of drugs rises the market becomes more competitive. There is evidence that this is supported by the equations; the coefficient on CD (the consumption of drugs) is negative and significantly different from zero, which suggests that as the consumption of drugs increases the price falls (i.e. the market becomes more competitive).

(e) The data are cross-sectional, hence there is no possibility of serial correlation, but one might expect heteroskedasticity since countries are going to be of very different sizes, and hence the error term is likely to mirror some of these values.

Question 7

The following estimates were calculated from a sample of 7,634 women respondents from the General Household Survey 1995. The dependent variable takes the value 1 if the woman was in paid employment, and 0 otherwise.

	OLS	Logit	Probit
high	0.093	0.423	0.259
	(0.015)	(0.071)	(0.043)
noqual	−0.210	−0.898	−0.554
	(0.013)	(0.056)	(0.035)
age	0.038	0.173	0.107
	(0.003)	(0.124)	(0.008)
age2	−0.051	−0.230	−0.142
	(0.003)	(0.069)	(0.009)
mar	0.024	0.103	0.063
	(0.009)	(0.057)	(0.035)
Constant	−0.068	−2.587	−1.593
	(0.049)	(0.225)	(0.137)

Where high is one if the respondent has a higher educational qualification, zero otherwise; noqual is one if the respondent has no qualifications, zero otherwise; age is age in years; age2 is (age × age) 100; mar is one if married, zero otherwise. Conventionally calculated standard errors are given in brackets for the ordinary least squares (OLS) results and asymptotic standard errors in brackets elsewhere.

(a) **Explain briefly how Probit estimates are calculated when the model has no intercept and only one explanatory variable.**

(6 marks)

(b) **For all three sets of estimates, test the null hypothesis that the coefficient of mar is zero. Which test statistics would you consider more reliable? Explain.**

(6 marks)

(c) **Using the OLS and Probit estimates, calculate the estimated probabilities of being in employment for a married woman aged 40 with a higher educational qualification. Comment on your results.**

(5 marks)

(d) **Test the null hypothesis that all the slope coefficients of the probit model are jointly equal to zero, given that:**

$$\ln L_R = -416.01$$

$$\ln L_U = -321.25$$

where $\ln L_R$ and $\ln L_U$ are the logs of the likelihood from the restricted and the unrestricted probit models, respectively.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) Chapters 10.1 (The linear probability model), 10.3 (Probit analysis) and 10.6 (An introduction to maximum likelihood estimation).

Dougherty, C. Subject guide (2011) Chapter 10 (Binary choice and limited dependent variable models, and maximum likelihood estimation).

Gujarati, D.N. and D.C. Porter *Basic Econometrics* (fifth edition): (ISBN 9780071276252), Chapters 15.2 (The Linear Probability Model (LPM)), 15.5 (The Logit Model) and 15.9 (The Probit Model).

Approaching the question

(a) The probit model uses the cumulative standardised normal distribution. The maximum likelihood technique is used to obtain the estimates of the parameters. Estimates have the standard maximum likelihood properties, i.e. the estimators are consistent, asymptotically efficient and asymptotically normally distributed.

[For technical details see Dougherty (fourth edition) Section 10.3].

(b) To test the null hypothesis the t test should be used. t tables are attached with the examination paper. Candidates should know how to look at the critical values. The null and alternative hypotheses should be clearly stated:

$$\mathrm{H}_0 : \text{coefficient of mar} = 0$$

$$\mathrm{H}_1 : \text{coefficient of mar} \neq 0.$$

t test statistics are 2.67 (OLS), 1.81 (logit), and 1.8 (probit).

The test based on the OLS estimates gives outright rejection. However, the standard errors for this test are wrongly calculated because of the heteroskedasticity of the error term. Therefore, we prefer probit or logit estimates.

(c) A large number of candidates wrongly thought that to obtain the probability in the case of probit a sophisticated calculator is needed. The probability in the case of probit can be read directly from the supplied statistical tables. We have:

$$0.093 + 0.038 \times 40 - 0.051 \times 16 + 0.024 \times 0.068 = 0.753 \qquad \text{(OLS)}$$

$$0.259 + 0.107 \times 40 - 0.142 \times 16 + 0.063 - 1.593 = 0.737 \qquad \text{(probit)}.$$

The statistical tables give a probability of 0.77, approximately.

They are all fairly close. The OLS estimates do not fall outside the probability bounds.

(d) To test the null hypothesis, the large sample likelihood-ratio test should be used. In large samples $-2(\ln L_R - \ln L_U)$ is distributed as a chi-squared distribution with degrees of freedom equal to the number of restrictions imposed by the null hypothesis. The null and alternative hypotheses should be clearly stated as:

$$\mathrm{H}_0 : \text{all the slope coefficients are} = 0$$

$$\mathrm{H}_1 : \text{at least one slope coefficient is} \neq 0.$$

We have:

$$-2(\ln L_R - \ln L_U) \sim \chi^2_5$$

and:

$$-2(-416.01 - (-321.25)) = 189.92.$$

The critical value of χ^2_5 at the 5% significance level is 11.07, hence we reject H_0.

Question 8

The relationship between a dependent variable Y and an explanatory variable X is given by the linear model

$$Y_i = \alpha + \beta X_i + u_i; \qquad i = 1, 2, \ldots, n$$

for an i.i.d. sample $\{Y_i, X_i\}$, but where the errors u_i are correlated with the explanatory variable, $\mathrm{E}(X_i u_i) = \sigma_{Xu} \neq 0$.

(a) **Assuming that** $\text{plim}\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2\right)=\sigma_{XX}$**, where** $\bar{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$**, show that the OLS estimator of** β **is inconsistent.**

(4 marks)

(b) **Three possible instrumental variables** z_1**,** z_2 **and** z_3 **are available in the data set. What properties do these instruments need to have in order to be able to use them to estimate** β **consistently by the Instrumental Variable, or the Two Stage Least Squares (2SLS), estimator?**

(4 marks)

(c) **Using all three instruments, describe in detail the 2SLS estimator.**

(4 marks)

(d) **Assuming conditional homoskedasticity of the errors,** $\text{E}(u_i^2\,|\,z_i)=\sigma^2$**, explain in detail how you would test whether** $\text{E}(z_iu_i)=0$**. What are the consequences for the 2SLS estimator if the test result implies a rejection of this hypothesis?**

(4 marks)

(e) **Explain in detail how would you test whether** $\text{E}(X_iu_i)=0$ **(Durbin–Wu–Hausman test). Which estimator of** β **would you use if this hypothesis does not get rejected?**

(4 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) Chapters 8.3 (Asymptotic properties of the OLS regression coefficients), 8.5 (Instrumental variables) and 9.3 (Instrumental variable estimation).

Dougherty, C. Subject guide Chapters 8 (Stochastic regressors and measurement errors) and 9 (Simultaneous equations estimation).

Approaching the question

The plim should be used to verify the consistency of the OLS estimators. Properties of the instrumental variables should be clearly explained. The answer is as follows.

(a) We have:

$$\begin{aligned}\text{plim}(\widehat{\beta}_{OLS}-\beta)=\text{plim}\left(\frac{\sum(X_i-\bar{X})u_i}{\sum(X_i-\bar{X})^2}\right) &= \text{plim}\left(\frac{\text{plim}\left(\frac{1}{n}\left(\sum_i X_iu_i-u_i\frac{1}{n}\sum_j X_j\right)\right)}{\text{plim}\left(\frac{1}{n}\sum(X_i-\bar{X})^2\right)}\right)\\ &= \frac{\left(1-\frac{1}{n}\right)\sigma_{Xu}}{\sigma_{XX}}\neq 0.\end{aligned}$$

(b) Candidates need to explain that z_i must be correlated with X_i but not with u_i, and does not already appear in the equation in its own right. Hence in the reduced form:

$$X_i=\pi_0+\pi_1z_{i1}+\pi_2z_{i2}+\pi_3z_{i3}+\omega_i$$

π_1, π_2 and π_3 should all not be zero (if one is zero then it should be dropped from the list). Furthermore, $\text{E}(z_iu_i)=0$.

(c) Estimate by OLS:

$$\widehat{X}_i=\widehat{\pi}_0+\widehat{\pi}_1z_{i1}+\widehat{\pi}_2z_{i2}+\widehat{\pi}_3z_{i3}.$$

Next, estimate by OLS:

$$Y_i=\alpha+\beta\widehat{X}_i+\varepsilon_i$$

in the second stage. $\widehat{X}$ only utilises exogenous variation due to the instruments, and hence identifies the causal effect β.

(d) nR^2 test: Regress the IV residuals:

$$\widehat{u}_{IV} = Y_i - \widehat{\alpha}_{IV} - \widehat{\beta}_{IV} X_i$$

on the instruments:

$$\widehat{u}_{IV} = \delta_0 + \delta_1 z_{i1} + \delta_2 z_{i2} + \delta_3 z_{i3} + \zeta_i.$$

nR^2 in this regression follows a chi-squared distribution with 2 degrees of freedom under the null hypothesis that $\mathrm{E}(z_i u_i) = 0$. If the test statistic is large, then reject the null hypothesis. If the test rejects the null hypothesis, then the IV estimator is not consistent.

(e) Add the first-stage residuals to the model:

$$Y_i = \alpha + \beta X_i + \gamma \widehat{\omega}_i + \xi_i.$$

Testing for exogeneity is a test of $\mathrm{H}_0 : \gamma = 0$. A t test is appropriate as standard errors are correct under the null hypothesis. If the test does not reject the null hypothesis use the OLS estimator of β.

Question 9

Answer the following questions:

(a) Explain what a trend stationary series and what a difference stationary series are. What is an important difference between the two types of stationarity?

(5 marks)

(b) Let the model be:

$$Y_t = \phi Y_{t-1} + u_t; \qquad t = 1, 2, \ldots, T$$

where u_t are independently and identically distributed as $N(0, \sigma^2)$. Explain how would you test $\mathrm{H}_0 : |\phi| = 1$ against the alternative of a zero mean, covariance stationary AR(1) process. Give the assumptions this test requires.

(5 marks)

(c) Consider an ADL(1,1) model:

$$Y_t = \alpha_1 + \alpha_2 Y_{t-1} + \alpha_3 X_t + \alpha_4 X_{t-1} + u_t$$

where both Y_t and X_t are I(1). Express the ADL(1,1) model in an error correction form and interpret the coefficients of the error correction model. Discuss the advantages of the error correction form.

(10 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) Chapters 13.1 (Stationarity and nontationarityy), 13.4 (Tests of nonstationarity) and 13.6 (Fitting models with nonstationary time series).

Dougherty, C. Subject guide Chapter 13 (Introduction to nonstationary time series).

Gujarati, D.N. and D.C. Porter *Basic Econometrics* (fifth edition) (ISBN 9780071276252): Chapters 21.5 (Trend Stationary (TS) and Difference Stationary (DS) Stochastic Processes) and 21.9 (The Unit Root Test).

Approaching the question

In part (a) the meaning of a trend stationary series and a difference stationaary series should be explained. Part (b) requires a discussion of the Dickey–Fuller test, and in part (c) the error correction model (ECM) should be derived and the interpretation of the coefficients of the ECM should be given. The answer is as follows.

(a) If after removing the trend from a nonstationary series the resulting variable becomes stationary, then the variable is called *trend stationary*. Let:

$$Z_t = X_t - \alpha_1 t = \alpha_0 + u_t$$

where $\mathrm{E}(u_t) = 0$, $\mathrm{var}(u_t) = \sigma^2$ and $\mathrm{E}(u_t u_{t-s}) = 0$ for all s and t, then:

$$\begin{aligned}
\mathrm{E}(Z_t) &= \mathrm{E}(\alpha_0 + u_t) = \alpha_0 \\
\mathrm{var}(Z_t) &= \mathrm{Var}(\alpha_0 + u_t) = \sigma^2 \\
\mathrm{cov}(Z_t, Z_{t-s}) &= \mathrm{E}[(Z_t - \mathrm{E}(Z_t))(Z_{t-s} - \mathrm{E}(Z_{t-s}))] = \mathrm{E}(u_t u_{t-s}) = 0.
\end{aligned}$$

This means that Z_t has constant mean and variance for all t, and covariance is zero for all $s > 0$. It implies that the series is trend stationary.

If a nonstationary process can be transformed into a stationary process by differencing then the series is said to be *difference stationary*.

Let X_t be a random walk with a drift:

$$X_t = \beta_0 + X_{t-1} + \varepsilon_t \qquad \text{(i)}$$

where $\mathrm{E}(\varepsilon_t) = 0$, $\mathrm{var}(\varepsilon_t) = \sigma^2$ and $\mathrm{E}(\varepsilon_t \varepsilon_s) = 0$ for all s and t, $s \neq t$.

Subtract X_{t-1} from both sides of (i) to get:

$$\Delta X_t = X_t - X_{t-1} = \beta_0 + \varepsilon_t.$$

It can be easily checked that $\mathrm{E}(\Delta X_t) = \beta_0$, $\mathrm{var}(\Delta X_t) = \sigma_\varepsilon^2$ and $\mathrm{cov}(\Delta X_t, \Delta X_{t-s}) = 0$ for all s and t, $s \neq t$. This means that ΔX_t is stationary. This implies that X_t is difference stationary. It is important to know whether a variable is difference stationary or trend stationary because for difference stationary variables shocks have a permanent effect whereas for trend stationary variables shocks are transitory.

(b) The standard test for a unit root is due to Dickey and Fuller. In order to test the null hypothesis of a random walk without drift against the alternative of a zero mean covariance stationary AR(1) process:

$$y_t = \phi y_{t-1} + \varepsilon_t$$

subtract y_{t-1} from both sides:

$$\begin{aligned}
y_t - y_{t-1} &= \phi y_t - y_{t-1} + \varepsilon_t \\
\Delta y_t &= (\phi - 1) y_{t-1} + \varepsilon_t = \rho y_{t-1} + \varepsilon_t.
\end{aligned}$$

Test $\mathrm{H}_0 : \rho = 0$ using:

$$\widehat{\tau} = \frac{\widehat{\rho} - 0}{\mathrm{s.e.}(\widehat{\rho})}.$$

We cannot use the standard t test procedure in this case because the distribution of the statistic is not a t distribution, so critical values have to be computed by Dickey and Fuller using Monte Carlo techniques (Dickey–Fuller tables). The test is sensitive to the presence of serial correlation in the error term so we need to take steps to remove the effects of this serial correlation – this is done by including lagged values of Δy_t in the regression. The statistic has the same asymptotic distribution as $\widehat{\tau}$.

(c) Consider a simple ADL(1,1) model (this is also known as ARDL(1, 1)):

$$Y_t = \alpha_1 + \alpha_2 Y_{t-1} + \alpha_3 X_t + \alpha_4 X_{t-1} + u_t \qquad \text{(i)}$$

Rewrite (i) as:

$$\begin{aligned}
Y_t - Y_{t-1} &= \alpha_1 + \alpha_2 Y_{t-1} - Y_{t-1} + \alpha_3 X_t - \alpha_3 X_{t-1} + \alpha_3 X_{t-1} + \alpha_4 X_{t-1} + u_t \\
\Delta Y_t &= \alpha_1 - (1 - \alpha_2) Y_{t-1} + \alpha_3 \Delta X_t + (\alpha_3 + \alpha_4) X_{t-1} + u_t \\
\Delta Y_t &= \alpha_3 \Delta X_t - (1 - \alpha_2) \left[Y_{t-1} - \frac{\alpha_1}{1 - \alpha_2} - \frac{\alpha_3 + \alpha_4}{1 - \alpha_2} X_{t-1} \right] + u_t \\
\Delta Y_t &= \alpha_3 \Delta X_t - (1 - \alpha_2)[Y_{t-1} - \beta_1 - \beta_2 X_{t-1}] + u_t
\end{aligned}$$

or:

$$\Delta Y_t = \alpha_3 \Delta X_t - \pi[Y_{t-1} - \beta_1 - \beta_2 X_{t-1}] + u_t \qquad \text{(ii)}$$

where:

$$\pi = 1 - \alpha_2, \qquad \beta_1 = \frac{\alpha_1}{1-\alpha_2} \qquad \text{and} \qquad \frac{\alpha_3 + \alpha_4}{1-\alpha_2}.$$

Equation (ii) is the ECM.

When the two variables Y and X are cointegrated, the ECM incorporates not only the short-run effects but also the long-run effects. The long-run equilibrium $Y_t - \beta_1 - \beta_2 X_{t-1}$ is included in the model together with the short-run effect captured by the differenced term.

All the terms in the ECM, given by (ii), are stationary. As Y and X are I(1), then ΔX and ΔY are I(0). As Y and X are cointegrated their linear combination:

$$u_t = Y_t - \beta_1 - \beta_2 X_{t-1}$$

is I(0).

The coefficient π provides us with the information about the speed of adjustment in cases of disequilibrium.

- If $\pi = 1$ then 100% of the adjustment takes place within the period. In other words adjustment is instantaneous and full.
- If $\pi = 0.5$ then 50% adjustment takes place each period.
- If $\pi = 0$ then there is no adjustment.

Question 10

(a) Consider a model:

$$\boldsymbol{Y_t = \beta X_t + u_t; \qquad t = 1, 2, \ldots, T}$$

where $\mathrm{E}(u_t) = 0$; $\mathrm{E}(u_t^2) = \sigma^2$ and $\mathrm{E}(u_s u_t) = 0$ if $s \neq t$ for all $s, t = 1, 2, \ldots, T$. X_t is fixed in repeated samples. The density function of u_t is given by:

$$\boldsymbol{f(u_t) = (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{u_t^2}{2\sigma^2}\right).}$$

i. Derive the likelihood function.

(6 marks)

ii. Obtain the maximum likelihood estimators of β and σ^2.

(8 marks)

(b) Consider a model:

$$\boldsymbol{Y_t = \alpha + \beta X_t + u_t; \qquad t = 1, 2, \ldots, T}$$

where $\mathrm{E}(u_t) = 0$; $\mathrm{E}(u_t^2) = \sigma^2$ and $\mathrm{E}(u_s u_t) = 0$ if $s \neq t$ for all $s, t = 1, 2, \ldots, T$. u is normally distributed. The parameters α and β have been estimated by maximum likelihood.

Explain how the hypothesis that the coefficients are jointly equal to zero can be tested.

(6 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fourth edition) Chapter 10.6 (An introduction to maximum likelihood estimation).

Dougherty, C. Subject guide Chapter 10 (Binary choice and limited dependent variable models, and maximum likelihood estimation).

Approaching the question

Part (a) involves derivation of the likelihood function. To obtain the maximum likelihood estimators the likelihood function should be differentiated with respect to the parameters and equated to zero. Part (b) involves the likelihood-ratio test. The solution is as follows.

(a) i. The likelihood function is:

$$L = (2\pi\sigma^2)^{-T/2} \exp\left(-\frac{\sum(Y_t - \beta X_t)^2}{2\sigma^2}\right).$$

ii. The log-likelihood function is:

$$\log L = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln(\sigma^2) - \frac{\sum(Y_t - \beta X_t)^2}{2\sigma^2}.$$

Differentiating, we have:

$$\frac{\partial \log L}{\partial \beta} = -\frac{2\sum(Y_t - \beta X_t)(-X_t)}{2\sigma^2} = 0 \qquad \Rightarrow \qquad \widehat{\beta}_{MLE} = \frac{\sum X_t Y_t}{\sum X_t^2}$$

and:

$$\frac{\partial \log L}{\partial \sigma^2} = -\frac{T}{2\sigma^2} + \frac{\sum(Y_t - \beta X_t)^2}{2\sigma^4} = 0 \qquad \Rightarrow \qquad \widehat{\sigma}^2_{MLE} = \frac{\sum(Y_t - \widehat{\beta}_{MLE} X_t)^2}{T}.$$

(b) This can be tested using the likelihood-ratio statistic $2(\log L - \log L_0)$. This is asymptotically distributed as a chi-squared distribution with degrees of freedom equal to the number of restrictions imposed by the null hypothesis, 2 in the present case.

Examiners' commentaries 2016

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2015–16. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone B

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly.

Section A

Answer all questions from this section.

Question 1

What do you understand by an instrumental variable (IV)? How you would estimate a model $y_t = \beta x_t + u_t$ using an IV z_t; $t = 1, 2, \ldots, T$? Examine the consistency of the IV estimator.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fourth edition): Chapter 8.5 (Instrumental variables).

Dougherty, C. Subject guide: Chapter 8.5 (Stochastic regressor and measurement errors).

Approaching the question

Consider the model:

$$y_t = \beta x_t + u_t, \qquad t = 1, 2, \ldots, T.$$

If x_t is not independently distributed of u_t then the OLS estimator of β will be inconsistent. In this situation an IV should be used. An IV should be:

- correlated with x
- should not be correlated with the disturbance term
- should not be an explanatory variable in its own right.

Let z be an IV. An estimator of β based on z is known as an IV estimator. It is defined as:

$$\widehat{\beta}_{IV} = \frac{\sum z_t y_t}{\sum z_t x_t}.$$

It can be shown that $\widehat{\beta}_{IV}$ is a consistent estimator of β.

$$\widehat{\beta}_{IV} = \frac{\sum z_t y_t}{\sum z_t x_t} = \frac{\sum z_t(\beta x_t + u_t)}{\sum z_t x_t} = \beta + \frac{\sum z_t u_t}{\sum z_t x_t}$$

and:

$$\text{plim}(\widehat{\beta}_{IV}) = \beta + \frac{\text{plim}\left(\sum z_t u_t/T\right)}{\text{plim}\left(\sum z_t x_t/T\right)} = \beta$$

hence $\widehat{\beta}_{IV}$ is a consistent estimator of β.

Note that:

$$\text{plim}\left(\sum z_t u_t/T\right) = 0 \qquad \text{and} \qquad \text{plim}\left(\sum z_t x_t/T\right) \neq 0.$$

Question 2

Derive the order of integration of x_t in the following models. Assume in each case that u_t is stationary, x_0 is fixed and $\mathrm{E}(u_s u_t) = 0$ if $s \neq t$.

(a) $x_t = \alpha_0 + u_t + u_{t-1}$; $t = 1, 2, \ldots, T$.

(3 marks)

(b) $x_t = \alpha_0 + \alpha_1 x_{t-1} + \alpha_2 t + u_t$; $|\alpha| < 1$; $t = 1, 2, \ldots, T$.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition): Chapter 13.1 (Stationarity and nonstationarity).

Dougherty, C. Subject guide: Chapter 13 (Introduction to nonstationary time series).

Gujarati, D.N. and D.C. Porter *Basic Econometrics* (fifth edition): (ISBN 9780071276252), Chapters 21.5 (Trend Stationary (TS) and Difference Stationary (DS) Stochastic Processes) and 21.6 (Integrated Stochastic Process).

Approaching the question

In both (a) and (b), it should be shown that the mean and the variance of the variable is constant and the covariances are independent of time but may depend on the length of the lag. The solution is as follows.

(a) $\mathrm{E}(x_t) = \alpha_0$ and $\mathrm{var}(x_t) = \mathrm{var}(u_t - u_{t-1}) = \mathrm{var}(u_t) + \mathrm{var}(u_{t-1}) = 2\sigma^2$ as $\mathrm{E}(u_s u_t) = 0$ if $s \neq t$ and $\mathrm{E}(x_t x_{t-s}) = \sigma^2$ if $s = 1$ and 0 if $s > 1$. Therefore, x_t is stationary or I(0).

(b) Lag this equation by one period to get:

$$x_{t-1} = \alpha_0 + \alpha_1 x_{t-2} + \alpha_2(t-1) + u_{t-1}.$$

Subtracting this from the first equation gives:

$$\Delta x_t = \alpha_1 \Delta x_{t-1} + \alpha_2 + u_t - u_{t-1}.$$

Since $|\alpha_1| < 1$, Δx_t follows a stationary AR(1) process.

As $\mathrm{E}(x_t) = \alpha_0 + \alpha_1 \mathrm{E}(x_{t-1}) + \alpha_2 t$, $\mathrm{E}(x_t)$ is clearly a function of time and $\mathrm{E}(x_t) \neq \mathrm{E}(x_{t-1})$. Therefore, x_t is non-stationary. Hence x_t is I(1) and it is trend stationary.

Question 3

Consider a model:

$$\boldsymbol{Y_i = \alpha + \beta X_i + u_i; \qquad i = 1, 2, \ldots, 6}$$

where $\mathrm{E}(u_i) = 0$; $\mathrm{E}(u_i^2) = \sigma^2$ and $\mathrm{E}(u_i u_j) = 0$ if $i \neq j$.

The observations on X_i are

X_1	X_2	X_3	X_4	X_5	X_6
1	2	3	4	5	6

The OLS estimator of β is $\widehat{\beta}$ and $\mathrm{V}(\widehat{\beta}) = \frac{\sigma^2}{17.5}$.

An alternative estimator of β is $\tilde{\beta} = \frac{1}{8}[Y_6 + Y_5 - Y_2 - Y_1]$.

Compare the sampling variance of $\tilde{\beta}$ with that of $\widehat{\beta}$.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (third edition) Chapters R.5, R.6, 1.4, 2.5, 12.3 and 12.4.

Gujarati, D.N. and D.C. Porter *Basic Econometrics* (fifth edition): (ISBN 9780071276252), Chapters 3.4 (Properties of Least Squares Estimators: The Gauss–Markov Theorem) and 3A.3 (Variances and Standard Errors of Least Squares Estimators).

Approaching the question

Replace the values of the X_is in the equation given by $\tilde{\beta}$. The variance of $\tilde{\beta}$ has to be calculated and compared with the variance of $\widehat{\beta}$ (the variance of $\widehat{\beta}$ is given). While deriving the variance, be careful to take into account that $\mathrm{E}(u_i u_j) = 0$ if $i \neq j$. The answer is as follows.

Replace the value of the X_is:

$$\begin{aligned}
\tilde{\beta} &= \frac{1}{8}[Y_6 + Y_5 - Y_2 - Y_1] \\
&= \frac{1}{8}[(\alpha + 6\beta + u_6) + (\alpha + 5\beta + u_5) - (\alpha + 2\beta + u_2) - (\alpha + \beta + u_1)] \\
&= \frac{1}{8}[8\beta + u_6 + u_5 - u_2 - u_1] \qquad \text{(i)}.
\end{aligned}$$

From (i), it is easy to see that:

$$\begin{aligned}
\text{var}(\tilde{\beta}) &= \text{var}\left[\frac{1}{8}(u_6 + u_5 - u_2 - u_1)\right] \qquad (\text{since var}(8\beta) = 0) \\
&= \frac{1}{8^2}[\text{var}(u_6) + \text{var}(u_5) + \text{var}(u_2) + \text{var}(u_1)] \qquad (\text{since E}(u_i u_j) = 0 \text{ if } i \neq j) \\
&= \frac{4\sigma^2}{64} \\
&= \frac{\sigma^2}{16} \qquad (\text{since var}(u_i) = \sigma^2 \text{ for all } i).
\end{aligned}$$

Hence $\text{var}(\tilde{\beta}) > \text{var}(\widehat{\beta})$ and so $\widehat{\beta}$ is more efficient in comparison to $\tilde{\beta}$.

Question 4

Consider a model:

$$\boldsymbol{y_t = \alpha y_{t-1} + u_t}$$

where

$$\boldsymbol{u_t = \theta \varepsilon_{t-1} + \varepsilon_t; \qquad t = 1, 2, \ldots, T}$$

and $\mathbf{E}(\varepsilon_t) = 0$, $\mathbf{E}(\varepsilon_t^2) = \sigma^2$ and $\mathbf{E}(\varepsilon_s \varepsilon_t) = 0$ if $s \neq t$ for all $s, t = 1, 2, \ldots, T$.

Show that the OLS estimator of α is inconsistent if $\theta \neq 0$.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition): Chapter 12 (Definition and consequences of autocorrelation).

Dougherty, C. Subject guide: Chapter 12 (Properties of regression models with time series data).

Approaching the question

The model contains a lagged dependent variable as an explanatory variable and also has autocorrelation. The ordinary least squares estimator will be inconsistent. To prove inconsistency, plim should be used. The answer is as follows.

The OLS estimator of α is:

$$\widehat{\alpha} = \frac{\sum y_t y_{t-1}}{\sum y_{t-1}^2} = \alpha + \frac{\sum u_t y_{t-1}}{\sum y_{t-1}^2} = \alpha + \frac{\sum \varepsilon_t y_{t-1}}{\sum y_{t-1}^2} + \frac{\theta \sum \varepsilon_{t-1} y_{t-1}}{\sum y_{t-1}^2}$$

where y_t has been substituted as $y_t = \alpha y_{t-1} + \varepsilon_t + \theta \varepsilon_{t-1}$.

Since $y_{t-1} = \alpha y_{t-2} + \varepsilon_{t-1} + \theta \varepsilon_{t-2}$ hence:

$$\text{plim} \frac{\sum \varepsilon_t y_{t-1}}{T} = 0.$$

It is easy to check that:

$$\text{plim} \frac{\sum \varepsilon_{t-1} y_{t-1}}{T} \neq 0.$$

It is also assumed:

$$\text{plim} \frac{\sum y_{t-1}^2}{T}$$

is greater than zero and finite.

Therefore, $\text{plim}(\widehat{\alpha}) \neq \alpha$, and so $\widehat{\alpha}$ is inconsistent.

Question 5

Consider an ADL(1, 1) model:

$$\boldsymbol{Y_t = \alpha_1 + \alpha_2 Y_{t-1} + \alpha_3 X_t + \alpha_4 X_{t-1} + u_t; \qquad t = 1, 2, \ldots, T.}$$

where both Y_t and X_t are I(1). u_t is the disturbance term where $\mathrm{E}(u_t) = 0$; $\mathrm{E}(u_t^2) = \sigma^2$ and $\mathrm{E}(u_s u_t) = 0$ for all $s \neq t$. Express the ADL(1, 1) model in an error correction form and interpret the coefficients of the error correction model.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition): Chapter 13.6 (Fitting models with nonstationary time series).

Gujarati, D.N. and D.C. Porter, *Basic Econometrics* (fifth edition): (ISBN 9780071276252), Chapter 21.11 (Cointegration: Regression of a Unit Root Time Series on Another Unit Root Time Series).

Approaching the question

The ADL(1,1) model should be expressed in error correction form. It involves some simple algebraic manipulations. A detailed interpretation of the coefficients of the error correction model should be given. The answer is as follows.

Consider a simple ADL(1,1) model (this is also known as ARDL(1, 1)):

$$Y_t = \alpha_1 + \alpha_2 Y_{t-1} + \alpha_3 X_t + \alpha_4 X_{t-1} + u_t \qquad \text{(i)}$$

Rewrite (i) as:

$$\begin{aligned}
Y_t - Y_{t-1} &= \alpha_1 + \alpha_2 Y_{t-1} - Y_{t-1} + \alpha_3 X_t - \alpha_3 X_{t-1} + \alpha_3 X_{t-1} + \alpha_4 X_{t-1} + u_t \\
\Delta Y_t &= \alpha_1 - (1-\alpha_2) Y_{t-1} + \alpha_3 \Delta X_t + (\alpha_3 + \alpha_4) X_{t-1} + u_t \\
\Delta Y_t &= \alpha_3 \Delta X_t - (1-\alpha_2)\left[Y_{t-1} - \frac{\alpha_1}{1-\alpha_2} - \frac{\alpha_3+\alpha_4}{1-\alpha_2} X_{t-1}\right] + u_t \\
\Delta Y_t &= \alpha_3 \Delta X_t - (1-\alpha_2)[Y_{t-1} - \beta_1 - \beta_2 X_{t-1}] + u_t
\end{aligned}$$

or:

$$\Delta Y_t = \alpha_3 \Delta X_t - \pi[Y_{t-1} - \beta_1 - \beta_2 X_{t-1}] + u_t \qquad \text{(ii)}$$

where:

$$\pi = 1 - \alpha_2, \qquad \beta_1 = \frac{\alpha_1}{1-\alpha_2} \qquad \text{and} \qquad \frac{\alpha_3+\alpha_4}{1-\alpha_2}.$$

Equation (ii) is the ECM.

When the two variables Y and X are cointegrated, the ECM incorporates not only the short-run effects but also the long-run effects. The long-run equilibrium $Y_t - \beta_1 - \beta_2 X_{t-1}$ is included in the model together with the short-run effect captured by the differenced term.

All the terms in the ECM, given by (ii), are stationary. As Y and X are I(1), then ΔX and ΔY are I(0). As Y and X are cointegrated their linear combination:

$$u_t = Y_t - \beta_1 - \beta_2 X_{t-1}$$

is I(0).

The coefficient π provides us with the information about the speed of adjustment in cases of disequilibrium.

- If $\pi = 1$ then 100% of the adjustment takes place within the period. In other words adjustment is instantaneous and full.
- If $\pi = 0.5$ then 50% adjustment takes place each period.
- If $\pi = 0$ then there is no adjustment.

Section B

Answer three questions from this section.

Question 6

(a) **Describe an adaptive expectations model.**

(6 marks)

(b) **Koyck investigated the relationship between investment in railcars and the volume of freight carried on the U.S. railroads using annual data for the period 1884–1939. Assuming that the desired stock of railcars in year t depended on the volume of the freight in year $t-1$ and $t-2$ and a time trend, and assuming that investment in railcars was subject to a partial adjustment process, he fitted the following regression using ordinary least squares:**

$$\widehat{I}_t = 0.077F_{t-1} + 0.017F_{t-2} - 0.003t - 0.110K_{t-1}; \qquad R^2 = 0.85; \quad t = 1, 2, \ldots, T$$

where $I_t = K_t - K_{t-1}$ is investment in railcars in year t (thousands), K_t is the stock of railcars at the end of year t (thousands), and F_t is the volume of freight handled in year t (ton-miles).

i. **Explain how Koyck's model can be derived from an adjustment equation $K_t - K_{t-1} = \lambda(K_t^* - K_{t-1})$ and a behavioural equation with dependent variable K_t^*, using variables F (lagged) and t (time).**

(6 marks)

ii. **Using Koyck's estimated equation estimate the parameters of your behavioural equation. What are the implications for the behaviour in the long run?**

(8 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) Chapter 11.4 (Models with a lagged dependent variable).

Dougherty, C. Subject guide: Chapter 11 (Models using time series data).

Gujarati, D.N. and D.C. Porter *Basic Econometrics* (fifth edition): (ISBN 9780071276252), Chapter 17.5 (Rationalization of the Koyck Model: The Adaptive Expectations Model).

Approaching the question

(a) An adaptive expectations model involves a learning process in which, in each time period, the actual value of the variable is compared with the value that has been expected. If the actual value is greater, the expected value is adjusted upwards for the next period. If it is lower, the expected value is adjusted downwards. The size of the adjustment is hypothesised to be proportional to the discrepancy between the actual and expected value.

If X is the variable in question, and X_t^e is the value expected in time period t given the information available at time period $t-1$, then:

$$X_{t+1}^e - X_t^e = \lambda(X_t - X_t^e) \qquad 0 \leq \lambda \leq 1$$

or:

$$X_{t+1}^e = \lambda X_t + (1-\lambda)X_t^e.$$

The model is derived to capture the changing nature of the formation of expectations, often in variables that are also changing with time. It is an attempt at a 'simple learning' solution to model building in order to forecast macroeconomic variables. Such variables include investment, savings and the demand for assets.

The model is estimated by repeated substitution for the expected variable, by its lagged variant which has known components of the previous period and the unobserved expectation lagged, until the term on the unobserved expectation $(1-\lambda)^s$ is so small as to be ignored resulting in a model where all the variables are observed, where s is the period lagged and λ is the speed of adjustment of expected and actual and λ is between 0 and 1. Technical details should be given.

(b) i. Given the information in the question, the model may be written as:

$$K_t^* = \beta_1 F_{t-1} + \beta_2 F_{t-2} + \beta_3 t + u_t.$$

It is given that:

$$K_t - K_{t-1} = I_t = \lambda(K_{t-1}^* - K_{t-1}).$$

Hence:

$$I_t = \lambda\beta_1 F_{t-1} + \lambda\beta_2 F_{t-2} + \lambda\beta_3 t - \lambda K_{t-1} + \lambda u_t.$$

ii. Estimates of the coefficients should be obtained and interpreted. The working and interpretations are given below.

From the fitted equation:

$$\widehat{\lambda} = 0.110, \quad \widehat{\beta}_1 = \frac{0.077}{0.110} = 0.70, \quad \widehat{\beta}_2 = \frac{0.017}{0.110} = 0.15 \quad \text{and} \quad \widehat{\beta}_3 = -\frac{0.0033}{0.110} = -0.030.$$

Hence the short-run effect of an increase of 1 million ton-miles of freight is to increase investment in rail-cars by 77 one year later and 17 two years later. It does not make much sense to talk of the short-run effect of a time trend.

In the long-run equilibrium, neglecting the effects of the disturbance term, K_t and K_t^* are both equal to the equilibrium value $\bar{K}$, and F_{t-1} and F_{t-2} are both equal to their equilibrium value $\bar{F}$. Hence:

$$\bar{K} = (\beta_1 + \beta_2)\bar{F} + \beta_3 t.$$

Therefore, an increase of one million ton-miles of freight will increase the stock of rail-cars by 850 and the time trend will be responsible for a secular decline of 33 rail-cars per year (obsolescence/damage).

Question 7

The following estimates were calculated from a sample of 7,634 women respondents from the General Household Survey 1995. The dependent variable takes the value 1 if the woman was in paid employment, and 0 otherwise.

	OLS	Logit	Probit
high	**0.093**	**0.423**	**0.259**
	(0.015)	**(0.071)**	**(0.043)**
noqual	**−0.210**	**−0.898**	**−0.554**
	(0.013)	**(0.056)**	**(0.035)**
age	**0.038**	**0.173**	**0.107**
	(0.003)	**(0.124)**	**(0.008)**
age2	**−0.051**	**−0.230**	**−0.142**
	(0.003)	**(0.069)**	**(0.009)**
mar	**0.024**	**0.103**	**0.063**
	(0.009)	**(0.057)**	**(0.035)**
Constant	**−0.068**	**−2.587**	**−1.593**
	(0.049)	**(0.225)**	**(0.137)**

Where high is one if the respondent has a higher educational qualification, zero otherwise; noqual is one if the respondent has no qualifications, zero otherwise; age is age in years; age2 is (age × age) 100; mar is one if married, zero otherwise. Conventionally calculated standard errors are given in brackets for the ordinary least squares (OLS) results and asymptotic standard errors in brackets elsewhere.

(a) Explain briefly how Probit estimates are calculated when the model has no intercept and only one explanatory variable.

(6 marks)

(b) For all three sets of estimates, test the null hypothesis that the coefficient of mar is zero. Which test statistics would you consider more reliable? Explain.

(6 marks)

(c) Using the OLS and Probit estimates, calculate the estimated probabilities of being in employment for a married woman aged 40 with a higher educational qualification. Comment on your results.

(5 marks)

(d) Test the null hypothesis that all the slope coefficients of the probit model are jointly equal to zero, given that:

$$\ln L_R = -416.01$$

$$\ln L_U = -321.25$$

where $\ln L_R$ and $\ln L_U$ are the logs of the likelihood from the restricted and the unrestricted probit models, respectively.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) Chapters 10.1 (The linear probability model), 10.3 (Probit analysis) and 10.6 (An introduction to maximum likelihood estimation).

Dougherty, C. Subject guide (2011) Chapter 10 (Binary choice and limited dependent variable models, and maximum likelihood estimation).

Gujarati, D.N. and D.C. Porter *Basic Econometrics* (fifth edition): (ISBN 9780071276252), Chapters 15.2 (The Linear Probability Model (LPM)), 15.5 (The Logit Model) and 15.9 (The Probit Model).

Approaching the question

(a) The probit model uses the cumulative standardised normal distribution. The maximum likelihood technique is used to obtain the estimates of the parameters. Estimates have the standard maximum likelihood properties, i.e. the estimators are consistent, asymptotically efficient and asymptotically normally distributed.

[For technical details see Dougherty (fourth edition) Section 10.3].

(b) To test the null hypothesis the t test should be used. t tables are attached with the examination paper. Candidates should know how to look at the critical values. The null and alternative hypotheses should be clearly stated:

$$\mathrm{H}_0 : \text{coefficient of mar} = 0$$

$$\mathrm{H}_1 : \text{coefficient of mar} \neq 0.$$

t test statistics are 2.67 (OLS), 1.81 (logit), and 1.8 (probit).

The test based on the OLS estimates gives outright rejection. However, the standard errors for this test are wrongly calculated because of the heteroskedasticity of the error term. Therefore, we prefer probit or logit estimates.

(c) A large number of candidates wrongly thought that to obtain the probability in the case of probit a sophisticated calculator is needed. The probability in the case of probit can be read directly from the supplied statistical tables. We have:

$$0.093 + 0.038 \times 40 - 0.051 \times 16 + 0.024 \times 0.068 = 0.753 \qquad \text{(OLS)}$$

$$0.259 + 0.107 \times 40 - 0.142 \times 16 + 0.063 - 1.593 = 0.737 \qquad \text{(probit)}.$$

The statistical tables give a probability of 0.77, approximately.

They are all fairly close. The OLS estimates do not fall outside the probability bounds.

(d) To test the null hypothesis, the large sample likelihood-ratio test should be used. In large samples $-2(\ln L_R - \ln L_U)$ is distributed as a chi-squared distribution with degrees of freedom equal to the number of restrictions imposed by the null hypothesis. The null and alternative hypotheses should be clearly stated as:

$$\mathrm{H}_0 : \text{all the slope coefficients are} = 0$$

$$\mathrm{H}_1 : \text{at least one slope coefficient is} \neq 0.$$

We have:

$$-2(\ln L_R - \ln L_U) \sim \chi^2_5$$

and:

$$-2(-416.01 - (-321.25)) = 189.92.$$

The critical value of χ^2_5 at the 5% significance level is 11.07, hence we reject H_0.

Question 8

Economists have tried to examine the catch-up hypothesis where it is predicted that poorer countries will grow faster than richer countries to 'catch-up' to the richer countries. To assess this hypothesis an economist regresses the growth rate of GDP (gross domestic product), gr_i, on log of GDP per capita, $\ln(gdp_i)$, as well as a regression including other variables which might influence growth rates. The following regression results come from regressions using data from 96 countries:

$$\begin{aligned} gr_i &= \underset{(0.023)}{-0.058} - \underset{(0.002)}{0.009} \ln(gdp_i) + e_{1i} \qquad \text{(A)} \end{aligned}$$

$$n = 96, R^2 = 0.122 \text{ and } s = 0.025$$

where estimated standard errors are given in brackets, gr_i is the growth rate of the ith country over the period 1985–1990, $\ln(gdp_i)$ is the logarithm of GDP for country i, e_{1i} is the residual and s is the estimated standard error of the residuals.

A second regression yields

$$\begin{aligned} gr_i &= -0.019 + 0.001\ln(gdp_i) + 0.12inv_i + e_{2i} \qquad \text{(B)} \\ &\qquad\quad (0.023) \qquad\qquad (0.002) \end{aligned}$$

$$n = 96, R^2 = 0.196 \text{ and } s = 0.024$$

where e_{2i} is the estimated residual and inv_i is the ratio of investment to GDP in 1985.

(a) Test, for each regression, the hypothesis that the coefficient on $\ln(gdp_i)$ is zero. What does the result of your tests say about the 'catch-up' hypothesis?

(7 marks)

(b) The R^2 measure is higher for equation (B) than for (A). Is this what you would expect? Explain.

(5 marks)

(c) The regression results may be affected by heteroskedasticity. Explain what you understand about heteroskedasticity, why the estimation results might be affected by heteroskedasticity and what the effects are if heteroskedasticity is present.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) Chapters 3.3 (Properties of multiple regression coefficients), 3.5 (Goodness of fit: R^2) and 7.1 (Heteroscedasticity and its implications).

Dougherty, C. Subject guide: Chapters 3 (Multiple regression analysis) and 7 (heteroscedasticity).

Gujarati, D.N. and D.C. Porter *Basic Econometrics* (fifth edition): (ISBN 9780071276252), Chapters 3.5 (The Coefficient of Determination R^2: A Measure of 'Goodness of Fit'), 11.1 (The Nature of Heteroscedasticity), 11.2 (OLS Estimation in the Presence of Heteroscedasticity) and 13.9 (Model Selection Criteria).

Approaching the question

(a) It is required to conduct tests of significance and relate the results of the tests to the 'catch-up' hypothesis.

The t-values are (A) 4.5 and (B) 0.043. The degrees of freedom are (A) 94 and (B) 93 giving 95% two-tailed critical values of ± 1.99. We reject the null hypothesis for (A), but do not reject for (B). A better answer would be to use a one-tailed test since the alternative hypothesis is that as GDP increases, the growth rate will fall, hence the coefficient will be negative, but the answer will be the same although the critical value is now -1.66. The results show that (A) supports the 'catch-up' hypothesis, but (B) does not. If (B) is the true model then (A) will suffer from omitted variable bias so that the 'catch-up' hypothesis is in doubt.

(b) It should be explained that R^2 is a non-decreasing function of the number of explanatory variables in the model. Candidates should show that as the number of explanatory variables is increased, R^2 will never decrease.

R^2 is defined as:

$$R^2 = \frac{\text{explained sum of squares}}{\text{total sum of squares}}.$$

As extra explanatory variables are added to an equation the explained sum of squares increases or does not change. Total sum of squares does not change, hence R^2 will never decrease. The result is as expected.

(c) The definition of heteroskedasticity is required. Candidates should explain the consequences of heteroskedasticity on the properties of ordinary least squares estimators.

The data are cross-sectional and deal with countries which are probably at very different levels of GDP, hence heteroskedasticity is to be expected. Heteroskedasticity is the condition where the error term variance is not constant and ordinary least squares produces unbiased and consistent, but inefficient, parameter estimates. Also standard errors are not correct hence making t and F tests invalid.

Question 9

(a) Explain the concept of the likelihood function and the maximum likelihood estimator.

(6 marks)

(b) Suppose that an event occurs with probability p. In a simple random sample of size n the event occurs m times.

i. Show that the maximum likelihood estimator (MLE) of p is $\frac{m}{n}$. Verify that the MLE maximises the likelihood function.

(8 marks)

ii. Derive the likelihood ratio statistic for the null hypothesis $p = p_0$. If $m = 40$ and $n = 100$, test the null hypothesis $p = 0.5$.

(6 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) Chapter 10.6 (An introduction to maximum likelihood estimation).

Dougherty, C. Subject guide: Chapter 10 (Binary choice and limited dependent variable models, and maximum likelihood estimation).

Gujarati, D.N. and D.C. Porter *Basic Econometrics* (fifth edition): (ISBN 9780071276252), Chapters 4.4 (The Method of Maximum Likelihood), 4A.1 (Maximum Likelihood Estimation of Two-Variable Regression Model) and 8A2 (Likelihood Ratio (LR) Test).

Approaching the question

(a) It is required to explain the concept of the likelihood function and the maximum likelihood estimator (MLE).

Definition: Let the probability density function (pdf) of X be $f(x;\theta)$ where θ is a parameter, then the likelihood function of a simple random sample $\{X_1, X_2, \ldots, X_n\}$ from X is the product of the individual densities of the X_is taken as a function of θ.

The joint pdf of $\{X_1, X_2, \ldots, X_n\}$ is:

$$f(x_1, x_2, \ldots, x_n; \theta) = f(x_1;\theta)\, f(x_2, \theta) \cdots f(x_n;\theta).$$

The likelihood function $L(\theta; x_1, x_2, \ldots, x_n)$ has the same formulation as the joint pdf, but now it is a function of θ.

Maximum Likelihood Estimator (MLE)

Let $X \sim f(x;\theta)$. The MLE $\widehat{\theta}$ of the parameter θ is an estimator that maximises the likelihood function.

To obtain the MLE set:

$$\frac{\partial L}{\partial \theta} = 0$$

and solve for $\widehat{\theta}$.

It is easier to take logs and then maximise, as in both the situations we will arrive at the same maximum, because:

$$\frac{\partial \ln L}{\partial \theta} = \frac{1}{L}\frac{\partial L}{\partial \theta} = 0 \qquad \Rightarrow \qquad \frac{\partial L}{\partial \theta} = 0.$$

To verify the maximum, the second derivative evaluated at the value of the MLE should be checked. It should be negative.

(b) To test the null hypothesis a large sample likelihood ratio test should be used.

i. p is the probability of the event occurring, hence $(1-p)$ is the probability that the event does not occur. The event occurs m times and does not occur $(n-m)$ times (n is the total number of observations).
The joint probability of the event occurring and not occurring is $p^m\,(1-p)^{n-m}$. The log-likelihood function is:

$$\log L(p) = m\log(p) + (n-m)\log(1-p).$$

Differentiating with respect to p we get:

$$\frac{\mathrm{d}\log p}{\mathrm{d}p} = \frac{m}{p} - \frac{n-m}{1-p}.$$

Equating it to zero we obtain the maximum likelihood estimator $\widehat{p}$ of p as:

$$\widehat{p} = \frac{m}{n}.$$

We should check that the second differential is negative to verify the maximum:

$$\frac{\mathrm{d}^2\log p}{\mathrm{d}p^2} = -\frac{m}{p^2} - \frac{n-m}{(1-p)^2}.$$

Evaluating it at $p = m/n$ we obtain:

$$\frac{\mathrm{d}^2\log p}{\mathrm{d}p^2} = -\frac{n^2}{m} - \frac{n-m}{(1-m/n)^2} < 0.$$

So we have chosen a value of p which maximises the likelihood function.

ii. The log-likelihood function is:

$$\log L(p) = m\log(p) + (n-m)\log(1-p).$$

The likelihood-ratio statistic is:

$$\begin{aligned} LR &= 2\left[\left(m\log\frac{m}{n} + (n-m)\log\left(1-\frac{m}{n}\right)\right) - (m\log p_0 + (n-m)\log(1-p_0))\right] \\ &= 2\left[m\log\left(\frac{m/n}{p_0}\right) + (n-m)\log\left(\frac{1-m/n}{1-p_0}\right)\right]. \end{aligned}$$

If $m = 40$ and $n = 100$ the LR statistic for $\mathrm{H}_0 : p = 0.5$ is:

$$LR = 2\left[40\log\left(\frac{0.4}{0.5}\right) + 60\log\left(\frac{1-0.4}{1-0.5}\right)\right] = 4.03.$$

The critical value of χ^2_1 at the 5% significance level is 3.84, and at the 1% significance level is 6.64. We would reject H_0 at the 5% but not at the 1% significance level.

Question 10

(a) **There are six missing values (denoted by x's) in the given Stata output**

i. **Model SS**

ii. **Model df**

iii. ***R*-squared**

iv. ***avetemp t* value**

v. ***harvrain* 95% Conf. Interval.**

Give the formulae for them and obtain their values based on the given output.

```
      Source |       SS       df       MS              Number of obs =      27
-------------+------------------------------           F(  3,    23) =   21.90
       Model |  xxxxxxxxxx    xx  2.58519261           Prob > F      =  0.0000
    Residual |  2.71468685    23  .118029863           R-squared     =  xxxxxx
-------------+------------------------------           Adj R-squared =  0.7069
       Total |  10.4702647    26  .402702488           Root MSE      =  .34355

------------------------------------------------------------------------------
     lnprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    wintrain |    .001282   .0005765     2.22   0.036     .0000894    .0024747
     avetemp |   .7123178   .1087674     xxxx   0.000     .4873154    .9373202
    harvrain |  -.0036242   .0009646    -3.76   0.001    xxxxxxxxx   xxxxxxxxx
       _cons |  -13.44433   1.969396    -6.83   0.000    -17.51834   -9.370326
------------------------------------------------------------------------------
```

where the explanatory variables are

***wintrain*: the level of winter rain (October – March) in millimetres**

***avetemp*: average temperature in the growing season (April – September)**

***harvrain*: level of harvest rain (August – September) in millimetres**

and the dependent variable (*lnprice*) is the price of mature red wine from the Bordeaux region of France at harvest time. The data relate to wines of different vintages (ages) in 1987.

(10 marks)

(b) **Interpret the regression results in (a).**

(3 marks)

(c) **If the age of the wine is included in the equation the estimates become:**

```
      Source |       SS       df       MS              Number of obs =      27
-------------+------------------------------           F(  4,    22) =   26.39
       Model |  8.66443586     4  2.16610897           Prob > F      =  0.0000
    Residual |  1.80582883    22  .082083129           R-squared     =  0.8275
-------------+------------------------------           Adj R-squared =  0.7962
       Total |  10.4702647    26  .402702488           Root MSE      =   .2865

------------------------------------------------------------------------------
     lnprice |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    wintrain |   .0011668    .000482     2.42   0.024     .0001671    .0021665
     avetemp |   .6163926   .0951755     6.48   0.000     .4190107    .8137745
    harvrain |  -.0038606   .0008075    -4.78   0.000    -.0055353   -.0021858
         age |   .0238474   .0071667     3.33   0.003    0.0089846    .0387103
       _cons |  -12.31227   1.677212    -7.34   0.000    -15.79059   -8.833945
------------------------------------------------------------------------------
```

i. **What is the interpretation of the coefficient of the variable age?**

(3 marks)

ii. If the true model is given by the specification in (c) do you think that the model in (a) should show evidence of specification error? Does it? Carefully explain your observations.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to Econometrics* (fourth edition) Chapters 4.2 (Logarithmic Transformations) and 6.2 (The effect of omitting a variable that ought to be included).

Dougherty, C. Subject guide: Chapters 4 (Transformation of variables) and 6 (Specification of regression variables).

Gujarati D.N. and D.C. Porter *Basic Econometrics* (fifth edition): (ISBN 9780071276252), Chapters 6.6 (Semilog Models: Log-Lin and Lin-Log Models) and 13.3 (Consequences of Model Specification Errors).

Approaching the question

(a) It is required to calculate the missing values. Formulae for each must be given.

i. Model SS = 7.7556.

ii. R-squared = 0.741.

iii. Model df = 3.

iv. *avetemp* t value = 6.55.

v. *harvrain* 95% CI = $[-0.0056, -0.0016]$.

(b) The dependent variable is in log form. The answer must explain the implication of this.

The critical t value is $t_{27-4} = 2.069$ hence all coefficients are significantly different from 0. *wintrain* increases the log price by 0.001 per millimetre, *avetemp* increases log price by 0.712 per degree and *harvrain* reduces log price by 0.0036 per millimetre.

(c) In answering part (ii), it should be mentioned that *age* appears to be part of the true model. Hence we have an omitted variable. The sign and size of omitted variables bias should be discussed.

The model in (a) has a specification error in that *age* appears to be part of the true model. The omitted variable bias will be related to the true coefficient and the correlation between the omitted variable and the remaining variables – the true coefficient is presumably non-zero and positive, but *age* is unlikely to be correlated to any of the included variables, hence one would not expect any measurable bias – and this is what you observe. The difference between the two models is very small.

This paper is not to be removed from the Examination Halls

UNIVERSITY OF LONDON **EC2020 ZA**

BSc degrees and Diplomas for Graduates in Economics, Management, Finance and the Social Sciences, the Diplomas in Economics and Social Sciences

Elements of Econometrics

Thursday, 18 May 2017 : 14:30 to 17:30

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Extracts from statistical tables are given after the final question on this paper.

Graph paper is provided at the end of this question paper. If used, it must be detached and fastened securely inside the answer book.

A calculator may be used when answering questions on this paper and it must comply in all respects with the specification given with your Admission Notice. The make and type of machine must be clearly stated on the front cover of the answer book.

If more questions are answered than requested, only the first answers attempted will be counted.

PLEASE TURN OVER

© University of London 2017

SECTION A

Answer all questions from this section

1. Discuss the drawbacks and advantages of using the Linear Probability Model when trying to explain a binary decision. In your answer clearly indicate what the Linear Probability Model is. **(8 marks)**

2. Discuss the consequences of measurement error.

   In your answer consider the following empirical study attempting to estimate the relationship between advertising and magazine circulation rates. A simple linear relationship is postulated between A_t, the advertising rate in magazine t, and C_t, the circulation figure for the magazine in question:

   $$A_t = \alpha + \beta C_t + \varepsilon_t, \quad t = 1, ..., T,$$

   The relation is estimated by least squares on data for $T = 75$ magazines. Unfortunately, there are considerable errors in measurement in the reported circulation figures. Critically discuss the following statement: 'The estimated coefficient $\widehat{\beta}$ will be too small.' Rigour of your answer will be rewarded. **(8 marks)**

3. Consider the following regression model:

   $$Y_i = \beta_0 + \beta_1 \frac{1}{X_i} + u_i, \qquad i = 1, ..., n.$$

   We assume that the errors $\{u_i\}_{i=1}^n$ are independent normal random variables with zero mean and variance σ^2/X_i^2. The regressor, $1/X_i$, is nonstochastic with positive sample variability and $X_i \neq 0$ for all i. You are interested in testing the hypothesis $H_0 : \beta_0 = 0$ against $H_1 : \beta_0 \neq 0$. You are advised to use the BLUE estimator of β_0 for this purpose.

   Discuss how you would obtain the BLUE estimator of β_0 (note, you are not asked to derive this estimator).

   Give two reasons why you would prefer using the BLUE estimator for β_0 instead of the OLS estimator $\hat{\beta}_{0,OLS} = \bar{Y} - \hat{\beta}_{1,OLS}\overline{(1/X)}$ when testing this hypothesis, where $\bar{Y} = \frac{1}{n}\sum Y_i$ and $\overline{(1/X)} = \frac{1}{n}\sum \frac{1}{X_i}$. **(8 marks)**

4. Consider the following non-stationary process:

$$y_t = \gamma_0 + \gamma_1 t + u_t, \text{ with } u_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

and ε_t i.i.d. $\left(0, \sigma^2\right)$. Indicate (with explanation) the source(s) of non-stationarity of y_t. Discuss how you would test whether y_t indeed is non-stationary? Clearly indicate the null and the alternative hypothesis, the test statistic and the rejection rule. What name do we give such a non-stationary process? **(8 marks)**

5. Suppose you are given a random sample $X_1, ..., X_n$ from the exponential distribution:

$$f(x) = \lambda \exp(-\lambda x), \quad x > 0, \ \lambda > 0.$$

According to this distribution $E(X_i) = 1/\lambda$ and $Var\left(X_i\right) = 1/\lambda^2$ for $i = 1, ..., n$.

Show that the maximum likelihood estimator for λ is $1/\bar{X}$ where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Is the estimator unbiased and/or consistent? Prove your claims. **(8 marks)**

SECTION B

Answer three questions from this section.

6. The following equation was estimated by Ordinary Least Squares based on second semester students in the fall term

$$\begin{aligned}
\widehat{\textit{trmgpa}} &= \underset{\substack{(.55)\\[.55]}}{-2.12} + \underset{\substack{(.175)\\[.166]}}{.900}\textit{crsgpa} + \underset{\substack{(.064)\\[.074]}}{.193}\textit{cumgpa} + \underset{\substack{(.0012)\\[.0012]}}{.0014}\textit{tothrs} \\
&\quad + \underset{\substack{(.0002)\\[.0002]}}{.0018}\textit{sat} - \underset{\substack{(.0018)\\[.0019]}}{.0039}\textit{hsperc} + \underset{\substack{(.085)\\[.079]}}{.351}\textit{female} - \underset{\substack{(.098)\\[.080]}}{.157}\textit{season} \\
n &= 269;\ R^2 = .465, \qquad (6.1)
\end{aligned}$$

where *trmgpa*$_i$ is the term GPA (grade point average) of individual i. *crsgpa*$_i$ is a measure of difficulty of courses taken by individual i (weighted overall average of GPA in selected courses), *cumgpa*$_i$ is the GPA of individual i prior to the current semester, *tothrs*$_i$ is the total credit hours of individual i prior to the semester, *sat*$_i$ is his/her SAT score (test taken for college admission in the USA), *hsperc*$_i$ is the graduating percentile of individual i in high school class, *female*$_i$ is a gender dummy, and *season*$_i$ is a dummy variable equal to unity if the student's sport is in season during the fall term. The usual standard errors are in parentheses and the White's heteroskedasticity-robust standard errors are in squared brackets.

(a) Explain the concept of heteroskedasticity and discuss the properties of the OLS estimator in the presence of heteroskedasticity. **(5 marks)**

(b) Do the variables *crsgpa, cumgpa* and *tothrs* have the expected estimated effects? Which of these variables are statistically significant at the 5% level? Does it matter which standard errors are used? **(5 marks)**

(c) In view of your concern about the presence of heteroskedasticity, provide the 95% confidence interval for β_{crsgpa} and use it to test the hypothesis $H_0 : \beta_{crsgpa} = 1$ against $H_A : \beta_{crsgpa} \neq 1$. Describe your conclusions. In view of your answer, indicate whether the p-value of the test is bigger or smaller than 5%. **(5 marks)**

(d) Discuss how you would conduct a test for heteroskedasticity in this setting. Clearly indicate the assumptions that underlie the test you suggest. Detail of your answer will be rewarded. **(5 marks)**

7. Consider the model:

$$y_t = \alpha_1 y_{t-1} + u_t,\ t = 1, ..., T$$

where $y_0 = 0$. $E(u_t) = 0$, $E(u_t^2) = \sigma^2$ and $E(u_t u_s) = 0$ when $s \neq t$, for all $s, t = 1, ..., T$

(a) Discuss what we mean by the concept of stationarity (more precisely "covariance stationarity") and indicate under what condition $\{y_t\}_{t=1}^T$ will be stationary. **(3 marks)**

(b) Discuss the Dickey-Fuller procedure used to test for the presence of a unit root in the above model. Clearly indicate the null and alternative hypothesis, test statistic and rejection rule. **(5 marks)**

(c) Consider a slight variation of the above model:

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + v_t,\ t = 1, ..., T,$$

where $y_0 = 0$. $E(v_t) = 0$, $E(v_t^2) = \sigma^2$ and $E(v_t v_s) = 0$ when $s \neq t$, for all $s, t = 1, ..., T$ What do we call such a process? Discuss the problem you will have when conducting your test as in (b). **(3 marks)**

(d) Instead of conducting the Dickey-Fuller procedure, you are told to apply the Augmented DF test. Indicate how you would conduct the test for the presence of a unit root here. Derivation of the test equation will be required for full marks. **(5 marks)**

(e) What are the potential problems associated with performing a regression with $I(1)$ variables? In your answer explain what it means for a variable to be $I(1)$. **(4 marks)**

8. For the US economy, let $gprice$ denote the monthly growth in the overall price level and let $gwage$ be the monthly growth in hourly wages. We have estimated the following distributed lag model (ADL(0,12)):

$$\begin{aligned}
\widehat{gprice}_t &= \underset{(.00057)}{-.00093} + \underset{(.052)}{.119}\, gwage_t + \underset{(.039)}{.097}\, gwage_{t-1} + \underset{(.039)}{.040}\, gwage_{t-2} \\
&\quad + \underset{(.039)}{.038}\, gwage_{t-3} + \underset{(.039)}{.081}\, gwage_{t-4} + \underset{(.039)}{.107}\, gwage_{t-5} + \underset{(.039)}{.095}\, gwage_{t-6} \\
&\quad + \underset{(.039)}{.104}\, gwage_{t-7} + \underset{(.039)}{.103}\, gwage_{t-8} + \underset{(.039)}{.159}\, gwage_{t-9} + \underset{(.039)}{.110}\, gwage_{t-10} \\
&\quad + \underset{(.039)}{.103}\, gwage_{t-11} + \underset{(.052)}{.016}\, gwage_{t-12}, \\
n &= 273,\ R^2 = .317,\ \bar{R}^2 = .283, DW = .99.
\end{aligned}$$

The usual standard errors are in parentheses.

(a) What is the estimated long-run effect (LRP)? Is it very different from one? Explain what the LRP tells us in this example. How does this differ from the short-run effect? **(4 marks)**

(b) We want to test whether the LRP is significantly smaller than one. Clearly indicating the null and the alternative hypothesis, give the test statistic and the rejection rule. What regression would you run to obtain the standard error of the LRP directly? **(6 marks)**

(c) Your result in (b) may be affected by the presence of autocorrelation. Briefly explain this and discuss how you could use the Durbin-Watson test to detect whether this is indeed the case. Indicate clearly the assumptions underlying the Durbin Watson test, the test statistic and the rejection rule. **(5 marks)**

Hint: For the critical values of the DW test, use the largest sample size and number of regressors available in the table.

(d) Assuming the Durbin-Watson test finds evidence of autocorrelation, discuss how you could resolve the problem by means of the iterated Cochrane-Orcutt procedure. For notational purposes, you may simplify your model to an ADL(0,1). **(5 marks)**

9. Let us consider the demand for fish. Using 97 daily price ($avgprc$) and quantity ($totqty$) observations on fish prices at the Fulton Fish Market in Manhattan, the following results were obtained by OLS:

$$\begin{aligned}\widehat{\log(totqty_t)} &= \underset{(.163)}{8.244} - \underset{(.176)}{.425}\log(avgprc_t) - \underset{(.226)}{.311}mon_t - \underset{(.223)}{.683}tues_t \\ &\quad - \underset{(.220)}{.533}wed_t + \underset{(.220)}{.067}thurs_t. \end{aligned} \tag{9.1}$$

The equation allows demand to differ across the days of the week, and Friday is the excluded dummy variable. The standard errors are in parentheses.

(a) Interpret the coefficient of $\log($*avgprc*$)$ and discuss whether it is significant. **(3 marks)**

(b) It is commonly thought that prices are jointly determined with quantity in equilibrium where demand equals supply. What are the consequences of this simultaneity for the properties of the OLS estimator? **(3 marks)**

(c) The variables $wave2_t$ and $wave3_t$ are measures of ocean wave heights over the past several days. In view of your answer in (b), what two assumptions do we need to make in order to use $wave2_t$ and $wave3_t$ as instruments for $\log(avgprc_t)$ in estimating the demand equation? Discuss whether these assumptions are reasonable. **(4 marks)**

(d) Below we report two sets of regression results, where the dependent variable is $\log(avgprc_t)$. Are $wave2_t$ and $wave3_t$ jointly significant? State the test statistic and rejection rule. How is your finding related to your answer in (c)? **(4 marks)**

	Regressors						
Dependent Variable $\log(avgprc_t)$	$constant$	$wave2$	$wave3$	*day-of-the-week dummies*	R^2	RSS	n
Regression (9.2)	$\underset{(.144)}{-1.022}$	$\underset{(.021)}{.094}$	$\underset{(.020)}{.053}$	yes	.3041	10.934	97
Regression (9.3)	$\underset{(.092)}{-.276}$	–	–	yes	.0088	15.576	97

(e) The following IV results were obtained in Stata:

$$\begin{aligned}\widehat{\log(totqty_t)} &= \underset{(.182)}{8.164} - \underset{(.327)}{.815}\log(avgprc_t) - \underset{(.229)}{.307}mon_t - \underset{(.226)}{.685}tues_t \\ &\quad - \underset{(.223)}{.521}wed_t + \underset{(.225)}{.095}thurs_t. \end{aligned} \tag{9.4}$$

Discuss how these results can be obtained using Two Stage Least Squares (2SLS) and briefly discuss how you would test whether the results in (9.1) and (9.4) are significantly different from each other. **(6 marks)**

10. Let us consider an analysis on recidivism (probability of re-arrest) among a group of young men in California who have at least one arrest prior to 1986. The dependent variable, *arr86,* is equal to unity if the man was arrested at least once during 1986, and zero otherwise.

	OLS A	OLS B	Logit A	Logit B	Logit B Marginal Effect
pcnv	$-.152$ (.021)	$-.162$ (.021)	$-.880$ (.122)	$-.901$ (.120)	$-.176$ (.023)
avgsen	.005 (.006)	.006 (.006)	.027 (.035)	.031 (.034)	.006 (.007)
tottime	$-.003$ (.005)	$-.002$ (.005)	$-.014$ (.028)	$-.010$ (.027)	$-.002$ (.005)
ptime86	$-.023$ (.005)	$-.022$ (.005)	$-.140$ (.031)	$-.127$ (.031)	$-.025$ (.006)
qemp86	$-.038$ (.005)	$-.043$ (.005)	$-.199$ (.028)	$-.216$ (.028)	$-.042$ (.005)
black	.170 (.024)	–	.823 (.117)	–	–
hispan	.096 (.021)	–	.522 (.109)	–	–
constant	.380 (.019)	.380 (.019)	$-.464$ (.095)	$-.169$ (.084)	
R^2	.068	.047			
$\log L$			-1512.35	-1541.24	

pcnv is the proportion of prior arrests that led to a conviction, *avgsen* is the average sentence served from prior convictions, *tottime* is the months spent in prison since age 18 prior to 1986, *ptime86* is months spent in prison in 1986, *qemp86* is the number of quarters the man was legally employed in 1986, while *black* and *hispan* are two race dummies (*white* the excluded dummy). The standard errors are reported in parentheses.

(a) When estimating the parameters by OLS, we are using the Linear Probability Model. Why might you then report heteroskedasticity-robust standard errors? **(2 marks)**

(b) Using the OLS B results, what is the estimated effect on the probability of arrest if *pcnv* goes from 0.25 to 0.75 holding everything else constant? **(4 marks)**

(question continues on next page)

(c) It is argued that the linear probability model is not appropriate for explaining the binary variable $arr86$ and a logit regression model has been estimated. Explain how the Logit estimates are obtained. **(5 marks)**

Hint: You may recall, that for the Logit model A, we will specify

$$\Pr(arr86_i = 1) = \Lambda\left(\beta_0 + \beta_1 \textit{pcnv} + \beta_2 \textit{avgsen} + ... + \beta_6 \textit{black} + \beta_7 \textit{hispan}\right),$$

where $\Lambda(z) = \frac{1}{1+\exp(-z)}$.

(d) Using the Logit model results, discuss whether *black* and *hispan* are jointly significant. Clearly indicate the null and alternative hypothesis, the test statistic and the rejection rule. **(3 marks)**

(e) An important distinction between the two approaches is that the marginal effect of *pcnv* on the probability of re-arrest is constant for the LPM unlike the marginal effect using the logit analysis. What this means for instance is that the estimated effect on the probability of arrest if *pcnv* goes from 0.25 to 0.75 will depend on the other characteristics.

i. Explain how the marginal effects evaluated at the mean values of the explanatory variables (reported in the last column) were obtained. Give a brief comment as to how they compare to the marginal effect of the associated LPM. **(3 marks)**

ii. Using the Logit B results, how would you obtain the estimated effect on the probability of arrest if *pcnv* goes from 0.25 to 0.75 for a *white* man, with characteristics *avgsen=1, tottime=1, ptime86=0 and qemp86=2.* A clear explanation of what calculations are required is sufficient. **(3 marks)**

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	**0·40**	0·6554	**0·80**	0·7881	**1·20**	0·8849	**1·60**	0·9452	**2·00**	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	**0·45**	0·6736	**0·85**	0·8023	**1·25**	0·8944	**1·65**	0·9505	**2·05**	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	**0·50**	0·6915	**0·90**	0·8159	**1·30**	0·9032	**1·70**	0·9554	**2·10**	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	**0·55**	0·7088	**0·95**	0·8289	**1·35**	0·9115	**1·75**	0·9599	**2·15**	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	**0·60**	0·7257	**1·00**	0·8413	**1·40**	0·9192	**1·80**	0·9641	**2·20**	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	**0·65**	0·7422	**1·05**	0·8531	**1·45**	0·9265	**1·85**	0·9678	**2·25**	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	**0·70**	0·7580	**1·10**	0·8643	**1·50**	0·9332	**1·90**	0·9713	**2·30**	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	**0·75**	0·7734	**1·15**	0·8749	**1·55**	0·9394	**1·95**	0·9744	**2·35**	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	**0·80**	0·7881	**1·20**	0·8849	**1·60**	0·9452	**2·00**	0·9772	**2·40**	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

When $x > 3·3$ the formula $1-\Phi(x) \doteqdot \frac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\frac{1}{x^2}+\frac{3}{x^4}-\frac{15}{x^6}+\frac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	5·0	1·6449	3·0	1·8808	2·0	2·0537	1·0	2·3263	0·10	3·0902
45	0·1257	4·8	1·6646	2·9	1·8957	1·9	2·0749	0·9	2·3656	0·09	3·1214
40	0·2533	4·6	1·6849	2·8	1·9110	1·8	2·0969	0·8	2·4089	0·08	3·1559
35	0·3853	4·4	1·7060	2·7	1·9268	1·7	2·1201	0·7	2·4573	0·07	3·1947
30	0·5244	4·2	1·7279	2·6	1·9431	1·6	2·1444	0·6	2·5121	0·06	3·2389
25	0·6745	4·0	1·7507	2·5	1·9600	1·5	2·1701	0·5	2·5758	0·05	3·2905
20	0·8416	3·8	1·7744	2·4	1·9774	1·4	2·1973	0·4	2·6521	0·01	3·7190
15	1·0364	3·6	1·7991	2·3	1·9954	1·3	2·2262	0·3	2·7478	0·005	3·8906
10	1·2816	3·4	1·8250	2·2	2·0141	1·2	2·2571	0·2	2·8782	0·001	4·2649
5	1·6449	3·2	1·8522	2·1	2·0335	1·1	2·2904	0·1	3·0902	0·0005	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

This paper is not to be removed from the Examination Halls

UNIVERSITY OF LONDON **EC2020 ZB**

BSc degrees and Diplomas for Graduates in Economics, Management, Finance and the Social Sciences, the Diplomas in Economics and Social Sciences

Elements of Econometrics

Thursday, 18 May 2017 : 14:30 to 17:30

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Extracts from statistical tables are given after the final question on this paper.

Graph paper is provided at the end of this question paper. If used, it must be detached and fastened securely inside the answer book.

A calculator may be used when answering questions on this paper and it must comply in all respects with the specification given with your Admission Notice. The make and type of machine must be clearly stated on the front cover of the answer book.

If more questions are answered than requested, only the first answers attempted will be counted.

PLEASE TURN OVER

© University of London 2017

SECTION A

Answer all questions from this section

1. Discuss the consequences of omitting relevant variables.

   In your answer consider the following empirical study attempting to estimate the social benefits of an increase in public spending on education. Using data from a cross-section survey of employees, the following regression is estimated:

   $$\log W_i = \alpha_1 + \alpha_2 S_i + \varepsilon_i,$$

   where W_i is the hourly wage rate and S_i is the number of years of schooling completed of employee i. The coefficient of S_i is found to be positive and strongly significant. What can be concluded from this? In your answer discuss the interpretation of α_2, and discuss whether we are likely to obtain an unbiased (consistent) estimator of it when using OLS. Rigour of your answer will be rewarded. **(8 marks)**

2. Suppose you are given a random sample $X_1, ..., X_n$ from the exponential distribution:

   $$f(x) = \lambda \exp(-\lambda x), \quad x > 0,\ \lambda > 0.$$

   According to this distribution $E(X_i) = 1/\lambda$ and $Var\,(X_i) = 1/\lambda^2$ for $i = 1, ..., n$.

   Show that the maximum likelihood estimator for λ is $1/\bar{X}$ where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Is the estimator unbiased and/or consistent? Prove your claims. **(8 marks)**

3. Consider the following non-stationary process:

   $$y_t = \gamma_0 + \gamma_1 t + u_t, \text{ with } u_t = \rho u_{t-1} + \varepsilon_t$$

   and ε_t i.i.d. $(0, \sigma^2)$. Indicate (with explanation) the source(s) of non-stationarity of y_t. Show that you can rewrite the model as:

   $$\Delta y_t = \beta_0 + \beta_1 t + \beta_2 y_{t-1} + \varepsilon_t.$$

   Clearly indicate the one-to-one relation between $(\gamma_0, \gamma_1, \rho)$ and $(\beta_0, \beta_1, \beta_2)$. What name do we give the non-stationary process y_t when $\rho = 1$? Briefly indicate how to test whether this is the case. **(8 marks)**

4. Explain the RESET test as a general test for functional form misspecification and discuss the drawbacks and advantages of this test.

   In your answer consider the following multiple linear regression model:

$$y_i = \gamma_1 + \gamma_2 x_{2i} + \gamma_3 x_{3i} + u_i \qquad\qquad i = 1, ..., n,$$

   where x_{2i} and x_{3i} are exogenous variables known to affect $E(y_i)$. **(8 marks)**

5. Consider the following regression model:

$$Y_i = \beta_0 + \beta_1 \frac{1}{X_i} + u_i, \qquad i = 1, ..., n.$$

   We assume that the errors $\{u_i\}_{i=1}^n$ are independent normal random variables with zero mean and variance σ^2/X_i^2. The regressor, $1/X_i$, is non-stochastic with positive sample variability. You are interested in testing the hypothesis $H_0 : \beta_0 = 0$ against $H_1 : \beta_0 \neq 0$. You are advised to use the BLUE estimator of β_0 for this purpose.

   Discuss how you would obtain the BLUE estimator of β_0 (note, you are not asked to derive this estimator).

   Give two reasons why you would prefer using the BLUE estimator for β_0 instead of the OLS estimator $\hat{\beta}_{0,OLS} = \bar{Y} - \hat{\beta}_{1,OLS}\overline{(1/X)}$ when testing this hypothesis, where $\bar{Y} = \frac{1}{n}\sum Y_i$ and $\overline{(1/X)} = \frac{1}{n}\sum \frac{1}{X_i}$. **(8 marks)**

SECTION B

Answer three questions from this section.

6. Let us consider a model for the sale price of Monet paintings. The data we have contains the sale prices, widths, and heights of 430 Monet paintings, which sold at auction for prices ranging from \$10,000 to \$33 million. A linear regression provided the following results:

$$\widehat{\ln \text{Price}}_i = \underset{(0.612)}{-8.427} + \underset{(0.091)}{1.334} \ln \text{Area}_i - \underset{(0.128)}{0.165} \text{Aspect Ratio}_i$$
$$N = 430, R^2 = 0.336,$$

where Area=Width$\times$Height and Aspect Ratio=Height/Width. The standard errors are given in parentheses.

(a) Test the joint significance of the regression. Discuss its relation to the goodness of fit measure: R^2. **(5 marks)**

(b) You want to test the hypothesis that auction prices are inelastic with respect to area. Specifically, you are asked to test $H_0 : \beta_{\ln \text{Area}} = 1$ against $H_0 : \beta_{\ln \text{Area}} > 1$. Perform this test, clearly indicating the test statistic and the rejection rule. In view of your answer, indicate whether the p-value of the test is bigger or smaller than 5%.**(5 marks)**

(c) A friend points out that you should be worried about the presence of heteroskedasticity. Explain the concept of heteroskedasticity and discuss the properties of the OLS estimator in the presence of heteroskedasticity. **(5 marks)**

(d) Discuss how you should modify the test in (b) if it is known that the variance of the error term is given by

$$\sigma_i^2 = \exp\left(\gamma_0 + \gamma_1 \ln \text{Area}_i + \gamma_2 \text{Aspect Ratio}_i\right),$$

where $(\gamma_0, \gamma_1, \gamma_2)$ are unknown parameters. You are told to use the FGLS estimator. Explain this estimator clearly. **(5 marks)**

7. In this question we look at a large data set on weekly hours worked by women having at least two children. Consider the following specification for labour supply, estimated by OLS

$$\begin{aligned}\widehat{hours} &= \underset{(6.503)}{-17.347} - \underset{(.116)}{2.242}kids + \underset{(.035)}{.938}educ + \underset{(.453)}{2.089}age \underset{(.008)}{-.028}\, age^2 - \underset{(.005)}{.075}nonmomi \\ n &= 31,857 \text{ and } R^2 = .052, \end{aligned} \tag{7.1}$$

where *kids* is the total number of children, *educ* is the years of schooling, *age* is the woman's age in years, *nonmomi* is income from sources other than the mother's wage income, and *hours* is the hours worked per week. The standard errors are given in parentheses.

(a) Interpret the coefficient of *kids* and discuss its statistical significance. **(3 marks)**

(b) It is commonly thought that the decision to have more children is correlated with unobserved factors that affect labour supply. What name do we give such a problem in econometrics, and what consequences will this have for the properties of the OLS estimator? Support your answer using a simple model. **(5 marks)**

(c) Consider the variables *multi2nd* and *samesex*, which are binary variables indicating whether the second birth was for multiple babies and whether the first two children are of the same gender. What properties do we need to assume for us to be able to use *multi2nd* and *samesex* as instruments for *kids*? Are these assumptions reasonable and can we test (if so how) these requirements? **(6 marks)**

You may assume that *educ, age* and *nonmomi* can be treated as being exogenous.

(d) The following IV results were obtained in Stata:

$$\begin{aligned}\widehat{hours} &= \underset{(7.020)}{-16.828} - \underset{(.134)}{2.504}kids + \underset{(.115)}{.916}educ + \underset{(.460)}{2.105}age \underset{(.008)}{-.028}\, age^2 - \underset{(.007)}{.076}nonmomi \\ n &= 31,857 \text{ and } R^2 = .052. \end{aligned} \tag{7.2}$$

Discuss how these results can be obtained using Two Stage Least Squares (2SLS) and briefly discuss how you would test whether the results in (7.1) and (7.2) are significantly different from each other. **(6 marks)**

8. Let us consider an analysis on recidivism (probability of re-arrest) among a group of young men in California who have at least one arrest prior to 1986. The dependent variable, *arr86,* is equal to unity if the man was arrested at least once during 1986, and zero otherwise.

	OLS A	OLS B	Logit A	Logit B	Logit B Marginal Effect
pcnv	$-.152$ (.021)	$-.162$ (.021)	$-.880$ (.122)	$-.901$ (.120)	$-.176$ (.023)
avgsen	.005 (.006)	.006 (.006)	.027 (.035)	.031 (.034)	.006 (.007)
tottime	$-.003$ (.005)	$-.002$ (.005)	$-.014$ (.028)	$-.010$ (.027)	$-.002$ (.005)
ptime86	$-.023$ (.005)	$-.022$ (.005)	$-.140$ (.031)	$-.127$ (.031)	$-.025$ (.006)
qemp86	$-.038$ (.005)	$-.043$ (.005)	$-.199$ (.028)	$-.216$ (.028)	$-.042$ (.005)
black	.170 (.024)	–	.823 (.117)	–	–
hispan	.096 (.021)	–	.522 (.109)	–	–
constant	.380 (.019)	.380 (.019)	$-.464$ (.095)	$-.169$ (.084)	
R^2	.068	.047			
$\log L$			-1512.35	-1541.24	

pcnv is the proportion of prior arrests that led to a conviction, *avgsen* is the average sentence served from prior convictions, *tottime* is the months spent in prison since age 18 prior to 1986, *ptime86* is months spent in prison in 1986, *qemp86* is the number of quarters the man was legally employed in 1986, while *black* and *hispan* are two race dummies (*white* the excluded dummy). The standard errors are reported in parentheses.

(a) When estimating the parameters by OLS, we are using the Linear Probability Model. Why might you then report heteroskedasticity-robust standard errors? **(2 marks)**

(b) Using the OLS B results, what is the estimated effect on the probability of arrest if *pcnv* goes from 0.25 to 0.75 holding everything else constant? **(4 marks)**

(question continues on next page)

(c) It is argued that the linear probability model is not appropriate for explaining the binary variable $arr86$ and a logit regression model has been estimated. Explain how the Logit estimates are obtained. **(5 marks)**

Hint: You may recall, that for the Logit model A, we will specify:

$$\Pr(arr86_i = 1) = \Lambda\left(\beta_0 + \beta_1 \textit{pcnv} + \beta_2 \textit{avgsen} + ... + \beta_6 \textit{black} + \beta_7 \textit{hispan}\right),$$

where $\Lambda(z) = \frac{1}{1+\exp(-z)}$.

(d) Using the Logit model results, discuss whether *black* and *hispan* are jointly significant. Clearly indicate the null and alternative hypothesis, the test statistic and the rejection rule. **(3 marks)**

(e) An important distinction between the two approaches is that the marginal effect of *pcnv* on the probability of re-arrest is constant for the LPM unlike the marginal effect using the logit analysis. What this means for instance is that the estimated effect on the probability of arrest if *pcnv* goes from 0.25 to 0.75 will depend on the other characteristics.

i. Explain how the marginal effects evaluated at the mean values of the explanatory variables (reported in the last column) were obtained. Give a brief comment as to how they compare to the marginal effect of the associated LPM. **(3 marks)**

ii. Using the Logit B results, how would you obtain the estimated effect on the probability of arrest if *pcnv* goes from 0.25 to 0.75 for a *white* man, with characteristics *avgsen=1, tottime=1, ptime86=0 and qemp86=2.* A clear explanation of what calculations are required is sufficient. **(3 marks)**

9. Let us consider a distributed lag model:

$$y_t = \beta_0 + \delta_0 z_t + \delta_1 z_{t-1} + ... + \delta_q z_{t-q} + \beta_1 x_t + u_t\,, \qquad t = 1, ..., T,$$

where u_t is independent of x_t, z_t, z_{t-1}, ..., and z_{t-q} with zero mean and constant variance. For simplicity, we will assume that there is no autocorrelation in the errors.

(a) Explain the concept of autocorrelation and indicate **for the above model** what the consequence of autocorrelation for our OLS estimator would be. **(4 marks)**

Let us consider the example of the effects of tax policy on the U.S. fertility rates. Let *gfr* denote the number of children born per 1,000 women aged 15-44, pe denotes the real value of the personal tax exemption, and $ww2$ and $pill$ are dummy variables (WW II, availability of the birth control pill). Using annual data, the following OLS results were obtained:

$$\begin{aligned} \widehat{gfr}_t \;=\; & \underset{(3.30)}{92.50} + \underset{(.126)}{.089} pe_t - \underset{(.153)}{.004} pe_{t-1} + \underset{(.165)}{.007} pe_{t-2} + \underset{(.154)}{.018} pe_{t-3} + \underset{(.105)}{.014} pe_{t-4} \\ & - \underset{(11.03)}{21.48} ww2_t - \underset{(3.94)}{31.25} pill_t, \qquad R^2 = .537, T = 68. \end{aligned} \tag{9.1}$$

The usual standard errors are reported in parentheses.

(b) A novice in econometrics argues that the result indicates that the effect of tax policy on the US fertility rate is ineffective because of the insignificance of the distributed lag coefficients (the coefficients on pe_t and its lags). Explain why he/she could be wrong. In your answer you may discuss what the short and long run effect of the tax policy are according to these results **(5 marks)**

(c) Your friend argues that only two lags should have been included. Using the same sample, he obtains the following result:

$$\begin{aligned} \widehat{gfr}_t \;=\; & \underset{(3.27)}{92.52} + \underset{(.119)}{.101} pe_t - \underset{(.147)}{.011} pe_{t-1} + \underset{(.119)}{.033} pe_{t-2} \\ & - \underset{(10.13)}{22.95} ww2_t - \underset{(3.76)}{30.83} pill_t, \qquad R^2 = .536, T = 68. \end{aligned} \tag{9.2}$$

Test for the joint significance of the third and fourth lag. Clearly indicate H_0 and H_1, the test statistic, the rejection rule and interpret your results. **(4 marks)**

(d) Your result in (c) may be affected by the presence of autocorrelation. Discuss how you would conduct the Breusch-Godfrey test for the presence of autocorrelation in (9.2). Clearly indicate H_0 and H_1, the test statistic, the rejection rule and interpret your results. Briefly indicate what you might want to do to try and remove the autocorrelation. **(7 marks)**

10. Consider the model:

$$y_t = \alpha_1 y_{t-1} + u_t,\ t = 1, ..., T,$$

where $y_0 = 0$. $E(u_t) = 0$, $E(u_t^2) = \sigma^2$ and $E(u_t u_s) = 0$ when $s \neq t$, for all $s, t = 1, ..., T$

(a) Discuss what we mean by the concept of stationarity (more precisely "covariance stationarity") and indicate under what condition $\{y_t\}_{t=1}^T$ will be stationary. **(3 marks)**

(b) Discuss the Dickey-Fuller procedure used to test for the presence of a unit root in the above model. Clearly indicate the null and alternative hypothesis, test statistic and rejection rule. **(5 marks)**

(c) Consider a slight variation of the above model:

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + v_t,\ t = 1, ..., T,$$

where $y_0 = 0$. $E(v_t) = 0$, $E(v_t^2) = \sigma^2$ and $E(v_t v_s) = 0$ when $s \neq t$, for all $s, t = 1, ..., T$ What do we call such a process? Discuss what problem you will have when conducting your test as in (b). **(3 marks)**

(d) Instead of conducting the Dickey-Fuller procedure, you are told to apply the Augmented DF test. Indicate how you would conduct the test for the presence of a unit root here. Derivation of the test equation will be required for full marks. **(5 marks)**

(e) What are the potential problems associated with performing a regression with $I(1)$ variables? In your answer explain what it means for a variable to be $I(1)$. **(4 marks)**

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	**0·40**	0·6554	**0·80**	0·7881	**1·20**	0·8849	**1·60**	0·9452	**2·00**	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	**0·45**	0·6736	**0·85**	0·8023	**1·25**	0·8944	**1·65**	0·9505	**2·05**	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	**0·50**	0·6915	**0·90**	0·8159	**1·30**	0·9032	**1·70**	0·9554	**2·10**	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	**0·55**	0·7088	**0·95**	0·8289	**1·35**	0·9115	**1·75**	0·9599	**2·15**	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	**0·60**	0·7257	**1·00**	0·8413	**1·40**	0·9192	**1·80**	0·9641	**2·20**	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	**0·65**	0·7422	**1·05**	0·8531	**1·45**	0·9265	**1·85**	0·9678	**2·25**	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	**0·70**	0·7580	**1·10**	0·8643	**1·50**	0·9332	**1·90**	0·9713	**2·30**	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	**0·75**	0·7734	**1·15**	0·8749	**1·55**	0·9394	**1·95**	0·9744	**2·35**	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	**0·80**	0·7881	**1·20**	0·8849	**1·60**	0·9452	**2·00**	0·9772	**2·40**	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	5·0	1·6449	3·0	1·8808	2·0	2·0537	1·0	2·3263	0·10	3·0902
45	0·1257	4·8	1·6646	2·9	1·8957	1·9	2·0749	0·9	2·3656	0·09	3·1214
40	0·2533	4·6	1·6849	2·8	1·9110	1·8	2·0969	0·8	2·4089	0·08	3·1559
35	0·3853	4·4	1·7060	2·7	1·9268	1·7	2·1201	0·7	2·4573	0·07	3·1947
30	0·5244	4·2	1·7279	2·6	1·9431	1·6	2·1444	0·6	2·5121	0·06	3·2389
25	0·6745	4·0	1·7507	2·5	1·9600	1·5	2·1701	0·5	2·5758	0·05	3·2905
20	0·8416	3·8	1·7744	2·4	1·9774	1·4	2·1973	0·4	2·6521	0·01	3·7190
15	1·0364	3·6	1·7991	2·3	1·9954	1·3	2·2262	0·3	2·7478	0·005	3·8906
10	1·2816	3·4	1·8250	2·2	2·0141	1·2	2·2571	0·2	2·8782	0·001	4·2649
5	1·6449	3·2	1·8522	2·1	2·0335	1·1	2·2904	0·1	3·0902	0·0005	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

Examiners' commentaries 2017

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2016–17. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

General remarks

Learning outcomes

At the end of the course, and having completed the Essential reading and activities, you should be able to:

- describe and apply the classical regression model and its application to cross-section data
- describe and apply the:
  - Gauss–Markov conditions and other assumptions required in the application of the classical regression model
  - reasons for expecting violations of these assumptions in certain circumstances
  - tests for violations
  - potential remedial measures, including, where appropriate, the use of instrumental variables
- recognise and apply the advantages of logit, probit and similar models over regression analysis when fitting binary choice models
- competently use regression, logit and probit analysis to quantify economic relationships using standard regression programmes (Stata and EViews) in simple applications
- describe and explain the principles underlying the use of maximum likelihood estimation
- apply regression analysis to time-series models using stationary time series, with awareness of some of the econometric problems specific to time series applications (for example, autocorrelation) and remedial measures
- recognise the difficulties that arise in the application of regression analysis to nonstationary time series, know how to test for unit roots, and know what is meant by cointegration.

Common mistakes committed by candidates

A large number of candidates are not able to clearly distinguish between sample variance and covariance, and population variance and covariance (this is happening year after year).

The use of $\mathrm{Cov}(X,Y)$ and $\mathrm{Var}(X)$ should be restricted to describing the population covariance and variances, respectively, with definitions:

$$\mathrm{Cov}(X,Y) = \mathrm{E}\left((X-\mathrm{E}(X))(Y-\mathrm{E}(Y))\right) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$$

and:

$$\mathrm{Var}(X) = \mathrm{E}((X-\mathrm{E}(X))^2) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$$

(you also may denote $\mathrm{Cov}(X,Y) = \sigma_{XY}$ and $\mathrm{Var}(X) = \sigma_X^2$). They are typically unknown, but fixed, quantities.

The sample covariance and variance are estimators of the population covariance and variance, respectively. They are defined as:

$$\text{Sample Cov}(X,Y) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})$$

and:

$$\text{Sample Var}(X) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

(you also may use $\widehat{\sigma}_{XY}$ and $\widehat{\sigma}_X^2$). You can compute them given the data.

With a slight abuse of notation, we often divide by n instead, which is irrelevant if we let n be large. The division by $n-1$ is a finite sample issue only (unbiasedness).

The sample covariance and variance show up in our definition of the OLS estimator of the slope in the simple linear regression model, not the population covariance and variance, as:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sum_{i=1}^{n}(X_i-\bar{X})^2} = \frac{\text{Sample Cov}(X,Y)}{\text{Sample Var}(X)} \neq \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}.$$

Treating them as being the same results in incorrect analyses and candidates losing significant marks.

Candidates should realise that $\frac{1}{n}\sum_{i=1}^{n} u_i$ is not the same as $\mathrm{E}(u_i)$. So, while we typically assume $\mathrm{E}(u_i) = 0$, this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} u_i = 0$. Also, while we may be happy to assume $\mathrm{E}(x_i u_i) = 0$ (uncorrelatedness between the errors and regressors), this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} x_i u_i = 0$. Note that:

- both $\frac{1}{n}\sum_{i=1}^{n} u_i$ and $\frac{1}{n}\sum_{i=1}^{n} x_i u_i$ are random variables, which take the value 0 with probability 0 (continuous random variables)!
- $\mathrm{E}(u_i) = 0$ and $\mathrm{E}(x_i u_i) = 0$ are fixed, not stochastic!

The differences between sample and population moments need to come across clearly when looking at unbiasedness and making consistency arguments. In both cases, we first simplify our estimator (plug in the true model) to obtain:

$$\widehat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i-\bar{X})u_i}{\sum_{i=1}^{n}(X_i-\bar{X})^2} = \beta + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \qquad \text{with } x_i = X_i - \bar{X}.$$

- For *unbiasedness*, clearly indicate that you want to show that $\mathrm{E}(\widehat{\beta}) = \beta$. Unbiasedness does not follow from $\sum_{i=1}^{n} x_i u_i = 0$, instead it follows from $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) = 0$.

  If we treat x_i as fixed, $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) \equiv \mathrm{E}\left(\sum_{i=1}^{n} d_i u_i\right) = \sum_{i=1}^{n} d_i \,\mathrm{E}(u_i)$ and then unbiasedness follows as $\mathrm{E}(u_i) = 0$.
- For *consistency*, clearly indicate that you want to show that $\mathrm{plim}(\widehat{\beta}) = \beta$. Using the plim properties, we show:

$$\begin{aligned}
\mathrm{plim}\ \widehat{\beta} = \beta + \mathrm{plim}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) &= \beta + \frac{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i u_i\right)}{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)} \\
&\equiv \beta + \frac{\mathrm{plim}\,(\text{Sample Cov}\,(x,u))}{\mathrm{plim}\,(\text{Sample Var}\,(x))} \\
&= \beta + \frac{\mathrm{Cov}\,(x,u)}{\mathrm{Var}\,(x)} \qquad \text{using the law of large numbers}
\end{aligned}$$

  where $\mathrm{Cov}(x,u) = 0$ and $\mathrm{Var}(x) > 0$, ensuring we get consistency.
- Remember, the law of large numbers ensures that sample averages converge to their population analogues.

Candidates struggled to give competent answers to the interpretation of empirical results. When interpreting an empirical result you should discuss the significance of the coefficients, magnitude and sign of the coefficients.

When conducting hypothesis tests, you should make sure that the Gauss–Markov conditions hold. The Gauss–Markov conditions have to be explicitly specified. Only writing that the Gauss–Markov conditions hold is not sufficient. As good practice, begin your examination by explicitly providing the Gauss–Markov conditions. You can then refer back to them thereafter. Moreover, ensure when conducting hypothesis testing that you clearly indicate the null and alternative hypotheses (in terms of the true parameters, say β_1), the test statistic (in terms of the parameter estimates, here $\widehat{\beta}_1$), its distribution (with degrees of freedom), the rejection rule (one-sided or-two sided) for a given significance level (typically 5%) with suitable critical values, and provide an interpretation of your result.

Just as last year, many candidates do not answer all parts of the question. Make sure you read the questions properly and provide all details that are requested. Not answering a question will automatically earn you a zero mark for that question.

Key steps to improvement

Essential reading for this course includes the subject guide and the following:

- Dougherty, C. *Introduction to econometrics.* (Oxford: Oxford University Press, 2016) 5th edition [ISBN 9780199676828]; http://oxfordtextbooks.co.uk/orc/dougherty5e/

Apart from the Essential readings you should do some supplementary reading. One very good book at the same level is:

- Gujarati, D.N. and D.C. Porter *Basic econometrics.* (McGraw–Hill, 2009, International edition) 5th edition [ISBN 9780071276252].

To understand the subject clearly it is important to supplement Dougherty's *Introduction to econometrics* (fifth edition) with the subject guide **EC2020 Elements of econometrics** (2016), especially Chapter 10 which covers maximum likelihood estimation. It is very important to carefully go through the subject guide. The subject guide contains solutions to the questions given in the main textbook and also some additional questions and solutions. Working through these will improve your understanding of the subject.

The chapter in the subject guide on maximum likelihood (Chapter 10) includes some additional theory which has not been covered in the main textbook. It is important to read the additional theory given in the subject guide to have a better understanding of the principles of maximum likelihood and tests based on the likelihood function.

Please check the VLE course page for resources for this subject such as a downloadable copy of the subject guide **EC2020 Elements of econometrics** (2016), PowerPoint slideshows that provide a graphical treatment of the topics covered in the textbook, datasets and statistical tables. Candidates should utilise datasets using standard regression programmes (STATA or EViews). This will help in the understanding of the subject.

Examination revision strategy

Many candidates are disappointed to find that their examination performance is poorer than they expected. This may be due to a number of reasons, but one particular failing is '**question spotting**', that is, confining your examination preparation to a few questions and/or topics which have come up in past papers for the course. This can have serious consequences.

We recognise that candidates might not cover all topics in the syllabus in the same depth, but you need to be aware that examiners are free to set questions on **any aspect** of the syllabus. This means that you need to study enough of the syllabus to enable you to answer the required number of examination questions.

The syllabus can be found in the Course information sheet available on the VLE. You should read the syllabus carefully and ensure that you cover sufficient material in preparation for the examination. Examiners will vary the topics and questions from year to year and may well set questions that have not appeared in past papers. Examination papers may legitimately include questions on any topic in the syllabus. So, although past papers can be helpful during your revision, you cannot assume that topics or specific questions that have come up in past examinations will occur again.

If you rely on a question-spotting strategy, it is likely you will find yourself in difficulties when you sit the examination. We strongly advise you not to adopt this strategy.

Examiners' commentaries 2017

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2016–17. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone A

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Section A

Answer all questions from this section.

Question 1

Discuss the drawbacks and advantages of using the Linear Probability Model when trying to explain a binary decision. In your answer clearly indicate what the Linear Probability Model is.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 10.1 (The linear probability model).

Subject guide (2016), Chapter 10.

Approaching the question

Candidates should clearly indicate that a linear probability model (LPM) is used to denote a model in which the dependent variable is binary, which takes the value 1 if the event occurs and

0 if it does not. It is estimated by ordinary least squares (OLS). Answers should discuss both drawbacks and advantages of using the LPM. A discussion of Logit/Probit and maximum likelihood estimation (MLE) is inappropriate here. The answer is as follows.

The LPM specifies $P_i = \Pr(Y_i = 1 \mid X_i) = \beta_1 + \beta_2 X_i$ and applies OLS to the model:

$$Y_i = \beta_1 + \beta_2 X_i + u_i$$

where $\mathrm{E}(u_i \mid X_i) = 0$. Recall, for the binary (discrete) random variable Y_i, we have:

$$\mathrm{E}(Y_i \mid X_i) = 1 \times \Pr(Y_i = 1 \mid X_i) + 0 \times \Pr(Y_i = 0 \mid X_i) = \Pr(Y_i = 1 \mid X_i)$$

hence we should interpret $\mathrm{E}(Y_i \mid X_i) = \beta_1 + \beta_2 X_i$ as the probability that the event will occur, given X_i.

Advantages: The results of the LPM are *easy to interpret* as marginal effects and/or *easy to estimate*. If we denote $\widehat{\beta}_0$ and $\widehat{\beta}_1$ as estimates of β_0 and β_1, respectively, then:

$$\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_i = \widehat{P}_i$$

is the estimated probability that the event will occur, and $\widehat{\beta}_1$ provides the marginal effect the explanatory variable X has on the probability of $Y = 1$, *ceteris paribus*.

Drawbacks: Since $\mathrm{E}(u_i) = 0$, we have:

$$\mathrm{E}(u_i) = (1 - \beta_0 - \beta_1 X_i) \underbrace{\Pr(Y_i = 1)}_{\beta_0 + \beta_1 X_i} + (-\beta_0 - \beta_1 X_i) \underbrace{\Pr(Y_i = 0)}_{1 - (\beta_0 + \beta_1 X_i)} = 0$$

and $\mathrm{Var}(u_i) = \mathrm{E}(u_i^2)$ exhibits *heteroskedasticity* rendering OLS *inefficient*. We have:

$$\begin{aligned}
\mathrm{E}(u_i^2) &= (1 - \beta_0 - \beta_1 X_i)^2 \underbrace{\Pr(Y_i = 1)}_{\beta_0 + \beta_1 X_i} + (-\beta_0 - \beta_1 X_i)^2 \underbrace{\Pr(Y_i = 0)}_{1 - (\beta_0 + \beta_1 X_i)} \\
&= (1 - \beta_0 - \beta_1 X_i)(\beta_1 + \beta_2 X_i)\left[(1 - \beta_0 - \beta_1 X_i) + (\beta_0 + \beta_1 X_i)\right] \\
&= (\beta_1 + \beta_2 X_i)(1 - \beta_0 - \beta_1 X_i) \\
&= \mathrm{E}(Y_i)(1 - \mathrm{E}(Y_i)) \\
&= P_i(1 - P_i) \qquad \text{for all } i.
\end{aligned}$$

In addition, as Y_i can take only two values, 1 or 0, u_i can only take the values $1 - \beta_0 - \beta_1 X_i$ when $Y_i = 1$ and $-\beta_0 - \beta_1 X_i$ when $Y_i = 0$ for given X. Therefore, the *errors are highly non-normal*. Finally, a major drawback with the LPM is that the *estimated probability can be negative or greater than 1, which are unreasonable results*. The fact that the marginal effects in the LPM are constant gives rise to this problem as OLS.

Question 2

Discuss the consequences of measurement error.

In your answer consider the following empirical study attempting to estimate the relationship between advertising and magazine circulation rates. A simple linear relationship is postulated between A_t, the advertising rate in magazine t, and C_t, the circulation figure for the magazine in question:

$$A_t = \alpha + \beta C_t + \varepsilon_t, \qquad t = 1, \ldots, T.$$

The relation is estimated by least squares on data for $T = 75$ magazines. Unfortunately, there are considerable errors in measurement in the reported circulation figures. Critically discuss the following statement: 'The estimated coefficient $\widehat{\beta}$ will be too small.' Rigour of your answer will be rewarded.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 8.4 (The consequences of measurement error).

Subject guide (2016), Chapter 8.

Approaching the question

Candidates should clearly discuss, with technical details (rigour), the consequences of measurement error. Inappropriate comments, for example stating that measurement error is caused by omitted variables, are penalised. It is important to clearly point out that measurement error in the explanatory variable (as is the case in our setting) induces correlation between the error and regressors – failure to point this out is serious. The answer is as follows.

Assuming we have classical measurement error, the measurement error on C_t causes our estimate of β to be *biased towards zero* – we also call this problem 'attenuation bias'. This is not the same as saying that the estimated coefficient will be too small.

Measurement in the circulation figures induces the problem of correlation between error and regressors, $\mathrm{E}(x_i\varepsilon_i) \neq 0$, giving rise to such inconsistency (bias).

Candidates should set up the model as follows. We are given the *true model*:

$$A_t = \alpha + \beta C_t + \varepsilon_t \qquad \text{where } \mathrm{Cov}(C_t, \varepsilon_t) = 0.$$

Unfortunately, we are told that C_t is *unobservable*. Instead we *observe* C_t^*, the circulation figures measured with error:

$$C_t^* = C_t + v_t \qquad \text{where } v_t \text{ is the } \textit{measurement error}, \quad v_t \text{ is i.i.d. } (0,\, \sigma_v^2).$$

The classical measurement assumptions ensure that:

$$\mathrm{Cov}(v_t, \varepsilon_t) = 0 \qquad \text{and} \qquad \mathrm{Cov}(v_t, C_t) = 0$$

with v_t independent of anything else in the model.

Therefore, the *estimable model* we can use to estimate β becomes:

$$A_t = \alpha + \beta C_t^* + u_t \qquad \text{with } u_t = \varepsilon_t - \beta v_t.$$

This estimator will be *inconsistent* as $\mathrm{Cov}(C_t^*, u_t) \neq 0$. Specifically, under the above assumptions:

$$\mathrm{Cov}(C_t^*, u_t) = \mathrm{Cov}(C_t + v_t, \varepsilon_t - \beta v_t) = -\beta\,\mathrm{Var}(v_t).$$

Candidates should clearly indicate the estimator, whose properties we need to discuss, as:

$$\widehat{\beta} = \frac{\sum(C_t^* - \bar{C}^*)(A_t - \bar{A})}{\sum(C_t^* - \bar{C}^*)^2} = \beta + \frac{\sum(C_t^* - \bar{C}^*)(u_t - \bar{u})}{\sum(C_t^* - \bar{C}^*)^2} = \beta + \frac{\text{Sample Cov}(C_t^*, u_t)}{\text{Sample Var}(C_t^*)}.$$

Using the *plim operator* and the *law of large numbers*, we then obtain:

$$\begin{aligned}
\mathrm{plim}(\widehat{\beta}) &= \beta + \frac{\mathrm{plim}\frac{1}{T}\sum(C_t^* - \bar{C}^*)(u_t - \bar{u})}{\mathrm{plim}\frac{1}{T}\sum(C_t^* - \bar{C}^*)^2} \\
&= \beta + \frac{\mathrm{plim}(\text{Sample Cov}(C_t^*, u_t))}{\mathrm{plim}(\text{Sample Var}(C_t^*))} \\
&= \beta + \frac{-\beta\,\mathrm{Var}(v_t)}{\mathrm{Var}(C_t^*)} \\
&= \beta\left(1 - \frac{\sigma_v^2}{\sigma_C^2 + \sigma_v^2}\right)
\end{aligned}$$

as $\text{Var}(C_t^*) = \text{Var}(C_t + v_t) = \sigma_C^2 + \sigma_v^2$.

Since $0 < \left(1 - \dfrac{\sigma_v^2}{\sigma_c^2 + \sigma_v^2}\right) < 1$, we have completed the proof of the attenuation bias.

Question 3

Consider the following regression model:

$$\boldsymbol{Y_i = \beta_0 + \beta_1 \frac{1}{X_i} + u_i, \qquad i = 1, \ldots, n.}$$

We assume that the errors $\{u_i\}_{i=1}^n$ are independent normal random variables with zero mean and variance σ^2/X_i^2. The regressor, $1/X_i$, is nonstochastic with positive sample variability and $X_i \neq 0$ for all i. You are interested in testing the hypothesis $\text{H}_0 : \beta_0 = 0$ against $\text{H}_1 : \beta_0 \neq 0$. You are advised to use the BLUE estimator of β_0 for this purpose.

Discuss how you would obtain the BLUE estimator of β_0 (note, you are not asked to derive this estimator).

Give two reasons why you would prefer using the BLUE estimator for β_0 instead of the OLS estimator $\widehat{\beta}_{0,OLS} = \bar{Y} - \widehat{\beta}_{1,OLS}\overline{(1/X)}$ when testing this hypothesis, where $\bar{Y} = \frac{1}{n}\sum Y_i$ and $\overline{(1/X)} = \frac{1}{n}\sum \frac{1}{X_i}$.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): R.10 (Type II error and the power of a test), Chapter 2.5 (The Gauss–Markov theorem), Chapter 2.6 (Testing hypotheses relating to the regression coefficients), Chapter 7.1 (Heteroskedasticity and its implications) and 7.3 (Remedies for heteroskedasticity).

Subject guide (2016), Chapter 7.

Approaching the question

Candidates should show a clear understanding of what it means for an estimator to be BLUE. Failure to recognise the presence of heteroskedasticity in this question is serious. Arguments that OLS on the regression itself would not be linear because of the form of the regressor $1/X_i$ are wrong. OLS on the regression model given is linear (it is linear in the parameters!), but due to the heteroskedasticity it is not efficient (BLUE)! Many candidates lost points simply because they did not discuss two reasons why, *for testing purposes,* one would prefer to use this BLUE estimator. Answer all parts of the question. The answer is as follows.

Since the only Gauss–Markov violation here is the presence of heteroskedasticity, we should propose to use weighted least squares (WLS) to make the problem go away. Regress:

$$Y_i X_i = \beta_0 X_i + \beta_1 + u_i X_i \qquad \text{for } i = 1, \ldots, n$$

or:

$$Y_i^* = \beta_1 + \beta_0 X_i + u_i^*.$$

This transformed model satisfies all the Gauss–Markov assumptions, $\text{E}(u_i^*) = \text{E}(u_i)\, X_i = 0$ and $\text{Var}(u_i^*) = \text{Var}(u_i X_i) = \text{Var}(u_i)\, X_i^2 = \sigma^2$. Do make sure you mention that the *Gauss–Markov theorem* ensures that OLS on this regression will yield our BLUEs of β_0 and β_1.

Benefits to using the BLUE of β_0 are (i) it is an efficient (more precise) estimator and because of this efficiency the test will have higher power (easier to reject the null hypothesis when it is false), and (ii) using the BLUE estimator allows us to directly use its standard error for inference. Had we used the OLS estimator instead, we would have needed to obtain heteroskedasticity-robust standard errors.

Question 4

Consider the following non-stationary process:

$$y_t = \gamma_0 + \gamma_1 t + u_t, \qquad \text{with } u_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$$

and ε_t i.i.d. $(0, \sigma^2)$. Indicate (with explanation) the source(s) of non-stationarity of y_t. Discuss how you would test whether y_t indeed is non-stationary. Clearly indicate the null and the alternative hypothesis, the test statistic and the rejection rule. What name do we give such a non-stationary process?

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 12.1 (Definition and consequences of autocorrelation), 13.1 (Stationarity and non-stationarity) and 13.5 (Tests of deterministic trends).

Subject guide (2016), Chapters 12 and 13.

Approaching the question

Candidates should show a clear understanding of what non-stationarity means. A discussion of a unit root test here is inappropriate as the dependence structure is given by an MA(2) process, not an AR process. The only source of non-stationarity is the presence of the deterministic trend, $\gamma_1 \neq 0$. Many candidates failed to recognise this. Answer all parts of the question, for example do not forget to answer how we call such a non-stationary process! The answer is as follows.

A process $\{y_t\}$ is (covariance) stationary if its mean and variance exist and do not depend on time and its covariance is a function of distance in time only (not location). Violation of either of these requirements renders the process non-stationary.

In this case $\{y_t\}$ is non-stationary because $\mathrm{E}(y_t) = \gamma_0 + \gamma_1 t$ depends on time when $\gamma_1 \neq 0$. Indeed, the only problem of non-stationarity here is the *presence of the deterministic trend*, as the dependence of $\{u_t\}$, and therefore $\{y_t\}$, is MA(2) and all finite-order moving average processes are (covariance) stationary.

Therefore, candidates should propose to test $\mathrm{H}_0 : \gamma_1 = 0$ against $\mathrm{H}_1 : \gamma_1 > 0$ (when trending upwards). For this we will use the t test $\widehat{\gamma}/\mathrm{SE}(\widehat{\gamma})$. Critical values are given by the $N(0, 1)$ distribution (since T is large). Reject H_0 if $\widehat{\gamma}/\mathrm{SE}(\widehat{\gamma}) > 1.645$ at the 5% significance level. (Suggesting a two-sided test is fine as well, but has less power.) Due to the presence of the dependence in the errors, *robust standard errors* should be used when implementing the test (this point was made by only a few candidates).

If we reject the null hypothesis, we call such a process $\{y_t\}$ *trend-stationary.*

Question 5

Suppose you are given a random sample $X_1, \ldots, X_n$ from the exponential distribution:

$$f(x) = \lambda \exp(-\lambda x), \qquad x > 0,\ \lambda > 0.$$

According to this distribution $\mathrm{E}(X_i) = 1/\lambda$ and $\mathrm{Var}(X_i) = 1/\lambda^2$ for $i = 1, \ldots, n$.

Show that the maximum likelihood estimator for λ is $1/\bar{X}$ where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Is the estimator unbiased and/or consistent? Prove your claims.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 4.2 (Logarithmic transformations) and Chapter 10.6 (An introduction to maximum likelihood estimation).

Subject guide (2016), Chapter 10.

Approaching the question

Candidates should clearly conduct the maximum likelihood procedure and ensure proper uses of the product and logarithm operators are displayed. Many candidates made the classic error of not recognising that $\mathrm{E}(1/\bar{X}) \neq 1/\mathrm{E}(\bar{X})$ and also failed to realise that the proof of consistency should have used the plim operator and the law of large numbers as $\mathrm{plim}(1/\bar{X}) = 1/(\mathrm{plim}\ \bar{X})$. Instead, many candidates attempted (unsuccessfully) to look at the sufficient conditions (what is $\mathrm{Var}(1/\bar{X})$?). Finally, candidates should make a clear distinction between the unknown (fixed) parameter λ and its MLE $\widehat{\lambda}$ (a random variable). The answer is as follows.

The likelihood function is given by the joint density of the data (which equals the product of the marginals given the independence of our observations):

$$L(\lambda) = \prod_{i=1}^{n} \lambda \exp(-\lambda X_i) = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} X_i\right).$$

The MLE is the value of λ which maximises this function, or by monotonicity, the log-likelihood function:

$$\ln L(\lambda) = \ln\left(\lambda^n \exp\left(-\lambda \sum_{i=1}^{n} X_i\right)\right) = n \ln \lambda - \lambda \sum_{i=1}^{n} X_i.$$

By the first-order condition, we require:

$$\left.\frac{\partial \ln L}{\partial \lambda}\right|_{\widehat{\lambda}_{MLE}} = \frac{n}{\widehat{\lambda}_{MLE}} - \sum_{i=1}^{n} X_i = 0 \qquad \text{or} \qquad \widehat{\lambda}_{MLE} = \frac{1}{\bar{X}} \qquad \text{with} \qquad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i.$$

(It should clearly state $\widehat{\lambda}_{MLE} = 1/\bar{X}$, saying $\lambda = 1/\bar{X}$ is incorrect.)

The MLE is *not unbiased*:

$$\mathrm{E}(\widehat{\lambda}_{MLE}) = \mathrm{E}\left(\frac{1}{\bar{X}}\right) \overset{\text{Jensen's Ineq}}{\neq} \frac{1}{\mathrm{E}(\bar{X})} = \frac{1}{\frac{1}{n}\sum \mathrm{E}\left(X_i\right)} = \lambda.$$

The MLE is *consistent*:

$$\mathrm{plim}\ \widehat{\lambda}_{MLE} \overset{\text{plim rules}}{=} \frac{1}{\mathrm{plim}(\bar{X})} \overset{\text{LLN}}{=} \frac{1}{\mathrm{E}(X_i)} = \lambda.$$

Section B

Answer three questions from this section.

Question 6

The following equation was estimated by Ordinary Least Squares based on second semester candidates in the fall term

$$\begin{aligned}\widehat{trmgpa} &= \underset{\substack{(.55)\\[.55]}}{-2.12} + \underset{\substack{(.175)\\[.166]}}{.900}\,crsgpa + \underset{\substack{(.064)\\[.074]}}{.193}\,cumgpa + \underset{\substack{(.0012)\\[.0012]}}{.0014}\,tothrs \\ &\quad + \underset{\substack{(.0002)\\[.0002]}}{.0018}\,sat - \underset{\substack{(.0018)\\[.0019]}}{.0039}\,hsperc + \underset{\substack{(.085)\\[.079]}}{.351}\,female - \underset{\substack{(.098)\\[.080]}}{.157}\,season \\ n &= 269;\ R^2 = .465 \qquad (6.1)\end{aligned}$$

where $trmgpa_i$ is the term GPA (grade point average) of individual i. $crsgpa_i$ is a measure of difficulty of courses taken by individual i (weighted overall average of GPA in selected courses), $cumgpa_i$ is the GPA of individual i prior to the current semester, $tothrs_i$ is the total credit hours of individual i prior to the semester, sat_i is his/her SAT score (test taken for college admission in the USA), $hsperc_i$ is the graduating percentile of individual i in high school class, $female_i$ is a gender dummy, and $season_i$ is a dummy variable equal to unity if the student's sport is in season during the fall term. The usual standard errors are in parentheses and the White's heteroskedasticity-robust standard errors are in squared brackets.

(a) **Explain the concept of heteroskedasticity and discuss the properties of the OLS estimator in the presence of heteroskedasticity.**

(5 marks)

(b) **Do the variables $crsgpa$, $cumgpa$ and $tothrs$ have the expected estimated effects? Which of these variables are statistically significant at the 5% level? Does it matter which standard errors are used?**

(5 marks)

(c) **In view of your concern about the presence of heteroskedasticity, provide the 95% confidence interval for β_{crsgpa} and use it to test the hypothesis $H_0 : \beta_{crsgpa} = 1$ against $H_1 : \beta_{crsgpa} \neq 1$. Describe your conclusions. In view of your answer, indicate whether the p-value of the test is bigger or smaller than 5%.**

(5 marks)

(d) **Discuss how you would conduct a test for heteroskedasticity in this setting. Clearly indicate the assumptions that underlie the test you suggest. Detail of your answer will be rewarded.**

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 2.6 (Testing hypotheses relating to the regression coefficients), R.12, and Chapter 7 (Heteroskedasticity).

Subject guide (2016), Chapter 7.

Approaching the question

In part (a), the concept of heteroskedasticity and the consequences on the OLS estimator need to be discussed. In part (b), the sign and significance of the coefficients should be explained. To conduct the test, a t test should be used where robust standard errors should be used because of

the presence of heteroskedasticity. In part (c), candidates should provide the 95% confidence interval and its interpretation. In part (d), a suitable test for heteroskedasticity should be provided. This question was answered by 85% of the candidates in Zone A. The answer is as follows.

(a) Discussion of the concept is standard bookwork, see Chapter 7.1 of Dougherty.

The properties of OLS are *inefficiency* (as heteroskedasticity is a violation of the Gauss–Markov assumptions), *unbiasedness* (heteroskedasticity does not violate the assumption that $\mathrm{E}(u \mid X) = 0$ needed for unbiasedness) and *consistency* (heteroskedasticity does not violate the assumption of contemperaneous uncorrelatedness needed for consistency). Moreover, the *usual standard errors of the OLS estimator will be invalid* (we will need to use White's heteroskedasticity-robust standard errors).

(b) These coefficients have the anticipated sign. If a student takes courses where grades are, on average, higher – as reflected by higher *crsgpa* – then his/her grades will be higher. The better the student has been in the past – as measured by *cumgpa* – the better the student does (on average) in the current semester. Finally, *tothrs* is a measure of experience, and its coefficient indicates an increasing return to experience.

We should use the t statistic to determine whether parameters are statistically significant: $\mathrm{H}_0 : \beta_i = 0$ against $\mathrm{H}_1 : \beta_i > 0$. Our test statistic is $\widehat{\beta}_i/\mathrm{SE}(\widehat{\beta}_i)$, which has a t_{n-8} distribution under H_0. (Exact distribution under Gauss–Markov assumptions + normality of the errors; approximate distribution when using robust standard errors in which case you may simply use $N(0, 1)$ critical values instead). Using a 5% significance level you should reject H_0 if $\widehat{\beta}_i/\mathrm{SE}(\widehat{\beta}_i) > 1.645$ with $i = \{crsgpa, cumgpa, tothrs\}$.

All parameters enter significantly. The t statistic for *crsgpa* is very large, over five using the usual standard error (which is the larger of the two). Using the robust standard error for *cumgpa*, its t statistic is about 2.61, which is also significant at the 5% significance level. The t statistic for *tothrs* is only about 1.17 using either standard error, so it is not significant. While the decision remains unchanged whichever standard errors are used, in the presence of heteroskedasticity we should use the robust standard errors as the other ones are invalid.

(c) The confidence interval, using the heteroskedasticity-robust standard errors, is given by:

$$\left[\widehat{\beta}_{crsgpa} - 1.96 \times \mathrm{SE}(\widehat{\beta}_{crsgpa}),\ \widehat{\beta}_{crsgpa} + 1.96 \times \mathrm{SE}(\widehat{\beta}_{crsgpa})\right] \qquad \text{or} \qquad [0.575,\ 1.255].$$

Since 1 lies in this 95% confidence interval, we do not reject the hypothesis that $\beta_{crgpga} = 1$. Conclusion: Everything else constant, candidates with a one-unit higher course GPA are expected to have a one-unit higher term GPA score. Many candidates were unable to provide this confidence interval and/or failed to interpret the conclusion from our test. Answers that conduct the test without reference to the confidence interval are incorrect.

The p-value is the lowest level of significance at which we want to reject the null hypothesis. Since we do not reject at the 5% significance level, the p-value will be bigger than 5%.

(d) Candidates can discuss here either the Goldfeld–Quandt test or White's test, see Chapter 7.2 of Dougherty.

Candidates should make it clear what the assumptions are that underlie either test. For the Goldfeldt–Quandt test they will have to specify the variable that enables them to divide the sample into large variance versus small variance observations (say *cumgpa*). When discussing White's test, on the other hand, no prior assumption is required about the form of heteroskedasticity.

Details of the test, test statistic, distribution and rejection rule should be provided (standard textbook answer).

Question 7

Consider the model:

$$y_t = \alpha_1 y_{t-1} + u_t, \qquad t = 1, \ldots, T$$

where $y_0 = 0$. $\mathrm{E}(u_t) = 0$, $\mathrm{E}(u_t^2) = \sigma^2$ and $\mathrm{E}(u_t u_s) = 0$ when $s \neq t$, for all $s, t = 1, \ldots, T$.

(a) **Discuss what we mean by the concept of stationarity (more precisely 'covariance stationarity') and indicate under what condition $\{y_t\}_{t=1}^T$ will be stationary.**

(3 marks)

(b) **Discuss the Dickey–Fuller procedure used to test for the presence of a unit root in the above model. Clearly indicate the null and alternative hypothesis, test statistic and rejection rule.**

(5 marks)

(c) **Consider a slight variation of the above model:**

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + v_t, \qquad t = 1, \ldots, T$$

where $y_0 = 0$. $\mathrm{E}(v_t) = 0$, $\mathrm{E}(v_t^2) = \sigma^2$ and $\mathrm{E}(v_t v_s) = 0$ when $s \neq t$, for all $s, t = 1, \ldots, T$. What do we call such a process? Discuss the problem you will have when conducting your test as in (b).

(3 marks)

(d) **Instead of conducting the Dickey–Fuller procedure, you are told to apply the Augmented DF test. Indicate how you would conduct the test for the presence of a unit root here. Derivation of the test equation will be required for full marks.**

(5 marks)

(e) **What are the potential problems associated with performing a regression with $I(1)$ variables? In your answer explain what it means for a variable to be $I(1)$.**

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 13.1 (Stationarity and nonstationarity), Chapter 13.4–13.5 (Tests of nonstationarity), and Chapter 13.6 (Cointegration).

Subject guide (2016), Chapter 13.

Approaching the question

In part (a), the concept of (covariance) stationarity needs to be given. In part (b), you need to provide the Dickey–Fuller test for unit roots and in part (d) you need to provide the Augmented Dickey–Fuller test. The latter allows us to deal with the fact that when conducting the Dickey–Fuller test we cannot have any autocorrelation, which would be the case if the dependence is not AR(1), but as is the case here AR(2). In part (e), a discussion of the spurious regression result is expected. This question was answered by 60% of the candidates in Zone A. The answer is as follows.

(a) Candidates should provide a textbook definition of (covariance) stationarity.

The requirement for (covariance) stationarity here is that $|\alpha_1| < 1$. (Stating $\alpha_1 < 1$ is permitted.) Here $\{y_t\}$ is described by an AR(1) process, which is stationary provided its coefficient, α_1, is smaller than 1 (in magnitude). If α_1 is equal to one we have a unit root, whereas if $|\alpha_1| > 1$ we have an explosive process.

(b) To test for a unit root, we want to estimate the following regression:

$$\Delta y_t = \gamma y_{t-1} + u_t \qquad \text{where } \gamma = (\alpha_1 - 1).$$

We test $\mathrm{H}_0 : \gamma = 0$ (nonstationarity) against $\mathrm{H}_1 : \gamma < 0$ (stationarity). (Alternatively, we estimate the original model and test $\mathrm{H}_0 : \alpha_1 = 1$ (nonstationarity, unit root) against $\mathrm{H}_1 : \alpha_1 < 1$ (stationary AR process).)

We use the Dickey–Fuller t test $\widehat{\gamma}/\mathrm{SE}(\widehat{\gamma}) = (\widehat{\alpha}_1 - 1)/\mathrm{SE}(\widehat{\alpha}_1)$. Because we have non-stationarity under the null hypothesis, this test is not standard and we have to use the Dickey–Fuller critical values. Given the alternative hypothesis, we reject H_0 when $\widehat{\gamma}/\mathrm{SE}(\widehat{\gamma}) < \tau$, where τ is the Dickey–Fuller critical value (no trend or intercept in the model). Candidates can, alternatively, discuss the Dickey–Fuller scaled coefficient test which is $T(\widehat{\alpha}_1 - 1) = T\widehat{\gamma}$.

If we reject the null hypothesis we have found evidence that our process is stationary (weakly dependent).

In the presence of a unit root, a shock to u_t will have an everlasting effect on the process $\{y_t\}$. Unit roots are persistent and strongly dependent.

(c) Here $\{y_t\}$ is an AR(2) process. (ADL(2, 0) is also acceptable.)

The problem with (b) will be serious, as *dependence (autocorrelation) in the error* will cause the error term u_t (which would then be equal to $\alpha_2 y_{t-2} + v_t$) and regressor y_{t-1} to be correlated and hence OLS will be inconsistent. Hence our test $\widehat{\gamma}/\mathrm{SE}(\widehat{\gamma})$ (or $T(\widehat{\alpha}_1 - 1)$) will be *invalid*.

(d) The Augmented Dickey–Fuller test suggests that we add further lags in our testing equation to remove the autocorrelation. In particular, we should estimate the following model:

$$\Delta y_t = \gamma_1 y_{t-1} + \gamma_2 \Delta y_{t-1} + v_t$$

and test the hypothesis $\mathrm{H}_0 : \gamma_1 = 0$ (non-stationarity) against $\mathrm{H}_1 : \gamma_1 < 0$ (stationarity). We should use the test given by $ADF = \widehat{\gamma}_1/\mathrm{SE}(\widehat{\gamma}_1)$ and we should reject H_0 when ADF is smaller than the critical value given by the Dickey–Fuller tables for a given significance level.

To derive this result observe that:

$$\begin{aligned}
\Delta y_t &= \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + v_t - y_{t-1} \\
&= (\alpha_1 - 1)\, y_{t-1} + \alpha_2 y_{t-2} + v_t \\
&= (\alpha_1 - 1)\, y_{t-1} + \alpha_2 \left(\underbrace{y_{t-2} - y_{t-1}}_{-\Delta y_{t-1}} + y_{t-1} \right) + v_t \\
&= (\alpha_1 + \alpha_2 - 1)\, y_{t-1} - \alpha_2 \Delta y_{t-1} + v_t
\end{aligned}$$

so $\gamma_1 = (\alpha_1 + \alpha_2 - 1)$ and $\gamma_2 = -\alpha_2$. Therefore, to test for a unit root in the AR(2) model we test whether $\alpha_1 + \alpha_2 < 1$.

(e) To say that a variable is $I(1)$ indicates that the variable has a unit root. We say that the variable then is integrated of order 1, revealing that by differencing the variable once we can make it stationary.

The potential problem associated with performing such a regression is that we may get a *spurious relation*. This is the setting where, due to the fact that both variables are trending, there is an appearance of a relationship that does not exist at all (high t statistics and a large R^2). To ensure that we have a meaningful (long-run) relationship between variables, we want to verify that instead we are dealing with a *cointegrating relationship*.

Question 8

For the US economy, let *gprice* denote the monthly growth in the overall price level and let *gwage* be the monthly growth in hourly wages. We have estimated the following distributed lag model (ADL(0, 12)):

$$\begin{aligned}\widehat{gprice}_t &= \underset{(.00057)}{-.00093} + \underset{(.052)}{.119}\,gwage_t + \underset{(.039)}{.097}\,gwage_{t-1} + \underset{(.039)}{.040}\,gwage_{t-2} \\ &\quad + \underset{(.039)}{.038}\,gwage_{t-3} + \underset{(.039)}{.081}\,gwage_{t-4} + \underset{(.039)}{.107}\,gwage_{t-5} + \underset{(.039)}{.095}\,gwage_{t-6} \\ &\quad + \underset{(.039)}{.104}\,gwage_{t-7} + \underset{(.039)}{.103}\,gwage_{t-8} + \underset{(.039)}{.159}\,gwage_{t-9} + \underset{(.039)}{.110}\,gwage_{t-10} \\ &\quad + \underset{(.039)}{.103}\,gwage_{t-11} + \underset{(.052)}{.016}\,gwage_{t-12}, \\ n &= 273,\ R^2 = .317,\ \bar{R}^2 = .283,\ DW = .99.\end{aligned}$$

The usual standard errors are in parentheses.

(a) **What is the estimated long-run effect (LRP)? Is it very different from one? Explain what the LRP tells us in this example. How does this differ from the short-run effect?**

(4 marks)

(b) **We want to test whether the LRP is significantly smaller than one. Clearly indicating the null and the alternative hypothesis, give the test statistic and the rejection rule. What regression would you run to obtain the standard error of the LRP directly?**

(6 marks)

(c) **Your result in (b) may be affected by the presence of autocorrelation. Briefly explain this and discuss how you could use the Durbin–Watson test to detect whether this is indeed the case. Indicate clearly the assumptions underlying the Durbin–Watson test, the test statistic and the rejection rule.**

(5 marks)

Hint: **For the critical values of the DW test, use the largest sample size and number of regressors available in the table.**

(d) **Assuming the Durbin–Watson test finds evidence of autocorrelation, discuss how you could resolve the problem by means of the iterated Cochrane–Orcutt procedure. For notational purposes, you may simplify your model to an ADL(0, 1).**

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 6.5 (Testing a linear restriction), Chapter 11.3 (Models with lagged explanatory variables), and Chapter 12.1–12.3 (Definition, consequences and detection of autocorrelation; Fitting a model subject to AR(1) autocorrelation).

Subject guide (2016), Chapters 6 and 11.

Approaching the question

In part (a), an interpretation of parameters in distributed lag models (short-run and long-run effects) needs to be given. In part (b), a one-sided t test for the LRP is required and candidates are expected to discuss a reparameterisation of the model that enables one to get the standard error of the LRP parameter. In part (c), the validity of our test in (b) is put into question due to the presence of autocorrelation and a discussion of the Durbin–Watson test is required. In part

(d), a discussion of the Cochrane–Orcutt procedure is required. Many candidates were unable to answer this question. (In view of the fact that the Cochrane–Orcutt procedure is not explicitly mentioned in Appendix A of the subject guide, with the approval of the external examiner, steps were undertaken to ensure candidates were not affected by this in their overall classification.) This question was answered by 25% of the candidates in Zone A. The answer is as follows.

(a) The estimated long-run effect is given by:

$$\widehat{\theta} = \widehat{\beta}_1 + \cdots + \widehat{\beta}_{13} = 1.172$$

where $\widehat{\beta}_i$ is the estimated coefficient on $gwage_{t-i+12}$ (numerical value irrelevant).

The LRP tells us what the long-run effect of *gwage* is on *gprice*. It shows the effect a permanent change in *gwage* with one unit has on *gprice* after all 12 periods (the last lagged response) have passed.

The short-run effect is $\widehat{\beta}_1$ and it tells us what the impact of current *gwage* on current *gprice* is.

(b) We want to test $\mathrm{H}_0 : \theta = 1$ against $\mathrm{H}_1 : \theta < 1$. By rejecting the null hypothesis against this one-sided alternative hypothesis, we may find evidence that the LRP is significantly smaller than 1!

The test statistic we should use is:

$$t = \frac{\widehat{\theta} - 1}{\mathrm{SE}(\widehat{\theta})} \sim t_{n-14} \qquad \text{under } \mathrm{H}_0 \text{ (Gauss–Markov assumptions + normality).}$$

Reject H_0 if $t < -1.645$ at the 5% significance level. (Important: candidates should recognise that $\mathrm{SE}(\widehat{\theta}) \neq \mathrm{SE}(\widehat{\beta}_1) + \cdots + \mathrm{SE}(\widehat{\beta}_{13})$.) If we reject H_0, we find evidence that the LRP is significantly smaller than one.

A candidate should discuss that one can rewrite (reparameterise) the model so that $\mathrm{SE}(\widehat{\theta})$ needed for this test can be obtained directly by our regression output (see Chapter 11.3 – Estimating long-run effects). Specifically:

$$gprice_t = \beta_0 + \theta gwage_t + \beta_2 \left(gwage_{t-1} - gwage_t\right) + \cdots + \beta_{13} \left(gwage_{t-12} - gwage_t\right) + u_t.$$

The model is exactly the same, observe $\theta - \beta_2 - \cdots - \beta_{13} = \beta_1$. Because both models are the same their R^2s are the same – it is only the interpretation of the parameters (here the parameter of $gwage_t$) that is different.

(c) In the presence of autocorrelation, we cannot trust the usual standard errors and, therefore, our test statistics will be invalid and our conclusions may be wrong.

The Durbin–Watson test allows us to test for the absence of autocorrelation against the alternative of an AR(1). The Durbin–Watson test requires the regressors to be deterministic (for example, cannot contain a lagged dependent variable). (An alternative test that does not suffer from this problem is the Breusch–Pagan test.)

Discussion of the Durbin–Watson test:

- $\mathrm{H}_0 : \rho = 0$ against $\mathrm{H}_1 : u_t = \rho u_{t-1} + \varepsilon_t$, for $|\rho| < 1$.
- $DW = \frac{\sum_{t=2}^{T} (\widehat{u}_t - \widehat{u}_{t-1})^2}{\sum_{t=2}^{T} \widehat{u}_t{}^2} \simeq 2(1 - \widehat{\rho})$, where $\widehat{u}_t$ are the OLS residuals from the regression in the question.
- Choose a significance level, α, (willingness to commit a Type I error) and find the critical values d_L and d_U.
- One-sided test: Reject H_0 in favour of positive autocorrelation if $DW < d_L$; test is inconclusive if it lies between d_L and d_U; do not reject H_0 when $DW > d_U$.
- k' is taken to be 5 (number of explanatory variables, here 13) and n is taken to be 100 (number of observations, here 273). Therefore, $d_L = 1.57$ and $d_U = 1.78$ with a 5% significance level. With $DW = 0.99$ we reject H_0 in favour of positive autocorrelation.

(d) Simplify the model to:

$$gprice_t = \beta_0 + \beta_1 gwage_t + \beta_2 gwage_{t-1} + u_t, \qquad t = 2, \ldots, T$$

and note that Durbin–Watson suggests $u_t = \rho u_{t-1} + \varepsilon_t$.

Discussion of the Cochrane–Orcutt procedure (see also Box 12.1 of Dougherty):

- Step 1: Obtain $\widehat{\beta}$ by running the original model by OLS, and compute the residuals $\widehat{u}_t = gprice_t - \widehat{\beta}_0 - \widehat{\beta}_1 gwage_t - \widehat{\beta}_2 gwage_{t-1}$.
- Step 2: Obtain an estimate of ρ using the residuals $\widehat{\rho} = \frac{\sum_{t=2}^{T} \widehat{u}_t \widehat{u}_{t-1}}{\sum_{t=2}^{T} \widehat{u}_t{}^2}$.
- Step 3: Estimate the transformed model:

$$\begin{aligned} & gprice_t - \widehat{\rho} gprice_{t-1} \\ = \quad & \beta_0 (1 - \widehat{\rho}) + \beta_1 (gwage_t - \widehat{\rho} gwage_{t-1}) + \beta_2 (gwage_{t-1} - \widehat{\rho} gwage_{t-2}) + v_t \\ & \text{for } t = 3, \ldots, T \qquad \text{(lose one observation).} \end{aligned}$$

  Use the new parameter estimates of β_0, β_1 and β_2 and recompute the residuals $\widehat{u}_t$.
- Step 4: Return to Step 2, until convergence.

If the sample is small, losing one observation is not a good idea, and it is preferable to use the Prais–Winsten estimator (Chapter 12.3 – Fitting a model subject to AR(1) correlation).

Alternative non-linear regressions to deal with the AR(1) correlation are available as well.

Question 9

Let us consider the demand for fish. Using 97 daily price ($avgprc$) and quantity ($totqty$) observations on fish prices at the Fulton Fish Market in Manhattan, the following results were obtained by OLS:

$$\begin{aligned} \log(\widehat{totqty_t}) \quad = \quad & \underset{(.163)}{8.244} - \underset{(.176)}{.425} \log(avgprc_t) - \underset{(.226)}{.311}\, mon_t - \underset{(.223)}{.683}\, tues_t \\ & - \underset{(.220)}{.533}\, wed_t + \underset{(.220)}{.067}\, thurs_t \end{aligned} \tag{9.1}$$

The equation allows demand to differ across the days of the week, and Friday is the excluded dummy variable. The standard errors are in parentheses.

(a) Interpret the coefficient of log($avgprc$) and discuss whether it is significant.

(3 marks)

(b) It is commonly thought that prices are jointly determined with quantity in equilibrium where demand equals supply. What are the consequences of this simultaneity for the properties of the OLS estimator?

(3 marks)

(c) The variables $wave2_t$ and $wave3_t$ are measures of ocean wave heights over the past several days. In view of your answer in (b), what two assumptions do we need to make in order to use $wave2_t$ and $wave3_t$ as instruments for log($avgprc_t$) in estimating the demand equation? Discuss whether these assumptions are reasonable.

(4 marks)

(d) Below we report two sets of regression results, where the dependent variable is log($avgprc_t$). Are $wave2_t$ and $wave3_t$ jointly significant? State the test statistic and rejection rule. How is your finding related to your answer in (c)?

(4 marks)

Dependent Variable $\log(avgprc_t)$	Regressors: *constant*	*wave2*	*wave3*	*day-of-the-week dummies*	R^2	*RSS*	*n*
Regression (9.2)	-1.022 (.144)	.094 (.021)	.053 (.020)	*yes*	.3041	10.934	97
Regression (9.3)	$-.276$ (.092)	–	–	*yes*	.0088	15.576	97

(e) The following IV results were obtained in Stata:

$$\begin{aligned}\log(\widehat{totqty_t}) &= \underset{(.182)}{8.164} - \underset{(.327)}{.815}\log(avgprc_t) - \underset{(.229)}{.307}\,mon_t - \underset{(.226)}{.685}\,tues_t \\ &\quad - \underset{(.223)}{.521}\,wed_t + \underset{(.225)}{.095}\,thurs_t \qquad (9.4)\end{aligned}$$

Discuss how these results can be obtained using Two Stage Least Squares (2SLS) and briefly discuss how you would test whether the results in (9.1) and (9.4) are significantly different from each other.

(6 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 6.5 (Testing a linear restriction), Chapter 8.3 (Instrumental variables), and Chapter 9 (Simultaneous equations estimation).

Subject guide (2016), Chapter 9.

Approaching the question

In part (a), a parameter interpretation is required (be specific as it is an elasticity here!) and a t test should be used to test for its significance. In part (b), a discussion of simultaneous equation bias is required. The joint determinacy in simultaneous equations induces an endogeneity problem: a problem where we have a correlation between the errors and regressors. This correlation renders the OLS parameter estimates inconsistent. Recognising that the IV estimator avoids this problem, in part (c) the requirements on the instruments need to be discussed as they relate to the particular example. In part (d), an F test needs to be proposed to test for the relevance of the instruments, and in part (e) the 2SLS estimator and the Durbin–Wu–Hausman test need to be discussed. This question was answered by 80% of the candidates in Zone A. The answer is as follows.

(a) We want to interpret the parameter as the *price elasticity of demand* after controlling for days-of-the-week differences. If the price increases by 1%, then quantity demanded decreases by 0.425%.

Use the t statistic $\widehat{\beta}/\mathrm{SE}(\widehat{\beta}) = -0.425/0.176 = -2.415$. It is significant as the 5% significance level critical value of t_{n-6} is 1.96 (Gauss–Markov assumptions + normality assumed).

(b) This is the *endogeneity problem*, whereby $\log(avgprc)$ will be correlated with the error term. Candidates making incorrect claims, for example it is a problem of multicollinearity, are penalised. If we were given the structural form equations for both supply and demand, we could explicitly derive this correlation by deriving the reduced form equation for $\log(avgprc)$. The error term in that reduced form would contain the error term of the demand equation, which induces this correlation.

OLS will result in inconsistent parameter estimates, or simply stated will cause simultaneous equation bias.

(c) To estimate the demand equation, we need at least one exogenous variable that appears in the supply equation (because we have one 'bad' variable) that does not also appear in the demand equation (relevance) and that variable needs to be uncorrelated with the error term in the demand equation (validity). We call such variables *instruments*.

Exclusion and validity: we need to assume that the instruments can properly be excluded from the demand equation and are uncorrelated with the demand error term. This may not be entirely reasonable – wave heights are determined partly by weather and demand at a local fish market could depend on weather.

Relevance: we need to make sure that the instruments are correlated with the 'bad' variable $\log(avgrpc)$. By ensuring that at least one of the instruments ($wave2_t$ and $wave3_t$) appears in the supply equation, we will be able to show that the reduced form for $\log(avgprc_t)$ will have these variables on the right-hand side, hence that there is a correlation between the instruments and the 'bad' regressor. Seems to be quite reasonable, and indeed there is indirect evidence of this in part (d), as the two variables are jointly significant in the reduced form for $\log(avgprc_t)$.

(d) We test the joint hypothesis $\mathrm{H}_0 : \beta_{wave2} = \beta_{wave3} = 0$ against $\mathrm{H}_1 : \beta_{wave2} \neq 0$ and/or $\beta_{wave3} \neq 0$. The F test is obtained as:

$$F = \frac{(RRSS - URSS)/2}{URSS/(97-7)} = \frac{(15.576 - 10.934)/2}{10.934/90} = 19.0.$$

Under the Gauss–Markov assumptions + normality, this gives us an $F_{2,\,90}$ random variable under the null hypothesis. For any reasonable significance level we will want to reject H_0 (at the 5% significance level the critical value is 3.10). An alternative formula that could have been used is:

$$\frac{(R_u^2 - R_r^2)/2}{(1 - R_u^2)/90}.$$

Conclusion: the test result ensures that our instrumental variables are indeed relevant. (Many candidates failed to mention this: answer all parts of the question.)

(e) Following a discussion of 2SLS (given below), a brief, non-technical, discussion of the Durbin–Wu–Hausman test should be provided.

The Durbin–Wu–Hausman test is a chi-squared test where the two sets of parameter estimates (OLS and 2SLS) are compared. Under the null hypothesis (there is no endogeneity problem and both 2SLS and OLS will be consistent) we should expect $\widehat{\beta}_{OLS} - \widehat{\beta}_{2SLS} \approx 0$, whereas under the alternative hypothesis (there is an endogeneity problem and only 2SLS is consistent) the parameter estimates can be quite different. The fact that under the null hypothesis OLS is efficient (the best instrument for something which is 'good' is always itself) makes it easy to work out the precision (variance) of $\widehat{\beta}_{OLS} - \widehat{\beta}_{2SLS}$ required for the test.

The 2SLS approach is as follows.

- Step 1: Requires us to make the 'bad' variable 'good' by regressing it (OLS) on all exogenous variables in our model:

$$\log(avgprc_t) = \delta_0 + \delta_1 mon_t + \cdots + \delta_4 thurs_t + \delta_5 wave2_t + \delta_6 wave3_t + e_t.$$

  We obtain the fitted values of this regression: $\log(\widehat{avgprc_t})$. All exogenous variables have to be included here!

- Step 2: Requires us to estimate the equation:

$$\log(totqty_t) = \beta_0 + \beta_1 \log(\widehat{avgprc_t}) + \beta_2 mon_t + \beta_3 tues_t + \beta_4 wed_t + \beta_5 thurs_t + u_t$$

  by OLS (hence *two* stage LS). The parameter estimates of this regression are our 2SLS estimates. Candidates may alternatively indicate that the second stage is an IV estimator, where $\log(\widehat{avgprc_t})$ (super instrument) is used as an instrument for $\log(avgprc_t)$.

Question 10

Let us consider an analysis on recidivism (probability of re-arrest) among a group of young men in California who have at least one arrest prior to 1986. The dependent variable, $arr86$, is equal to unity if the man was arrested at least once during 1986, and zero otherwise.

	OLS A	OLS B	Logit A	Logit B	Logit B Marginal Effect
pcnv	−.152	−.162	−.880	−.901	−.176
	(.021)	(.021)	(.122)	(.120)	(.023)
avgsen	.005	.006	.027	.031	.006
	(.006)	(.006)	(.035)	(.034)	(.007)
tottime	−.003	−.002	−.014	−.010	−.002
	(.005)	(.005)	(.028)	(.027)	(.005)
ptime86	−.023	−.022	−.140	−.127	−.025
	(.005)	(.005)	(.031)	(.031)	(.006)
qemp86	−.038	−.043	−.199	−.216	−.042
	(.005)	(.005)	(.028)	(.028)	(.005)
black	.170	–	.823	–	–
	(.024)		(.117)		
hispan	.096	–	.522	–	–
	(.021)		(.109)		
constant	.380	.380	−.464	−.169	
	(.019)	(.019)	(.095)	(.084)	
R^2	.068	.047			
$\log L$			−1512.35	−1541.24	

pcnv is the proportion of prior arrests that led to a conviction, *avgsen* is the average sentence served from prior convictions, *tottime* is the months spent in prison since age 18 prior to 1986, *ptime*86 is months spent in prison in 1986, *qemp*86 is the number of quarters the man was legally employed in 1986, while *black* and *hispan* are two race dummies (*white* the excluded dummy). The standard errors are reported in parentheses.

(a) When estimating the parameters by OLS, we are using the Linear Probability Model. Why might you then report heteroskedasticity-robust standard errors?

(2 marks)

(b) Using the OLS B results, what is the estimated effect on the probability of arrest if *pcnv* goes from 0.25 to 0.75 holding everything else constant?

(4 marks)

(c) It is argued that the linear probability model is not appropriate for explaining the binary variable $arr86$ and a logit regression model has been estimated. Explain how the Logit estimates are obtained.

(5 marks)

Hint: You may recall, that for the Logit model A, we will specify

$$\Pr(arr86_i = 1) = \Lambda\left(\beta_0 + \beta_1 pcnv + \beta_2 avgsen + \cdots + \beta_6 black + \beta_7 hispan\right)$$

where $\Lambda(z) = \frac{1}{1+\exp(-z)}$.

(d) Using the Logit model results, discuss whether *black* and *hispan* are jointly significant. Clearly indicate the null and alternative hypothesis, the test statistic and the rejection rule.

(3 marks)

(e) An important distinction between the two approaches is that the marginal effect of *pcnv* on the probability of re-arrest is constant for the LPM unlike the marginal effect using the logit analysis. What this means for instance is that the estimated effect on the probability of arrest if *pcnv* goes from 0.25 to 0.75 will depend on the other characteristics.

i. Explain how the marginal effects evaluated at the mean values of the explanatory variables (reported in the last column) were obtained. Give a brief comment as to how they compare to the marginal effect of the associated LPM.

(3 marks)

ii. Using the Logit B results, how would you obtain the estimated effect on the probability of arrest if *pcnv* goes from 0.25 to 0.75 for a *white* man, with characteristics $avgsen = 1$, $tottime = 1$, $ptime86 = 0$ and $qemp86 = 2$. A clear explanation of what calculations are required is sufficient.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 10.1 (The linear probability model), Chapter 10.2–10.3 (Logit and Probit analysis) and Chapter 10.6 (An introduction to maximum likelihood estimation).

Subject guide (2016), Chapter 10.

Approaching the question

In parts (a) and (b), a discussion of the LPM and an interpretation of its parameters – constant marginal effects – need to be given. In parts (c) and (e), a related discussion needs to be given for the logit model. Unlike in the LPM where we imposed linearity, the marginal effects in the logit model are no longer constant due to the fact that in the logit model we use a non-linear, $\Lambda(z)$, specification of the probabilities. In part (c), a discussion of the maximum likelihood estimator used for our logit parameter estimates is required, and in part (d) a test for joint significance using the likelihood ratio test is needed. This question was answered by 40% of the candidates in Zone A. The answer is as follows.

(a) The linear probability model suffers from heteroskedasticity. Specifically, the heteroskedasticity takes the form $\text{Var}(arr86_i \mid X_i) = p_i\,(1 - p_i)$ with:

$$p_i = \text{E}(arr86_i \mid X_i) = \beta_0 + \beta_1 pcnv_i + \beta_2 avgsen_i + \cdots + \beta_6 black_i + \beta_7 hispan_i.$$

Since this invalidates the standard errors, we will want to report heteroskedasticity-robust standard errors instead.

(b) Candidates should clearly indicate that $\widehat{P}_i = \widehat{\beta}_0 + \widehat{\beta}_1 pcnv_i + \widehat{\beta}_2 avgsen_i + \cdots + \widehat{\beta}_5 qemp86_i$.

The estimated effect is:

$$\widehat{\beta}_1(0.75 - 0.25) = -0.162 \times 0.5 = -0.081.$$

Therefore, the probability of re-arrest decreases by 0.081 or by 8.1 percentage points. (An 8.1% decrease is different from an 8.1 percentage point decrease!)

(c) Candidates are expected to indicate that we use the maximum likelihood estimator (MLE) to estimate the β parameters. Make sure you answer the question, you were asked how the estimates are obtained! Few candidates were able to give relevant details.

The log-likelihood function that will be maximised is given by:

$$\begin{aligned}
\log L(\beta) &= \sum_{i=1}^{n} arr86_i \times \log(\Pr(arr86_i = 1)) + (1 - arr86_i) \times \log(\Pr(arr86_i = 0)) \\
&= \sum_{i=1}^{n} arr86_i \times \log(\Lambda(z_i)) + (1 - arr86_i) \times \log(1 - \Lambda(z_i)) \\
\text{where } z_i &= \beta_0 + \beta_1 pcnv_i + \beta_2 avgsen_i + \cdots + \beta_6 black_i + \beta_7 hispan_i.
\end{aligned}$$

The MLE will set the partial derivatives equal to zero:

$$\left.\frac{\partial \log L(\beta)}{\partial \beta_i}\right|_{\widehat{\beta}_{MLE}} = 0 \qquad \text{for } i = 0, \ldots, 7$$

i.e. it chooses those values of $\widehat{\beta}$ that will set the derivatives equal to zero. Heuristically, these parameter estimates ensure that the predicted $\Pr(arr86_i = 1)$ is large for individuals for whom $arr86_i = 1$, and the predicted $\Pr(arr86_i = 0)$ is large for individuals for whom $arr86_i = 0$. This ensures that the 'likelihood' of observing the data is largest for these values of β.

(d) Candidates should indicate they want to test $\mathrm{H}_0 : \beta_{black} = \beta_{hispan} = 0$ against $\mathrm{H}_1 : \beta_{black} \neq 0$ and/or $\beta_{hispan} \neq 0$.

The likelihood ratio (LR) test statistic equals $-2(\ln L^R - \ln L^U) \overset{a}{\sim} \chi^2_2$ under H_0, where $\ln L^R$ is the log-likelihood function of the restricted model (Logit B) and $\ln L^U$ is the log-likelihood function of the unrestricted model (Logit A). $LR = -2 \times (-1541.24 - -1512.35) = 57.78$.

With a 5% significance level our critical value is given by 5.99, so we clearly reject the null hypothesis, rendering ethnicity an important factor in explaining re-offending rates.

(e) i. The marginal effects of interest describe how $\Pr(arr86_i)$ changes as a result of the explanatory variables.

For continuous variables that means, for example:

$$\begin{aligned} \frac{\partial \Pr(arr86_i)}{\partial pcnv} &= f(z_i) \times \beta_1 \qquad \text{with } f(z) = \frac{\mathrm{d}\Lambda(z)}{\mathrm{d}z} = \frac{\mathrm{e}^{-z}}{(1+\mathrm{e}^{-z})^2} \\ \text{with } z_i &= \beta_0 + \beta_1 pcnv_i + \beta_2 avgsen_i + \cdots + \beta_6 black_i + \beta_7 hispan_i \end{aligned}$$

(no need to give the derivative itself). Therefore, the marginal effects depend on the characteristics of the individual. The marginal effects reported are for an individual with average characteristics:

$$\bar{z} = \beta_0 + \beta_1 \overline{pcnv} + \beta_2 \overline{avgsen} + \cdots + \beta_6 \overline{black} + \beta_7 \overline{hispan}.$$

(Whether this individual exists is another matter. Indeed, it may be more interesting to report the average of these marginal effects over all individuals.)

The marginal effects for the average person are quite comparable to those obtained by the LPM (an easy mark, so do not forget to mention this – answer all parts of the question!).

Candidates making these points clearly in (c) were rewarded here.

ii. Candidates here will need to observe that we need to compare predicted probabilities using the logit specification of the probabilities: $\Lambda(z) = \dfrac{1}{1+\exp(-z)}$.

$\widehat{\Pr}(arr86 = 1 \,|\, pcnv = 0.25, avgsen = 1, tottime = 1, ptime86 = 0, qemp86 = 2) = \Lambda(z_1)$, where:

$$z_1 = 0.25 \times (-0.901) + 1 \times (0.031) + 1 \times (-0.010) + 0 \times (-0.127) + 2 \times (-0.216) = -0.636$$

and $\Lambda(z_1) = \dfrac{1}{1+\mathrm{e}^{0.636}} = 0.346$.

$\widehat{\Pr}(arr86 = 1 \,|\, pcnv = 0.75, avgsen = 1, tottime = 1, ptime86 = 0, qemp86 = 2) = \Lambda(z_2)$, where:

$$z_2 = 0.75 \times (-0.901) + 1 \times (0.031) + 1 \times (-0.010) + 0 \times (-0.127) + 2 \times (-0.216) = -1.087$$

and $\Lambda(z_2) = \dfrac{1}{1+\mathrm{e}^{1.087}} = 0.252$.

So we see a drop in the probability of re-arrest equalling $0.346 - 0.252 = 0.094$.

Candidates are not expected to provide explicit numbers. Full marks can be obtained for clarity of approach. Failure to recognise the importance of Λ is serious.

Examiners' commentaries 2017

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2016–17. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone B

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Section A

Answer all questions from this section.

Question 1

Discuss the consequences of omitting relevant variables.

In your answer consider the following empirical study attempting to estimate the social benefits of an increase in public spending on education. Using data from a cross-section survey of employees, the following regression is estimated:

$$\log W_i = \alpha_1 + \alpha_2 S_i + \varepsilon_i,$$

where W_i is the hourly wage rate and S_i is the number of years of schooling completed of employee i. The coefficient of S_i is found to be positive and strongly significant. What can be concluded from this? In your answer discuss the interpretation of α_2, and discuss whether we are likely to obtain an unbiased (consistent) estimator of it when using OLS. Rigour of your answer will be rewarded.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 6.2 (The effect of omitting a variable that ought to be included).

Subject guide (2016), Chapter 6.

Approaching the question

Candidates should clearly discuss, with technical details (rigour), the consequences of omitting relevant variables. Inappropriate comments, for example relating this to multicollinearity are penalised. It is important to clearly point out that omitting relevant variables only causes serious problems if the omitted variable is correlated with included regressors. Only then will omitting relevant variables induce a correlation between the error (which then contains this omitted variable) and the included regressors (which cause inconsistency). Failure to point this out is serious. Candidates should answer all parts of the question and explain why the model in question is likely to suffer from this problem. The answer is as follows.

The interpretation of α_2 is as the returns to education – it provides the proportional change in wage per unit change in schooling, in other words $100 \times \alpha_2$ provides the percentage change in wage per unit change in schooling. If $\alpha_2 = 0.06$, this represents a 6% return. Many candidates stated α_2%, which is wrong!

It is indeed likely in this case that we have omitted relevant variables, such as ability, which in general generates *omitted variable bias* (OVB) for our incorrectly-specified model when using OLS as ability will be correlated with education.

Derivation of OVB:

Provide estimator: The model we consider here is:

$$\log W_i = \alpha_1 + \alpha_2 S_i + \varepsilon_i$$

and our interest is in the OLS estimator of α_2, which is given by:

$$\widehat{\alpha}_2 = \frac{\sum(S_i - \bar{S})(\log W_i - \overline{\log W})}{\sum(S_i - \bar{S})^2} = \frac{\text{Sample Cov}(S_i, \log W_i)}{\text{Sample Var}(S_i)}.$$

Proof of bias/inconsistency: Plugging in the true model:

$$\log W_i = \alpha_1 + \alpha_2 S_i + \gamma \mathit{Ability}_i + \varepsilon_i$$

we obtain:

$$\widehat{\alpha}_2 = \alpha_2 + \frac{\text{Sample Cov}(S_i, \mathit{Abilty}_i)}{\text{Sample Var}(S_i)}\gamma + \frac{\text{Sample Cov}(S_i, \varepsilon_i)}{\text{Sample Var}(S_i)}.$$

To consider the bias, we take expectations. *Assuming that S_i and Abilty_i are non-stochastic:*

$$\text{E}(\widehat{\alpha}_2) = \alpha_2 + \underbrace{\frac{\text{Sample Cov}(S_i, \mathit{Abilty}_i)}{\text{Sample Var}(S_i)}\gamma}_{\text{Omitted Variable Bias}} \neq \alpha_2$$

that is, $\widehat{\alpha}_2$ is a biased estimator of α_2. This bias does not go away when the sample size tends to infinity $\rightarrow$ the estimator is inconsistent as well.

Candidates may directly establish the inconsistency as well, recognising that most likely S_i and Abilty_i *are stochastic*:

$$\begin{aligned}
\text{plim } \widehat{\alpha}_2 &= \alpha_2 + \frac{\text{plim Sample Cov}(S_i, \mathit{Abilty}_i)}{\text{plim Sample Var}(S_i)}\gamma + \frac{\text{plim Sample Cov}(S_i, \varepsilon_i)}{\text{plim Sample Var}(S_i)} \\
&= \alpha_2 + \frac{\text{Cov}(S_i, \mathit{Abilty}_i)}{\text{Var}(S_i)}\gamma + \frac{\text{Cov}(S_i, \varepsilon_i)}{\text{Var}(S_i)} \\
&= \alpha_2 + \frac{\text{Cov}(S_i, \mathit{Abilty}_i)}{\text{Var}(S_i)}\gamma
\end{aligned}$$

and:

$$\frac{\mathrm{Cov}(S_i, Abilty_i)}{\mathrm{Var}(S_i)}\gamma > 0.$$

Candidates need to clearly indicate (in proof or verbally) that OVB necessitates that our omitted variable is correlated with variables included in our incorrectly-specified regression, and here we expect that to be the case.

(Intuitively, we note that $\widehat{\alpha}_2$ incorporates the indirect effect $Ability_i$ has on $\log W_i$ through its correlation with S_i in addition to the direct effect S_i has on $\log W_i$. That is, $\widehat{\alpha}_2$ will take on some of the explanatory effect of $Ability_i$ due to the correlation between S_i and $Ability_i$.

The t test used for signalling significance of α_2 in fact is also invalid due to the fact that another serious consequence of omitting a relevant variable is that the standard errors of the coefficients and the test statistics in general are invalid.

Note: The severe econometric consequences of omitting relevant variables contrasts sharply with the less severe consequences of including irrelevant variables where unbiasedness remains and the test statistics also remain valid. The efficiency of our estimator is affected though. General result: Imposing valid linear restrictions allows us to obtain unbiased, less variable (more efficient) parameter estimates (lower standard errors) which enhance the power of our tests.)

Question 2

Suppose you are given a random sample $X_1, \ldots, X_n$ from the exponential distribution:

$$f(x) = \lambda \exp(-\lambda x), \qquad x > 0,\ \lambda > 0.$$

According to this distribution $\mathrm{E}(X_i) = 1/\lambda$ and $\mathrm{Var}(X_i) = 1/\lambda^2$ for $i = 1, \ldots, n$.

Show that the maximum likelihood estimator for λ is $1/\bar{X}$ where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. Is the estimator unbiased and/or consistent? Prove your claims.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 4.2 (Logarithmic transformations) and Chapter 10.6 (An introduction to maximum likelihood estimation).

Subject guide (2016), Chapter 10.

Approaching the question

Candidates should clearly conduct the maximum likelihood procedure and ensure proper uses of the product and logarithm operators are displayed. Many candidates made the classic error of not recognising that $\mathrm{E}(1/\bar{X}) \neq 1/\mathrm{E}(\bar{X})$ and also failed to realise that the proof of consistency should have used the plim operator and the law of large numbers as $\mathrm{plim}(1/\bar{X}) = 1/(\mathrm{plim}\ \bar{X})$. Instead, many candidates attempted (unsuccessfully) to look at the sufficient conditions (what is $\mathrm{Var}(1/\bar{X})$?). Finally, candidates should make a clear distinction between the unknown (fixed) parameter λ and its MLE $\widehat{\lambda}$ (a random variable). The answer is as follows.

The likelihood function is given by the joint density of the data (which equals the product of the marginals given the independence of our observations):

$$L(\lambda) = \prod_{i=1}^{n} \lambda \exp(-\lambda X_i) = \lambda^n \exp\left(-\lambda \sum_{i=1}^{n} X_i\right).$$

The MLE is the value of λ which maximises this function, or by monotonicity, the log-likelihood function:

$$\ln L(\lambda) = \ln\left(\lambda^n \exp\left(-\lambda \sum_{i=1}^n X_i\right)\right) = n \ln \lambda - \lambda \sum_{i=1}^n X_i.$$

By the first-order condition, we require:

$$\left.\frac{\partial \ln L}{\partial \lambda}\right|_{\widehat{\lambda}_{MLE}} = \frac{n}{\widehat{\lambda}_{MLE}} - \sum_{i=1}^n X_i = 0 \qquad \text{or} \qquad \widehat{\lambda}_{MLE} = \frac{1}{\bar{X}} \qquad \text{with} \qquad \bar{X} = \frac{1}{n}\sum_{i=1}^n X_i.$$

(It should clearly state $\widehat{\lambda}_{MLE} = 1/\bar{X}$, saying $\lambda = 1/\bar{X}$ is incorrect.)

The MLE is *not unbiased*:

$$\mathrm{E}(\widehat{\lambda}_{MLE}) = \mathrm{E}\left(\frac{1}{\bar{X}}\right) \overset{\text{Jensen's Ineq}}{\neq} \frac{1}{\mathrm{E}(\bar{X})} = \frac{1}{\frac{1}{n}\sum \mathrm{E}\left(X_i\right)} = \lambda.$$

The MLE is *consistent*:

$$\operatorname{plim} \widehat{\lambda}_{MLE} \overset{\text{plim rules}}{=} \frac{1}{\operatorname{plim}(\bar{X})} \overset{\text{LLN}}{=} \frac{1}{\mathrm{E}(X_i)} = \lambda.$$

Question 3

Consider the following non-stationary process:

$$\boldsymbol{y_t = \gamma_0 + \gamma_1 t + u_t, \qquad \text{with } u_t = \rho u_{t-1} + \varepsilon_t}$$

and ε_t i.i.d. $(0, \sigma^2)$. Indicate (with explanation) the source(s) of non-stationarity of y_t. Show that you can rewrite the model as:

$$\boldsymbol{\Delta y_t = \beta_0 + \beta_1 t + \beta_2 y_{t-1} + \varepsilon_t.}$$

Clearly indicate the one-to-one relation between $(\gamma_0, \gamma_1, \rho)$ and $(\beta_0, \beta_1, \beta_2)$. What name do we give the non-stationary process y_t when $\rho = 1$? Briefly indicate how to test whether this is the case.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 12.1 (Definition and consequences of autocorrelation), Chapter 13.1 (Stationarity and non-stationarity), and Chapter 13.4–13.5 (Tests of non-stationarity).

Subject guide (2016), Chapter 13.

Approaching the question

Candidates should show a clear understanding of what non-stationarity means. Here there are two possible sources of non-stationarity: the presence of the deterministic trend, $\gamma_1 \neq 0$, and the presence of a unit root, $\rho = 1$. The dependence structure is given by an AR(1) which may exhibit unit roots giving rise to long dependence. A discussion of the Dickey–Fuller test for the presence of a unit root is required. Answer all parts of the question: many candidates failed to answer how we call such a non-stationary process when $\rho = 1$! The answer is as follows.

A process $\{y_t\}$ is (covariance) stationary if its mean and variance exist and do not depend on time and its covariance is a function of distance in time only (not location). Violation of either of these requirements renders the process non-stationary.

In this case $\{y_t\}$ is non-stationary, because $\mathrm{E}(y_t) = \gamma_0 + \gamma_1 t$ depends on time when $\gamma_1 \neq 0$. If $\rho = 1$, there is a more serious non-stationarity problem as that indicates the presence of a unit root: in that case the variance will depend on time as well, and there will be strong dependence.

$\{u_t\}$ and, therefore, $\{y_t\}$ is an AR(1) process, which is (trend) stationary only if $|\rho| < 1$.

To obtain the testing equation, observe that we need to subtract from:

$$y_t = \gamma_0 + \gamma_1 t + u_t$$

the equation:

$$\rho y_{t-1} = \rho\gamma_0 + \rho\gamma_1 (t-1) + \rho u_{t-1}$$

to yield:

$$y_t - \rho y_{t-1} = \gamma_0 (1-\rho) + \rho\gamma_1 + \gamma_1 (1-\rho) t + \underbrace{u_t - \rho u_{t-1}}_{\varepsilon_t}.$$

By rewriting we get:

$$\Delta y_t = \underbrace{\gamma_0 (1-\rho) + \rho\gamma_1}_{\beta_0} + \underbrace{\gamma_1 (1-\rho)}_{\beta_1} t + \underbrace{(\rho - 1)}_{\beta_2} y_{t-1} + \varepsilon_t.$$

A *common mistake* is not realising that when lagging the original equation by one period that t becomes $t-1$.

If $\rho = 1$ then we call the process *difference stationary*, or integrated of order 1, $I(1)$. Observe, we would be left with:

$$\Delta y_t = \gamma_1 + \varepsilon_t \qquad \text{or} \qquad y_t = \gamma_1 + y_{t-1} + \varepsilon_t$$

i.e. a random walk with drift. Realise that if $\rho = 1$ we get $\beta_2 = 0$ and $\beta_1 = 0$.

If $\rho < 1$ while $\gamma_1 \neq 0$ then we call the process *trend stationary*. Observe, we would be left with:

$$y_t = \beta_0 + \beta_1 t + \rho y_{t-1} + \varepsilon_t.$$

Here y_t has a deterministic trend with short-run fluctuations represented by a stationary AR(1).

We can use a Dickey–Fuller test to test $\mathrm{H}_0 : \rho = 1$ (difference stationary) against $\mathrm{H}_1 : \rho < 1$ (trend stationary). Using the test statistic $\widehat{\beta}_2/\mathrm{SE}(\widehat{\beta}_2)$ and the Dickey–Fuller critical values (with trend and constant), we reject H_0 if $\widehat{\beta}_2/\mathrm{SE}(\widehat{\beta}_2) < \tau$.

Question 4

Explain the RESET test as a general test for functional form misspecification and discuss the drawbacks and advantages of this test.

In your answer consider the following multiple linear regression model:

$$\boldsymbol{y_i = \gamma_1 + \gamma_2 x_{2i} + \gamma_3 x_{3i} + u_i \qquad i = 1, \ldots, n,}$$

where x_{2i} and x_{3i} are exogenous variables known to affect $\mathrm{E}(y_i)$.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 4.3 (Models with quadratic and interactive variables).

Subject guide (2016), Chapter 4.

Approaching the question

Candidates should show a clear understanding of what the RESET test is and what the drawbacks and advantages of the test are. Many candidates were unable to answer this question. (In view of the fact that the RESET test is not explicitly mentioned in Appendix A of the subject guide, with the approval of the external examiner, steps were undertaken to ensure candidates were not affected by this in their overall classification.)

To perform the RESET test, we first perform OLS on our model and obtain the fitted values:

$$\widehat{y}_i = \widehat{\gamma}_1 + \widehat{\gamma}_2 x_{2i} + \widehat{\gamma}_3 x_{3i} \qquad \text{for } i = 1, \ldots, n.$$

Next we obtain $\widehat{y}_i^2 = (\widehat{\gamma}_1 + \widehat{\gamma}_2 x_{2i} + \widehat{\gamma}_3 x_{3i})^2$. This variable will pick up quadratic and interactive non-linearities, if present, without necessarily being highly correlated with any of the x variables and consuming only 1 degrees of freedom.

Next, we perform the following regression:

$$y_i = \gamma_1 + \gamma_2 x_{2i} + \gamma_3 x_{3i} + \gamma_4 \widehat{y}_i^2 + e_i \qquad \text{for } i = 1, \ldots, n$$

and test for the significance of γ_4 with $\mathrm{H}_0 : \gamma_4 = 0$ against $\mathrm{H}_1 : \gamma_4 \neq 0$. For this we use the t test where $\widehat{\gamma}_4/\mathrm{SE}(\widehat{\gamma}_4) \sim t_{n-4}$ under H_0. If we reject H_0 then we find evidence of some type of non-linearity. (The RESET test may include $\widehat{y}_i^3$ as well, in which case it becomes an F test.)

Drawbacks: The test does not indicate the actual form of the non-linearity and it may fail to detect other types of non-linearity.

Advantages: Easy to implement, without incurring a serious loss in degrees of freedom or inducing problems of near multicollinearity.

Question 5

Consider the following regression model:

$$\boldsymbol{Y_i = \beta_0 + \beta_1 \frac{1}{X_i} + u_i, \qquad i = 1, \ldots, n.}$$

We assume that the errors $\{u_i\}_{i=1}^n$ are independent normal random variables with zero mean and variance σ^2/X_i^2. The regressor, $1/X_i$, is non-stochastic with positive sample variability. You are interested in testing the hypothesis $\mathrm{H}_0 : \beta_0 = 0$ against $\mathrm{H}_1 : \beta_0 \neq 0$. You are advised to use the BLUE estimator of β_0 for this purpose.

Discuss how you would obtain the BLUE estimator of β_0 (note, you are not asked to derive this estimator).

Give two reasons why you would prefer using the BLUE estimator for β_0 instead of the OLS estimator $\widehat{\beta}_{0,OLS} = \bar{Y} - \widehat{\beta}_{1,OLS}\overline{(1/X)}$ when testing this hypothesis, where $\bar{Y} = \frac{1}{n}\sum Y_i$ and $\overline{(1/X)} = \frac{1}{n}\sum \frac{1}{X_i}$.

(8 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): R.10 (Type II error and the power of a test), Chapter 2.5 (The Gauss–Markov theorem), Chapter 2.6 (Testing hypotheses relating to the regression coefficients), Chapter 7.1 (Heteroskedasticity and its implications) and 7.3 (Remedies for heteroskedasticity).

Subject guide (2016), Chapter 7.

Approaching the question

Candidates should show a clear understanding of what it means for an estimator to be BLUE. Failure to recognise the presence of heteroskedasticity in this question is serious. Arguments that OLS on the regression itself would not be linear because of the form of the regressor $1/X_i$ are wrong. OLS on the regression model given is linear (it is linear in the parameters!), but due to the heteroskedasticity it is not efficient (BLUE)! Many candidates lost points simply because they did not discuss two reasons why, *for testing purposes,* one would prefer to use this BLUE estimator. Answer all parts of the question. The answer is as follows.

Since the only Gauss–Markov violation here is the presence of heteroskedasticity, we should propose to use weighted least squares (WLS) to make the problem go away. Regress:

$$Y_i X_i = \beta_0 X_i + \beta_1 + u_i X_i \qquad \text{for } i = 1, \ldots, n$$

or:

$$Y_i^* = \beta_1 + \beta_0 X_i + u_i^*.$$

This transformed model satisfies all the Gauss–Markov assumptions, $\mathrm{E}(u_i^*) = \mathrm{E}(u_i)\, X_i = 0$ and $\mathrm{Var}(u_i^*) = \mathrm{Var}(u_i X_i) = \mathrm{Var}(u_i)\, X_i^2 = \sigma^2$. Do make sure you mention that the *Gauss–Markov theorem* ensures that OLS on this regression will yield our BLUEs of β_0 and β_1.

Benefits to using the BLUE of β_0 are (i) it is an efficient (more precise) estimator and because of this efficiency the test will have higher power (easier to reject the null hypothesis when it is false), and (ii) using the BLUE estimator allows us to directly use its standard error for inference. Had we used the OLS estimator instead, we would have needed to obtain heteroskedasticity-robust standard errors.

Section B

Answer three questions from this section.

Question 6

Let us consider a model for the sale price of Monet paintings. The data we have contains the sale prices, widths, and heights of 430 Monet paintings, which sold at auction for prices ranging from \$10,000 to \$33 million. A linear regression provided the following results:

$$\widehat{\ln \mathbf{Price}}_i = \underset{(0.612)}{-8.427} + \underset{(0.091)}{1.334} \ln \mathbf{Area}_i - \underset{(0.128)}{0.165} \mathbf{Aspect\ Ratio}_i$$

$$N = 430, R^2 = 0.336$$

where Area = Width × Height and Aspect Ratio = Height/Width. The standard errors are given in parentheses.

(a) **Test the joint significance of the regression. Discuss its relation to the goodness of fit measure: R^2.**

(5 marks)

(b) **You want to test the hypothesis that auction prices are inelastic with respect to area. Specifically, you are asked to test $\mathrm{H}_0 : \beta_{\ln \mathrm{Area}} = 1$ against $\mathrm{H}_1 : \beta_{\ln \mathrm{Area}} > 1$. Perform this test, clearly indicating the test statistic and the rejection rule. In view of your answer, indicate whether the p-value of the test is bigger or smaller than 5%.**

(5 marks)

(c) **A friend points out that you should be worried about the presence of heteroskedasticity. Explain the concept of heteroskedasticity and discuss the properties of the OLS estimator in the presence of heteroskedasticity.**

(5 marks)

(d) Discuss how you should modify the test in (b) if it is known that the variance of the error term is given by:

$$\boldsymbol{\sigma_i^2 = \exp(\gamma_0 + \gamma_1 \ln \text{Area}_i + \gamma_2 \text{Aspect Ratio}_i)},$$

where $(\gamma_0, \gamma_1, \gamma_2)$ are unknown parameters. You are told to use the FGLS estimator. Explain this estimator clearly.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 2.6 (Testing hypothesis relating to the regression coefficients), Chapter 3.5 (Goodness of fit: R^2), Chapter 7.1 (Heteroskedasticity and its implications) and Chapter 7.3 (Remedies for heteroskedasticity).

Subject guide (2016), Chapter 7.

Approaching the question

In part (a), an F test should be used to test the significance of the regression and its relation with the goodness of fit measure, R^2, should be discussed. In part (b), a one-sided t test should be used. In part (c), the concept of heteroskedasticity and the consequences on the OLS estimator need to be discussed. In part (d), we are told that there is evidence of heteroskedasticity. Failure to recognise this is serious. Candidates are asked to discuss the (feasible) weighted least squares estimator. Unfortunately, many candidates were confused because of the request to use the FGLS estimator, which is not discussed in the textbook or subject guide. The issue was raised with the external examiner, and steps were undertaken to ensure candidates were not affected by this in their overall classification. This question was answered by 85% of the candidates in Zone B. The answer is as follows.

(a) We need to test $\text{H}_0 : \beta_{\ln area} = \beta_{aspect} = 0$ against H_1: $\beta_{\ln area} \neq 0$ and/or $\beta_{aspect} \neq 0$, assuming all Gauss–Markov assumptions + normality, we use the F test:

$$\frac{(RRSS - URSS)/2}{URSS/(430-3)} = \frac{R^2}{1-R^2} \times \frac{427}{2} = \frac{0.336}{1-0.336} \times \frac{427}{2} = 108.04 \sim F_{2,\,427}$$

under H_0. At the 5% significance level the critical value is 3.00, so we strongly reject finding evidence that the regressors are important.

If the goodness of fit $R^2 = \dfrac{ESS}{TSS} = 1 - \dfrac{RSS}{TSS}$ is large, then our explanatory variables are important (help to explain the variation in the dependent variable), hence we should reject the null hypothesis that neither variable is important.

(b) Given the Gauss–Markov assumptions + normality, we will use the t statistic, here:

$$\frac{\widehat{\beta}_{\ln area} - 1}{\text{SE}(\widehat{\beta}_{\ln area})} = \frac{1.334 - 1}{0.091} = 3.6703$$

which is t_{427} under H_0. At the 5% significance level we need to reject H_0 when it exceeds 1.645 (one-sided test), which it does. Therefore, we find evidence that auction prices are elastic.

The p-value is the lowest significance level at which we want to reject H_0. Since we reject H_0 at the 5% significance level, the p-value will be smaller than 5%.

(c) Discussion of the concept is standard bookwork, see Chapter 7.1 of Dougherty.

The properties of OLS are *inefficiency* (as heteroskedasticity is a violation of the Gauss–Markov assumptions), *unbiasedness* (heteroskedasticity does not violate the assumption that $\text{E}(u \mid X) = 0$ needed for unbiasedness) and *consistency* (heteroskedasticity does not violate the assumption of contemporaneous uncorrelatedness needed for consistency). Moreover, the *usual standard errors of the OLS estimator will be invalid* (we will need to use White's heteroskedasticity-robust standard errors).

(d) Candidates should recognise that we are now told that there is evidence of *heteroskedasticity*, $\sigma_i^2 \neq \sigma^2$, which means that we should use weighted least squares instead for efficiency purposes. That is, we should perform OLS on:

$$\frac{\ln \text{Price}_i}{\sigma_i} = \beta_0 \frac{1}{\sigma_i} + \beta_1 \frac{\ln \text{Area}_i}{\sigma_i} + \beta_2 \frac{\text{Aspect Ratio}_i}{\sigma_i} + e_i.$$

This will give us the BLUE of our parameters, where:

$$\sigma_i = \exp(0.5(\gamma_0 + \gamma_1 \ln \text{Area}_i + \gamma_2 \text{Aspect Ratio}_i)).$$

A problem with its implementation is that we cannot run the above regression if we *do not know* γ_0, γ_1 and γ_2. This is why we want to propose FGLS (or *feasible weighted least squares*) – not in the syllabus. Once (consistent) estimates of γ_0, γ_1 and γ_2 are obtained, we can use them to estimate the weights:

$$\widehat{\sigma}_i^2 = \exp(\widehat{\gamma}_0 + \widehat{\gamma}_1 \ln \text{Area}_i + \widehat{\gamma}_2 \text{Aspect Ratio}_i)$$

and perform OLS on:

$$\frac{\ln \text{Price}_i}{\widehat{\sigma}_i} = \beta_0 \frac{1}{\widehat{\sigma}_i} + \beta_1 \frac{\ln \text{Area}_i}{\widehat{\sigma}_i} + \beta_2 \frac{\text{Aspect Ratio}_i}{\widehat{\sigma}_i} + \text{error}_i.$$

These estimates are the FGLS estimates which, if the sample size is large, are efficient.

Candidates were asked to indicate that these estimates, and their associated standard errors, should be used instead when we apply the test in (b). Alternatively, the robust standard errors should have been used in (b).

To estimate the γ parameters, the following regression can be run:

$$\ln \widehat{\varepsilon}_i^2 = \gamma_0 + \gamma_1 \ln \text{Area}_i + \gamma_2 \text{Aspect Ratio}_i + v_i$$

where $\widehat{\varepsilon}_i$ are the OLS residuals! Note: if there is heteroskedasticity, the OLS residuals should display this problem. As $\ln \sigma_i^2 = \gamma_0 + \gamma_1 \ln\text{Area}_i + \gamma_2 \text{Aspect Ratio}_i$, the above equation enables us to get these estimates needed to make our WLS feasible.

Question 7

In this question we look at a large data set on weekly hours worked by women having at least two children. Consider the following specification for labour supply, estimated by OLS

$$\begin{aligned} \widehat{hours} &= \underset{(6.503)}{-17.347} - \underset{(.116)}{2.242 kids} + \underset{(.035)}{.938\, educ} + \underset{(.453)}{2.089 age} \underset{(.008)}{-.028\, age^2} - \underset{(.005)}{.075\, nonmomi} \\ n &= 31,857 \text{ and } R^2 = .052 \end{aligned} \tag{7.1}$$

where *kids* is the total number of children, *educ* is the years of schooling, *age* is the woman's age in years, *nonmomi* is income from sources other than the mother's wage income, and *hours* is the hours worked per week. The standard errors are given in parentheses.

(a) Interpret the coefficient of *kids* and discuss its statistical significance.

(3 marks)

(b) It is commonly thought that the decision to have more children is correlated with unobserved factors that affect labour supply. What name do we give such a problem in econometrics, and what consequences will this have for the properties of the OLS estimator? Support your answer using a simple model.

(5 marks)

(c) Consider the variables *multi2nd* and *samesex*, which are binary variables indicating whether the second birth was for multiple babies and whether the first two children are of the same gender. What properties do we need to assume for us to be able to use *multi2nd* and *samesex* as instruments for *kids*? Are these assumptions reasonable and can we test (if so how) these requirements?

(6 marks)

You may assume that *educ*, *age* and *nonmomi* can be treated as being exogenous.

(d) The following IV results were obtained in Stata:

$$\begin{aligned} \widehat{hours} &= \underset{(7.020)}{-16.828} - \underset{(.134)}{2.504}kids + \underset{(.115)}{.916}\,educ + \underset{(.460)}{2.105}age \underset{(.008)}{-.028}\,age^2 - \underset{(.007)}{.076}\,nonmomi \\ n &= 31,857 \text{ and } R^2 = .052 \end{aligned} \tag{7.2}$$

Discuss how these results can be obtained using Two Stage Least Squares (2SLS) and briefly discuss how you would test whether the results in (7.1) and (7.2) are significantly different from each other.

(6 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 6.5 (Testing a linear restriction), Chapter 8.3 (Instrumental variables), and Chapter 9 (Simultaneous equations estimation).

Subject guide (2016), Chapter 9.

Approaching the question

In part (a), a parameter interpretation is required and a t test should be used to test for its significance. In part (b), a discussion of simultaneous equation bias is required. The joint determinacy in simultaneous equations induces an endogeneity problem: a problem where we have a correlation between the errors and regressors. This correlation renders the OLS parameter estimates inconsistent. Technical details need to be provided. Recognising that the IV estimator avoids this problem, in part (c) the requirements on the instruments need to be discussed as it relates to the particular example. In part (d), the 2SLS estimator and the Durbin–Wu–Hausman test need to be discussed. This question was answered by 80% of the candidates in Zone B. The answer is as follows.

(a) For the population of women who have at least two children, having an additional child lowers expected hours worked by about 2.2 hours per week on average, holding other factors fixed.

The resulting t statistic for the hypothesis $\mathrm{H}_0 : \beta_{kids} = 0$ against $\mathrm{H}_1 : \beta_{kids} \neq 0$ is $\widehat{\beta}_{kids}/\mathrm{SE}(\widehat{\beta}_{kids})$, which is very large in magnitude (thanks to the large sample size), and the parameter is highly significant.

(b) This is the problem called *endogeneity*.

The consequences for our OLS estimator are severe, as we will not be able to establish consistency (neither unbiasedness for that matter).

Candidates are expected to prove in a simple model that correlation between the error and regressors yields inconsistency. Need to make use of plim rules and the law of large numbers.

(c) Instruments need to satisfy the requirements of validity (they need to be uncorrelated with the error term ($\mathrm{E}(samesex_i u_i) = \mathrm{E}(multi2nd_i u_i) = 0$)) – typically not easy to test (can consider an overidentification test, but do not discuss). Reasonable: Having a multiple birth with the second pregnancy, and having the first two children of the same gender, are random events in the sense that they cannot be influenced (or, at least currently, are not). This

suggests that the dummy variables *samesex* and *multi2nd* might be exogenous to the labour supply decision.

Instruments also need to be relevant – related to *kids*. The latter can be tested easily by looking at the joint significance of *multi2nd* and *samesex* in the (reduced form) equation:

$$kids_i = \delta_0 + \delta_1 educ_i + \delta_2 age_i + \delta_3 age_i^3 + \delta_4 nonmom_i + \delta_5 multi2nd_i + \delta_6 samesex_i + e_i$$

(that is testing $\mathrm{H}_0 : \delta_5 = \delta_6 = 0$) for which you can use the F test. Reasonable: Clearly having multiple births for the second pregnancy will increase the number of children and the interest of having children of different genders might indicate that when the first two births have the same gender that the parents might have tried again (i.e. increase total number of children).

(d) Discuss the 2SLS approach:

- Step 1: Requires us to make the 'bad' variable 'good' by regressing it on all exogenous variables in our model:

$$kids_i = \delta_0 + \delta_1 educ_i + \delta_2 age_i + \delta_3 age_i^3 + \delta_4 nonmom_i + \delta_5 multi2nd_i + \delta_6 samesex_i + e_i.$$

  We obtain the fitted values of this regression: $\widehat{kids}$. All exogenous variables have to be included here!

- Step 2: Requires us to estimate the equation:

$$hours_i = \beta_0 + \beta_1 \widehat{kids_i} + \beta_2 educ_i + \beta_3 age_i + \beta_4 age_i^2 + \beta_5 nonmom_i + u_i.$$

  The parameter estimates of this regression are our 2SLS estimates.

Candidates are expected to give a (not too technical) discussion of the Durbin–Wu–Hausman test. Requires to compare the IV and OLS estimates: we want to reject the null hypothesis of absence of endogeneity if they are quite different (relative to their precision). Since both estimators are consistent under the null hypothesis, both estimates should be similar if there is no endogeneity. Explicit form of the test need not be given for full marks.

Question 8

Let us consider an analysis on recidivism (probability of re-arrest) among a group of young men in California who have at least one arrest prior to 1986. The dependent variable, *arr*86, is equal to unity if the man was arrested at least once during 1986, and zero otherwise.

	OLS A	OLS B	Logit A	Logit B	Logit B Marginal Effect
pcnv	−.152	−.162	−.880	−.901	−.176
	(.021)	(.021)	(.122)	(.120)	(.023)
avgsen	.005	.006	.027	.031	.006
	(.006)	(.006)	(.035)	(.034)	(.007)
tottime	−.003	−.002	−.014	−.010	−.002
	(.005)	(.005)	(.028)	(.027)	(.005)
ptime86	−.023	−.022	−.140	−.127	−.025
	(.005)	(.005)	(.031)	(.031)	(.006)
qemp86	−.038	−.043	−.199	−.216	−.042
	(.005)	(.005)	(.028)	(.028)	(.005)
black	.170	–	.823	–	–
	(.024)		(.117)		
hispan	.096	–	.522	–	–
	(.021)		(.109)		
constant	.380	.380	−.464	−.169	
	(.019)	(.019)	(.095)	(.084)	
R^2	.068	.047			
log L			−1512.35	−1541.24	

pcnv **is the proportion of prior arrests that led to a conviction,** ***avgsen*** **is the average sentence served from prior convictions,** ***tottime*** **is the months spent in prison since age 18 prior to 1986,** ***ptime*****86 is months spent in prison in 1986,** ***qemp*****86 is the number of quarters the man was legally employed in 1986, while** ***black*** **and** ***hispan*** **are two race dummies (*****white*** **the excluded dummy). The standard errors are reported in parentheses.**

(a) **When estimating the parameters by OLS, we are using the Linear Probability Model. Why might you then report heteroskedasticity-robust standard errors?**

(2 marks)

(b) **Using the OLS B results, what is the estimated effect on the probability of arrest if** ***pcnv*** **goes from 0.25 to 0.75 holding everything else constant?**

(4 marks)

(c) **It is argued that the linear probability model is not appropriate for explaining the binary variable** ***arr*****86 and a logit regression model has been estimated. Explain how the Logit estimates are obtained.**

(5 marks)

Hint: **You may recall, that for the Logit model A, we will specify**

$$\Pr(arr86_i = 1) = \Lambda\left(\beta_0 + \beta_1 pcnv + \beta_2 avgsen + \cdots + \beta_6 black + \beta_7 hispan\right)$$

where $\Lambda(z) = \frac{1}{1+\exp(-z)}$.

(d) **Using the Logit model results, discuss whether** ***black*** **and** ***hispan*** **are jointly significant. Clearly indicate the null and alternative hypothesis, the test statistic and the rejection rule.**

(3 marks)

(e) **An important distinction between the two approaches is that the marginal effect of** ***pcnv*** **on the probability of re-arrest is constant for the LPM unlike the marginal effect using the logit analysis. What this means for instance is that the estimated effect on the probability of arrest if** ***pcnv*** **goes from 0.25 to 0.75 will depend on the other characteristics.**

i. **Explain how the marginal effects evaluated at the mean values of the explanatory variables (reported in the last column) were obtained. Give a brief comment as to how they compare to the marginal effect of the associated LPM.**

(3 marks)

ii. **Using the Logit B results, how would you obtain the estimated effect on the probability of arrest if** ***pcnv*** **goes from 0.25 to 0.75 for a** ***white*** **man, with characteristics** $avgsen = 1$, $tottime = 1$, $ptime86 = 0$ **and** $qemp86 = 2$. **A clear explanation of what calculations are required is sufficient.**

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 10.1 (The linear probability model), Chapter 10.2–10.3 (Logit and Probit analysis) and Chapter 10.6 (An introduction to maximum likelihood estimation).

Subject guide (2016), Chapter 10.

Approaching the question

In parts (a) and (b), a discussion of the LPM and an interpretation of its parameters – constant marginal effects – need to be given. In parts (c) and (e), a related discussion needs to be given

for the logit model. Unlike in the LPM where we imposed linearity, the marginal effects in the logit model are no longer constant due to the fact that in the logit model we use a non-linear, $\Lambda(z)$, specification of the probabilities. In part (c), a discussion of the maximum likelihood estimator used for our logit parameter estimates is required, and in part (d) a test for joint significance using the likelihood ratio test is needed. This question was answered by 39% of the candidates in Zone B. The answer is as follows.

(a) The linear probability model suffers from heteroskedasticity. Specifically, the heteroskedasticity takes the form $\mathrm{Var}(arr86_i \mid X_i) = p_i\,(1-p_i)$ with:

$$p_i = \mathrm{E}(arr86_i \mid X_i) = \beta_0 + \beta_1 pcnv_i + \beta_2 avgsen_i + \cdots + \beta_6 black_i + \beta_7 hispan_i.$$

Since this invalidates the standard errors, we will want to report heteroskedasticity-robust standard errors instead.

(b) Candidates should clearly indicate that $\widehat{P}_i = \widehat{\beta}_0 + \widehat{\beta}_1 pcnv_i + \widehat{\beta}_2 avgsen_i + \cdots + \widehat{\beta}_5 qemp86_i$.

The estimated effect is:

$$\widehat{\beta}_1(0.75 - 0.25) = -0.162 \times 0.5 = -0.081.$$

Therefore, the probability of re-arrest decreases by 0.081 or by 8.1 percentage points. (An 8.1% decrease is different from an 8.1 percentage point decrease!)

(c) Candidates are expected to indicate that we use the maximum likelihood estimator (MLE) to estimate the β parameters. Make sure you answer the question, you were asked how the estimates are obtained! Few candidates were able to give relevant details.

The log-likelihood function that will be maximised is given by:

$$\begin{aligned}
\log L(\beta) &= \sum_{i=1}^{n} arr86_i \times \log(\Pr(arr86_i = 1)) + (1 - arr86_i) \times \log(\Pr(arr86_i = 0)) \\
&= \sum_{i=1}^{n} arr86_i \times \log(\Lambda(z_i)) + (1 - arr86_i) \times \log(1 - \Lambda(z_i)) \\
\text{where } z_i &= \beta_0 + \beta_1 pcnv_i + \beta_2 avgsen_i + \cdots + \beta_6 black_i + \beta_7 hispan_i.
\end{aligned}$$

The MLE will set the partial derivatives equal to zero:

$$\left.\frac{\partial \log L(\beta)}{\partial \beta_i}\right|_{\widehat{\beta}_{MLE}} = 0 \qquad \text{for } i = 0, \ldots, 7$$

i.e. it chooses those values of $\widehat{\beta}$ that will set the derivatives equal to zero. Heuristically, these parameter estimates ensure that the predicted $\Pr(arr86_i = 1)$ is large for individuals for whom $arr86_i = 1$, and the predicted $\Pr(arr86_i = 0)$ is large for individuals for whom $arr86_i = 0$. This ensures that the 'likelihood' of observing the data is largest for these values of β.

(d) Candidates should indicate they want to test $\mathrm{H}_0 : \beta_{black} = \beta_{hispan} = 0$ against $\mathrm{H}_1 : \beta_{black} \neq 0$ and/or $\beta_{hispan} \neq 0$.

The likelihood ratio (LR) test statistic equals $-2(\ln L^R - \ln L^U) \overset{a}{\sim} \chi^2_2$ under H_0, where $\ln L^R$ is the log-likelihood function of the restricted model (Logit B) and $\ln L^U$ is the log-likelihood function of the unrestricted model (Logit A). $LR = -2 \times (-1541.24 - -1512.35) = 57.78$.

With a 5% significance level our critical value is given by 5.99, so we clearly reject the null hypothesis, rendering ethnicity an important factor in explaining re-offending rates.

(e) i. The marginal effects of interest describe how $\Pr(arr86_i)$ changes as a result of the explanatory variables.

For continuous variables that means, for example:

$$\begin{aligned}
\frac{\partial \Pr(arr86_i)}{\partial pcnv} &= f(z_i) \times \beta_1 \qquad \text{with } f(z) = \frac{\mathrm{d}\Lambda(z)}{\mathrm{d}z} = \frac{\mathrm{e}^{-z}}{(1+\mathrm{e}^{-z})^2} \\
\text{with } z_i &= \beta_0 + \beta_1 pcnv_i + \beta_2 avgsen_i + \cdots + \beta_6 black_i + \beta_7 hispan_i
\end{aligned}$$

(no need to give the derivative itself). Therefore, the marginal effects depend on the characteristics of the individual. The marginal effects reported are for an individual with average characteristics:

$$\bar{z} = \beta_0 + \beta_1 \overline{pcnv} + \beta_2 \overline{avgsen} + \cdots + \beta_6 \overline{black} + \beta_7 \overline{hispan}.$$

(Whether this individual exists is another matter. Indeed, it may be more interesting to report the average of these marginal effects over all individuals.)

The marginal effects for the average person are quite comparable to those obtained by the LPM (an easy mark, so do not forget to mention this – answer all parts of the question!).

Candidates making these points clearly in (c) were rewarded here.

ii. Candidates here will need to observe that we need to compare predicted probabilities using the logit specification of the probabilities: $\Lambda(z) = \dfrac{1}{1+\exp(-z)}$.

$\widehat{\Pr}(arr86 = 1 \,|\, pcnv = 0.25, avgsen = 1, tottime = 1, ptime86 = 0, qemp86 = 2) = \Lambda(z_1)$, where:

$$z_1 = 0.25 \times (-0.901) + 1 \times (0.031) + 1 \times (-0.010) + 0 \times (-0.127) + 2 \times (-0.216) = -0.636$$

and $\Lambda(z_1) = \dfrac{1}{1+\mathrm{e}^{0.636}} = 0.346$.

$\widehat{\Pr}(arr86 = 1 \,|\, pcnv = 0.75, avgsen = 1, tottime = 1, ptime86 = 0, qemp86 = 2) = \Lambda(z_2)$, where:

$$z_2 = 0.75 \times (-0.901) + 1 \times (0.031) + 1 \times (-0.010) + 0 \times (-0.127) + 2 \times (-0.216) = -1.087$$

and $\Lambda(z_2) = \dfrac{1}{1+\mathrm{e}^{1.087}} = 0.252$.

So we see a drop in the probability of re-arrest equalling $0.346 - 0.252 = 0.094$.

Candidates are not expected to provide explicit numbers. Full marks can be obtained for clarity of approach. Failure to recognise the importance of Λ is serious.

Question 9

Let us consider a distributed lag model:

$$y_t = \beta_0 + \delta_0 z_t + \delta_1 z_{t-1} + \cdots + \delta_q z_{t-q} + \beta_1 x_t + u_t\,, \qquad t = 1, \ldots, T$$

where u_t is independent of $x_t, z_t, z_{t-1}, \ldots$, and z_{t-q} with zero mean and constant variance. For simplicity, we will argue that there is no autocorrelation in the errors.

1. **Explain the concept of autocorrelation and indicate for the above model what the consequence of autocorrelation for our OLS estimator would be.**

(4 marks)

Let us consider the example of the effects of tax policy on the U.S. fertility rates. Let gfr denote the number of children born per 1,000 women aged 15–44, pe denotes the real value of the personal tax exemption, and $ww2$ and $pill$ are dummy variables (WW II, availability of the birth control pill). Using annual data, the following OLS results were obtained:

$$\begin{aligned}\widehat{gfr}_t \;=\; & \underset{(3.30)}{92.50} + \underset{(.126)}{.089}\,pe_t - \underset{(.153)}{.004}\,pe_{t-1} + \underset{(.165)}{.007}\,pe_{t-2} + \underset{(.154)}{.018}\,pe_{t-3} + \underset{(.105)}{.014}\,pe_{t-4} \\ & - \underset{(11.03)}{21.48}\,ww2_t - \underset{(3.94)}{31.25}\,pill_t, \qquad R^2 = .537, T = 68.\end{aligned} \tag{9.1}$$

The usual standard errors are reported in parentheses.

(b) A novice in econometrics argues that the result indicates that the effect of tax policy on the US fertility rate is ineffective because of the insignificance of the distributed lag coefficients (the coefficients on pe_t and its lags). Explain why he/she could be wrong. In your answer you may discuss what the short and long run effect of the tax policy are according to these results.

(5 marks)

(c) Your friend argues that only two lags should have been included. Using the same sample, he obtains the following result:

$$\begin{aligned}\widehat{gfr}_t &= \underset{(3.27)}{92.52} + \underset{(.119)}{.101}pe_t - \underset{(.147)}{.011}pe_{t-1} + \underset{(.119)}{.033}pe_{t-2} \\ &\quad - \underset{(10.13)}{22.95}ww2_t - \underset{(3.76)}{30.83}pill_t, \qquad R^2 = .536,\ T = 68.\end{aligned} \tag{9.2}$$

Test for the joint significance of the third and fourth lag. Clearly indicate H_0 and H_1, the test statistic, the rejection rule and interpret your results.

(4 marks)

(d) Your result in (c) may be affected by the presence of autocorrelation. Discuss how you would conduct the Breusch–Godfrey test for the presence of autocorrelation in (9.2). Clearly indicate H_0 and H_1, the test statistic, the rejection rule and interpret your results. Briefly indicate what you might want to do to try and remove the autocorrelation.

(7 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 3.4 (Multicollinearity), Chapter 6.5 (Testing a linear restriction), Chapter 11.5 (Assumption C.7 and the properties of estimators), Chapter 11.3 (Models with lagged explanatory variables) and Chapter 12.1–12.3 (Definition and consequences of autocorrelation; Fitting a model subject to AR(1) autocorrelation).

Subject guide (2016), Chapters 11 and 12.

Approaching the question

In part (a), the concept of autocorrelation and the consequences on the OLS estimator need to be discussed. In part (b), candidates are expected to discuss the problem associated with multicollinearity. An interpretation of the parameters in distributed lag models (short- and long-run effects) can be given for partial credit. In part (b), a one sided t test for the LRP is required and candidates are expected to discuss a reparameterisation of the model that enables one to get the standard error of the LRP parameter. In part (c), an F test should be used to test for the joint significance of the third and fourth lag. Implementing a test for significance of the regression here is wrong – answer the question! In part (d), the Breusch–Godfrey test for the presence of autocorrelation should be discussed, and candidates should discuss ways of removing the problem of autocorrelation. This question was answered by 36% of the candidates in Zone B. The answer is as follows.

(a) Discussion of the concept is standard bookwork, see Chapter 12.1 of Dougherty.

The properties of OLS include *inefficiency* (autocorrelation is a violation of the Gauss–Markov assumptions). The OLS parameters will be *biased* (we do not have independence of the errors with all (future) values of the regressors – see Chapter 11.5 of Dougherty), but *consistent* (because of the presence of autocorrelation, it is important to point out that there is no lagged dependent variable here; consistency requires the uncorrelatedness between the error and regressors which are satisfied) and *invalid standard errors* and we will need to use HAC-robust standard errors.

(b) The distributed lag coefficients are $\delta_0, \delta_1, \ldots, \delta_4$ in this case. They tell us how if pe (z) increases by one unit today, but then falls back to its original level, y will change in each future period. Observe:

$$\begin{aligned}
\mathbf{y_t} &= \beta_0 + \boxed{\delta_0 \mathbf{z_t}} + \delta_1 z_{t-1} + \cdots + \delta_4 z_{t-4} + \beta_1 x_t + u_t \text{ (impact on } y_t\text{; contemporaneous)} \\
\mathbf{y_{t+1}} &= \beta_0 + \delta_0 z_{t+1} + \boxed{\delta_1 \mathbf{z_t}} + \cdots + \delta_4 z_{t-3} + \beta_1 x_{t+1} + u_{t+1} \text{ (impact on } y_{t+1}\text{)} \\
&\vdots \\
\mathbf{y_{t+4}} &= \beta_0 + \delta_0 z_{t+4} + \delta_1 z_{t+3} + \cdots + \boxed{\delta_4 \mathbf{z_t}} + \beta_1 x_{t+1} + u_{t+1} \text{ (impact on } y_{t+4}\text{).}
\end{aligned}$$

The novice is wrong, because the individual insignificance can be caused by the *near multicollinearity* of $pe_t, pe_{t-1}, \ldots, pe_{t-4}$. (Jointly, they are likely to be significant.)

By removing lags, more significant coefficients may be found.

The short-run impact is given by δ_0, while the long-run impact is given by $\delta_0 + \delta_1 + \cdots + \delta_4$. The LRP typically is significant even if the distributed lag coefficients are not.

(c) Candidates should test $\mathrm{H}_0 : \beta_{pe_{-3}} = 0$ and $\beta_{pe_{-4}} = 0$ against $\mathrm{H}_1 : \beta_{pe_{-3}} \neq 0$ and/or $\beta_{pe_{-4}} \neq 0$.

We want to use the F test here, which is distributed (asymptotically) as $F_{2,\,60}$, and at the 5% significance level we would reject H_0 if it exceeds 3.15.

With R_u^2 denoting the R^2 of the unrestricted model and R_r^2 denoting the R^2 of the restricted model, the F test is given by:

$$F = \frac{(R_u^2 - R_r^2)/2}{(1 - R_u^2)/60} = \frac{(0.537 - 0.536)/2}{(1 - 0.537)/60} = 0.065.$$

So this suggests that two lags would suffice, which would reduce the evidence of multicollinearity.

(d) When testing for autocorrelation, let us assume that:

$$u_t = \rho u_{t-1} + e_t \qquad \text{with } |\rho| < 1 \text{ and } e_t \text{ white noise (i.i.d. } (0,\, \sigma^2)).$$

A test for autocorrelation then becomes:

$$\mathrm{H}_0 : \rho = 0 \text{ (no autocorrelation)} \qquad \text{vs.} \qquad \mathrm{H}_1 : \rho \neq 0 \text{ (autocorrelation).}$$

As with a test of heteroskedasticity, a test of autocorrelation makes use of the OLS residuals – if there is autocorrelation, our OLS residuals will display this relationship.

The testing equation we should use is:

$$\widehat{u}_t = \gamma_0 + \rho \widehat{u}_{t-1} + \gamma_1 pe_t + \gamma_2 pe_{t-1} + \gamma_3 pe_{t-2} + \gamma_4 ww2_t + \gamma_1 pill_t + v_t.$$

We need to use the t test (and F test if general AR(p) or MA(p) is assumed) and we should reject the null hypothesis of zero autocorrelation if the test statistic $\widehat{\rho}/\mathrm{SE}(\widehat{\rho})$ in absolute value is larger than 1.96 (the asymptotic critical value).

To remove autocorrelation you may want to *introduce more dynamics in the model*, for example introduce lagged dependent variables. The Cochrane–Orcutt procedure (or other non-linear approach) can be suggested as well, but that requires us to know the form of autocorrelation which we may not (some details need to be provided).

Question 10

Consider the model:

$$\boldsymbol{y_t = \alpha_1 y_{t-1} + u_t, \qquad t = 1, \ldots, T}$$

where $y_0 = 0$. $\mathrm{E}(u_t) = 0$, $\mathrm{E}(u_t^2) = \sigma^2$ and $\mathrm{E}(u_t u_s) = 0$ when $s \neq t$, for all $s, t = 1, \ldots, T$.

(a) Discuss what we mean by the concept of stationarity (more precisely 'covariance stationarity') and indicate under what condition $\{y_t\}_{t=1}^{T}$ will be stationary.

(3 marks)

(b) Discuss the Dickey–Fuller procedure used to test for the presence of a unit root in the above model. Clearly indicate the null and alternative hypothesis, test statistic and rejection rule.

(5 marks)

(c) Consider a slight variation of the above model:

$$y_t = \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + v_t, \qquad t = 1, \ldots, T$$

where $y_0 = 0$. $\mathrm{E}(v_t) = 0$, $\mathrm{E}(v_t^2) = \sigma^2$ and $\mathrm{E}(v_t v_s) = 0$ when $s \neq t$, for all $s, t = 1, \ldots, T$. What do we call such a process? Discuss what problem you will have when conducting your test as in (b).

(3 marks)

(d) Instead of conducting the Dickey–Fuller procedure, you are told to apply the Augmented DF test. Indicate how you would conduct the test for the presence of a unit root here. Derivation of the test equation will be required for full marks.

(5 marks)

(e) What are the potential problems associated with performing a regression with $I(1)$ variables? In your answer explain what it means for a variable to be $I(1)$.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 13.1 (Stationarity and nonstationarity), Chapter 13.4–13.5 (Tests of nonstationarity), and Chapter 13.6 (Cointegration).

Subject guide (2016), Chapter 13.

Approaching the question

In part (a), the concept of (covariance) stationarity needs to be given. In part (b), you need to provide the Dickey–Fuller test for unit roots and in (d) you need to provide the Augmented Dickey–Fuller test. The latter allows us to deal with the fact that when conducting the Dickey–Fuller test we cannot have any autocorrelation, which would be the case if the dependence is not AR(1), but as is the case here AR(2). In (e), a discussion of the spurious regression result is expected. This question was answered by 45% of the candidates in Zone B. The answer is as follows.

(a) Candidates should provide a textbook definition of (covariance) stationarity.

The requirement for (covariance) stationarity here is that $|\alpha_1| < 1$. (Stating $\alpha_1 < 1$ is permitted.) Here $\{y_t\}$ is described by an AR(1) process, which is stationary provided its coefficient, α_1, is smaller than 1 (in magnitude). If α_1 is equal to one we have a unit root, whereas if $|\alpha_1| > 1$ we have an explosive process.

(b) To test for a unit root, we want to estimate the following regression:

$$\Delta y_t = \gamma y_{t-1} + u_t \qquad \text{where } \gamma = (\alpha_1 - 1).$$

We test $\mathrm{H}_0 : \gamma = 0$ (nonstationarity) against $\mathrm{H}_1 : \gamma < 0$ (stationarity). (Alternatively, we estimate the original model and test $\mathrm{H}_0 : \alpha_1 = 1$ (nonstationarity, unit root) against $\mathrm{H}_1 : \alpha_1 < 1$ (stationary AR process).)

We use the Dickey–Fuller t test $\widehat{\gamma}/\mathrm{SE}(\widehat{\gamma}) = (\widehat{\alpha}_1 - 1)/\mathrm{SE}(\widehat{\alpha}_1)$. Because we have non-stationarity under the null hypothesis, this test is not standard and we have to use the

Dickey–Fuller critical values. Given the alternative hypothesis, we reject H_0 when $\widehat{\gamma}/\mathrm{SE}(\widehat{\gamma}) < \tau$, where τ is the Dickey–Fuller critical value (no trend or intercept in the model). Candidates can, alternatively, discuss the Dickey–Fuller scaled coefficient test which is $T(\widehat{\alpha}_1 - 1) = T\widehat{\gamma}$.

If we reject the null hypothesis we have found evidence that our process is stationary (weakly dependent).

In the presence of a unit root, a shock to u_t will have an everlasting effect on the process $\{y_t\}$. Unit roots are persistent and strongly dependent.

(c) Here $\{y_t\}$ is an AR(2) process. (ADL(2, 0) is also acceptable.)

The problem with (b) will be serious, as *dependence (autocorrelation) in the error* will cause the error term u_t (which would then be equal to $\alpha_2 y_{t-2} + v_t$) and regressor y_{t-1} to be correlated and hence OLS will be inconsistent. Hence our test $\widehat{\gamma}/\mathrm{SE}(\widehat{\gamma})$ (or $T(\widehat{\alpha}_1 - 1)$) will be *invalid*.

(d) The Augmented Dickey–Fuller test suggests that we add further lags in our testing equation to remove the autocorrelation. In particular, we should estimate the following model:

$$\Delta y_t = \gamma_1 y_{t-1} + \gamma_2 \Delta y_{t-1} + v_t$$

and test the hypothesis $\mathrm{H}_0 : \gamma_1 = 0$ (non-stationarity) against $\mathrm{H}_1 : \gamma_1 < 0$ (stationarity). We should use the test given by $ADF = \widehat{\gamma}_1/\mathrm{SE}(\widehat{\gamma}_1)$ and we should reject H_0 when ADF is smaller than the critical value given by the Dickey–Fuller tables for a given significance level.

To derive this result observe that:

$$\begin{aligned}
\Delta y_t &= \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + v_t - y_{t-1} \\
&= (\alpha_1 - 1)\, y_{t-1} + \alpha_2 y_{t-2} + v_t \\
&= (\alpha_1 - 1)\, y_{t-1} + \alpha_2 \left(\underbrace{y_{t-2} - y_{t-1}}_{-\Delta y_{t-1}} + y_{t-1} \right) + v_t \\
&= (\alpha_1 + \alpha_2 - 1)\, y_{t-1} - \alpha_2 \Delta y_{t-1} + v_t
\end{aligned}$$

so $\gamma_1 = (\alpha_1 + \alpha_2 - 1)$ and $\gamma_2 = -\alpha_2$. Therefore, to test for a unit root in the AR(2) model we test whether $\alpha_1 + \alpha_2 < 1$.

(e) To say that a variable is $I(1)$ indicates that the variable has a unit root. We say that the variable then is integrated of order 1, revealing that by differencing the variable once we can make it stationary.

The potential problem associated with performing such a regression is that we may get a *spurious relation*. This is the setting where, due to the fact that both variables are trending, there is an appearance of a relationship that does not exist at all (high t statistics and a large R^2). To ensure that we have a meaningful (long-run) relationship between variables, we want to verify that instead we are dealing with a *cointegrating relationship*.

This paper is not to be removed from the Examination Hall

UNIVERSITY OF LONDON **EC2020 ZA**

BSc degrees and Diplomas for Graduates in Economics, Management, Finance and the Social Sciences, the Diplomas in Economics and Social Sciences

Elements of Econometrics

Thursday, 24 May 2018 : 14:30 to 17:30

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Extracts from statistical tables are given after the final question on this paper.

Graph paper is provided at the end of this question paper. If used, it must be detached and fastened securely inside the answer book.

A calculator may be used when answering questions on this paper and it must comply in all respects with the specification given with your Admission Notice. The make and type of machine must be clearly stated on the front cover of the answer book.

If more questions are answered than requested, only the first answers attempted will be counted.

PLEASE TURN OVER

© University of London 2018

SECTION A

Answer all questions from this section

1. We are interested in investigating the factors governing the precision of regression coefficients. Consider the model

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$$

with OLS parameter estimates $\hat{\beta}_1, \hat{\beta}_2$ and $\hat{\beta}_3$. Under the Gauss Markov assumptions, we have

$$Var\left(\hat{\beta}_2\right) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}\left(X_{2i} - \bar{X}_2\right)^2} \times \frac{1}{1 - r_{X_2X_3}^2},$$

where σ_ε^2 is the variance of ε and $r_{X_2X_3}$ is the sample correlation between X_2 and X_3.

(a) **(4 marks)** Provide four factors that help with obtaining more precise parameter estimates for, say, $\hat{\beta}_2$.

(b) **(4 marks)** In light of your answer to (a), discuss the concept of near multicollinearity. What consequences does its presence have when considering single and joint significance testing of our slope parameters?

2. Consider the linear regression model

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 Y_{t-1} + u_t, \qquad t = 1, ..., T$$

where the errors u_t are distributed independently of the regressors X_t and $|\beta_2| < 1$. You suspect that the, mean zero, errors exhibit autocorrelation.

(a) **(2 marks)** Explain what we mean by the concept of autocorrelation.

(b) Assume that u_t follows an AR(1) process.

i. **(3 marks)** Discuss, for the given model, the consequences for the ordinary least squares estimator. Support your answers with suitable arguments.

ii. **(3 marks)** Discuss how you would detect the presence of autocorrelation in the errors in this model. Clearly indicate the null and alternative hypothesis, the test statistic, and rejection rule.

3. For the population of men who grew up with disadvantaged backgrouns, let *poverty* be a dummy variable equal to one if a man is currently living below the poverty line, and zero otherwise. The variable *age* is age and *educ* is total years of schooling. Let *vocat* be an indicator equal to unity if a man's high school offered vocational training. Using a random sample of 850 men, you obtain

$$\widehat{\Pr\left(\textit{poverty=1}|\textit{educ},\textit{age},\textit{vocat}\right)} = \Lambda\left(0.453 - 0.016\ \textit{age} - 0.087\ \textit{educ} - 0.049\ \textit{vocat}\right)$$

where $\Lambda(z) = \exp(z)/(1+\exp(z))$ is the logit function.

(a) **(5 marks)** It is argued that using the logit regression model is better than using the linear probability model when explaining the binary variable *poverty*. Discuss the benefits/drawback of using the logit regression model when trying to explain a binary variable.

(b) **(3 marks)** For a 40-year old man, with 12 years of education, what is the estimated effect of having vocational training available in high school on the probability of currently living in poverty?
Hint: Clarity of computations required is enough, no need to give an exact number.

4. The following model jointly determines monthly child support payments and monthly visitation rights for divorced couples with children:

$$\textit{support} = \alpha_1 + \alpha_2\ \textit{visits} + \alpha_3\ \textit{finc} + \alpha_4\ \textit{fremarr} + \alpha_5\ \textit{dist} + \varepsilon_1$$
$$\textit{visits} = \beta_1 + \beta_2\ \textit{support} + \beta_3\ \textit{mremarr} + \beta_4\ \textit{dist} + \varepsilon_2$$

We assume that children live with their mothers, so that fathers pay child support. Thus, the first equation is the father's "reaction function": it describes the amount of child support paid for any given level of visitation rights and the other exogenous variables *finc* (father's income), *fremarr* (binary indicator if father remarried), and *dist* (miles currently between the mother and father's residence). Similarly the second equation is the mother's reaction function: it describes visitation rights for a given amount of child support; *mremarr* is a binary indicator for whether the woman is remarried.

(a) **(3 marks)** Examine the identification of each structural equation.

(b) **(5 marks)** Your friend suggests you should implement the IV estimator to estimate the β parameters consistently. He tells you to use *finc* as instrument for *support.* Provide a critical discussion of this suggestion.

5. Consider the simple linear regression model

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

under the classical linear regression model assumptions, where X_i is fixed under repeated sampling. The usual OLS estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased for their respective population parameters. Let $\tilde{\beta}_1$ be the estimator of β_1 obained by assuming the intercept is zero.

(a) **(4 marks)** Show that the **restricted least squares** estimator of β_1 is given by

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}.$$

(b) **(4 marks)** Find $E\left(\tilde{\beta}_1\right)$ in terms of the X_i, β_0 and β_1. Verify that $\tilde{\beta}_1$ is unbiased for β_1 when the population intercept is zero. Are there other cases where $\tilde{\beta}_1$ is unbiased?

SECTION B

Answer three questions from this section.

6. Let us consider the estimation of a hedonic price function for houses. The hedonic price refers to the implicit price of a house given certain attributes (e.g., the number of bedrooms). The data contains the sale price of 546 houses sold in the summer of 1987 in Canada along with their important features. The following characteristics are available: the lot size of the property in square feet (*lotsize*), the numbers of bedrooms (*bedrooms*), the number of full bathrooms (*bathrooms*), and a dummy indicating the presence of airconditioning (*airco)*.

   Consider the following ordinary least squares results

$$\begin{aligned}\widehat{\log(\textit{price})}_i &= \underset{\substack{(.232)\\[.233]}}{7.094} + \underset{\substack{(.028)\\[.028]}}{0.400}\log(\textit{lotsize})_i + \underset{\substack{(.015)\\[.017]}}{0.078}\,\textit{bedrooms}_i + \\ &\quad \underset{\substack{(.023)\\[.024]}}{0.216}\,\textit{bathrooms}_i + \underset{\substack{(.024)\\[.023]}}{0.212}\,\textit{airco}_i, \qquad n = 546,\ RSS = 32.622\end{aligned} \tag{6.1}$$

   The usual standard errors are in parentheses, the heteroskedasticity robust standard errors are in square brackets, and RSS measures the residual sum of squares.

   (a) **(5 marks)** Interpret the parameter estimates on $\log(\textit{lotsize})$, *bedrooms*, and *airco*. Briefly discuss the statistical significance of the results.

   (b) **(5 marks)** Suppose that lot size was measured in square metres rather than square feet. How would this affect the parameter estimates of the slopes and intercept? How would this affect the fitted values? *Note*: the conversion (approximate) $1m^2 = 10ft^2$.

   (c) **(5 marks)** We are interested in testing the hypothesis $H_0 : \beta_{\textit{bedrooms}} = \beta_{\textit{bathrooms}}$ against the alternative $H_A : \beta_{\textit{bedrooms}} \neq \beta_{\textit{bathrooms}}$. Discuss a test for this hypothesis that makes use of the following restricted regression result

$$\begin{aligned}\widehat{\log(\textit{price})}_i &= \underset{(.234)}{6.994} + \underset{(.282)}{0.408}\log(\textit{lotsize})_i + \underset{(.011)}{0.127}\,\textit{bbrooms}_i + \underset{(.024)}{0.215}\,\textit{airco}_i, \\ &\quad n = 546,\ RSS = 33.758\end{aligned} \tag{6.2}$$

   where *bbrooms* $=$ *bedrooms* $+$ *bathrooms*. Clearly indicate the assumptions you are making for this test to be valid.

   (d) **(5 marks)** You are interested in testing for the presence of heteroskedasticity. Say you are told that the variance is increasing with $\log(\textit{lotsize})$. Discuss how you would test for the presence of heteroskedasticity. What is the name of the test you are proposing?

7. The following question concerns the effects of background characteristics and admission assessment scores on the performance of students in the final university examinations in a UK university. The following equation was estimated by Ordinary Least Squares:

$$\begin{aligned} \widehat{\textit{finalavg}} &= \underset{(2.78)}{53.89} + \underset{(.04)}{0.03}\ \textit{tst_reas} + \underset{(.02)}{0.05}\ \textit{tst_quan} + \underset{(.02)}{0.06}\ \textit{interview} \\ &\quad \underset{(.78)}{-0.04}\ \textit{indep} + \underset{(.66)}{0.67}\ \textit{male} + \underset{(.97)}{0.06}\ \textit{indep*male} \\ n &= 325;\ R^2 = .06, \end{aligned} \tag{7.1}$$

where *finalavg* is the average finals score (the outcome), *tst_reas* and *tst_quan* are the pre-admission reasoning and quantitative test scores respectively, *interview* is the pre-admission interview score, *indep* indicates whether the student attended an independent school (1=yes, 0=no), and *male* indicates whether the student is male (1=yes, 0=no). The usual standard errors are in parentheses.

(a) **(5 marks)** We want to test whether gender has a significant impact on students' finals performance. Clearly indicating the null and the alternative hypothesis, provide the test statistic and the rejection rule. Discuss what information you would need to enable you to implement this test. You are expected to provide the assumptions which underlie your test.

(b) **(5 marks)** If we do not include the interaction term *indep*male* in our regression model, what are we implicitly assuming about the effect of gender and school background on finals performance?

(c) **(5 marks)** Suppose students who did not attend an independent school, attended a state school. Using the results in (7.1), provide the parameter estimates you would obtain if you had applied Ordinary Least Squares to the equation

$$\begin{aligned} \textit{finalavg} &= \beta_0 + \beta_1\ \textit{tst_reas} + \beta_2\ \textit{tst_quan} + \beta_3\ \textit{interview} \\ &\quad \beta_4\ \textit{state} + \beta_5\ \textit{male} + \beta_6\ \textit{state*male} + \varepsilon, \end{aligned} \tag{7.2}$$

where *state* indicates whether the student attended a state school (1=yes, 0=no).

(d) **(5 marks)** Discuss any problem you may have in estimating the model if all males in your sample have attended an independent school prior to attending university. What name does this problem have and what can you do to mitigate this problem?

8. An OLS regression of y_t on x_t and x_{t-1} gives the following results (with the standard errors given in parentheses)

$$\widehat{y}_t = \underset{(2.30)}{8.88} + \underset{(3.01)}{5.07} x_t - \underset{(3.01)}{3.18}\ x_{t-1};\ R^2 = .095,\ T = 209 \tag{8.1}$$

(a) **(4 marks)** What are the estimates of the short-run and long-run effect of x_t on y_t? Interpret these estimates.

(b) **(4 marks)** Test the hypothesis that a one unit increase in x results in a ten unit increase in y in the same year. Under what assumptions is this test valid?

Let e_t be the OLS residuals from the above regression. An OLS regression of e_t on e_{t-1} yields

$$e_t = \underset{(2.12)}{0.55} + \underset{(0.18)}{0.44} e_{t-1} + \underset{(2.18)}{2.16} x_t - \underset{(.99)}{1.09} x_{t-1};\ R^2 = .175,\ T = 208 \tag{8.2}$$

(c) **(5 marks)** Using this result, test for evidence of autocorrelation, clearly indicating the null and alternative hypotheses, the test statistic, rejection rule and assumptions underlying the test. What name do we give this test?

(d) You are interested in testing whether the long-run effect of x_t on y_t is statistically significant.

i. **(4 marks)** Discuss how to reparameterize (8.1) to ensure that your regression output will provide you with a standard error for the long-run effect.

ii. **(3 marks)** Discuss the problem of implementing your test using the standard error obtained in (d)i. when you do find evidence of autocorrelation in (8.1). Briefly indicate how you proceed with your test.

9. Consider the model

$$\begin{aligned} y_t &= \alpha + \beta t + \varepsilon_t,\ t = 1, .., T \\ \varepsilon_t &= \rho \varepsilon_{t-1} + v_t, \text{ and} \end{aligned} \tag{9.1}$$

v_t is an i.i.d.$\left(0, \sigma^2\right)$ innovation which is independent of the past. Let $|\rho| \leq 1$.

(a) **(4 marks)** What name do we give the ε_t process given above? Provide the condition(s) that ensures that ε_t is stationary? In your answer discuss what we mean by the concept of stationarity (more precisely "covariance stationarity").

(b) It will be important to distinguish between the above process for y_t being "trend stationary" as opposed to "difference stationary".

i. **(4 marks)** Explain these concepts clearly. Why is it important to distinguish between these two types of non-stationarity?

ii. **(2 marks)** Show that under the condition you provided in (a) that y_t is trend stationary.

iii. **(2 marks)** Show that if ε_t is difference stationary then y_t is difference stationary.

(c) **(4 marks)** Show that you can rewrite the above model in the following form

$$\Delta y_t = \gamma_1 + \gamma_2 t + \gamma_3 y_{t-1} + v_t. \tag{9.2}$$

Clearly indicate the relation between $(\gamma_1, \gamma_2, \gamma_3)$ and (α, β, ρ).

(d) **(4 marks)** What problem do you see here with using 9.2 to conducting the Dickey Fuller Test to distinguish between trend and difference stationarity when v_t exhibits autocorrelation? What solution do you suggest we adopt.

10. Let *math*10 denote the percentage of students at a high school receiving passing score on a standardized math test. We are interested in estimating the effect of per student spending on math performance. A simple model is

$$\textit{math}10_i = \beta_0 + \beta_1 \log(\textit{expend}_i) + \beta_2 \log(\textit{enroll}_i) + \beta_3\, \textit{poverty}_i + u_i \tag{10.1}$$

where, for each high school i, *poverty*$_i$ is the percentage of students living in poverty, *expend*$_i$ is the spending per student and *enroll*$_i$ the number of registered students. You may assume that this model satisfies all Gauss-Markov assumptions.

You are faced with the fact that data is unavailable on a key variable: *poverty*.

(a) **(5 marks)** Discuss the properties (unbiasedness and consistency) of the estimators when you drop the variable *poverty?* Explain your answers.

You do have information available on a closely related variable: the percentage of students eligible for the federally funded school lunch program, *lnchprg*$_i$. Let us consider using *lnchprg*$_i$ as a proxy for *poverty*$_i$.

(b) **(2 marks)** Briefly discuss why *lnchprg*$_i$ is a sensible proxy variable for the unobserved variable *poverty*$_i$.

(c) **(5 marks)** It is unlikely that *lnchprg*$_i$ is an ideal proxy, in the sense that there is an exact linear relationship between them, instead, we will assume that

$$\textit{poverty}_i = \alpha_0 + \alpha_1\, \textit{lnchprg}_i + v_i,\ \alpha_1 \neq 0 \tag{10.2}$$

Discuss the assumptions you need to make to enable consistent parameter estimators of β_1 and β_2 using your estimable equation

$$\textit{math}10_i = \gamma_0 + \gamma_1 \log(\textit{expend}_i) + \gamma_2 \log(\textit{enroll}_i) + \gamma_3\, \textit{lnchprg}_i + e_i,$$

Hint: Consider the relation between the γ and the β parameters and express e_i in terms of u_i and v_i.

(question continues on next page)

(d) The OLS results with and without *lnchprg*$_i$ as an explanatory variable are given by (standard errors in parentheses):

$$\widehat{math10}_i = \underset{(26.72)}{-69.24} + \underset{(3.30)}{11.13}\log(expend_i) + \underset{(0.615)}{0.022}\log(enroll_i),$$
$$N = 428, R^2 = 0.0297$$

$$\widehat{math10}_i = \underset{(24.99)}{-23.14} + \underset{(3.04)}{7.75}\log(expend_i) - \underset{(0.58)}{1.26}\log(enroll_i) - \underset{(0.036)}{0.324}\, lnchprg_i$$
$$N = 428, R^2 = 0.1893$$

i. **(4 marks)** Interpret the coefficient on *lnchprg*. What does this parameter tell us regarding the parameter of interest β_3.

ii. **(4 marks)** Give an intuitive discussion explaining why the effect of expenditures on *math*10_i is lower in the regression where *lnchprg*$_i$ is included than where it is excluded.

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	**0·40**	0·6554	**0·80**	0·7881	**1·20**	0·8849	**1·60**	0·9452	**2·00**	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	**0·45**	0·6736	**0·85**	0·8023	**1·25**	0·8944	**1·65**	0·9505	**2·05**	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	**0·50**	0·6915	**0·90**	0·8159	**1·30**	0·9032	**1·70**	0·9554	**2·10**	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	**0·55**	0·7088	**0·95**	0·8289	**1·35**	0·9115	**1·75**	0·9599	**2·15**	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	**0·60**	0·7257	**1·00**	0·8413	**1·40**	0·9192	**1·80**	0·9641	**2·20**	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	**0·65**	0·7422	**1·05**	0·8531	**1·45**	0·9265	**1·85**	0·9678	**2·25**	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	**0·70**	0·7580	**1·10**	0·8643	**1·50**	0·9332	**1·90**	0·9713	**2·30**	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	**0·75**	0·7734	**1·15**	0·8749	**1·55**	0·9394	**1·95**	0·9744	**2·35**	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	**0·80**	0·7881	**1·20**	0·8849	**1·60**	0·9452	**2·00**	0·9772	**2·40**	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

When $x > 3·3$ the formula $1-\Phi(x) \doteqdot \frac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\frac{1}{x^2}+\frac{3}{x^4}-\frac{15}{x^6}+\frac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	5·0	1·6449	3·0	1·8808	2·0	2·0537	1·0	2·3263	0·10	3·0902
45	0·1257	4·8	1·6646	2·9	1·8957	1·9	2·0749	0·9	2·3656	0·09	3·1214
40	0·2533	4·6	1·6849	2·8	1·9110	1·8	2·0969	0·8	2·4089	0·08	3·1559
35	0·3853	4·4	1·7060	2·7	1·9268	1·7	2·1201	0·7	2·4573	0·07	3·1947
30	0·5244	4·2	1·7279	2·6	1·9431	1·6	2·1444	0·6	2·5121	0·06	3·2389
25	0·6745	4·0	1·7507	2·5	1·9600	1·5	2·1701	0·5	2·5758	0·05	3·2905
20	0·8416	3·8	1·7744	2·4	1·9774	1·4	2·1973	0·4	2·6521	0·01	3·7190
15	1·0364	3·6	1·7991	2·3	1·9954	1·3	2·2262	0·3	2·7478	0·005	3·8906
10	1·2816	3·4	1·8250	2·2	2·0141	1·2	2·2571	0·2	2·8782	0·001	4·2649
5	1·6449	3·2	1·8522	2·1	2·0335	1·1	2·2904	0·1	3·0902	0·0005	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

This paper is not to be removed from the Examination Hall

UNIVERSITY OF LONDON **EC2020 ZB**

BSc degrees and Diplomas for Graduates in Economics, Management, Finance and the Social Sciences, the Diplomas in Economics and Social Sciences

Elements of Econometrics

Thursday, 24 May 2018 : 14:30 to 17:30

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Extracts from statistical tables are given after the final question on this paper.

Graph paper is provided at the end of this question paper. If used, it must be detached and fastened securely inside the answer book.

A calculator may be used when answering questions on this paper and it must comply in all respects with the specification given with your Admission Notice. The make and type of machine must be clearly stated on the front cover of the answer book.

If more questions are answered than requested, only the first answers attempted will be counted.

PLEASE TURN OVER

© University of London 2018

SECTION A

Answer all questions from this section

1. Consider the consumption function

$$C_t = \alpha + \lambda Y_t + \varepsilon_t, \tag{1.1}$$

where C_t is aggregate consumption at t, λ is marginal propensity to consume ($0 < \lambda < 1$) and Y_t is aggregate income at t defined as

$$Y_t = C_t + A_t,$$

where A_t is the sum of investment and government consumption at t. Assume that A_t is uncorrelated with ε_t and that the shock ε_t is mean zero i.i.d across t. A random sample of size n containing Y_t, C_t and A_t is available.

(a) **(2 marks)** Provide the reduced form equation for Y_t.

(b) **(6 marks)** Show that the OLS estimator of λ in (1.1) is inconsistent. You are asked to indicate the direction of this inconsistency.
Note: you are not expected to derive the OLS estimator.

2. Consider the simple linear regression model

$$Y_t = \beta X_t + u_t, \qquad t = 1, ..., T$$

where the errors u_t are distributed independently of the regressors X_t. You suspect that the, mean zero, errors exhibit autocorrelation.

(a) **(2 marks)** Explain what we mean by the concept of autocorrelation.

(b) Assume you are told that u_t follows an MA(1) process.

i. **(3 marks)** Discuss whether the OLS estimator $\hat{\beta}$ is a consistent estimator for β. Justify your answers with suitable technical derivations.
Note: you are not expected to derive the OLS estimator.

ii. **(3 marks)** Suppose you want to test $H_0 : \beta = 1$ against $H_A : \beta < 1$. Discuss how you would conduct this test based on the OLS estimator, recognizing the presence of autocorrelation in the error.

3. Consider the simple linear regression model

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

under the classical linear regression model assumptions, where X_i is fixed under repeated sampling. The usual OLS estimators $\hat{\beta}_0$ and $\hat{\beta}_1$ are unbiased for their respective population parameters. Let $\tilde{\beta}_0$ be the estimator of β_0 when β_1 equals 1.

(a) **(4 marks)** Show that the restricted least squares estimator of β_0 is given by

$$\tilde{\beta}_0 = \bar{Y} - \bar{X}$$

where $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ and $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

(b) **(4 marks)** Find $E\left(\tilde{\beta}_0\right)$ in terms of the X_i, β_0 and β_1. Verify that $\tilde{\beta}_0$ is unbiased for β_0 when $\beta_1 = 1$. Are there other cases where $\tilde{\beta}_0$ is unbiased?

4. We are interested in investigating the factors governing the precision of regression coefficients. Consider the model

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$$

with OLS parameter estimates $\hat{\beta}_1, \hat{\beta}_2$ and $\hat{\beta}_3$. Under the Gauss Markov assumptions, we have

$$Var\left(\hat{\beta}_2\right) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}\left(X_{2i} - \bar{X}_2\right)^2} \times \frac{1}{1 - r_{X_2X_3}^2},$$

where σ_ε^2 is the variance of ε and $r_{X_2X_3}$ is the sample correlation between X_2 and X_3.

(a) **(4 marks)** Provide four factors that help with obtaining more precise parameter estimates for, say, $\hat{\beta}_2$.

(b) **(4 marks)** Assume that the true value of $\beta_3 = 0$, so that the above model includes an irrelevant variable. Discuss the effect of including this irrelevant variable on the unbiasedness and precision of $\hat{\beta}_2$.

5. A probit model to explain whether a firm is taken over by another firm during a given year postulates

$$\begin{aligned}\Pr(\textit{takeover=1}|x) &= \Phi(\beta_0 + \beta_1\ \textit{avgprof} + \beta_2\ \textit{mtkval} + \beta_3\ \textit{debtearn} + \beta_4\ \textit{ceoten} \\ &\quad + \beta_5\ \textit{ceosal} + \beta_6\ \textit{ceoage})\end{aligned}$$

where $\Phi(z)$ is the cumulative standardized normal distribution. *takeover* is a binary response variable, *avgprof* is the firm's average profit margin over several prior years, *mktval* is the market value of the firm, *debtearn* is the debt-to-earnings ratio, and *ceoten, ceosal,* and *ceoage* are the tenure, annual salary, and age of the chief executive officer, respectively.

(a) **(5 marks)** It is argued that using the probit regression model is better than using the linear probability model when explaining the binary variable *takeover*. Discuss the benefits/drawback of using the probit regression model when trying to explain a binary variable.

(b) **(3 marks)** Discuss how you would implement the LR test that variables related to the CEO have no effect on the probability of takeover, other factors being equal. Clearly indicate the null, alternative, test statistic and rejection rule.

SECTION B

Answer three questions from this section.

6. The following question concerns the effects of background characteristics on student's performance in the SAT (Scholastic Assessment Test). The SAT test is used for college admissions in the US.

$$\begin{aligned}\widehat{sat} &= \underset{(6.29)}{1,028.10} \underset{(3.83)}{+19.3}\ hsize \underset{(.53)}{-2.19}\ hsize^2 \underset{(4.29)}{-45.09}\ female \underset{(12.71)}{-169.81}\ black \\ &\quad \underset{(12.71)}{+62.31} female * black \\ n &= 4,127;\ R^2 = .0858.\end{aligned}$$

The variable *hsize* is the size of the student's high school graduating class, in hundreds, *female* is a gender dummy variable (1=female, 0=male), and *black* is a race dummy variable (1=black, 0=otherwise). The standard errors are in parentheses.

(a) **(5 marks)** What is the economic rationale for including $hsize^2$ in the above regression? Using this equation, determine for a given gender and race, what the graduating class size would be at which the predicted SAT scores are maximized.

(b) **(5 marks)** Holding *hsize* fixed, what is the estimated difference in SAT scores between nonblack females and nonblack males? Is this difference statistically significant? Interpret this result.

(c) **(5 marks)** What is the estimated difference in SAT score between black females and nonblack females? What would you need to do to test whether the difference is statistically significant?

(d) **(5 marks)** Discuss any problem you may have in estimating the model if all females in your sample are black. What name does this problem have and what can you do to mitigate this problem?

7. Let us consider the estimation of a hedonic price function for houses. The hedonic price refers to the implicit price of a house given certain attributes (e.g., the number of bedrooms). The data contains the sale price of 546 houses sold in the summer of 1987 in Canada along with their important features. The following characteristics are available: the lot size of the property in square feet (*lotsize*), the numbers of bedrooms (*bedrooms*), the number of full bathrooms (*bathrooms*), and a dummy indicating the presence of airconditioning (*airco)*.

Consider the following ordinary least squares results

$$\begin{aligned}\widehat{\log(\textit{price})}_i &= \underset{\substack{(.232)\\[.233]}}{7.094} + \underset{\substack{(.028)\\[.028]}}{0.400}\log(\textit{lotsize})_i + \underset{\substack{(.015)\\[.017]}}{0.078}\,\textit{bedrooms}_i + \\ &\quad \underset{\substack{(.023)\\[.024]}}{0.216}\,\textit{bathrooms}_i + \underset{\substack{(.024)\\[.023]}}{0.212}\,\textit{airco}_i, \qquad n = 546,\ RSS = 32.622 \end{aligned} \tag{7.1}$$

The usual standard errors are in parentheses, the heteroskedasticity robust standard errors are in square brackets, and RSS measures the residual sum of squares.

(a) **(5 marks)** Interpret the parameter estimates on $\log(\textit{lotsize})$, *bedrooms*, and *airco*. Briefly discuss the statistical significance of the results.

(b) **(5 marks)** Suppose that lot size was measured in square metres rather than square feet. How would this affect the parameter estimates of the slopes and intercept? How would this affect the fitted values? *Note*: the conversion (approximate) $1m^2 = 10ft^2$.

(c) **(5 marks)** We are interested in testing the hypothesis $H_0 : \beta_{\textit{bedrooms}} = \beta_{\textit{bathrooms}}$ against the alternative $H_A : \beta_{\textit{bedrooms}} \neq \beta_{\textit{bathrooms}}$. Discuss a test for this hypothesis that makes use of the following restricted regression result

$$\begin{aligned}\widehat{\log(\textit{price})}_i &= \underset{(.234)}{6.994} + \underset{(.282)}{0.408}\log(\textit{lotsize})_i + \underset{(.011)}{0.127}\,\textit{bbrooms}_i + \underset{(.024)}{0.215}\,\textit{airco}_i, \\ &\quad n = 546,\ RSS = 33.758 \end{aligned} \tag{7.2}$$

where *bbrooms* $=$ *bedrooms* $+$ *bathrooms*. Clearly indicate the assumptions you are making for this test to be valid.

(d) **(5 marks)** You are interested in testing for the presence of heteroskedasticity. Say you are told that the variance is increasing with $\log(\textit{lotsize})$. Discuss how you would test for the presence of heteroskedasticity. What is the name of the test you are proposing?

8. Let *math*10 denote the percentage of students at a high school receiving passing score on a standardized math test. We are interested in estimating the effect of per student spending on math performance. A simple model is

$$\textit{math}10_i = \beta_0 + \beta_1 \log(\textit{expend}_i) + \beta_2 \log(\textit{enroll}_i) + \beta_3\, \textit{poverty}_i + u_i \tag{8.1}$$

where, for each high school i, *poverty*$_i$ is the percentage of students living in poverty, *expend*$_i$ is the spending per student and *enroll*$_i$ the number of registered students. You may assume that this model satisfies all Gauss-Markov assumptions.

You are faced with the fact that data is unavailable on a key variable: *poverty*.

(a) **(5 marks)** Discuss the properties (unbiasedness and consistency) of the estimators when you drop the variable *poverty?* Explain your answers.

You do have information available on a closely related variable: the percentage of students eligible for the federally funded school lunch program, *lnchprg*$_i$. Let us consider using *lnchprg*$_i$ as a proxy for *poverty*$_i$.

(b) **(2 marks)** Briefly discuss why *lnchprg*$_i$ is a sensible proxy variable for the unobserved variable *poverty*$_i$.

(c) **(5 marks)** It is unlikely that *lnchprg*$_i$ is an ideal proxy, in the sense that there is an exact linear relationship between them, instead, we will assume that

$$\textit{poverty}_i = \alpha_0 + \alpha_1\, \textit{lnchprg}_i + v_i,\ \alpha_1 \neq 0 \tag{8.2}$$

Discuss the assumptions you need to make to enable consistent parameter estimators of β_1 and β_2 using your estimable equation

$$\textit{math}10_i = \gamma_0 + \gamma_1 \log(\textit{expend}_i) + \gamma_2 \log(\textit{enroll}_i) + \gamma_3\, \textit{lnchprg}_i + e_i,$$

Hint: Consider the relation between the γ and the β parameters and express e_i in terms of u_i and v_i.

(question continues on next page)

(d) The OLS results with and without *lnchprg*$_i$ as an explanatory variable are given by (standard errors in parentheses):

$$\widehat{math10}_i = \underset{(26.72)}{-69.24} + \underset{(3.30)}{11.13}\log(expend_i) + \underset{(0.615)}{0.022}\log(enroll_i),$$
$$N = 428, R^2 = 0.0297$$

$$\widehat{math10}_i = \underset{(24.99)}{-23.14} + \underset{(3.04)}{7.75}\log(expend_i) - \underset{(0.58)}{1.26}\log(enroll_i) - \underset{(0.036)}{0.324}\, lnchprg_i$$
$$N = 428, R^2 = 0.1893$$

i. **(4 marks)** Interpret the coefficient on *lnchprg*. What does this parameter tell us regarding the parameter of interest β_3.

ii. **(4 marks)** Give an intuitive discussion explaining why the effect of expenditures on *math*10_i is lower in the regression where *lnchprg*$_i$ is included than where it is excluded.

9. Let us consider monthly data on the short-term interest rate (the three month Treasury Bill rate) and on the AAA corporate bond yield in the USA. The data run from January 1950 to December 1999. Let *DUS3MT* denote the changes in three-month Treasury Bill rate, and *DAAA* denote the changes in AAA bond rate. We consider the following results (with the standard errors given in parentheses)

$$\widehat{DAAA}_t = \underset{(.007)}{0.006} + \underset{(.015)}{0.275}\, DUS3MT_t, \qquad t = 1, ..., 600 \tag{9.1}$$
$$RSS = 17.486;\ DW = 1.447$$

where RSS is the residual sum of squares and DW is the Durbin Watson test.
A researcher interpreting the residuals suggests that the errors show a positive correlation over time.

(a) **(5 marks)** What are the consequences of this correlation for the above regression results?

(b) **(5 marks)** Use the results above to test for the presence of first order positive autocorrelation. Clearly specify the null and alternative hypothesis, test statistic, assumptions underlying the test, and the acceptance/rejection rule.

(c) **(5 marks)** In an attempt to remove the autocorrelation you consider the following specification

$$\widehat{DAAA}_t = \underset{(.007)}{0.005} + \underset{(.015)}{0.252}\, DUS3MT_t - \underset{(.018)}{0.080}\, DUS3MT_{t-1} + \underset{(.040)}{0.290} DAAA_{t-1}, \tag{9.2}$$
$$RSS = 16.087;\ DW = 1.897$$

Comment on the following statement "The Durbin Watson statistic is closer to 2, indicating that we have succesfully removed the autocorrelation". If you disagree with this statement, suggest what you would need to do instead.

(d) **(5 marks)** Discuss the Common Factor Test as a model specification suitable for this model. What extra information do you need to conduct this test.

10. Consider the model

$$y_t = \alpha + \beta x_t + \varepsilon_t \,,\ t = 1, ..., T \tag{10.1}$$

where y_t and x_t are both integrated of order one.

(a) **(6 marks)** Explain what it means to say that y_t is integrated of order one. Discuss how you would test for this. In your answer make sure that it is clear how to implement your test.

(b) **(2 marks)** Give an example of an economic variable that is potentially integrated of order one and give an intuitive explanation why you expect this process to be integrated of order 1.

(c) It will be important to distinguish whether the above relationship is "spurious" as opposed to "cointegrating".

i. **(4 marks)** Explain what it means to say that y_t and x_t have a cointegrating relationship and how does that contrast to a spurious relationship.

ii. **(4 marks)** Discuss how you can test for evidence of a cointegrating relationship.

(d) **(4 marks)** Suppose that

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t,\ \ |\rho| < 1,$$

and v_t is an i.i.d.$\left(0, \sigma^2\right)$ innovation which is independent of ε_{t-1}. Show that you can rewrite equation (10.1) in terms of an error correction model:

$$\Delta y_t = \delta_1 \Delta x_t + \delta_2 (y_{t-1} - \alpha - \beta x_{t-1}) + v_t.$$

Clearly indicate the relation between (δ_1, δ_2) and (α, β, ρ). Give an economic intuition behind this result.

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	**0·40**	0·6554	**0·80**	0·7881	**1·20**	0·8849	**1·60**	0·9452	**2·00**	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	**0·45**	0·6736	**0·85**	0·8023	**1·25**	0·8944	**1·65**	0·9505	**2·05**	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	**0·50**	0·6915	**0·90**	0·8159	**1·30**	0·9032	**1·70**	0·9554	**2·10**	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	**0·55**	0·7088	**0·95**	0·8289	**1·35**	0·9115	**1·75**	0·9599	**2·15**	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	**0·60**	0·7257	**1·00**	0·8413	**1·40**	0·9192	**1·80**	0·9641	**2·20**	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	**0·65**	0·7422	**1·05**	0·8531	**1·45**	0·9265	**1·85**	0·9678	**2·25**	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	**0·70**	0·7580	**1·10**	0·8643	**1·50**	0·9332	**1·90**	0·9713	**2·30**	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	**0·75**	0·7734	**1·15**	0·8749	**1·55**	0·9394	**1·95**	0·9744	**2·35**	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	**0·80**	0·7881	**1·20**	0·8849	**1·60**	0·9452	**2·00**	0·9772	**2·40**	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	5·0	1·6449	3·0	1·8808	2·0	2·0537	1·0	2·3263	0·10	3·0902
45	0·1257	4·8	1·6646	2·9	1·8957	1·9	2·0749	0·9	2·3656	0·09	3·1214
40	0·2533	4·6	1·6849	2·8	1·9110	1·8	2·0969	0·8	2·4089	0·08	3·1559
35	0·3853	4·4	1·7060	2·7	1·9268	1·7	2·1201	0·7	2·4573	0·07	3·1947
30	0·5244	4·2	1·7279	2·6	1·9431	1·6	2·1444	0·6	2·5121	0·06	3·2389
25	0·6745	4·0	1·7507	2·5	1·9600	1·5	2·1701	0·5	2·5758	0·05	3·2905
20	0·8416	3·8	1·7744	2·4	1·9774	1·4	2·1973	0·4	2·6521	0·01	3·7190
15	1·0364	3·6	1·7991	2·3	1·9954	1·3	2·2262	0·3	2·7478	0·005	3·8906
10	1·2816	3·4	1·8250	2·2	2·0141	1·2	2·2571	0·2	2·8782	0·001	4·2649
5	1·6449	3·2	1·8522	2·1	2·0335	1·1	2·2904	0·1	3·0902	0·0005	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

Examiners' commentaries 2018

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2017–18. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

General remarks

Learning outcomes

At the end of the course, and having completed the Essential reading and activities, you should be able to:

- describe and apply the classical regression model and its application to cross-section data
- describe and apply the:
  - Gauss–Markov conditions and other assumptions required in the application of the classical regression model
  - reasons for expecting violations of these assumptions in certain circumstances
  - tests for violations
  - potential remedial measures, including, where appropriate, the use of instrumental variables
- recognise and apply the advantages of logit, probit and similar models over regression analysis when fitting binary choice models
- competently use regression, logit and probit analysis to quantify economic relationships using standard regression programmes (Stata and EViews) in simple applications
- describe and explain the principles underlying the use of maximum likelihood estimation
- apply regression analysis to time-series models using stationary time series, with awareness of some of the econometric problems specific to time series applications (for example, autocorrelation) and remedial measures
- recognise the difficulties that arise in the application of regression analysis to nonstationary time series, know how to test for unit roots, and know what is meant by cointegration.

Common mistakes committed by candidates

A large number of candidates are not able to clearly distinguish between sample variance and covariance, and population variance and covariance (this is happening year after year).

The use of $\mathrm{Cov}(X,Y)$ and $\mathrm{Var}(X)$ should be restricted to describing the population covariance and variances, respectively, with definitions:

$$\mathrm{Cov}(X,Y) = \mathrm{E}\left((X-\mathrm{E}(X))(Y-\mathrm{E}(Y))\right) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$$

and:

$$\mathrm{Var}(X) = \mathrm{E}((X-\mathrm{E}(X))^2) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$$

(you also may denote $\mathrm{Cov}(X,Y) = \sigma_{XY}$ and $\mathrm{Var}(X) = \sigma_X^2$). They are typically unknown, but fixed, quantities.

The sample covariance and variance are estimators of the population covariance and variance, respectively. They are defined as:

$$\text{Sample Cov}(X,Y) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})$$

and:

$$\text{Sample Var}(X) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

(you also may use $\widehat{\sigma}_{XY}$ and $\widehat{\sigma}_X^2$). You can compute them given the data.

With a slight abuse of notation, we often divide by n instead, which is irrelevant if we let n be large. The division by $n-1$ is a finite sample issue only (unbiasedness).

The sample covariance and variance show up in our definition of the OLS estimator of the slope in the simple linear regression model, not the population covariance and variance, as:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sum_{i=1}^{n}(X_i-\bar{X})^2} = \frac{\text{Sample Cov}(X,Y)}{\text{Sample Var}(X)} \neq \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}.$$

Treating them as being the same results in incorrect analyses and candidates losing significant marks.

Candidates should realise that $\frac{1}{n}\sum_{i=1}^{n} u_i$ is not the same as $\mathrm{E}(u_i)$. So, while we typically assume $\mathrm{E}(u_i) = 0$, this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} u_i = 0$. Also, while we may be happy to assume $\mathrm{E}(x_i u_i) = 0$ (uncorrelatedness between the errors and regressors), this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} x_i u_i = 0$. Note that:

- both $\frac{1}{n}\sum_{i=1}^{n} u_i$ and $\frac{1}{n}\sum_{i=1}^{n} x_i u_i$ are random variables, which take the value 0 with probability 0 (continuous random variables)!
- $\mathrm{E}(u_i) = 0$ and $\mathrm{E}(x_i u_i) = 0$ are fixed, not stochastic!

The differences between sample and population moments need to come across clearly when looking at unbiasedness and making consistency arguments. In both cases, we first simplify our estimator (plug in the true model) to obtain:

$$\widehat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i-\bar{X})u_i}{\sum_{i=1}^{n}(X_i-\bar{X})^2} = \beta + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \qquad \text{with } x_i = X_i - \bar{X}.$$

- For *unbiasedness*, clearly indicate that you want to show that $\mathrm{E}(\widehat{\beta}) = \beta$. Unbiasedness does not follow from $\sum_{i=1}^{n} x_i u_i = 0$, instead it follows from $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) = 0$.

  If we treat x_i as fixed, $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) \equiv \mathrm{E}\left(\sum_{i=1}^{n} d_i u_i\right) = \sum_{i=1}^{n} d_i \,\mathrm{E}(u_i)$ and then unbiasedness follows as $\mathrm{E}(u_i) = 0$.
- For *consistency*, clearly indicate that you want to show that $\mathrm{plim}(\widehat{\beta}) = \beta$. Using the plim properties, we show:

$$\begin{aligned}
\mathrm{plim}\ \widehat{\beta} = \beta + \mathrm{plim}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) &= \beta + \frac{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i u_i\right)}{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)} \\
&\equiv \beta + \frac{\mathrm{plim}\,(\text{Sample Cov}\,(x,u))}{\mathrm{plim}\,(\text{Sample Var}\,(x))} \\
&= \beta + \frac{\mathrm{Cov}\,(x,u)}{\mathrm{Var}\,(x)} \qquad \text{using the law of large numbers}
\end{aligned}$$

  where $\mathrm{Cov}(x,u) = 0$ and $\mathrm{Var}(x) > 0$, ensuring we get consistency.
- Remember, the law of large numbers ensures that sample averages converge to their population analogues.

Candidates struggled to give competent answers to the interpretation of empirical results. When interpreting an empirical result you should discuss the significance of the coefficients, magnitude and sign of the coefficients.

When conducting hypothesis tests, you should make sure that the Gauss–Markov conditions hold. The Gauss–Markov conditions have to be explicitly specified. Only writing that the Gauss–Markov conditions hold is not sufficient. As good practice, begin your examination by explicitly providing the Gauss–Markov conditions. You can then refer back to them thereafter. Moreover, ensure when conducting hypothesis testing that you clearly indicate the null and alternative hypotheses (in terms of the true parameters, say β_1), the test statistic (in terms of the parameter estimates, here $\widehat{\beta}_1$), its distribution (with degrees of freedom), the rejection rule (one-sided or-two sided) for a given significance level (typically 5%) with suitable critical values, and provide an interpretation of your result.

Just as last year, many candidates do not answer all parts of the question. Make sure you read the questions properly and provide all details that are requested. Not answering a question will automatically earn you a zero mark for that question.

Key steps to improvement

Essential reading for this course includes the subject guide and the following:

- Dougherty, C. *Introduction to econometrics.* (Oxford: Oxford University Press, 2016) 5th edition [ISBN 9780199676828]; http://oxfordtextbooks.co.uk/orc/dougherty5e/

Apart from the Essential readings you should do some supplementary reading. One very good book at the same level is:

- Gujarati, D.N. and D.C. Porter *Basic econometrics.* (McGraw–Hill, 2009, International edition) 5th edition [ISBN 9780071276252].

To understand the subject clearly it is important to supplement Dougherty's *Introduction to econometrics* (fifth edition) with the subject guide **EC2020 Elements of econometrics** (2016), especially Chapter 10 which covers maximum likelihood estimation. It is very important to carefully go through the subject guide. The subject guide contains solutions to the questions given in the main textbook and also some additional questions and solutions. Working through these will improve your understanding of the subject.

The chapter in the subject guide on maximum likelihood (Chapter 10) includes some additional theory which has not been covered in the main textbook. It is important to read the additional theory given in the subject guide to have a better understanding of the principles of maximum likelihood and tests based on the likelihood function.

Please check the VLE course page for resources for this subject such as a downloadable copy of the subject guide **EC2020 Elements of econometrics** (2016), PowerPoint slideshows that provide a graphical treatment of the topics covered in the textbook, datasets and statistical tables. Candidates should utilise datasets using standard regression programmes (STATA or EViews). This will help in the understanding of the subject.

Examination revision strategy

Many candidates are disappointed to find that their examination performance is poorer than they expected. This may be due to a number of reasons, but one particular failing is '**question spotting**', that is, confining your examination preparation to a few questions and/or topics which have come up in past papers for the course. This can have serious consequences.

We recognise that candidates might not cover all topics in the syllabus in the same depth, but you need to be aware that examiners are free to set questions on **any aspect** of the syllabus. This means that you need to study enough of the syllabus to enable you to answer the required number of examination questions.

The syllabus can be found in the Course information sheet available on the VLE. You should read the syllabus carefully and ensure that you cover sufficient material in preparation for the examination. Examiners will vary the topics and questions from year to year and may well set questions that have not appeared in past papers. Examination papers may legitimately include questions on any topic in the syllabus. So, although past papers can be helpful during your revision, you cannot assume that topics or specific questions that have come up in past examinations will occur again.

If you rely on a question-spotting strategy, it is likely you will find yourself in difficulties when you sit the examination. We strongly advise you not to adopt this strategy.

Examiners' commentaries 2018

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2016–17. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone A

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Section A

Answer all questions from this section.

Question 1

We are interested in investigating the factors governing the precision of regression coefficients. Consider the model

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i$$

with OLS parameter estimates $\widehat{\beta}_1$, $\widehat{\beta}_2$ and $\widehat{\beta}_3$. Under the Gauss–Markov assumptions, we have

$$\text{Var}\left(\widehat{\beta}_2\right) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}(X_{2i} - \bar{X}_2)^2} \times \frac{1}{1 - r_{X_2X_3}^2},$$

where σ_ε^2 is the variance of ε and $r_{X_2X_3}$ is the sample correlation between X_2 and X_3.

(a) Provide four factors that help with obtaining more precise parameter estimates for, say, $\widehat{\beta}_2$.

(4 marks)

(b) In light of your answer to (a), discuss the concept of near multicollinearity. What consequences does its presence have when considering single and joint significance testing of our slope parameters?

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 3.3 (Properties of the multiple linear regression coefficients), 3.4 (Multicollinearity) and Chapter 3.5, Relationship between F statistic and t statistic).

Dougherty, C. Subject guide (2016): Chapter 3.

Approaching the question

(a) Candidates should note that the expression of the variance can be rewritten as:

$$\mathrm{Var}(\widehat{\beta}_2) = \frac{\sigma_\varepsilon^2}{n \times MSD(X_2) \times (1 - r^2_{X_2X_3})}$$

where $MSD(X_2) \equiv n^{-1} \sum (X_{2i} - \bar{X}_2)^2$. The four factors that affect the precision then are: n, $MSD(X_2)$, $r^2_{X_2X_3}$ and σ^2_ε. Therefore, to obtain more precise parameter estimates of β_2 it is desirable to have: (i) small error variance, (ii) large sample size, (iii) large sample variability of the regressor X_2, and (iv) small correlation among the regressors X_2 and X_3.

Some candidates gave a discussion of the Gauss–Markov assumptions which is not the answer.

(b) The issue of near multicollinearity is associated with the setting where $r^2_{X_2X_3}$ is close to 1, yielding very imprecise parameter estimates (large variance). Concept and consequences of multicollinearity are standard bookwork.

Question 2

Consider the linear regression model

$$\boldsymbol{Y_t = \beta_0 + \beta_1 X_t + \beta_2 Y_{t-1} + u_t, \quad t = 1, \ldots, T}$$

where the errors u_t are distributed independently of the regressors X_t and $|\beta_2| < 1$. You suspect that the, mean zero, errors exhibit autocorrelation.

(a) Explain what we mean by the concept of autocorrelation.

(2 marks)

(b) Assume that u_t follows an AR(1) process.

i. Discuss, for the given model, the consequences for the ordinary least squares estimator. Support your answers with suitable arguments.

(3 marks)

ii. Discuss how you would detect the presence of autocorrelation in the errors in this model. Clearly indicate the null and alternative hypothesis, the test statistic, and rejection rule.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 12.3 (Fitting a model subject to AR(1) autocorrelation, and Chapter 12.2 (Detection of autocorrelation).

Dougherty, C. Subject guide (2016): Chapter 12.

Approaching the question

(a) Candidates should clearly indicate what autocorrelation is (standard bookwork) and indicate that in the presence of the lagged endogenous variable Y_{t-1} this yields inconsistency as $\text{Cov}(Y_{t-1}, u_t) \neq 0$.

(b) i. Specifically, since ε_t is uncorrelated with X_{t-1}, Y_{t-2} and u_{t-1} and since the errors u_t are distributed independently of the regressors.

$$\begin{aligned}\text{Cov}(Y_{t-1}, u_t) &= \text{Cov}(\beta_0 + \beta_1 X_{t-1} + \beta_2 Y_{t-2} + u_{t-1}, \rho u_{t-1} + \varepsilon_t) \\ &= \beta_2 \rho \text{Cov}(Y_{t-2}, u_{t-1}) + \rho \text{Cov}(u_{t-1}, u_{t-1}).\end{aligned}$$

Using the fact that $\text{Cov}(Y_{t-2}, u_{t-1}) = \text{Cov}(Y_{t-1}, u_t)$ by covariance stationarity we get:

$$\text{Cov}(Y_{t-1}, u_t) = (1 - \beta_2 \rho)^{-1} \rho \sigma_u^2 \neq 0.$$

ii. To detect autocorrelation the Breusch–Godfrey test should be proposed, which requires us to run the following auxiliary regression:

$$\widehat{u}_t = \gamma_0 + \gamma_1 X_t + \gamma_2 Y_{t-1} + \rho \widehat{u}_{t-1} + v_t.$$

The use of the Durbin–Watson test is incorrect, a Durbin h-test may be proposed as well. Candidates should clearly indicate H_0 and H_1, the test statistic and the rejection rule.

Question 3

For the population of men who grew up with disadvantaged backgrounds, let *poverty* be a dummy variable equal to one if a man is currently living below the poverty line, and zero otherwise. The variable *age* is age and *educ* is total years of schooling. Let *vocat* be an indicator equal to unity if a man's high school offered vocational training. Using a random sample of 850 men, you obtain

$$\mathbf{Pr}(\boldsymbol{poverty} = 1 \,|\, \widehat{\boldsymbol{educ}}, \boldsymbol{age}, \boldsymbol{vocat}) = \boldsymbol{\Lambda}(0.453 - 0.016\boldsymbol{age} - 0.087\boldsymbol{educ} - 0.049\boldsymbol{vocat})$$

where $\boldsymbol{\Lambda}(z) = \exp(z) = (1 + \exp(z))$ is the logit function.

(a) It is argued that using the logit regression model is better than using the linear probability model when explaining the binary variable *poverty*. Discuss the benefits/drawback of using the logit regression model when trying to explain a binary variable.

(5 marks)

(b) For a 40-year old man, with 12 years of education, what is the estimated effect of having vocational training available in high school on the probability of currently living in poverty?

***Hint*: Clarity of computations required is enough, no need to give an exact number.**

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 10.1 (the linear probability model), Chapter 10.2 (logit analysis), and Chapter 10.6 (introduction to maximum likelihood estimation.

Dougherty, C. Subject guide (2016): Chapter 10.

Approaching the question

(a) The logit model has two main advantages over the linear probability model (LPM): predicted probabilities are restricted to lie in $[0, 1]$ and MLE is (asymptotically) efficient whereas OLS (LPM) will be inefficient given the inherent presence of heteroskedasticity.

The main drawbacks of the logit model relative to the linear probability model are that the coefficients cannot be directly interpreted as the marginal effects of the regressor(s) of interest and it is also computationally more complicated.

(b) Candidates will here need to observe that we need to compare predicted probabilities using the logit specification of the probabilities:

$$\Lambda(z) = \frac{1}{1 + \exp(-z)}.$$

For a 40-year old man with 12 years of education with vocational training the estimated probability of living in poverty is given by:

$$\Pr(y_i = 1 \,|\, age_i = 40, \widehat{educ}_i = 12, vocat_i = 1) = \frac{\exp(z_1)}{1 + \exp(z_1)} \approx 0.218$$

where $z_1 = 0.453 - 0.016 \times 40 - 0.087 \times 12 - 0.049 \times 1 \approx -1.28$.

The estimated probability of living in poverty for the same man without the vocational training is given by:

$$\Pr(y_i = 1 \,|\, age_i = 40; \widehat{educ}_i = 12; vocat_i = 0) = \frac{\exp(z_2)}{1 + \exp(z_2)} \approx 0.226$$

where $z_2 = z_1 + 0.049 \times 1 \approx -1.23$.

Therefore, for a 40-year old man with 12 years of education having vocational training in high school decreases the probability of living in poverty by 0.8 percentage points.

Question 4

The following model jointly determines monthly child support payments and monthly visitation rights for divorced couples with children:

$$\boldsymbol{support = \alpha_1 + \alpha_2 visits + \alpha_3 finc + \alpha_4 fremarr + \alpha_5 dist + \varepsilon_1}$$

$$\boldsymbol{visits = \beta_1 + \beta_2 support + \beta_3 mremarr + \beta_4 dist + \varepsilon_2}$$

We assume that children live with their mothers, so that fathers pay child support. Thus, the first equation is the father's 'reaction function': it describes the amount of child support paid for any given level of visitation rights and the other exogenous variables *finc* (father's income), *fremarr* (binary indicator if father remarried), and *dist* (miles currently between the mother and father's residence). Similarly the second equation is the mother's reaction function: it describes visitation rights for a given amount of child support; *mremarr* is a binary indicator for whether the woman is remarried.

(a) Examine the identification of each structural equation.

(3 marks)

(b) Your friend suggests you should implement the IV estimator to estimate the β parameters consistently. He tells you to use *finc* as instrument for *support*. Provide a critical discussion of this suggestion.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 8.3 (Instrumental variables), and Chapter 9 (Simultaneous equations estimation).

Dougherty, C. Subject guide (2016): Chapter 9.

Approaching the question

(a) The first equation is *exactly identified* since there is one exogenous variable ($mremarr_i$) available as instrument to the endogenous regressor ($visits_i$).

The second equation is *overidentified* since there are two exogenous variables ($finc_i$ and $fremarr_i$) available as instruments to the endogenous regressor ($support_i$).

Let G denote the number of equations in the system of simultaneous equations. It is true that in both equations there are $G-1$ endogenous variables (but you may be given a setting where there are fewer endogenous variables), and in each case you have excluded enough exogenous variables (k) you can use as instruments for these bad variables. Answers tend to be very vague.

(b) The IV approach suggested by the friend would yield consistent estimates for β parameters since $finc_i$ is exogenous and correlated with support (assuming $\alpha_3 \neq 0$). However, given this equation is overidentified, we could obtain more efficient estimates by using a two-stage least squares approach in which both $fremarr_i$ and $finc_i$ are included in the vector of instruments.

Question 5

Consider the simple linear regression model

$$Y_i = \beta_0 + \beta_1 X_i + u_i$$

under the classical linear regression model assumptions, where X_i is fixed under repeated sampling. The usual OLS estimators $\widehat{\beta}_0$ and $\widehat{\beta}_1$ are unbiased for their respective population parameters. Let $\tilde{\beta}_1$ be the estimator of β_1 obtained by assuming the intercept is zero.

(a) Show that the restricted least squares estimator of β_1 is given by

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}.$$

(4 marks)

(b) Find $\mathrm{E}(\tilde{\beta}_1)$ in terms of the X_i, β_0 and β_1. Verify that $\tilde{\beta}_1$ is unbiased for β_1 when the population intercept is zero. Are there other cases where $\tilde{\beta}_1$ is unbiased?

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 1.3 (Derivation of the regression coefficients) and Chapter 2.3 (The random components and unbiasedness of the OLS regression coefficients).

Dougherty, C. Subject guide (2016): Chapters 1 and 2.

Approaching the question

(a) Formally, restricted least squares estimates of β_0 and β_1 solve the following problem:

$$(\tilde{\beta}_0, \tilde{\beta}_1) = \min_{b_0, b_1} \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2, \quad \text{subject to } b_0 = 0.$$

This is the same as performing OLS on the model where $\beta_0 = 0$, that is performing OLS while leaving out the intercept.

Therefore, $\tilde{\beta}_0 = 0$ and $\tilde{\beta}_1 : \min_{b_1} \sum_{i=1}^{n} (Y_i - b_1 X_i)^2$. The first-order condition is given by:

$$-2 \sum_{i=1}^{n} (Y_i - \tilde{\beta}_1 X_i) X_i = 0 \qquad \Leftrightarrow \qquad \tilde{\beta}_1 = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2}.$$

(b) While candidates found the discussion of restricted least squares difficult, there was no reason not to answer the second part which was standard. Plug in the true model and use properties of sums to obtain:

$$\begin{aligned}
\tilde{\beta}_1 &= \left(\sum_{i=1}^{n} X_i^2 \right)^{-1} \left(\sum_{i=1}^{n} X_i (\beta_0 + \beta_1 X_i + u_i) \right) \\
&= \left(\sum_{i=1}^{n} X_i^2 \right)^{-1} \left(\sum_{i=1}^{n} X_i \right) \beta_0 + \beta_1 + \left(\sum_{i=1}^{n} X_i^2 \right)^{-1} \left(\sum_{i=1}^{n} X_i u_i \right).
\end{aligned}$$

Taking expectations using the fact that X_is are fixed and $\mathrm{E}(u_i) = 0$, we have:

$$\begin{aligned}
\mathrm{E}(\tilde{\beta}_1) &= \left(\sum_{i=1}^{n} X_i^2 \right)^{-1} \left(\sum_{i=1}^{n} X_i \right) \beta_0 + \beta_1 + \left(\sum_{i=1}^{n} X_i^2 \right)^{-1} \left(\sum_{i=1}^{n} X_i \mathrm{E}(u_i) \right) \\
&= \left(\sum_{i=1}^{n} X_i^2 \right)^{-1} \left(\sum_{i=1}^{n} X_i \right) \beta_0 + \beta_1.
\end{aligned}$$

Therefore, $\tilde{\beta}_1$ is unbiased if either: (i) $\beta_0 = 0$ or (ii) $\sum_{i=1}^{n} X_i = 0$.

Section B

Answer three questions from this section.

Question 6

Let us consider the estimation of a hedonic price function for houses. The hedonic price refers to the implicit price of a house given certain attributes (e.g., the number of bedrooms). The data contains the sale price of 546 houses sold in the summer of 1987 in Canada along with their important features. The following characteristics are available: the lot size of the property in square feet (*lotsize*), the numbers of bedrooms (*bedrooms*), the number of full bathrooms (*bathrooms*), and a dummy indicating the presence of airconditioning (*airco*).

Consider the following ordinary least squares results

$$\begin{aligned}
\widehat{\log(price)}_i = \underset{\substack{(.232)\\[.233]}}{7.094} + \underset{\substack{(.028)\\[.028]}}{0.400} \log(lotsize)_i + \underset{\substack{(.015)\\[.017]}}{0.078} bedrooms_i + \qquad (6.1)\\
\underset{\substack{(.023)\\[.024]}}{0.216} bathrooms_i + \underset{\substack{(.024)\\[.023]}}{0.212} airco_i \quad n = 546, \; RSS = 32.622
\end{aligned}$$

The usual standard errors are in parentheses, the heteroskedasticity robust standard errors are in square brackets, and RSS measures the residual sum of squares.

(a) **Interpret the parameter estimates on $\log(lotsize)$, $bedrooms$, and $airco$. Briefly discuss the statistical significance of the results.**

(5 marks)

(b) **Suppose that lot size was measured in square metres rather than square feet. How would this affect the parameter estimates of the slopes and intercept? How would this affect the fitted values? *Note*: the conversion (approximate) $1m^2 = 10ft^2$.**

(5 marks)

(c) **We are interested in testing the hypothesis $\mathrm{H}_0 : \beta_{bedrooms} = \beta_{bathrooms}$ against the alternative $H_1 : \beta_{bedrooms} \neq \beta_{bathrooms}$. Discuss a test for this hypothesis that makes use of the following restricted regression result**

$$\widehat{\log(price)}_i = \underset{(.234)}{6.994} + \underset{(.282)}{0.408}\log(lotsize)_i + \underset{(.011)}{0.127}bbrooms_i + \underset{(.024)}{0.215}airco_i \qquad (6.2)$$

$$n = 546, \ RSS = 33.758$$

where $bbrooms = bedrooms + bathrooms$. Clearly indicate the assumptions you are making for this test to be valid.

(5 marks)

(d) **You are interested in testing for the presence of heteroskedasticity. Say you are told that the variance is increasing with $\log(lotsize)$. Discuss how you would test for the presence of heteroskedasticity. What is the name of the test you are proposing?**

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 1.4 (Interpretation of a regression equation – units of measurement), Chapter 4.2 (Logarithmic transformations), Chapter 2.6 (Testing hypotheses relating to the regression coefficients), Chapter 7.1 (Heteroskedasticity and its implications), and Chapter 7.2 (Detection of heteroskedasticity).

Dougherty, C. Subject guide (2016): Chapter 7.

Approaching the question

(a) Clear discussion of interpretation required (units not always clear): On average, holding the remaining variables in the regression constant, (i) a 1% increase in lot size is associated with a 0.4% increase in house price, (ii) each extra bedroom is associated with a 7.8% increase in house price, and (iii) houses with air conditioning are 21.2% more expensive than those without. All estimates are statistically significant at 5% significance levels. Candidates should clearly indicate H_0 and H_1, the test statistic and the rejection rule.

(b) Let $lotsize_i$ be the lot size in square feet and $\widetilde{lotsize}_i$ be the lot size in square metres. We have that $\widetilde{lotsize}_i = (10)^{-1}lotsize_i$ and $\log(\widetilde{lotsize}_i) = \log((10)^{-1}) + \log(lotsize_i)$. Since this is an additive transformation of one of the explanatory variables we have that (i) the regression slopes will not be affected, (ii) the intercept will change to $7.094 - \log((10)^{-1}) \times 0.4$, and (iii) the fitted values will also not be affected. Many candidates made an error here, ignoring the fact that the variable whose measurement was changed entered in log form.

(c) Candidates would have to recognise that (6.2) is a restricted version of (6.1) where $\beta_{bedrooms} = \beta_{bathrooms}$ is imposed. We therefore need to use the F test. Test statistic:

$$F = \frac{RRSS - URSS}{URSS} \times \frac{n-K}{J} = \frac{33.758 - 32.622}{32.622} \times \frac{541}{1} \approx 18.84.$$

Assuming the Gauss–Markov assumptions plus normality of the error term hold: under H_0, $F \sim F_{1,541}$. At the 5% signicance level we reject H_0 since $F > 3.86$. Conclusion: The effect of one extra bathroom is different from the effect of one extra bedroom. Some candidates did not recognise this and were proposing a test on the coefficient of *bbrooms* equalling zero which is wrong.

(d) Assuming Gauss–Markov assumptions plus the normality of the error hold, he can use the Goldfeld–Quandt test for heteroskedasticity. For that purpose, we should first order the 546 observations by the magnitude of $\log(lotsize_i)$. Fit one regression for the first n^* observations and another for the last n^* observations (usually n^* equals one-third of the sample). Let RSS_1 and RSS_2 denote the sum of squared residuals in each of these regressions, respectively.

- $H_0 : \sigma_2^2 = \sigma_1^2$ vs. $H_1 : \sigma_2^2 > \sigma_1^2$.
- Test statistic: $GQ = RSS_2/RSS_1$.
- Under H_0, $GQ \sim F_{n^*-k,\, n^*-k}$.
- Reject if GQ is greater than the 95th percentile of the F distribution above.

Candidates need to be careful not to state $H_0 : RSS_2 = RSS_1$ vs. $H_1 : RSS_2 > RSS_1$. Both RSS_1 and RSS_2 are random variables. Because the sample sizes are identical it is also correct to state $H_0 : RSS_1$ and RSS_2 are not statistically different. Note that simply by having one sample larger than the other, you could also have a larger residual sum of squares.

Question 7

The following question concerns the effects of background characteristics and admission assessment scores on the performance of students in the final university examinations in a UK university. The following equation was estimated by Ordinary Least Squares:

$$\begin{aligned}\widehat{finalavg} = \underset{(2.78)}{53.89} + \underset{(.04)}{0.03}tst_reas + \underset{(.02)}{0.05}tst_quan + \underset{(.02)}{0.06}interview \\ -\underset{(.78)}{0.04}indep + \underset{(.66)}{0.67}male + \underset{(.97)}{0.06}indep*male \\ n = 325,\ R^2 = .06, \end{aligned} \tag{7.1}$$

where *finalavg* is the average finals score (the outcome), *tst_reas* and *tst_quan* are the pre-admission reasoning and quantitative test scores respectively, *interview* is the pre-admission interview score, *indep* indicates whether the student attended an independent school (1 = yes, 0 = no), and *male* indicates whether the student is male (1 = yes, 0 = no). The usual standard errors are in parentheses.

(a) We want to test whether gender has a significant impact on students' finals performance. Clearly indicating the null and the alternative hypothesis, provide the test statistic and the rejection rule. Discuss what information you would need to enable you to implement this test. You are expected to provide the assumptions which underlie your test.

(5 marks)

(b) If we do not include the interaction term *indep*male* in our regression model, what are we implicitly assuming about the effect of gender and school background on finals performance?

(5 marks)

(c) Suppose students who did not attend an independent school, attended a state school. Using the results in (7.1), provide the parameter estimates you would obtain if you had applied Ordinary Least Squares to the equation

$$\begin{aligned} finalavg = \beta_0 + \beta_1 tst_reas + \beta_2 tst_quan + \beta_3 interview \\ \beta_4 state + \beta_5 male + \beta_6 state*male + \varepsilon, \end{aligned} \tag{7.2}$$

where *state* indicates whether the student attended a state school (1 = yes, 0 = no).

(5 marks)

(d) Discuss any problem you may have in estimating the model if all males in your sample have attended an independent school prior to attending university. What name does this problem have and what can you do to mitigate this problem?

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 5.1–5.3 (Dummy variables).

Dougherty, C. Subject guide (2016): Chapter 5.

Approaching the question

(a) For given values of the remaining explanatory variables, the gender performance gap measured as the difference between male and female average final scores is given by β_5 if $indep_i = 0$ or $\beta_5 + \beta_6$ if $indep_i = 1$. Therefore, we want to perform the following test:

- $\mathrm{H}_0 : \beta_5 = 0$ and $\beta_6 = 0$ vs. $\mathrm{H}_1 : \beta_5 \neq 0$ or $\beta_6 \neq 0$.
- Need the R^2 of a regression of *finalavg* on a constant, *tst_reas*, *tst_quant*, *interview* and *indep* (the restricted model).
- Compute the test statistic:

$$F = \frac{R^2_{ur} - R^2_r}{1 - R^2_{ur}} \times \frac{n - K}{J} = \frac{0.06 - R^2_r}{1 - 0.06} \times \frac{325 - 7}{2}.$$

- Under Gauss–Markov assumptions plus normality, $F \sim F_{2,\,318}$.
- At the 5% signicance level, reject H_0 if $F > 3.07$.

(b) If we do not include the interaction term we are implicitly assuming that the effect of gender on performance is the same in independent and non-independent schools. We are also assuming that the effect of attending an independent school on final performance is the same for both male and female students.

(c) Let $\widehat{\theta}$s indicate the estimated parameters from the initial regression. From the original regression we have:

$$\begin{aligned}
\widehat{y}_i &= \widehat{\theta}_0 + \cdots + \widehat{\theta}_4 indep_i + \widehat{\theta}_5 male_i + \widehat{\theta}_6 indep_i \times male_i \\
&= \widehat{\theta}_0 + \cdots + \widehat{\theta}_4 (1 - state_i) + \widehat{\theta}_5 male_i + \widehat{\theta}_6 (1 - state_i) \times male_i \\
&= (\widehat{\theta}_0 + \widehat{\theta}_4) + \cdots - \widehat{\theta}_4 state_i + (\widehat{\theta}_5 + \widehat{\theta}_6) male_i - \widehat{\theta}_6 state_i \times male_i
\end{aligned}$$

where $\cdots = \widehat{\theta}_1 tst_reas_i + \widehat{\theta}_2 tst_quant_i + \widehat{\theta}_3 interview_i$. Mapping original estimates into estimated parameters in (7.2) yields:

- $\widehat{\beta}_0 = \widehat{\theta}_0 + \widehat{\theta}_4 = 53.89 - 0.04 = 53.85$
- $\widehat{\beta}_4 = -\widehat{\theta}_4 = -(-0.04) = 0.04$
- $\widehat{\beta}_6 = -\widehat{\theta}_6 = -0.06$
- $\widehat{\beta}_5 = \widehat{\theta}_5 + \widehat{\theta}_6 = 0.67 + 0.06 = 0.73$
- $\widehat{\beta}_2$, $\widehat{\beta}_3$ and $\widehat{\beta}_4$ remain the same from (7.1).

(d) If all males attended an independent school, then $male_i = male_i \times indep_i \; \forall \; i$, so the variables will be perfectly collinear and he would not be able to obtain the OLS estimates from (7.1). In this case, we could either drop the interaction term from the regression or obtain more data such that some men that have not attended independent school will also be in the sample.

Question 8

An OLS regression of y_t on x_t and x_{t-1} gives the following results (with the standard errors given in parentheses)

$$\widehat{y}_t = \underset{(2.30)}{8.88} + \underset{(3.01)}{5.07}\,x_t - \underset{(3.01)}{3.18}\,x_{t-1}; \quad R^2 = .095,\ T = 209 \tag{8.1}$$

(a) **What are the estimates of the short-run and long-run effect of x_t on y_t? Interpret these estimates.**

(4 marks)

(b) **Test the hypothesis that a one unit increase in x results in a ten unit increase in y in the same year. Under what assumptions is this test valid?**

(4 marks)

Let e_t be the OLS residuals from the above regression. An OLS regression of e_t on e_{t-1} yields

$$e_t = \underset{(2.12)}{0.55} + \underset{(0.18)}{0.44}\,e_{t-1} + \underset{(2.18)}{2.16}\,x_t - \underset{(.99)}{1.09}x_{t-1}; \quad R^2 = .175,\ T = 208 \tag{8.2}$$

(c) **Using this result, test for evidence of autocorrelation, clearly indicating the null and alternative hypotheses, the test statistic, rejection rule and assumptions underlying the test. What name do we give this test?**

(5 marks)

(d) **You are interested in testing whether the long-run effect of x_t on y_t is statistically significant.**

i. **Discuss how to reparameterise (8.1) to ensure that your regression output will provide you with a standard error for the long-run effect.**

(4 marks)

ii. **Discuss the problem of implementing your test using the standard error obtained in (d)i. when you do find evidence of autocorrelation in (8.1). Briefly indicate how you proceed with your test.**

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 6.5 (Testing a linear restriction), Chapter 11.3 (Models with lagged explanatory variables), and Chapter 12.1–12.3 (Definition, consequences and detection of autocorrelation; Fitting a model subject to AR(1) autocorrelation).

Dougherty, C. Subject guide (2016): Chapters 6 and 12.

Approaching the question

(a) Short-run effect: It shows the immediate effect a change in x with 1 unit has on y. Increases with 5.07 units.

Long-run effect: It shows the effect a permanent change in x with one unit has on y after 1 period (the last lagged response) has passed. Increases with $5.07 - 3.18 = 1.89$ units.

(b) Asked to perform the t test for $\mathrm{H}_0 : \beta_1 = 10$ vs. $\mathrm{H}_1 : \beta_1 \neq 10$ (standard bookwork). Validity of Gauss–Markov + normality.

(c) The test equation tells us that we have to perform the Breusch–Godfrey test for first-order autocorrelation. Let ρ denote the coefficient associated with e_{t-1}.

- $\mathrm{H}_0 : \rho = 0$ (No autocorrelation) vs. $\mathrm{H}_1 : \rho \neq 0$ (Autocorrelation).
- Test statistic: $LM = nR^2 = 28 \times 0.175 \approx 4.9$.

- Under H_0, $LM \overset{a}{\sim} \chi^2_1$.
- Reject H_0 in favour of autocorrelation since $LM > 3.84$.

(d) i. We need to reparameterise (8.1) to obtain the standard error of $\widehat{\beta}_1 + \widehat{\beta}_2$ directly:

$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_{t-1} + u_t$$
$$y_t = \beta_0 + \beta_1 x_t + \beta_2 x_t - \beta_2 x_t + \beta_2 x_{t-1} + u_t$$
$$y_t = \beta_0 + (\beta_1 + \beta_2) x_t - \beta_2 \Delta x_t + u_t.$$

Regress y_t on a constant, x_t and Δx_t and the standard error of the estimated coefficient on x_t is the standard error of the long-run effect.

ii. The usual standard errors will be invalidated so we need to use HAC standard errors instead. Given the robust standard errors, one could compute the test statistic robust to autocorrelation as $t^{robust} = \widehat{\theta}/se(\widehat{\theta})^{HAC}$ and conclude the long-run effect is statistically significant at the 5% significance level if $|t^{robust}| > 1.96$.

Question 9

Consider the model

$$\boldsymbol{y_t = \alpha + \beta t + \varepsilon_t, \quad t = 1, \ldots, T} \qquad \textbf{(9.1)}$$
$$\boldsymbol{\varepsilon_t = \rho \varepsilon_{t-1} + v_t, \quad \text{and}}$$

v_t is an i.i.d. $(0, \sigma^2)$ innovation which is independent of the past. Let $|\rho| \leq 1$.

(a) What name do we give the ε_t process given above? Provide the condition(s) that ensures that ε_t is stationary. In your answer discuss what we mean by the concept of stationarity (more precisely 'covariance stationarity').

(4 marks)

(b) It will be important to distinguish between the above process for y_t being 'trend stationary' as opposed to 'difference stationary'.

i. Explain these concepts clearly. Why is it important to distinguish between these two types of non-stationarity?

(4 marks)

ii. Show that under the condition you provided in (a) that y_t is trend stationary.

(2 marks)

iii. Show that if ε_t is difference stationary then y_t is difference stationary.

(2 marks)

(c) Show that you can rewrite the above model in the following form

$$\boldsymbol{\Delta y_t = \gamma_1 + \gamma_2 t + \gamma_3 y_{t-1} + v_t.} \qquad \textbf{(9.2)}$$

Clearly indicate the relation between $(\gamma_1, \gamma_2, \gamma_3)$ and (α, β, ρ).

(4 marks)

(d) What problem do you see here with using (9.2) to conducting the Dickey–Fuller Test to distinguish between trend and difference stationarity when v_t exhibits autocorrelation? What solution do you suggest we adopt?

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 12.1 (Definition and consequences of autocorrelation), Chapter 13.1 (Stationarity and nonstationarity), and Chapter 13.4–13.5 (Tests of nonstationarity).

Dougherty, C. Subject guide (2016): Chapter 13.

Approaching the question

(a) Standard bookwork.

(b) i. A process $\{y_t\}_{t=-\infty}^{\infty}$ is said to be difference stationary if its first-difference is stationary ($\Delta y_t \sim I(0)$). It is said to be trend stationary if the process $y_t - \beta t$ is stationary. It is important to distinguish between the two because the source of non-stationarity has different implications on how we proceed to obtain a stationary time-series to use for regression analysis and for statistical inference.

ii. If $|\rho| < 1$ then ε_t is stationary, since $y_t - \beta t = \alpha + \varepsilon_t$ is also stationary. Therefore, y_t is trend stationary.

iii. Suppose that ε_t is difference stationary ($\Delta\varepsilon_t \sim I(0)$). Taking first differences of (9.1) we obtain $\Delta y_t = \beta + \Delta\varepsilon_t$, since β is just a constant we have that $\Delta y_t \sim I(0)$, that is, y_t is difference stationary.

(c) Lagging (9.1) by one period and multiplying both sides by ρ we obtain:

$$\rho y_{t-1} = \rho\alpha + \rho\beta(t-1) + \rho\varepsilon_{t-1}.$$

Subtracting the above from (9.1) yields:

$$y_t - \rho y_{t-1} = (1-\rho)\alpha + \beta t - \rho\beta t + \rho\beta + \varepsilon_t - \rho\varepsilon_{t-1}.$$

Let $v_t \equiv \varepsilon_t - \rho\varepsilon_{t-1}$ and rearrange:

$$y_t = ((1-\rho)\alpha + \rho\beta) + \beta(1-\rho)t + \rho y_{t-1} + v_t.$$

Finally, subtract y_{t-1} on both sides to obtain:

$$\Delta y_t = \underbrace{(1-\rho)\alpha + \rho\beta}_{\gamma_1} + \underbrace{\beta(1-\rho)}_{\gamma_2}t + \underbrace{\rho}_{\gamma_3} y_{t-1} + v_t.$$

(d) If v_t exhibits autocorrelation it must be eliminated before running the test regression otherwise the Dickey–Fuller test will not be valid. To eliminate it we should include the lags of Δy_t in the test equation (9.2). This test is then known as the augmented Dickey–Fuller test.

Question 10

Let $math10$ denote the percentage of students at a high school receiving a passing score on a standardised math test. We are interested in estimating the effect of per student spending on math performance. A simple model is

$$math10_i = \beta_0 + \beta_1 \log(expend_i) + \beta_2 \log(enroll_i) + \beta_3 poverty_i + u_i \qquad (10.1)$$

where, for each high school i; $poverty_i$ is the percentage of students living in poverty, $expend_i$ is the spending per student and $enroll_i$ the number of registered students. You may assume that this model satisfies all Gauss–Markov assumptions.

You are faced with the fact that data is unavailable on a key variable: $poverty$.

(a) Discuss the properties (unbiasedness and consistency) of the estimators when you drop the variable poverty. Explain your answers.

(5 marks)

You do have information available on a closely related variable: the percentage of students eligible for the federally funded school lunch program, $lnchprg_i$. Let us consider using $lnchprg_i$ as a proxy for $poverty_i$.

(b) Briefly discuss why $lnchprg_i$ is a sensible proxy variable for the unobserved variable $poverty_i$.

(2 marks)

(c) It is unlikely that $lnchprg_i$ is an ideal proxy, in the sense that there is an exact linear relationship between them, instead, we will assume that

$$poverty_i = \alpha_0 + \alpha_1 lnchprg_i + v_i, \quad \alpha_1 \neq 0 \qquad (10.2)$$

Discuss the assumptions you need to make to enable consistent parameter estimators of β_1 and β_2 using your estimable equation

$$math10_i = \gamma_0 + \gamma_1 \log(expend_i) + \gamma_2 \log(enroll_i) + \gamma_3 lnchprg_i + e_i,$$

***Hint*: Consider the relation between the γ and the β parameters and express e_i in terms of u_i and v_i.**

(5 marks)

(d) The OLS results with and without $lnchprg_i$ as an explanatory variable are given by (standard errors in parentheses):

$$\widehat{math10}_i = -\underset{(26.72)}{69.24} + \underset{(3.30)}{11.13} \log expend_i + \underset{(0.615)}{0.022} \log emroll_i$$
$$N = 428, \; R^2 = 0.0297$$

$$\widehat{math10}_i = -\underset{(24.99)}{23.14} + \underset{(3.04)}{7.75} \log expend_i - \underset{(0.58)}{1.26} \log emroll_i - \underset{(0.036)}{0.324} lnchprg_i$$
$$N = 428, \; R^2 = 0.1893$$

i. Interpret the coefficient on $lnchprg$. What does this parameter tell us regarding the parameter of interest β_3?

(4 marks)

ii. Give an intuitive discussion explaining why the effect of expenditures on $math10_i$ is lower in the regression where $lnchprg_i$ is included than where it is excluded.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 6.2 (The effect of omitting a variable that ought to be included), Chapter 6.4 (Proxy variables).

Dougherty, C. Subject guide (2016): Chapter 6.

Approaching the question

(a) Consider rewriting (10.1) as:

$$math10_i = \beta_0 + \beta_1 \log(expend_i) + \beta_2 \log(enroll_i) + \varepsilon_i$$

where $\varepsilon_i = \beta_3 poverty_i + u_i$. Assuming $\beta_3 \neq 0$, if poverty is correlated with either (log) expenditures and/or (log) enrollment the model will suffer from endogeneity due to omitted variables and OLS will be biased and inconsistent. That is likely to be the case since schools that have smaller expenditures tend to be located in poorer neighbourhoods and hence to have more students living in poverty conditions.

(b) It is a sensible proxy because it is likely to be correlated with poverty and to capture some of the effect of poverty since usually students eligible for the lunch program tend to be those with low levels of family income.

(c) Plug in (10.2) for poverty in (10.1) to obtain:

$$math10_i = \underbrace{(\beta_0 + \beta_3\alpha_0)}_{\gamma_0} + \underbrace{\beta_1}_{\gamma_1} \log(expend_i) + \underbrace{\beta_2}_{\gamma_2} \log(enroll_i) + \underbrace{\beta_3\alpha_1}_{\gamma_3} lnchprg_i + \underbrace{(\beta_3 v_i + u_i)}_{\varepsilon_i}.$$

Given that the model (10.1) satisfies all the Gauss–Markov assumptions, to obtain consistent estimates of β_1 and β_2 we need that: (i) v_i is uncorrelated with $\log(expend_i)$, $\log(enroll_i)$ and $lnchprg_i$, and (ii) $lnchprg_i$ is uncorrelated with u_i.

(d) i. On average, holding expenditure and enrollment constant, a 1 percentage point increase in the number of students eligible for the lunch program is associated with a 0.324 percentage point fall in the percentage of students receiving a passing score in the standardised math test. Since $\gamma_3 = \alpha_1\beta_3$ and assuming $\alpha_1 > 0$, the direction of the effect (sign) of poverty of $math10$ is the same as the effect of $lnchprg$ on $math10$, in this case, with a negative coefficient on $lnchprg$ we can infer that poverty has a negative effect on $math10$.

ii. Omitting relevant variables will result in the remaining parameters attempting to pick up its effect through the correlation these omitted variables have with the included regressors. We, therefore, expect the effect to be smaller, as part of the effect we attribute to expenditure in the short regression is actually coming from the fact that high schools that have larger expenditures tend to have fewer students eligible for the lunch program and those students tend to perform worse in the standardised math test.

Examiners' commentaries 2018

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2017–18. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone B

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Section A

Answer all questions from this section.

Question 1

Consider the consumption function

$$C_t = \alpha + \lambda Y_t + \varepsilon_t \tag{1.1}$$

where C_t is aggregate consumption at t, λ is marginal propensity to consume ($0 < \lambda < 1$) and Y_t is aggregate income at t defined as

$$Y_t = C_t + A_t,$$

where A_t is the sum of investment and government consumption at t. Assume that A_t is uncorrelated with ε_t and that the shock ε_t is mean zero i.i.d. across t. A random sample of size n containing Y_t, C_t and A_t is available.

(a) Provide the reduced form equation for Y_t.

(2 marks)

(b) Show that the OLS estimator of λ in (1.1) is inconsistent. You are asked to indicate the direction of this inconsistency.

***Note*: you are not expected to derive the OLS estimator.**

(6 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 9.1 (Structural and reduced form equations) and Chapter 9.2 (Simultaneous equation bias).

Dougherty, C. Subject guide (2016): Chapter 9.

Approaching the question

(a) Plugging in the consumption function into the aggregate income identity:

$$Y_t = (\alpha + \lambda Y_t + \varepsilon_t) + A_t.$$

Rearranging yields, $Y_t = (1-\lambda)^{-1}\alpha + (1-\lambda)^{-1}A_t + (1-\lambda)^{-1}\varepsilon_t$.

(b) The OLS estimator of λ is given by:

$$\widehat{\lambda} = \frac{\sum_{t=1}^{T}(C_t - \bar{C})(Y_t - \bar{Y})}{\sum_{t=1}^{T}(Y_t - \bar{Y})^2} = \frac{T^{-1}\sum_{t=1}^{T}(C_t - \bar{C})(Y_t - \bar{Y})}{T^{-1}\sum_{t=1}^{T}(Y_t - \bar{Y})^2}.$$

Taking probability limits and using a suitable LLN:

$$\text{plim}(\widehat{\lambda}) = \frac{\text{Cov}(C_t, Y_t)}{\text{Var}(Y_t)} = \frac{\text{Cov}(\alpha + \lambda Y_t + \varepsilon_t, Y_t)}{\text{Var}(Y_t)} = \lambda + \frac{\text{Cov}(\varepsilon_t, Y_t)}{\text{Var}(Y_t)}.$$

Using the reduced form from part (a) we can show OLS is inconsistent since:

$$\text{Cov}(\varepsilon_t, Y_t) = (1-\lambda)^{-1}\text{Cov}(\varepsilon_t, \alpha + A_t + \varepsilon_t) = (1-\lambda)^{-1}\text{Var}(\varepsilon_t) \neq 0$$

where the second equality uses that A_t is uncorrelated with ε_t. Finally, since $\lambda \in (0,1)$ and both $\text{Var}(Y_t)$ and $\text{Var}(\varepsilon_t)$ are positive, there is a positive inconsistency in OLS estimator, in other words, propensity to consume is overestimated in (1.1).

Question 2

Consider the simple linear regression model

$$\boldsymbol{Y_t = \beta X_t + u_t, \qquad t = 1, \ldots, T}$$

where the errors u_t are distributed independently of the regressors X_t. You suspect that the, mean zero, errors exhibit autocorrelation.

(a) Explain what we mean by the concept of autocorrelation.

(2 marks)

(b) Assume you are told that u_t follows an MA(1) process.

i. Discuss whether the OLS estimator $\widehat{\beta}$ is a consistent estimator for β. Justify your answers with suitable technical derivations.

***Note*: you are not expected to derive the OLS estimator.**

(3 marks)

ii. Suppose you want to test $H_0 : \beta = 1$ against $H_1 : \beta < 1$. Discuss how you would conduct this test based on the OLS estimator, recognising the presence of autocorrelation in the error.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 6.5 (Testing a linear restriction), Chapter 12.1–12.3 (Definition, consequences and detection of autocorrelation; Fitting a model subject to AR(1) autocorrelation).

Dougherty, C. Subject guide (2016): Chapters 6 and 12.

Approaching the question

(a) Candidates should indicate clearly what autocorrelation is (standard bookwork).

(b) i. The OLS estimator is given by:

$$\widehat{\beta} = \frac{\sum_{t=1}^{T} X_t Y_t}{\sum_{t=1}^{T} X_t^2} = \frac{T^{-1}\sum_{t=1}^{T} X_t Y_t}{T^{-1}\sum_{t=1}^{T} X_t^2}.$$

To analyse consistency we take probability limits and apply a LLN to obtain:

$$\text{plim}(\widehat{\beta}) = \frac{\text{E}(X_t Y_t)}{\text{E}(X_t^2)} = \frac{\text{E}(X_t(\beta X_t + u_t))}{\text{E}(X_t^2)} = \beta + \frac{\text{E}(X_t u_t)}{\text{E}(X_t^2)}.$$

The OLS estimator is consistent as long as $\text{E}(X_t u_t) = 0$ (and $\text{E}(X_t^2) \neq 0$). This is satisfied since the zero mean errors u_t are assumed to be independent (and hence uncorrelated) of the regressors X_t.

ii. Recognising the presence of autocorrelation in the error when testing the single linear restriction we need to make use of HAC standard errors. This fact was ignored by many. The test statistic we use is $t = (\widehat{\beta} - 1)/se(\widehat{\beta})^{HAC}$ and we should reject H_0 when t is too small. Under H_0, $t \overset{a}{\sim} N(0, 1)$ (also accepted would be $t \sim t_{n-k}$). At the 5% significance level we reject H_0 if $t < -1.645$.

Question 3

Consider the simple linear regression model

$$\boldsymbol{Y_i = \beta_0 + \beta_1 X_i + u_i}$$

under the classical linear regression model assumptions, where X_i is fixed under repeated sampling. The usual OLS estimators $\widehat{\beta}_0$ and $\widehat{\beta}_1$ are unbiased for their respective population parameters. Let $\tilde{\beta}_0$ be the estimator of β_0 when β_1 equals 1.

(a) Show that the restricted least squares estimator of β_0 is given by

$$\boldsymbol{\tilde{\beta}_0 = \bar{Y} - \bar{X}}$$

where $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$ and $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.

(4 marks)

(b) Find $\text{E}(\tilde{\beta}_0)$ in terms of the X_i, β_0 and β_1. Verify that $\tilde{\beta}_0$ is unbiased for β_0 when $\beta_1 = 1$. Are there other cases where $\tilde{\beta}_0$ is unbiased?

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 1.3 (Derivation of the regression coefficients) and Chapter 2.3 (The random components and unbiasedness of the OLS regression coefficients).

Dougherty, C. Subject guide (2016): Chapters 1 and 2.

Approaching the question

(a) Formally, restricted least squares estimates of β_0 and β_1 solve the following problem:

$$(\tilde{\beta}_0, \tilde{\beta}_1) = \min_{b_0, b_1} \sum_{i=1}^{n} (Y_i - b_0 - b_1 X_i)^2, \quad \text{subject to } b_1 = 1.$$

This is the same as performing OLS on the model where $\beta_1 = 1$.

Therefore, $\tilde{\beta}_1 = 1$ and $\tilde{\beta}_0 = \min_{b_0} \sum_{i=1}^{n} (Y_i - b_1 X_i)^2$. The first-order condition is given by:

$$-2 \sum_{i=1}^{n} (Y_i - \tilde{\beta}_0 - X_i) = 0 \qquad \Leftrightarrow \qquad \tilde{\beta}_0 = \bar{Y} - \bar{X}.$$

(b) While candidates found the discussion of restricted least squares difficult, there was no reason not to answer the second part which was standard. Using our estimator, notice that averaging the true model yields $\bar{Y} = \beta_0 + \beta_1 \bar{X} + \bar{u}$. Plugging in our estimator $\tilde{\beta}_0$, we have:

$$\tilde{\beta}_0 = (\beta_0 + \beta_1 \bar{X} + \bar{u}) - \bar{X} = \beta_0 + (\beta_1 - 1)\bar{X} + \bar{u}.$$

Taking expectations using the fact that X_is are fixed and $\mathrm{E}(u_i) = 0$, we have:

$$\begin{aligned} E(\tilde{\beta}_0) &= \beta_0 + (\beta_1 - 1)\bar{X} + n^{-1} \sum_{i=1}^{n} \mathrm{E}(u_i) \\ &= \beta_0 + (\beta_1 - 1)\bar{X}. \end{aligned}$$

Therefore, $\tilde{\beta}_0$ is unbiased if either: (i) $\beta_1 = 1$, or (ii) $\bar{X} = 0$.

Question 4

We are interested in investigating the factors governing the precision of regression coefficients. Consider the model

$$\boldsymbol{Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + \varepsilon_i}$$

with OLS parameter estimates $\widehat{\beta}_1$, $\widehat{\beta}_2$ and $\widehat{\beta}_3$. Under the Gauss–Markov assumptions, we have

$$\mathbf{Var}\left(\widehat{\beta}_2\right) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n} (X_{2i} - \bar{X}_2)^2} \times \frac{1}{1 - r_{X_2 X_3}^2},$$

where σ_ε^2 is the variance of ε and $r_{X_2 X_3}$ is the sample correlation between X_2 and X_3.

(a) Provide four factors that help with obtaining more precise parameter estimates for, say, $\widehat{\beta}_2$.

(4 marks)

(b) Assume that the true value of $\beta_3 = 0$, so that the above model includes an irrelevant variable. Discuss the effect of including this irrelevant variable on the unbiasedness and precision of $\widehat{\beta}_2$.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 3.3 (Properties of the multiple linear regression coefficients), Chapter 6.3 (The effect of including a variable that ought not to be included).

Dougherty, C. Subject guide (2016): Chapter 3.

Approaching the question

(a) Candidates should note that the expression of the variance can be rewritten as:

$$\mathrm{Var}(\widehat{\beta}_2) = \frac{\sigma_\varepsilon^2}{n \times MSD(X_2) \times (1 - r^2_{X_2X_3})}$$

where $MSD(X_2) \equiv n^{-1}\sum(X_{2i} - \bar{X}_2)^2$. The four factors that affect the precision then are: n, $MSD(X_2)$, $r^2_{X_2X_3}$ and σ_ε^2. Therefore, to obtain more precise parameter estimates of β_2 it is desirable to have: (i) small error variance, (ii) large sample size, (iii) large sample variability of the regressor X_2, and (iv) small correlation among the regressors X_2 and X_3.

Some candidates gave a discussion of the Gauss–Markov assumptions which is not the answer.

(b) Concept and consequences of including irrelevant variables are standard bookwork – unbiased but less precise because the irrelevant variable typically is correlated with the included regressor, i.e. $r^2_{X_2X_3} \neq 0$.

More precisely: Let $\tilde{\beta}_2$ be the OLS estimator in a regression without including X_3 and $\widehat{\beta}_2$ be the OLS estimator in a regression with X_3. If $r_{X_2,X_3} \neq 0$ we have:

$$\mathrm{Var}\left(\widehat{\beta}_2\right) = \frac{\sigma_\varepsilon^2}{\sum\limits_{i=1}^{n}(X_{2i} - \bar{X}_2)^2} \times \frac{1}{1 - r^2_{X_2X_3}} > \frac{\sigma_\varepsilon^2}{\sum\limits_{i=1}^{n}(X_{2i} - \bar{X}_2)^2} \equiv \mathrm{Var}(\tilde{\beta}_2).$$

Question 5

A probit model to explain whether a firm is taken over by another firm during a given year postulates

$$\mathbf{Pr}(\boldsymbol{takeover} = 1 \mid \boldsymbol{x}) = \boldsymbol{\Phi}(\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1\boldsymbol{avgprof} + \boldsymbol{\beta}_2\boldsymbol{mtkval} + \boldsymbol{\beta}_3\boldsymbol{debtearn} + \boldsymbol{\beta}_4\boldsymbol{ceoten} + \boldsymbol{\beta}_5\boldsymbol{ceosal} + \boldsymbol{\beta}_6\boldsymbol{ceoage})$$

where $\boldsymbol{\Phi}(z)$ is the cumulative standardised normal distribution. *takeover* is a binary response variable, *avgprof* is the firm's average profit margin over several prior years, *mktval* is the market value of the firm, *debtearn* is the debt-to-earnings ratio, and *ceoten*, *ceosal*, and *ceoage* are the tenure, annual salary, and age of the chief executive officer, respectively.

(a) It is argued that using the probit regression model is better than using the linear probability model when explaining the binary variable *takeover*. Discuss the benefits/drawback of using the probit regression model when trying to explain a binary variable.

(5 marks)

(b) Discuss how you would implement the LR test that variables related to the CEO have no effect on the probability of takeover, other factors being equal. Clearly indicate the null, alternative, test statistic and rejection rule.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 10.1 (The linear probability model), Chapter 10.3 (Probit analysis), and Chapter 10.6 (Introduction to maximum likelihood estimation.

Dougherty, C. Subject guide (2016): Chapter 10.

Approaching the question

(a) The probit model has two main advantages over the linear probability model (LPM): predicted probabilities are restricted to lie in $[0, 1]$ and MLE is (asymptotically) efficient whereas OLS (LPM) will be inefficient given the inherent presence of heteroskedasticity.

The main drawbacks of the probit model relative to the linear probability model is that the coefficients cannot be directly interpreted as the marginal effects of the regressor(s) of interest and it is also computationally more complicated.

(b) Let $\ln L^U$ be the maximised value of the log-likelihood of the unrestricted model and $\ln L^R$ be the maximised value of the log-likelihood of a model that excludes all the CEO-related variables (*ceoten*, *ceosal* and *ceoage*). Perform a likelihood-ratio test.

- $\mathrm{H}_0 : \beta_4 = \beta_5 = \beta_6 = 0$ vs. $\mathrm{H}_1 : \beta_4 \neq 0$ or $\beta_5 \neq 0$ or $\beta_6 \neq 0$.
- Test statistic: $LR = 2 \times (\ln L^U - \ln L^R)$.
- Under H_0, $LR \overset{a}{\sim} \chi^2_3$.
- At the 5% significance level, reject H_0 if $LR > 7.815$.

Section B

Answer three questions from this section.

Question 6

The following question concerns the effects of background characteristics on student's performance in the SAT (Scholastic Assessment Test). The SAT test is used for college admissions in the US.

$$\widehat{sat} = \underset{(6.29)}{1,028.10} + \underset{(3.83)}{19.3}\, hsize + \underset{(.53)}{-2.19} hsize^2 - \underset{(4.29)}{45.09} female - \underset{(12.71)}{169.81} black$$
$$+ \underset{(12.71)}{62.31} female{*}black$$

$$n = 4,127, \;\; R^2 = .0858$$

The variable *hsize* is the size of the student's high school graduating class, in hundreds, *female* is a gender dummy variable (1 = female, 0 = male), and *black* is a race dummy variable (1 = black, 0 = otherwise). The standard errors are in parentheses.

(a) What is the economic rationale for including $hsize^2$ in the above regression? Using this equation, determine for a given gender and race, what the graduating class size would be at which the predicted SAT scores are maximised.

(5 marks)

(b) Holding *hsize* fixed, what is the estimated difference in SAT scores between nonblack females and nonblack males? Is this difference statistically significant? Interpret this result.

(5 marks)

(c) What is the estimated difference in SAT score between black females and nonblack females? What would you need to do to test whether the difference is statistically significant?

(5 marks)

(d) Discuss any problem you may have in estimating the model if all females in your sample are black. What name does this problem have and what can you do to mitigate this problem?

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 4.3 (Models with quadratic variables), Chapter 6.5 (Testing a linear restriction), Chapter 5.1–5.3 (Dummy variables).

Dougherty, C. Subject guide (2016): Chapter 5.

Approaching the question

(a) The economic rationale for including a quadratic term in class size is that it allows for diminishing effects of class size on student's performance. The optimal class size is given by:

$$\left.\frac{\partial \widehat{SAT}}{\partial hsize}\right]_{hsize^*} = 0 \quad \Leftrightarrow \quad 19.3 - 2 \times 2.19 \times hsize^* = 0 \quad \Leftrightarrow \quad hsize^* = \frac{19.3}{4.38} \approx 4.4.$$

The optimal graduating class size is about 440 students.

(b) The expected difference in SAT scores between nonblack females and nonblack males is given by:

$$\mathrm{E}(sat \,|\, hsize, black = 0, female = 1) - \mathrm{E}(sat \,|\, hsize, black = 0, female = 0) = \beta_3.$$

The estimated difference is -45.03 which means that on average, holding the size of graduating class constant, nonblack females have a SAT score that is 45 points below nonblack males. To test the significance of this difference we perform a hypothesis test:

- $\mathrm{H}_0 : \beta_3 = 0$ vs. $\mathrm{H}_1 : \beta_3 \neq 0$.
- Test statistic: $t = \widehat{\beta}_3 / se(\widehat{\beta}_3) = -45.09/4.29 \approx -10.5$.
- Under H_0 and assuming Gauss–Markov plus normality, $t \sim t_{4121}$.
- At the 5% significance level, we reject H_0 since $|t| > 1.96$.

(c) The expected difference in SAT scores between black females and nonblack females is given by:

$$\mathrm{E}(sat \,|\, hsize, black = 1, female = 1) - \mathrm{E}(sat \,|\, hsize, black = 0, female = 1) = \beta_4 + \beta_5.$$

The estimated difference is given by $-169.81 + 62.31 = -107.5$ which means that, holding size of graduating class constant, on average black females have a sat score that is 107.5 points lower than their white counterparts. To perform the test of $\mathrm{H}_0 : \beta_4 + \beta_5 = 0$ vs. $\mathrm{H}_1 : \beta_4 + \beta_5 \neq 0$ one would need either: (i) $\widehat{\mathrm{Cov}(\widehat{\beta}_4, \widehat{\beta}_5)}$ to compute the t statistic, or (ii) the R^2 of the restricted model to compute the F statistic.

(d) If all females are black, then $female_i = female_i \times black_i \;\forall\; i$, so the variables will be perfectly collinear and he would not be able to obtain the OLS estimates. In this case, we could either drop the interaction term from the regression at the expense of not being able to identify any heterogeneity in the effects or obtain more data such that the sample contains some white females.

Question 7

Let us consider the estimation of a hedonic price function for houses. The hedonic price refers to the implicit price of a house given certain attributes (e.g., the number of bedrooms). The data contains the sale price of 546 houses sold in the summer of 1987 in Canada along with their important features. The following characteristics are available: the lot size of the property in square feet ($lotsize$), the numbers of bedrooms ($bedrooms$), the number of full bathrooms ($bathrooms$), and a dummy indicating the presence of airconditioning ($airco$).

Consider the following ordinary least squares results

$$\widehat{\log(price)}_i = \underset{\substack{(.232)\\[.233]}}{7.094} + \underset{\substack{(.028)\\[.028]}}{0.400}\log(lotsize)_i + \underset{\substack{(.015)\\[.017]}}{0.078}bedrooms_i + \underset{\substack{(.023)\\[.024]}}{0.216}bathrooms_i + \underset{\substack{(.024)\\[.023]}}{0.212}airco_i \quad n = 546,\ RSS = 32.622 \tag{7.1}$$

The usual standard errors are in parentheses, the heteroskedasticity robust standard errors are in square brackets, and RSS measures the residual sum of squares.

(a) **Interpret the parameter estimates on $\log(lotsize)$, $bedrooms$, and $airco$. Briefly discuss the statistical significance of the results.**

(5 marks)

(b) **Suppose that lot size was measured in square metres rather than square feet. How would this affect the parameter estimates of the slopes and intercept? How would this affect the fitted values? *Note*: the conversion (approximate) $1m^2 = 10ft^2$.**

(5 marks)

(c) **We are interested in testing the hypothesis $H_0 : \beta_{bedrooms} = \beta_{bathrooms}$ against the alternative $H_1 : \beta_{bedrooms} \neq \beta_{bathrooms}$. Discuss a test for this hypothesis that makes use of the following restricted regression result**

$$\widehat{\log(price)}_i = \underset{(.234)}{6.994} + \underset{(.282)}{0.408}\log(lotsize)_i + \underset{(.011)}{0.127}bbrooms_i + \underset{(.024)}{0.215}airco_i \tag{7.2}$$
$$n = 546,\ RSS = 33.758$$

where $bbrooms = bedrooms + bathrooms$. Clearly indicate the assumptions you are making for this test to be valid.

(5 marks)

(d) **You are interested in testing for the presence of heteroskedasticity. Say you are told that the variance is increasing with $\log(lotsize)$. Discuss how you would test for the presence of heteroskedasticity. What is the name of the test you are proposing?**

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 1.4 (Interpretation of a regression equation – units of measurement), Chapter 4.2 (Logarithmic transformations), Chapter 2.6 (Testing hypotheses relating to the regression coefficients), Chapter 7.1 (Heteroskedasticity and its implications), and Chapter 7.2 (Detection of heteroskedasticity).

Dougherty, C. Subject guide (2016): Chapter 7.

Approaching the question

(a) Clear discussion of interpretation required (units not always clear): On average, holding the remaining variables in the regression constant, (i) a 1% increase in lot size is associated with a 0.4% increase in house price, (ii) each extra bedroom is associated with a 7.8% increase in house price, and (iii) houses with air conditioning are 21.2% more expensive than those without. All estimates are statistically significant at 5% significance levels. Candidates should clearly indicate H_0 and H_1, the test statistic and the rejection rule.

(b) Let $lotsize_i$ be the lot size in square feet and $\widetilde{lotsize}_i$ be the lot size in square metres. We have that $\widetilde{lotsize}_i = (10)^{-1} lotsize_i$ and $\log(\widetilde{lotsize}_i) = \log((10)^{-1}) + \log(lotsize_i)$. Since this is an additive transformation of one of the explanatory variables we have that (i) the regression slopes will not be affected, (ii) the intercept will change to $7.094 - \log((10)^{-1}) \times 0.4$, and (iii) the fitted values will also not be affected. Many candidates made an error here, ignoring the fact that the variable whose measurement was changed entered in log form.

(c) Candidates would have to recognise that (6.2) is a restricted version of (6.1) where $\beta_{bedrooms} = \beta_{bathrooms}$ is imposed. We therefore need to use the F test. Test statistic:

$$F = \frac{RRSS - URSS}{URSS} \times \frac{n-K}{J} = \frac{33.758 - 32.622}{32.622} \times \frac{541}{1} \approx 18.84.$$

Assuming the Gauss–Markov assumptions plus normality of the error term hold: under H_0, $F \sim F_{1,\,541}$. At the 5% signicance level we reject H_0 since $F > 3.86$. Conclusion: The effect of one extra bathroom is different from the effect of one extra bedroom. Some candidates did not recognise this and were proposing a test on the coefficient of *bbrooms* equalling zero which is wrong.

(d) Assuming Gauss–Markov assumptions plus the normality of the error hold, he can use the Goldfeld–Quandt test for heteroskedasticity. For that purpose, we should first order the 546 observations by the magnitude of $\log(lotsize_i)$. Fit one regression for the first n^* observations and another for the last n^* observations (usually n^* equals one-third of the sample). Let RSS_1 and RSS_2 denote the sum of squared residuals in each of these regressions, respectively.

- $H_0 : \sigma_2^2 = \sigma_1^2$ vs. $H_1 : \sigma_2^2 > \sigma_1^2$.
- Test statistic: $GQ = RSS_2/RSS_1$.
- Under H_0, $GQ \sim F_{n^*-k,\,n^*-k}$.
- Reject if GQ is greater than the 95th percentile of the F distribution above.

Candidates need to be careful not to state $H_0 : RSS_2 = RSS_1$ vs. $H_1 : RSS_2 > RSS_1$. Both RSS_1 and RSS_2 are random variables. Because the sample sizes are identical it is also correct to state $H_0 : RSS_1$ and RSS_2 are not statistically different. Note that simply by having one sample larger than the other, you could also have a larger residual sum of squares.

Question 8

Let *math*10 denote the percentage of students at a high school receiving a passing score on a standardised math test. We are interested in estimating the effect of per student spending on math performance. A simple model is

$$math10_i = \beta_0 + \beta_1 \log(expend_i) + \beta_2 \log(enroll_i) + \beta_3 poverty_i + u_i \qquad (8.1)$$

where, for each high school i; $poverty_i$ is the percentage of students living in poverty, $expend_i$ is the spending per student and $enroll_i$ the number of registered students. You may assume that this model satisfies all Gauss–Markov assumptions.

You are faced with the fact that data is unavailable on a key variable: *poverty*.

(a) **Discuss the properties (unbiasedness and consistency) of the estimators when you drop the variable poverty. Explain your answers.**

(5 marks)

You do have information available on a closely related variable: the percentage of students eligible for the federally funded school lunch program, $lnchprg_i$. Let us consider using $lnchprg_i$ as a proxy for $poverty_i$.

(b) **Briefly discuss why $lnchprg_i$ is a sensible proxy variable for the unobserved variable $poverty_i$.**

(2 marks)

(c) **It is unlikely that $lnchprg_i$ is an ideal proxy, in the sense that there is an exact linear relationship between them, instead, we will assume that**

$$poverty_i = \alpha_0 + \alpha_1 lnchprg_i + v_i, \quad \alpha_1 \neq 0 \qquad (8.2)$$

Discuss the assumptions you need to make to enable consistent parameter estimators of β_1 and β_2 using your estimable equation

$$math10_i = \gamma_0 + \gamma_1 \log(expend_i) + \gamma_2 \log(enroll_i) + \gamma_3 lnchprg_i + e_i,$$

***Hint*: Consider the relation between the γ and the β parameters and express e_i in terms of u_i and v_i.**

(5 marks)

(d) **The OLS results with and without $lnchprg_i$ as an explanatory variable are given by (standard errors in parentheses):**

$$\widehat{math10}_i = \underset{(26.72)}{-69.24} + \underset{(3.30)}{11.13} \log expend_i + \underset{(0.615)}{0.022} \log emroll_i$$
$$N = 428, \; R^2 = 0.0297$$

$$\widehat{math10}_i = \underset{(24.99)}{-23.14} + \underset{(3.04)}{7.75} \log expend_i - \underset{(0.58)}{1.26} \log emroll_i - \underset{(0.036)}{0.324} lnchprg_i$$
$$N = 428, \; R^2 = 0.1893$$

i. **Interpret the coefficient on $lnchprg$. What does this parameter tell us regarding the parameter of interest β_3?**

(4 marks)

ii. **Give an intuitive discussion explaining why the effect of expenditures on $math10_i$ is lower in the regression where $lnchprg_i$ is included than where it is excluded.**

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 6.2 (The effect of omitting a variable that ought to be included), Chapter 6.4 (Proxy variables).

Dougherty, C. Subject guide (2016): Chapter 6.

Approaching the question

(a) Consider rewriting (8.1) as:

$$math10_i = \beta_0 + \beta_1 \log(expend_i) + \beta_2 \log(enroll_i) + \varepsilon_i$$

where $\varepsilon_i = \beta_3 poverty_i + u_i$. Assuming $\beta_3 \neq 0$, if poverty is correlated with either (log) expenditures and/or (log) enrollment the model will suffer from endogeneity due to omitted variables and OLS will be biased and inconsistent. That is likely to be the case since schools that have smaller expenditures tend to be located in poorer neighbourhoods and hence to have more students living in poverty conditions.

(b) It is a sensible proxy because it is likely to be correlated with poverty and to capture some of the effect of poverty since usually students eligible for the lunch program tend to be those with low levels of family income.

(c) Plug in (8.2) for poverty in (8.1) to obtain:

$$math10_i = \underbrace{(\beta_0 + \beta_3\alpha_0)}_{\gamma_0} + \underbrace{\beta_1}_{\gamma_1} \log(expend_i) + \underbrace{\beta_2}_{\gamma_2} \log(enroll_i) + \underbrace{\beta_3\alpha_1}_{\gamma_3} lnchprg_i + \underbrace{(\beta_3 v_i + u_i)}_{\varepsilon_i}.$$

Given that the model (8.1) satisfies all the Gauss–Markov assumptions, to obtain consistent estimates of β_1 and β_2 we need that: (i) v_i is uncorrelated with $\log(expend_i)$, $\log(enroll_i)$ and $lnchprg_i$, and (ii) $lnchprg_i$ is uncorrelated with u_i.

(d) i. On average, holding expenditure and enrollment constant, a 1 percentage point increase in the number of students eligible for the lunch program is associated with a 0.324 percentage point fall in the percentage of students receiving a passing score in the standardised math test. Since $\gamma_3 = \alpha_1\beta_3$ and assuming $\alpha_1 > 0$, the direction of the effect (sign) of poverty of $math10$ is the same as the effect of $lnchprg$ on $math10$, in this case, with a negative coefficient on $lnchprg$ we can infer that poverty has a negative effect on $math10$.

ii. Omitting relevant variables will result in the remaining parameters attempting to pick up its effect through the correlation these omitted variables have with the included regressors. We, therefore, expect the effect to be smaller, as part of the effect we attribute to expenditure in the short regression is actually coming from the fact that high schools that have larger expenditures tend to have fewer students eligible for the lunch program and those students tend to perform worse in the standardised math test.

Question 9

Let us consider monthly data on the short-term interest rate (the three month Treasury Bill rate) and on the AAA corporate bond yield in the USA. The data run from January 1950 to December 1999. Let $DUS3MT$ denote the changes in three-month Treasury Bill rate, and $DAAA$ denote the changes in AAA bond rate. We consider the following results (with the standard errors given in parentheses)

$$\widehat{DAAA}_t = \underset{(.007)}{0.006} + \underset{(.015)}{0.275} DUS3MT_t \quad t = 1, \ldots, 600 \tag{9.1}$$

$$RSS = 17.486; \; DW = 1.447$$

where RSS is the residual sum of squares and DW is the Durbin–Watson test.

A researcher interpreting the residuals suggests that the errors show a positive correlation over time.

(a) What are the consequences of this correlation for the above regression results?

(5 marks)

(b) Use the results above to test for the presence of first-order positive autocorrelation. Clearly specify the null and alternative hypothesis, test statistic, assumptions underlying the test, and the acceptance/rejection rule.

(5 marks)

(c) In an attempt to remove the autocorrelation you consider the following specification

$$\widehat{DAAA}_t = \underset{(.007)}{0.005} + \underset{(.015)}{0.252} DUS3MT_t - \underset{(.018)}{0.080} DUS3MT_{t-1} + \underset{(.040)}{0.290} DAAA_{t-1} \tag{9.2}$$

$$RSS = 16.087; \; DW = 1.897$$

Comment on the following statement 'The Durbin–Watson statistic is closer to 2, indicating that we have succesfully removed the autocorrelation'. If you disagree with this statement, suggest what you would need to do instead.

(5 marks)

(d) Discuss the Common Factor Test as a model specification suitable for this model. What extra information do you need to conduct this test.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 12.1–12.3 (Definition, consequences and detection of autocorrelation; Fitting a model subject to AR(1) autocorrelation).

Dougherty, C. Subject guide (2016): Chapters 12.

Approaching the question

(a) Standard bookwork discussion of autocorrelation in the setting where there are no lagged dependent variables.

(b) Discussion of Durbin–Watson test expected – bookwork. Candidates should clearly indicate the H_0 and H_1, test statistic and rejection rule. Assumptions underlying the test: tests only AR(1) autocorrelation, no lagged dependent variables (deterministic regressors only) in presence of intercept.

(c) We disagree with the statement because the Durbin–Watson test is not valid in the presence of lagged dependent variables since in this case regressors cannot be strictly exogenous. Instead, we should use the Breusch–Godfrey test which is asymptotically valid with predetermined regressors a weaker requirement than strict exogeneity. Alternatively, the Durbin h test can be proposed.

(d) Assuming AR(1) autocorrelation and letting $Y_t = DAAA_t$ and $X_t = DUS3MT_t$ we can remove the autocorrelation by rewriting the model as:

$$Y_t = (1-\rho)\beta_1 + \rho Y_{t-1} + \beta_2 X_t - \beta_2 \rho X_{t-1} + \varepsilon_t.$$

This is a restricted version of a more general ADL(1,1) model:

$$Y_t = \lambda_1 + \lambda_2 Y_{t-1} + \lambda_3 X_t + \lambda_4 X_{t-1} + \varepsilon_t$$

under the restriction $\lambda_4 = -\lambda_2\lambda_3$. We can test this restriction using the Common Factor Test. Let RSS_r and RSS_u be the restricted sums of squares of the restricted and unrestricted model, respectively.

- Test statistic: $CF = n \times \log(RSS_r/RSS_u)$.
- Under H_0, $CF \overset{a}{\sim} \chi^2_1$.
- Extra information needed: RSS_r (since RSS_u is given in (9.2)).

Question 10

Consider the model

$$\boldsymbol{y_t = \alpha + \beta x_t + \varepsilon_t, \quad t = 1, \ldots, T} \qquad \textbf{(10.1)}$$

where y_t and x_t are both integrated of order one.

(a) Explain what it means to say that y_t is integrated of order one. Discuss how you would test for this. In your answer make sure that it is clear how to implement your test.

(6 marks)

(b) Give an example of an economic variable that is potentially integrated of order one and give an intuitive explanation why you expect this process to be integrated of order 1.

(2 marks)

(c) It will be important to distinguish whether the above relationship is 'spurious' as opposed to 'cointegrating'.

i. Explain what it means to say that y_t and x_t have a cointegrating relationship and how does that contrast to a spurious relationship.

(4 marks)

ii. Discuss how you can test for evidence of a cointegrating relationship.

(4 marks)

(d) Suppose that

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t, \quad |\rho| < 1,$$

and v_t is an i.i.d. $(0, \sigma^2)$ innovation which is independent of ε_{t-1}. Show that you can rewrite equation (10.1) in terms of an error correction model:

$$\Delta y_t = \delta_1 \Delta x_t + \delta_2(y_{t-1} - \alpha - \beta x_{t-1}) + v_t.$$

Clearly indicate the relation between (δ_1, δ_2) and (α, β, ρ). Give an economic intuition behind this result.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 13.1 (Stationarity and nonstationarity), Chapter 13.4–13.5 (Tests of nonstationarity), and Chapter 13.6 (Cointegration).

Dougherty, C. Subject guide (2016): Chapter 13.

Approaching the question

(a) Integrated of order one (or simply I(1)) means that the process can be made stationary by differencing once. One could test for this using a Dickey–Fuller or an Augmented Dickey–Fuller test. Standard bookwork. For instance, suppose $y_t = \rho y_{t-1} + \varepsilon_t$. We can conduct the Dickey–Fuller test by running the auxiliary regression of the form:

$$\Delta y_t = \gamma y_{t-1} + \varepsilon_t$$

where $\gamma \equiv \rho - 1$. Provide test statistic and discuss rejection rule, clearly indicating H_0 and H_1.

(b) One example could be CPI due to its clearly trending behaviour over time it can be suggestive of a time-series integrated of order 1.

(c) i. Two variables integrated of order one are said to be cointegrated if there exists a linear combination of them that is integrated of order zero. In those cases, a regression of (10.1) will capture the long-run relationship between y_t and x_t and we should prefer an Error Correction Model to capture both the short-term and the long-term relationships among the two variables. In contrast, if the two series are not cointegrated a regression like (10.1) is likely to indicate a very strong relationship between the two variables even when there is no relationship at all. Such a regression is called a spurious regression.

ii. If y_t and x_t are both found to be integrated of order one, one could test for cointegration as follows:

* Run a regression of y_t on a constant and x_t and collect residuals.
* Perform a DF or ADF test on the residual series.
* If reject the null of non-stationarity conclude the series are cointegrated.

(d) Lag (10.1) by one period and multiply both sides by ρ to obtain:

$$\rho y_{t-1} = \rho\alpha + \rho\beta x_{t-1} + \rho\varepsilon_{t-1}.$$

Subtract the above from (10.1), hence:

$$y_t - \rho y_{t-1} = (1-\rho)\alpha + \beta x_t - \rho\beta x_{t-1} + \underbrace{\varepsilon_t - \rho\varepsilon_{t-1}}_{v_t}.$$

Rearranging:

$$y_t = \alpha + \rho(y_{t-1} - \alpha - \beta x_{t-1}) + \beta x_t + v_t.$$

Subtract $y_{t-1} = \alpha + \beta x_{t-1} + \varepsilon_{t-1}$, hence:

$$\Delta y_t = \rho(y_{t-1} - \alpha - \beta x_{t-1}) + \beta(x_t - x_{t-1}) + v_t - \varepsilon_{t-1}.$$

Finally, using $\varepsilon_{t-1} = y_{t-1} - \alpha - \beta x_{t-1}$ we obtain:

$$\Delta y_t = \beta\Delta x_t + (\rho - 1)(y_{t-1} - \alpha - \beta x_{t-1}) + v_t.$$

Therefore, $\delta_1 = \beta$ and $\delta_2 = \rho - 1$. Economic intuition: this is a process that reverts to the long-run equilibrium after a shock (δ_1 captures the short-run dynamics and δ_2 captures the speed of the reversal to the long-run equilibrium).

THIS PAPER IS NOT TO BE REMOVED FROM THE EXAMINATION HALL

UNIVERSITY OF LONDON

EC2020 ZA

BSc DEGREES AND GRADUATE DIPLOMAS IN ECONOMICS, MANAGEMENT, FINANCE AND THE SOCIAL SCIENCES, THE DIPLOMA IN ECONOMICS AND SOCIAL SCIENCES AND THE CERTIFICATE IN EDUCATION IN SOCIAL SCIENCES

Elements of Econometrics

Thursday 23 May 2019: 10.00 – 13.00

Time allowed: 3 hours

DO NOT TURN OVER UNTIL TOLD TO BEGIN

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Extracts from statistical tables are given after the final question on this paper.

Graph paper is provided at the end of this question paper. If used, it must be detached and fastened securely inside the answer book.

A handheld calculator may be used when answering questions on this paper and it must comply in all respects with the specification given with your Admission Notice. The make and type of machine must be clearly stated on the front cover of the answer book.

If more questions are answered than requested, only the first answers attempted will be counted.

© University of London 2019

UL19/0091

SECTION A

Answer all questions from this section

1. Consider the following regression model

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 t + \varepsilon_t, \qquad t = 1, ..., T.$$

Both $\{Y_t\}_{t=1}^T$ and $\{X_t\}_{t=1}^T$ are trend stationary processes. The errors $\{\varepsilon_t\}_{t=1}^T$ are independent random variables with zero mean and constant variance.

(a) **(4 marks)** Discuss the concept "trend stationarity" and contrast it to the concept "difference stationarity". In your answer make sure you also explain what stationarity means.

(b) **(4 marks)** Provide a clear interpretation of the parameter β_1. You are told that you can obtain the estimator for β_1 using "detrended" variables only. Discuss this statement.

2. Consider the following ADL(1,1) model relating the crime rate in a particular province, *crime*$_t$, to the clear-up rate (percentage of crimes resulting in a conviction)

$$\textit{crime}_t = \alpha + \rho \textit{crime}_{t-1} + \delta_1 \textit{clearup}_t + \delta_2 \textit{clearup}_{t-1} + u_t, \text{ with } |\rho| < 1$$

where u_t is white noise, an i.i.d. innovation that is uncorrelated to anything in the past.

(a) **(2 marks)** Briefly indicate whether OLS will provide unbiased and consistent parameter estimators.

(b) **(2 marks)** Derive the long run relationship between *crime* and *clearup*.

(c) **(4 marks)** Rewrite the model above in terms of an error correction model (ECM) and interpret its coefficients.

3. Consider the simple linear regression model

$$Y_i = \beta X_i + u_i, \qquad i = 1, ..., n.$$

We assume that the errors $\{u_i\}_{i=1}^n$ are independent normal random variables with zero mean. The regressor $\{X_i\}_{i=1}^n$ is non-stochastic (fixed under repeated sampling). You suspect that the errors exhibit heteroskedasticity.

(a) **(3 marks)** Explain what we mean by the concept of heteroskedasticity. Enhance your answer with the help of a graphical illustration.

(b) **(5 marks)** Suppose you want to test $H_0 : \beta = 1$ against $H_1 : \beta > 1$. Discuss how you would conduct this test based on the OLS estimator, recognizing the presence of heteroskedasticity. Please provide a detailed answer.

4. Consider the simple linear regression model

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \qquad i = 1, ..., n.$$

We assume that the errors $\{u_i\}_{i=1}^n$ are independent random variables with zero mean. The regressor $\{X_i\}_{i=1}^n$ is non-stochastic (fixed under repeated sampling). Under these conditions, the OLS estimator for β_1, $\hat{\beta}_1$, is unbiased. (You are not asked to derive $\hat{\beta}_1$).

(a) **(2 marks)** Explain the concept of unbiasedness of an estimator.

(b) **(4 marks)** Let us consider two other estimators for the slope β_1 :

$$\hat{\beta}_1^{\circ} = \frac{\sum_{i=1}^n (Z_i - \bar{Z}) Y_i}{\sum_{i=1}^n (Z_i - \bar{Z}) X_i}; \quad \hat{\beta}_1^{*} = \frac{\sum_{i=1}^n (Z_i - \bar{Z}) Y_i}{\sum_{i=1}^n (Z_i - \bar{Z}) Z_i}$$

where $Z_i = \sqrt{X_i}$ for all i and $\bar{Z} = \frac{1}{n}\sum_{i=1}^n Z_i$. Please indicate whether $\hat{\beta}_1^{\circ}$ and $\hat{\beta}_1^{*}$ are unbiased estimators for β_1. Clearly show your derivations.

(c) **(2 marks)** Briefly indicate how you would choose between the three estimators, $\hat{\beta}_1$, $\hat{\beta}_1^{\circ}$ and $\hat{\beta}_1^{*}$.

5. Consider the OLS estimator for β in the linear regression model

$$Y_i = \beta X_i + \varepsilon_i, \qquad i = 1, ..., n,$$

where $\{(Y_i, X_i)\}_{i=1}^n$ form an i.i.d. sample from a population and the errors are drawn from an **unknown** distribution with mean zero and variance σ^2. You are told that X_i and ε_i are **uncorrelated** (not necessarily independent therefore!).

(a) **(4 marks)** Discuss the importance of convergence in probability.

(b) **(4 marks)** Discuss the importance of convergence in distribution.

SECTION B

Answer three questions from this section.

6. Let us consider how workplace smoking bans affect the incidence of smoking. Below, we use data on 10,000 US indoor workers from 1991 to 1993 taken from "Do Workplace Smoking Bans Reduce Smoking", by Evans et al. (*American Economic Review*, 1999). Let *smoker* be a dummy variable indicating whether a worker smokes (1=yes, 0=no) and *smkban* a dummy variable indicating whether there is a ban on smoking in the workplace (1=yes, 0=no).

   (a) **(6 marks)** The following OLS regression results were obtained

$$\begin{aligned} \widehat{smoker} &= \underset{(0.007)}{0.290} - \underset{(0.009)}{0.078}smkban \qquad (6.1) \\ n &= 10000,\ R^2 = 0.0078,\ RSS = 1821.59 \end{aligned}$$

   The standard errors (SEs) are in parentheses. Interpret the parameter estimates of the coefficient on *smkban*. Provide the (approximate) 95% confidence interval for the coefficient on *smkban.* How can we use this confidence interval to test the hypothesis that $\beta_{smkban} = 0$?

A further specification was considered that included other characteristics of the worker: the age (in years), gender (male/female), ethnicity (black/hispanic/white), and level of education (E1=highschool dropout, E2=highschool graduate, E3=some college, E4=college graduate, E5 Master degree or above). The following OLS regression results were obtained for this multiple linear regression model

$$\begin{aligned} \widehat{smoker} &= \underset{(0.019)}{0.201} - \underset{(0.009)}{0.045}smkban - \underset{(.009)}{0.033}female - \underset{(.0003)}{0.001}age - \underset{(.016)}{0.027}black \qquad (6.2) \\ &\quad -\underset{(.014)}{0.104}hispanic + \underset{(.019)}{0.310}E1 + \underset{(.012)}{0.224}E2 + \underset{(.012)}{0.156}E3 + \underset{(.012)}{0.042}E4 \\ n &= 10000,\ R^2 = 0.0526,\ RSS = 1736.81 \end{aligned}$$

The SEs are in parentheses.

   (b) **(3 marks)** Compare the coefficient estimates on *smkban* from the simple and multiple regression model in (6.1) and (6.2) and explain why the estimates differ.

   (c) **(5 marks)** Interpret the estimated parameter on E2 (highschool graduate) in (6.2) and indicate how you can obtain its p-value and what information the p-value provides.

   (d) **(6 marks)** Both the simple and multiple regression model suffer from heteroskedasticity. Explain why. What are the implications of heteroskedasticy for the parameter estimates and the standard errors in (6.1) and (6.2)? What can you do to resolve this problem? Explain your answer.

7. In question 6 we considered the linear regression model to study how workplace smoking bans affect the incidence of smoking. Here we consider the results from applying a probit regression of *smoker* (1=yes, 0=no) on *smkban* (1=yes, 0=no) and the other explanatory variables:

```
. probit smoker smkban female age black hispanic E1 E2 E3 E4

Iteration 0:   log likelihood = -5537.1662
Iteration 1:   log likelihood = -5255.1526
Iteration 2:   log likelihood =  -5252.349
Iteration 3:   log likelihood = -5252.3489

Probit regression                                 Number of obs     =     10,000
                                                  LR chi2(9)        =     569.63
                                                  Prob > chi2       =     0.0000
Log likelihood = -5252.3489                       Pseudo R2         =     0.0514
```

smoker	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
smkban	-.1517626	.0289268	-5.25	0.000	-.208458	-.0950671
female	-.1106249	.0287785	-3.84	0.000	-.1670298	-.05422
age	-.0042031	.0011748	-3.58	0.000	-.0065057	-.0019006
black	-.07969	.0525369	-1.52	0.129	-.1826604	.0232804
hispanic	-.3327039	.0476677	-6.98	0.000	-.4261308	-.2392769
E1	1.094231	.0714121	15.32	0.000	.9542663	1.234197
E2	.8518588	.0594747	14.32	0.000	.7352906	.9684271
E3	.6492566	.0606989	10.70	0.000	.530289	.7682241
E4	.2224747	.0649939	3.42	0.001	.0950891	.3498603
_cons	-.9842425	.0756055	-13.02	0.000	-1.132427	-.8360584

(a) **(5 marks)** It is argued that using the probit regression model is better than using the linear probability model when explaining the binary variable *smoker*. Discuss the benefits/drawbacks of using the Probit model when trying to explain a binary variable.

(b) **(5 marks)** Explain briefly how the Probit estimates are obtained and discuss the properties of the parameter estimates.

Hint: You may recall that for the Probit model, we will specify

$$\Pr(\textit{smoker} = 1) = \Phi\left(\beta_0 + \beta_1 \textit{smkban} + \beta_2 \textit{female} + ... + \beta_8 E3 + \beta_9 E4\right)$$

where Φ is the standard normal CDF (cumulative distribution function).

(c) **(5 marks)** Explain how you can estimate the effect of the smoking ban on the probability of smoking for a 50-year old white, college graduated man. You are not expected to use your calculator, clarity of the computations required is enough.

(d) **(5 marks)** Discuss how you could test the joint significance of the worker's characteristics (gender, age, ethnicity and level of education) using the likelihood ratio test. Clearly indicate the test statistic, its distribution, the rejection rule and the additional information you would need to implement it.

8. An economist is interested in estimating the production function for widgets which is postulated to follow a Cobb-Douglas specification:

$$Y_i = \exp(\beta_0) L_i^{\beta_L} K_i^{\beta_K} \exp(u_i),$$

where Y_i is a measure of output for firm i, L_i is labor, K_i is capital stock and u_i is an unobserved term that captures technological or managerial efficiency and other external factors (e.g., weather). The parameters to be estimated are $(\beta_0, \beta_L, \beta_K)$. Taking logs,

$$\ln Y_i = \beta_0 + \beta_L \ln L_i + \beta_K \ln K_i + u_i.$$

(a) **(5 marks)** Provide the interpretation of the parameter β_L. What effect, if any, will changing the units of measurement of labor have on the parameter estimates for $(\beta_0, \beta_L, \beta_K)$? Explain your answer.

(b) **(5 marks)** Assume that you have a cross-section of firms and that more productive firms hire less workers (labor). Explain why OLS would not provide consistent estimates for $(\beta_0, \beta_L, \beta_K)$. Would it over- or underestimate β_L on average? Clearly explain your answer.

Instead of applying OLS, the economist decides to use the average wage paid by firm i, W_i, as an instrument for the (log) quantity of labor employed by that firm, $\ln L_i$.

(c) **(7 marks)** Describe in detail how you would estimate the parameters of the production function using Two Stage Least Squares (TSLS). What restrictions would be necessary for this researcher to successfully use this instrumental variable in the estimation of the parameters $(\beta_0, \beta_L, \beta_K)$ and what would you need to assume about capital stock?

(d) **(3 marks)** If average wages per firm do not vary much by firm (potentially because of unionization or high mobility of the labor force), how would this affect the properties of the estimation procedure suggested in (c)? Explain your answer.

9. Let us consider the expectations augmented Phillips curve (see also Mankiw, 1994):

$$\textit{infl}_t - \textit{infl}_t^e = \beta_1 \left(\textit{unem}_t - \mu_0\right) + e_t,$$

where μ_0 is the natural rate of unemployment (assumed to be constant over time) and $\textit{infl}_t^e$ is the expected rate of inflation formed in $t-1$.

This model suggests that there is a tradeoff between unanticipated inflation ($\textit{infl}_t$-$\textit{infl}_t^e$) and cyclical unemployment (difference between actual unemployment and the natural rate of unemployment). We assume that e_t (also called supply shock) is an i.i.d. random variable with zero mean.

(a) **(2 marks)** You are told that expectations are formed as follows

$$\textit{infl}_t^e - \textit{infl}_{t-1}^e = \lambda \left(\textit{infl}_{t-1} - \textit{infl}_{t-1}^e\right)$$

What name do we give such a process and how should we interpret λ?

(b) **(7 marks)** Show that you can rewrite the model as

$$\Delta\textit{infl}_t = \gamma_0 + \gamma_1 \textit{unem}_t + \gamma_2 \textit{unem}_{t-1} + v_t, \tag{9.1}$$

where $\Delta\textit{infl}_t = \textit{infl}_t - \textit{infl}_{t-1}$. Clearly indicate the relation between $(\gamma_0, \gamma_1, \gamma_2)$ and $(\mu_0, \beta_1, \lambda)$ and show that

$$v_t = e_t - (1-\lambda)\, e_{t-1}.$$

Hint: If you want you may use the following shorthand notation in your derivations: $y_t = \textit{infl}_t$, $y_t^e = \textit{infl}_t^e$ and $x_t = \textit{unem}_t$.

(c) **(3 marks)** Discuss what assumptions you would like to make about e_t (the supply shock) that will guarantee that the OLS estimator for the parameters in (9.1) is consistent. *Hint*: You may want to give the assumptions you need to make about v_t (composite error term) first.

(d) **(3 marks)** Show how you can obtain a consistent estimator for λ using your consistent estimates for $(\gamma_0, \gamma_1, \gamma_2)$? Provide a proof of its consistency. [*Note:* If you did not manage to get an explicit relation between λ and $(\gamma_0, \gamma_1, \gamma_2)$, consider $\lambda = g(\gamma_0, \gamma_1, \gamma_2)$ where $g(\cdot)$ is some continuous function].

(e) **(5 marks)** One of the assumptions provided in (c) is rather unreasonable (*Hint*: future unemployment may be related to current supply shocks). Discuss how you could use IV (TSLS) to obtain a consistent estimator for the parameters in (9.1). Discuss what conditions you instruments need to satisfy and propose a suitable instrument in this setting.

10. This question is based on "Capital Accumulation and Growth: A New Look at the Empirical Evidence", by Bond et al. (*Journal of Applied Econometrics*, 2010). In this article, the authors are interested in a regression model for the (logarithm of) output-per-capita y_t in a given country and time period t similar to:

$$y_t = \alpha + \rho y_{t-1} + \beta x_t + \gamma t + \varepsilon_t, \ \ |\rho| < 1, \tag{10.1}$$

where x_t is (the logarithm of) investment-per output. Imagine that investment rates are also affected by current output-per-capita, so that

$$x_t = \varphi_0 + \varphi_1 y_t + u_t. \tag{10.2}$$

Equations (10.1) and (10.2) then form a simultaneous equation model. Both errors ε_t and u_t have zero mean.

For (a) and (b), we start by assuming that ε_t and u_t are not serially correlated.

(a) **(5 marks)** Obtain the reduced form equations for y_t and x_t (Note: the variables t and y_{t-1} are exogenous). Are equations (10.1) and (10.2) identified? Discuss.

(b) **(5 marks)** Assume that $x_t = \varphi_0 + u_t$ (that is $\varphi_1 = 0$). Under what conditions is the OLS estimator for the parameters in the structural equation for output-per-capita (i.e., equation (10.1)) consistent? Discuss.

(c) **(5 marks)** Assume that $x_t = \varphi_0 + u_t$ (that is $\varphi_1 = 0$), u_t is not serially correlated, and $\gamma = 0$. How would you test whether ε_t is serially correlated?

(d) **(5 marks)** Assume that $x_t = \varphi_0 + u_t$ (that is $\varphi_1 = 0$), $\rho = 0$ and $\gamma = 0$. Using the Dickey-Fuller test, the authors fail to reject the hypothesis of a unit root in y_t and x_t for most of the countries in their sample at usual significance levels (5%). (Note: we are no longer assuming ε_t or u_t to not be serially correlated). Explaining your answers, what can you say about the OLS estimator for
and
applied to equation (10.1)

i. if ε_t is non-stationary and strongly dependent?
i. if ε_t is stationary and weakly dependent?

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	5·0	1·6449	3·0	1·8808	2·0	2·0537	1·0	2·3263	0·10	3·0902
45	0·1257	4·8	1·6646	2·9	1·8957	1·9	2·0749	0·9	2·3656	0·09	3·1214
40	0·2533	4·6	1·6849	2·8	1·9110	1·8	2·0969	0·8	2·4089	0·08	3·1559
35	0·3853	4·4	1·7060	2·7	1·9268	1·7	2·1201	0·7	2·4573	0·07	3·1947
30	0·5244	4·2	1·7279	2·6	1·9431	1·6	2·1444	0·6	2·5121	0·06	3·2389
25	0·6745	4·0	1·7507	2·5	1·9600	1·5	2·1701	0·5	2·5758	0·05	3·2905
20	0·8416	3·8	1·7744	2·4	1·9774	1·4	2·1973	0·4	2·6521	0·01	3·7190
15	1·0364	3·6	1·7991	2·3	1·9954	1·3	2·2262	0·3	2·7478	0·005	3·8906
10	1·2816	3·4	1·8250	2·2	2·0141	1·2	2·2571	0·2	2·8782	0·001	4·2649
5	1·6449	3·2	1·8522	2·1	2·0335	1·1	2·2904	0·1	3·0902	0·0005	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

Examiners' commentaries 2019

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2018–19. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

General remarks

Learning outcomes

At the end of the course, and having completed the Essential reading and activities, you should be able to:

- describe and apply the classical regression model and its application to cross-section data
- describe and apply the:
  - Gauss–Markov conditions and other assumptions required in the application of the classical regression model
  - reasons for expecting violations of these assumptions in certain circumstances
  - tests for violations
  - potential remedial measures, including, where appropriate, the use of instrumental variables
- recognise and apply the advantages of logit, probit and similar models over regression analysis when fitting binary choice models
- competently use regression, logit and probit analysis to quantify economic relationships using standard regression programmes (Stata and EViews) in simple applications
- describe and explain the principles underlying the use of maximum likelihood estimation
- apply regression analysis to time-series models using stationary time series, with awareness of some of the econometric problems specific to time series applications (for example, autocorrelation) and remedial measures
- recognise the difficulties that arise in the application of regression analysis to nonstationary time series, know how to test for unit roots, and know what is meant by cointegration.

Common mistakes committed by candidates

A large number of candidates are not able to clearly distinguish between sample variance and covariance, and population variance and covariance (this is happening year after year).

The use of $\mathrm{Cov}(X,Y)$ and $\mathrm{Var}(X)$ should be restricted to describing the population covariance and variances, respectively, with definitions:

$$\mathrm{Cov}(X,Y) = \mathrm{E}\left((X-\mathrm{E}(X))(Y-\mathrm{E}(Y))\right) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$$

and:

$$\mathrm{Var}(X) = \mathrm{E}((X-\mathrm{E}(X))^2) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$$

(you also may denote $\mathrm{Cov}(X,Y) = \sigma_{XY}$ and $\mathrm{Var}(X) = \sigma_X^2$). They are typically unknown, but fixed, quantities.

The sample covariance and variance are estimators of the population covariance and variance, respectively. They are defined as:

$$\text{Sample Cov}(X,Y) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})$$

and:

$$\text{Sample Var}(X) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

(you also may use $\widehat{\sigma}_{XY}$ and $\widehat{\sigma}_X^2$). You can compute them given the data.

With a slight abuse of notation, we often divide by n instead, which is irrelevant if we let n be large. The division by $n-1$ is a finite sample issue only (unbiasedness).

The sample covariance and variance show up in our definition of the OLS estimator of the slope in the simple linear regression model, not the population covariance and variance, as:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{\sum_{i=1}^{n}(X_i-\bar{X})^2} = \frac{\text{Sample Cov}(X,Y)}{\text{Sample Var}(X)} \neq \frac{\mathrm{Cov}(X,Y)}{\mathrm{Var}(X)}.$$

Treating them as being the same results in incorrect analyses and candidates losing significant marks.

Candidates should realise that $\frac{1}{n}\sum_{i=1}^{n} u_i$ is not the same as $\mathrm{E}(u_i)$. So, while we typically assume $\mathrm{E}(u_i) = 0$, this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} u_i = 0$. Also, while we may be happy to assume $\mathrm{E}(x_i u_i) = 0$ (uncorrelatedness between the errors and regressors), this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} x_i u_i = 0$. Note that:

- both $\frac{1}{n}\sum_{i=1}^{n} u_i$ and $\frac{1}{n}\sum_{i=1}^{n} x_i u_i$ are random variables, which take the value 0 with probability 0 (continuous random variables)!
- $\mathrm{E}(u_i) = 0$ and $\mathrm{E}(x_i u_i) = 0$ are fixed, not stochastic!

The differences between sample and population moments need to come across clearly when looking at unbiasedness and making consistency arguments. In both cases, we first simplify our estimator (plug in the true model) to obtain:

$$\widehat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i-\bar{X})u_i}{\sum_{i=1}^{n}(X_i-\bar{X})^2} = \beta + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \quad \text{with } x_i = X_i - \bar{X}.$$

- For *unbiasedness*, clearly indicate that you want to show that $\mathrm{E}(\widehat{\beta}) = \beta$. Unbiasedness does not follow from $\sum_{i=1}^{n} x_i u_i = 0$, instead it follows from $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) = 0$.

  If we treat x_i as fixed, $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) \equiv \mathrm{E}\left(\sum_{i=1}^{n} d_i u_i\right) = \sum_{i=1}^{n} d_i \,\mathrm{E}(u_i)$ and then unbiasedness follows as $\mathrm{E}(u_i) = 0$.
- For *consistency*, clearly indicate that you want to show that $\mathrm{plim}(\widehat{\beta}) = \beta$. Using the plim properties, we show:

$$\begin{aligned}
\mathrm{plim}\ \widehat{\beta} = \beta + \mathrm{plim}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) &= \beta + \frac{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i u_i\right)}{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)} \\
&\equiv \beta + \frac{\mathrm{plim}\left(\text{Sample Cov}\,(x,u)\right)}{\mathrm{plim}\left(\text{Sample Var}\,(x)\right)} \\
&= \beta + \frac{\mathrm{Cov}\,(x,u)}{\mathrm{Var}\,(x)} \quad \text{using the law of large numbers}
\end{aligned}$$

  where $\mathrm{Cov}(x,u) = 0$ and $\mathrm{Var}(x) > 0$, ensuring we get consistency.
- Remember, the law of large numbers ensures that sample averages converge to their population analogues.

Candidates struggled to give competent answers to the interpretation of empirical results. When interpreting an empirical result you should discuss the significance of the coefficients, magnitude and sign of the coefficients.

Candidates particularly struggled with the difference between percentage points and changes when it comes to the interpretation of coefficients in binary choice models. A coefficient in such a model gives the percentage point difference in the outcome variable when the regressor changes by 1 unit. The % change is a *relative* change which depends on the baseline value of the outcome variable (for example, the realisation when a regressor equals zero in the case of a dummy variable regressor). In the case of dummy variables as regressors, candidates often missed that the coefficients have to be interpreted relative to the base case which is the left out dummy variable (see appropriate questions in the examinations for details).

When conducting hypothesis tests, you should make sure that the Gauss–Markov conditions hold. The Gauss–Markov conditions have to be explicitly specified. Only writing that the Gauss–Markov conditions hold is not sufficient. As good practice, begin your examination by explicitly providing the Gauss–Markov conditions. You can then refer back to them thereafter. Moreover, ensure when conducting hypothesis testing that you clearly indicate the null and alternative hypotheses (in terms of the true parameters, say β_1), the test statistic (in terms of the parameter estimates, here $\widehat{\beta}_1$), its distribution (with degrees of freedom), the rejection rule (one-sided or-two sided) for a given significance level (typically 5%) with suitable critical values, and provide an interpretation of your result.

Exogeneity and endogeneity are often very simply explained as variables 'outside of the model' or 'inside the model', respectively. While this is not wrong, it is not a particularly clear or helpful definition, in particular in general contexts which are not related to simultaneous equations models. Defining these concepts in terms of the relationship between regressor and error term is much clearer and naturally leads to a way to show exogeneity or endogeneity by considering the covariance between regressors and error terms.

Candidates often missed that all exogenous regressors and instruments have to be included in the first stage when performing TSLS estimation.

It was very common to misunderstand the central limit theorem. While it provides the distribution of the *estimator* $\widehat{\beta}$, no matter what the true distribution of the errors is, it does not say that the distribution of the error term is approximately normal in large samples.

When asked about identification of structural form equations, candidates often compared the number of excluded regressors to the number of endogenous variables in general, without pointing out which variables are the excluded regressors and could, therefore, be used as an instrument.

Just as last year, many candidates do not answer all parts of the question. Make sure you read the questions properly and provide all details that are requested. Not answering a question will automatically earn you a zero mark for that question.

Key steps to improvement

Essential reading for this course includes the subject guide and the following:

- Dougherty, C. *Introduction to econometrics*. (Oxford: Oxford University Press, 2016) 5th edition [ISBN 9780199676828]; http://oxfordtextbooks.co.uk/orc/dougherty5e/

Apart from the Essential readings you should do some supplementary reading. One very good book at the same level is:

- Gujarati, D.N. and D.C. Porter *Basic econometrics*. (McGraw–Hill, 2009, International edition) 5th edition [ISBN 9780071276252].

To understand the subject clearly it is important to supplement Dougherty's *Introduction to econometrics* (fifth edition) with the subject guide **EC2020 Elements of econometrics** (2016), especially Chapter 10 which covers maximum likelihood estimation. It is very important to carefully go through the subject guide. The subject guide contains solutions to the questions given in the main textbook and also some additional questions and solutions. Working through these will improve your understanding of the subject.

The chapter in the subject guide on maximum likelihood (Chapter 10) includes some additional theory which has not been covered in the main textbook. It is important to read the additional theory given in the subject guide to have a better understanding of the principles of maximum likelihood and tests based on the likelihood function.

Please check the VLE course page for resources for this subject such as a downloadable copy of the subject guide **EC2020 Elements of econometrics** (2016), PowerPoint slideshows that provide a graphical treatment of the topics covered in the textbook, datasets and statistical tables. Candidates should utilise datasets using standard regression programmes (STATA or EViews). This will help in the understanding of the subject.

Examination revision strategy

Many candidates are disappointed to find that their examination performance is poorer than they expected. This may be due to a number of reasons, but one particular failing is '**question spotting**', that is, confining your examination preparation to a few questions and/or topics which have come up in past papers for the course. This can have serious consequences.

We recognise that candidates might not cover all topics in the syllabus in the same depth, but you need to be aware that examiners are free to set questions on **any aspect** of the syllabus. This means that you need to study enough of the syllabus to enable you to answer the required number of examination questions.

The syllabus can be found in the Course information sheet available on the VLE. You should read the syllabus carefully and ensure that you cover sufficient material in preparation for the examination. Examiners will vary the topics and questions from year to year and may well set questions that have not appeared in past papers. Examination papers may legitimately include questions on any topic in the syllabus. So, although past papers can be helpful during your revision, you cannot assume that topics or specific questions that have come up in past examinations will occur again.

If you rely on a question-spotting strategy, it is likely you will find yourself in difficulties when you sit the examination. We strongly advise you not to adopt this strategy.

Examiners' commentaries 2019

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2018–19. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone A

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Section A

Answer all questions from this section.

Question 1

Consider the following regression model:

$$Y_t = \beta_0 + \beta_1 X_t + \beta_2 t + \varepsilon_t, \quad t = 1, \ldots, T.$$

Both $\{Y_t\}_{t=1}^T$ and $\{X_t\}_{t=1}^T$ are trend stationary processes. The errors $\{\varepsilon_t\}_{t=1}^T$ are independent random variables with zero mean and constant variance.

(a) **Discuss the concept 'trend stationarity' and contrast it to the concept 'difference stationarity'. In your answer make sure you also explain what stationarity means.**

(4 marks)

(b) **Provide a clear interpretation of the parameter β_1. You are told that you can obtain the estimator for β_1 using 'detrended' variables only. Discuss this statement.**

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 13.1 and 13.2.

Dougherty, C. Subject guide (2016): Chapter 13.

Approaching the question

(a) *Trend stationarity* relates to a process which has a deterministic trend ($y_t = \alpha + \beta t + \varepsilon_t$ where ε_t is weakly dependent). It is a type of non-stationarity which can be removed by detrending. Once we include a trend in the regression, we can use standard statistical inference methods: the errors (and, therefore, Y_t) will be weakly dependent and consistent estimates can be obtained.

Difference stationarity relates to a process which can be made stationary by taking first differences: ΔY_t, examples are random walk and random walk with drift. Here Y_t is persistent, strongly dependent, and estimation of the process Y_t on Y_{t-1} would fail to produce consistent parameter estimates.

By stationarity we mean *covariance stationarity*, which requires constant (finite) means and variances, and covariances which only depend on distance in time, i.e. $\mathrm{E}(Y_t, Y_{t-s})$ only depends on s, not on location t.

(b) β_1 is the marginal effect of X_t on Y_t after controlling for a linear time trend, that is holding time constant (*ceteris paribus*).

Frisch–Waugh–Lovell (regression anatomy) argue that including the time trend $\beta_2 t$ explicitly in the regression is the same as working with detrended data where we control for the trend beforehand (implicitly).

Since both Y_t and X_t have a linear time trend, we need to include it in the regression in order to obtain consistent estimates of β_1. Omitting the time trend in this setting would give rise to omitted variable bias (OVB).

Question 2

Consider the following ADL(1, 1) model relating the crime rate in a particular province, $crime_t$, to the clear-up rate (percentage of crimes resulting in a conviction):

$$crime_t = \alpha + \rho crime_{t-1} + \delta_1 clearup_t + \delta_2 clearup_{t-1} + u_t, \quad \text{with } |\rho| < 1$$

where u_t is white noise, an i.i.d. innovation that is uncorrelated to anything in the past.

(a) Briefly indicate whether OLS will provide unbiased and consistent parameter estimators.

(2 marks)

(b) Derive the long run relationship between *crime* and *clearup*.

(2 marks)

(c) Rewrite the model above in terms of an error correction model (ECM) and interpret its coefficients.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 11.3, 11.4 and 11.5.

Dougherty, C. Subject guide (2016): Chapter 11.

Approaching the question

(a) OLS will be consistent because the error is uncorrelated with the regressors. OLS will be biased because of the presence of the lagged endogenous variable. The latter is due to the fact that u_t will be correlated with $crime_{t+s}$ for any $s \geq 0$ (regressor $crime_{t-1}$ in future periods).

(b) Let in equilibrium $X_t = X_{t-1} = \tilde{X}$ for both variables in the model, so that:

$$\widetilde{crime} = \alpha + \rho \widetilde{crime} + \delta_1 \widetilde{clearup} + \delta_2 \widetilde{clearup}.$$

After rewriting we obtain the long-run relationship:

$$crime_t = \frac{\alpha}{1-\rho} + \frac{\delta_1 + \delta_2}{1-\rho} clearup_t.$$

(c) The error correction model states that the change in *crime* in any period will be governed by the change in *clearup* and an error correction mechanism which reveals the speed with which *crime* changes in response to a disequilibrium. Here, we get:

$$\Delta crime_t = \delta_1 \Delta clearup_t + (\rho - 1) diseq_{t-1} + u_t$$

where:

$$diseq_{t-1} = crime_{t-1} - \frac{\alpha}{1-\rho} - \frac{\delta_1 + \delta_2}{1-\rho} clearup_{t-1}$$

and $(\rho - 1)$ indicates the speed of adjustment.

Question 3

Consider the simple linear regression model:

$$\boldsymbol{Y_i = \beta X_i + u_i, \quad i = 1, \ldots, n.}$$

We assume that the errors $\{u_i\}_{i=1}^n$ are independent normal random variables with zero mean. The regressor $\{X_i\}_{i=1}^n$ is non-stochastic (fixed under repeated sampling). You suspect that the errors exhibit heteroskedasticity.

(a) Explain what we mean by the concept of heteroskedasticity. Enhance your answer with the help of a graphical illustration.

(3 marks)

(b) Suppose you want to test $\mathrm{H_0} : \beta = 1$ against $\mathrm{H_1} : \beta > 1$. Discuss how you would conduct this test based on the OLS estimator, recognising the presence of heteroskedasticity. Please provide a detailed answer.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 7.1 and 7.2.

Dougherty, C. Subject guide (2016): Chapters 2 and 7.

Approaching the question

(a) Heteroskedasticity means that the error variance depends on the regressors, it violates the assumption that $\mathrm{E}(u_i^2) = \sigma^2$.

A graphical illustration should show data points and some depiction of the variance – for example, drawing the distribution for the errors at different levels of X. The graph should be clearly labelled.

(b) We need to propose the t test where the standard error on β needs to be estimated taking the heteroskedasticity into account (the usual t statistic is invalid because of the presence of heteroskedastity):

$$T = \frac{\widehat{\beta} - 1}{\text{s.e.}(\widehat{\beta})} \sim t_{n-1} \text{ under } \mathrm{H}_0.$$

The White robust standard error we use for this is the square root of the following heteroskedasticity-robust variance estimate:

$$\widehat{\mathrm{Var}}(\widehat{\beta}) = \frac{\sum_{i=1}^{n} X_i^2 \widehat{u}_i^2}{\left(\sum_{i=1}^{n} X_i^2\right)^2}.$$

The test then proceeds in the standard way. Using the 5% significance level we reject H_0 if the realisation of our test statistic exceeds the critical value: reject if $t > t_{0.05,\, n-1}$ (with n large, we reject H_0 if $t > 1.645$). Otherwise, fail to reject H_0.

A good number of candidates proposed a Goldfeld–Quandt test of heteroskedasticity instead of a discussion of how to conduct hypothesis testing given heteroskedasticity which is what the question asked. It is important to carefully read what the question asks!

Question 4

Consider the simple linear regression model:

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \quad i = 1, \ldots, n.$$

We assume that the errors $\{u_i\}_{i=1}^n$ are independent random variables with zero mean. The regressor $\{X_i\}_{i=1}^n$ is non-stochastic (fixed under repeated sampling). Under these conditions, the OLS estimator for β_1, $\widehat{\beta}_1$, is unbiased. (You are not asked to derive $\widehat{\beta}_1$).

(a) Explain the concept of unbiasedness of an estimator.

(2 marks)

(b) Let us consider two other estimators for the slope β_1:

$$\widehat{\beta}_1^{\circ} = \frac{\sum_{i=1}^{n} (Z_i - \bar{Z}) Y_i}{\sum_{i=1}^{n} (Z_i - \bar{Z}) X_i} \quad \text{and} \quad \widehat{\beta}_1^{*} = \frac{\sum_{i=1}^{n} (Z_i - \bar{Z}) Y_i}{\sum_{i=1}^{n} (Z_i - \bar{Z}) Z_i}$$

where $Z_i = \sqrt{X_i}$ for all i and $\bar{Z} = \frac{1}{n} \sum_{i=1}^{n} Z_i$. Please indicate whether $\widehat{\beta}_1^{\circ}$ and $\widehat{\beta}_1^{*}$ are unbiased estimators for β_1. Clearly show your derivations.

(4 marks)

(c) Briefly indicate how you would choose between the three estimators, $\widehat{\beta}_1$, $\widehat{\beta}_1^{\circ}$ and $\widehat{\beta}_1^{*}$.

(2 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 2.3 and 2.5.

Dougherty, C. Subject guide (2016): Chapter 2.

Approaching the question

(a) Unbiasedness means $\mathrm{E}(\widehat{\beta}_1) = \beta_1$. It means that the expected value of the estimator is the true parameter; that is, we are correct on average in repeated samples. This ensures that we will not make systematic errors when estimating β.

(b) In answering this question, it is easiest to derive results using Z_i rather than working with $Z_i = \sqrt{X_i}$. It is important to note furthermore that $\sum(Z_i - \bar{Z}) = 0$.

When plugging in the true model ($Y_i = \beta_0 + \beta_1 X_i + u_i$), we can then write:

$$\widehat{\beta}_1^{\circ} = \beta_1 + \frac{\sum(Z_i - \bar{Z})u_i}{\sum(Z_i - \bar{Z})X_i}$$

and:

$$\mathrm{E}(\widehat{\beta}_1^{\circ}) = \beta_1 + \frac{\sum(Z_i - \bar{Z})\mathrm{E}(u_i)}{\sum(Z_i - \bar{Z})X_i} = \beta_1$$

(X and Z are assumed to be non-stochastic). Unbiasedness follows as $\mathrm{E}(u_i) = 0$.

For the second estimator, when plugging in the true model, we can write:

$$\widehat{\beta}_1^{*} = \beta_1 + \frac{\sum(Z_i - \bar{Z})X_i}{\sum(Z_i - \bar{Z})Z_i} + \frac{\sum(Z_i - \bar{Z})u_i}{\sum(Z_i - \bar{Z})Z_i}$$

and:

$$\mathrm{E}(\widehat{\beta}_1^{*}) = \beta_1 + \frac{\sum(Z_i - \bar{Z})Z_i^2}{\sum(Z_i - \bar{Z})Z_i} \neq \beta_1$$

using again the fact that Z is non-stochastic and $\mathrm{E}(u_i) = 0$, hence $\widehat{\beta}_1^{*}$ is biased.

Side notes: (i) $\widehat{\beta}_1^{\circ}$ essentially is an IV estimator where Z is used as an instrument for X. Both Z and X are exogenous here. (ii) $\widehat{\beta}_1^{*}$, on the other hand, is the slope parameter of a regression of Z instead of X on Y, and hence should not give us an unbiased estimator of β_1.

Many candidates failed to realise that $\sum(Z_i - \bar{Z}) = 0$ and wrongly concluded that the first estimator $\widehat{\beta}_1^{\circ}$ is biased.

(c) When choosing between $\widehat{\beta}_1$ and $\widehat{\beta}_1^{\circ}$ (both unbiased), efficiency considerations matter. Since the model satisfies the Gauss–Markov assumptions, we know that by the Gauss–Markov theorem, $\widehat{\beta}_1$ is BLUE and should, therefore, be chosen over $\widehat{\beta}_1^{\circ}$. Nevertheless, there may be a trade-off between bias and variance, and the mean squared error of $\widehat{\beta}_1^{*}$ could be smaller than that of $\widehat{\beta}_1$ (which is unbiased).

The bias–variance trade-off was mentioned only by a few candidates.

Question 5

Consider the OLS estimator for β in the linear regression model:

$$\boldsymbol{Y_i = \beta X_i + \varepsilon_i, \quad i = 1, \ldots, n}$$

where $\{(Y_i, X_i)\}_{i=1}^{n}$ form an i.i.d. sample from a population and the errors are drawn from an <u>unknown</u> distribution with mean zero and variance σ^2. You are told that X_i and ε_i are <u>uncorrelated</u> (not necessarily independent therefore!).

(a) Discuss the importance of convergence in probability.

(4 marks)

(b) Discuss the importance of convergence in distribution.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections R.14 and R.15.

Dougherty, C. Subject guide (2016): Chapter 8.

Approaching the question

(a) Convergence in probability plays an important role when proving *consistency* of estimators. A consistent estimator converges in probability to the true parameter and we either write $\text{plim}(\widehat{\beta}) = \beta$ or $\widehat{\beta} \xrightarrow{p} \beta$.

Proofs of consistency (a large sample property) are particularly important in settings where we cannot show that our estimator is unbiased (a finite sample property). For the OLS estimator:

$$\text{plim}(\widehat{\beta}) = \text{plim}(\beta) + \frac{\text{plim}(\frac{1}{n}\sum X_i\varepsilon_i)}{\text{plim}(\frac{1}{n}\sum X_i^2)} = \beta + \frac{\text{E}(X_i\varepsilon_i)}{\text{E}(X_i^2)} = \beta$$

where we use the properties of the plim operator and the Law of Large Numbers to argue that the sample averages converge in probability to their population equivalents.

(b) Convergence in distribution plays an important role in settings where we do not know the exact sampling distribution of our estimator, such as in the linear regression setting when the distribution of u is unknown. We need a distributional result in order for us to conduct hypothesis testing.

The central limit theorem allows us to obtain an approximate (asymptotic) distribution for $\widehat{\beta}$ even without knowing the true distribution of u. The asymptotic distribution provides a good approximation of the true sampling distribution as long as the sample is sufficiently large. This powerful result enables us to conduct hypothesis testing which is valid assuming our sample is large.

Section B

Answer three questions from this section.

Question 6

Let us consider how workplace smoking bans affect the incidence of smoking. Below, we use data on 10,000 US indoor workers from 1991 to 1993 taken from 'Do Workplace Smoking Bans Reduce Smoking', by Evans et al. (*American Economic Review*, 1999).

Let *smoker* be a dummy variable indicating whether a worker smokes (1 = yes, 0 = no) and *smkban* a dummy variable indicating whether there is a ban on smoking in the workplace (1 = yes, 0 = no).

(a) The following OLS regression results were obtained:

$$\widehat{smoker} = \underset{(0.007)}{0.290} - \underset{(0.009)}{0.078}\, smkban \tag{6.1}$$

$$n = 10000,\ R^2 = 0.0078,\ RSS = 1821.59$$

The standard errors (SEs) are in parentheses. Interpret the parameter estimates of the coefficient on *smkban*. Provide the (approximate) 95% confidence interval for the coefficient on *smkban*. How can we use this confidence interval to test the hypothesis that $\beta_{smkban} = 0$?

(6 marks)

A further specification was considered that included other characteristics of the worker: the age (in years), gender (male/female), ethnicity (black/hispanic/white), and level of education ($E1$ = highschool dropout, $E2$ = highschool graduate, $E3$ = some college, $E4$ = college graduate, $E5$ = Master degree or above). The following OLS regression results were obtained for this multiple linear regression model:

$$\widehat{smoker} = \underset{(0.019)}{0.201} - \underset{(0.009)}{0.045}\,smkban - \underset{(.009)}{0.033}female - \underset{(.0003)}{0.001}\,age - \underset{(.016)}{0.027}black \tag{6.2}$$

$$-\underset{(.014)}{0.104}hispanic + \underset{(.019)}{0.310}E1 + \underset{(.012)}{0.224}E2 + \underset{(.012)}{0.156}E3 + \underset{(.012)}{0.042}E4$$

$$n = 10000,\ R^2 = 0.0526,\ RSS = 1736.81$$

The SEs are in parentheses.

(b) Compare the coefficient estimates on *smkban* from the simple and multiple regression model in (6.1) and (6.2) and explain why the estimates differ.

(3 marks)

(c) Interpret the estimated parameter on $E2$ (highschool graduate) in (6.2) and indicate how you can obtain its p-value and what information the p-value provides.

(5 marks)

(d) Both the simple and multiple regression model suffer from heteroskedasticity. Explain why. What are the implications of heteroskedasticy for the parameter estimates and the standard errors in (6.1) and (6.2)? What can you do to resolve this problem? Explain your answer.

(6 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 2.6, 3.2, 6.2, 7.3 and 10.1.

Dougherty, C. Subject guide (2016): Chapters 5, 6 and 10.

Approaching the question

(a) This is the linear probability model, where:

$$\mathrm{E}(smoker = 1 \mid X) = P(smoker = 1 \mid X) = \beta_0 + \beta_1 smkban.$$

If the smoking ban is introduced, the probability of a worker smoking decreases by 7.8 percentage points (not 7.8%).

An (approximate) 95% confidence interval is:

$$(\widehat{\beta}_1 - z_{crit,\,0.025} \times \text{s.e.}(\widehat{\beta}_1),\ \widehat{\beta}_1 + z_{crit,\,0.025} \times \text{s.e.}(\widehat{\beta}_1)$$
$$= (-0.078 - 1.96 \times 0.009,\ -0.078 + 1.96 \times 0.009) \approx (-0.096,\ -0.060)$$

where z relates to the standard normal distribution (for large degrees of freedom the t distribution converges to the standard normal).

We can use confidence intervals for testing. For a two-sided $\mathrm{H}_0 : \beta_1 = 0$ test, we reject H_0 at the 5% significance level if zero does not lie in the 95% confidence interval.

(b) The reason why the coefficient estimates on *smkban* differ in (6.1) and (6.2) is due to the *omitted variable bias problem.* When we omit relevant variables (for example, education and gender) which are related to the included variable *smkban*), then the parameter estimates on *smkban* will estimate not only the direct effect that *smkban* has on smoking (the parameter of interest) but the indirect effect of these omitted variables as well. We also call these omitted variables confounders.

The more negative effect found in (a) signals that once the effect of education on smoking (gender, age and ethnicity) is controlled for, the true effect of a smoking ban on the probability of smoking is smaller.

(c) The coefficient on $E2$ means that highschool graduates are 22.4 percentage points more likely to smoke relative to people with a Master's degree or above since $E5$ *is the left out base category*. (Many candidates failed to recognise that the left out category was required to interpret the parameter clearly). The t statistic for this effect is:

$$t = \frac{\widehat{\beta}}{\text{s.e.}(\widehat{\beta})} = \frac{0.224}{0.012} = 18.667$$

which has a p-value smaller than 0.05 (we would reject H_0 at the 5% significance level). The p-value is the lowest level of significance at which we can reject the null hypothesis. Since the sample size is large, we can use the standard normal distribution to obtain the p-value (candidates may provide a graphical discussion instead). For a one-sided test, $p = 1 - \Phi(t)$ and for a two-sided test, $p = 2 \times (1 - \Phi(t))$, assuming $t \geq 0$.

(d) In the linear probability model (LPM) we have to deal with the problem of heteroskedasticity. Conditional on X, $\mathrm{Var}(y \,|\, X) = p(X)(1 - p(X)) \equiv \mathrm{Var}(u \,|\, X)$. The implication of heteroskedasticity for the estimated standard errors is that they are wrong (they rely on homoskedasticity) and inference based on them would be invalid. The parameter estimates themselves remain unbiased and consistent. For inference, we should use heteroskedasticity-robust standard errors or apply WLS instead.

Question 7

In question 6 we considered the linear regression model to study how workplace smoking bans affect the incidence of smoking. Here we consider the results from applying a probit regression of *smoker* (1 = yes, 0 = no) on *smkban* (1 = yes, 0 = no) and the other explanatory variables:

```
. probit smoker smkban female age black hispanic E1 E2 E3 E4

Iteration 0:   log likelihood = -5537.1662
Iteration 1:   log likelihood = -5255.1526
Iteration 2:   log likelihood =  -5252.349
Iteration 3:   log likelihood = -5252.3489

Probit regression                               Number of obs     =     10,000
                                                LR chi2(9)        =     569.63
                                                Prob > chi2       =     0.0000
Log likelihood = -5252.3489                     Pseudo R2         =     0.0514
```

smoker	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
smkban	-.1517626	.0289268	-5.25	0.000	-.208458	-.0950671
female	-.1106249	.0287785	-3.84	0.000	-.1670298	-.05422
age	-.0042031	.0011748	-3.58	0.000	-.0065057	-.0019006
black	-.07969	.0525369	-1.52	0.129	-.1826604	.0232804
hispanic	-.3327039	.0476677	-6.98	0.000	-.4261308	-.2392769
E1	1.094231	.0714121	15.32	0.000	.9542663	1.234197
E2	.8518588	.0594747	14.32	0.000	.7352906	.9684271
E3	.6492566	.0606989	10.70	0.000	.530289	.7682241
E4	.2224747	.0649939	3.42	0.001	.0950891	.3498603
_cons	-.9842425	.0756055	-13.02	0.000	-1.132427	-.8360584

(a) It is argued that using the probit regression model is better than using the linear probability model when explaining the binary variable *smoker*. Discuss

the benefits/drawbacks of using the Probit model when trying to explain a binary variable.

(5 marks)

(b) Explain briefly how the Probit estimates are obtained and discuss the properties of the parameter estimates.

Hint: **You may recall that for the Probit model, we will specify:**

$$\mathbf{Pr}(\boldsymbol{smoker} = \mathbf{1}) = \boldsymbol{\Phi}(\boldsymbol{\beta_0} + \boldsymbol{\beta_1 smkban} + \boldsymbol{\beta_2 female} + \cdots + \boldsymbol{\beta_8 E3} + \boldsymbol{\beta_9 E4})$$

where $\boldsymbol{\Phi}$ is the standard normal CDF (cumulative distribution function).

(5 marks)

(c) Explain how you can estimate the effect of the smoking ban on the probability of smoking for a 50-year old white, college graduated man. You are not expected to use your calculator, clarity of the computations required is enough.

(5 marks)

(d) Discuss how you could test the joint significance of the worker's characteristics (gender, age, ethnicity and level of education) using the likelihood ratio test. Clearly indicate the test statistic, its distribution, the rejection rule and the additional information you would need to implement it.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 10.3 and 10.6.

Dougherty, C. Subject guide (2016): Chapter 10.

Approaching the question

(a) The probit model has three main advantages over the linear probability model (LPM): (i) predicted probabilities are restricted to lie in $[0, 1]$, (ii) maximum likelihood estimators are (asymptotically) efficient whereas OLS (LPM) estimators will be inefficient, and (iii) maximum likelihood estimators automatically deal with heteroskedasticity.

The main drawbacks of the probit model relative to the LPM are that (i) the coefficients cannot be directly interpreted as the marginal effects of the regressor(s) of interest, so we need to compute predicted probabilities using the probit specification, and (ii) it is also computationally more complicated.

(b) The parameters are estimated by maximum likelihood estimation, where the log-likelihood function is given by:

$$\begin{aligned}
\log L(\beta) &= \sum_{i=1}^{n} \{smoker_i \log(P(smoker_i = 1 \,|\, X) + (1 - smoker_i) \log(P(smoker_i = 0 \,|\, X)\} \\
&= \sum_{i=1}^{n} \{smoker_i \log(\Phi(\beta_0 + \beta_1 smkban_i + \beta_2 female_i + \cdots + \beta_9 E4_i) + \\
&\quad (1 - smoker_i) \log(1 - \Phi(\beta_0 + \beta_1 smkban_i + \beta_2 female_i + \cdots + \beta_9 E4_i))\}.
\end{aligned}$$

To obtain the parameter estimates the first-order conditions are solved (numerically as no explicit formulae exist).

Under suitable regularity conditions, the estimates are consistent, asymptotically normally, and asymptotically efficient.

(c) The effect is the difference in predicted probabilities between the man with the characteristics given, with $smkban = 1$ versus $smkban = 0$:

$$\widehat{P}(y_i = 1 \mid X, smkban = 1) - \widehat{P}(y_i = 1 \mid X, smkban = 0)$$

$$= \Phi\left(\sum_k x_{ki}\widehat{\beta}_k + \widehat{\beta}_1\right) - \Phi\left(\sum_k x_{ki}\widehat{\beta}_k\right)$$

$$= \Phi(-0.984 + 0.222 - 0.004 \times 50 - 0.152) - \Phi(-0.984 + 0.222 - 0.004 \times 50).$$

A discussion of marginal effects (ignoring the fact that *smkban* is a dummy variable) is acceptable as well.

(d) We would like to use the LR test, to test:

$$\mathrm{H}_0 : \beta_j = 0 \ \forall \ j \in \{\text{worker characteristics}\}$$

$$\mathrm{H}_1 : \beta_j \neq 0 \text{ for at least one } j \in \{\text{worker characteristics}\}.$$

Estimate the restricted (R) and the unrestricted probit model (U). The restricted model imposes H_0. The unrestricted model is just the originally estimated model above. The LR test makes use of the difference between the two estimated log-likelihood functions (the log ratio of the likelihood functions). The test statistic is:

$$LR = 2(\log L^U - \log L^R) \sim \chi^2_J$$

where J denotes the degrees of freedom which is equal to the number of restrictions; that is, $J = 8$ here. For a given significance level we reject H_0 if its realisation exceeds the critical value given by the χ^2_8 distribution, which equals 15.51 for the 5% significance level.

Question 8

An economist is interested in estimating the production function for widgets which is postulated to follow a Cobb–Douglas specification:

$$\boldsymbol{Y_i = \exp(\beta_0) L_i^{\beta_L} K_i^{\beta_K} \exp(u_i)}$$

where Y_i is a measure of output for firm i; L_i is labour, K_i is capital stock and u_i is an unobserved term that captures technological or managerial efficiency and other external factors (e.g., weather). The parameters to be estimated are $(\beta_0, \beta_L, \beta_K)$. Taking logs:

$$\boldsymbol{\ln Y_i = \beta_0 + \beta_L \ln L_i + \beta_K \ln K_i + u_i.}$$

(a) Provide the interpretation of the parameter β_L. What effect, if any, will changing the units of measurement of labour have on the parameter estimates for $(\beta_0, \beta_L, \beta_K)$? Explain your answer.

(5 marks)

(b) Assume that you have a cross-section of firms and that more productive firms hire less workers (labour). Explain why OLS would not provide consistent estimates for $(\beta_0, \beta_L, \beta_K)$. Would it over- or underestimate β_L on average? Clearly explain your answer.

(5 marks)

Instead of applying OLS, the economist decides to use the average wage paid by firm i, W_i, as an instrument for the (log) quantity of labour employed by that firm, $\ln L_i$.

(c) Describe in detail how you would estimate the parameters of the production function using Two Stage Least Squares (TSLS). What restrictions would be necessary for this researcher to successfully use this instrumental variable in the estimation of the parameters $(\beta_0, \beta_L, \beta_K)$ and what would you need to assume about capital stock?

(7 marks)

(d) If average wages per firm do not vary much by firm (potentially because of unionisation or high mobility of the labour force), how would this affect the properties of the estimation procedure suggested in (c)? Explain your answer.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 4.2, 6.2 and 9.3.

Dougherty, C. Subject guide (2016): Chapters 4 and 8.

Approaching the question

(a) Since the specification is in logs, β_L is the output *elasticity* with respect to labour. It measures the % change in output when labour is increased by 1%, *ceteris paribus*. Since this is a relative change measure, changing the units of measurement will not change the parameter estimate. The same holds true for β_K, the output elasticity with respect to capital. However, the intercept will be changed by a change in the units of labour. If labour is measured in new units $L_i^* = c \times L_i$, the log specification would give rise to $\ln L_i^* = \ln c \times L_i = \ln c + \ln L_i$ which means the intercept is shifted by $\beta_L \ln c$ units.

(b) The problem with using OLS on the above regression is that there will be correlation between the error u_i and the regressor $\ln L_i$ as more productive firms (higher u_i) are associated with hiring fewer workers (lower L_i). This correlation will make the parameter estimates inconsistent. This problem is a result of the *omission of relevant variables* which creates OVB. The parameter estimates for β_L will be underestimated as it captures the fact that firms with higher technological or managerial efficiency require less labour for the same output.

(c) Two-stage least squares (TSLS) proceeds in two stages. In the first stage, $\ln L_i$ is regressed on all other exogenous regressors and the instrument w. This yields the fitted values:

$$\widehat{\ln L_i} = \widehat{\pi}_0 + \widehat{\pi}_1 w_i + \widehat{\pi}_2 \ln K_i.$$

In the second stage, the original regression is run where the predicted values from the first stage above $\widehat{\ln L_i}$ are used instead of the original regressor $\ln L_i$. We run:

$$\ln Y_i = \beta_0 + \beta_L \widehat{\ln L_i} + \beta_K \ln K_i + u_i.$$

Alternatively, the second stage can be described as performing IV on the original regression where we use $\widehat{\ln L_i}$ as an instrument for $\ln L_i$.

The instrumental variable approach requires the instrument to be relevant and exogenous. Instrument relevance means that the instrument must be correlated with the original regressor, that is $\text{Cov}(\ln L_i, w_i) \neq 0$. Instrument exogeneity requires the instrument to be uncorrelated with the error term; that is, $\text{Cov}(w_i, u_i) = 0$.

For consistent parameter estimates, we require capital stock to also be exogenous, i.e. uncorrelated with the error term.

(d) Candidates should recall the variance of the IV estimator in the simple linear regression setting, revealing the importance that the instrument w is highly correlated with the endogenous regressor $\ln L$ for its precision. If there is not much variation in wages, the instrument is likely not to be very correlated with $\ln L_i$. This is a problem we also call the 'weak instrument problem' as we will have a weak first stage ($\ln L_i$ is not predicted well by the instrument and other exogenous regressors). This problem results in imprecise TSLS estimators.

Question 9

Let us consider the expectations augmented Phillips curve (see also Mankiw, 1994):

$$infl_t - infl_t^e = \beta_1(unem_t - \mu_0) + e_t$$

where μ_0 is the natural rate of unemployment (assumed to be constant over time) and $infl_t^e$ is the expected rate of inflation formed in $t-1$.

This model suggests that there is a trade-off between unanticipated inflation ($infl_t - infl_t^e$) and cyclical unemployment (difference between actual unemployment and the natural rate of unemployment). We assume that e_t (also called supply shock) is an i.i.d. random variable with zero mean.

(a) **You are told that expectations are formed as follows:**

$$infl_t^e - infl_{t-1}^e = \lambda(infl_{t-1} - infl_{t-1}^e).$$

What name do we give such a process and how should we interpret λ?

(2 marks)

(b) **Show that you can rewrite the model as:**

$$\Delta infl_t = \gamma_0 + \gamma_1 unem_t + \gamma_2 unem_{t-1} + v_t \tag{9.1}$$

where $\Delta infl_t = infl_t - infl_{t-1}$. Clearly indicate the relation between $(\gamma_0, \gamma_1, \gamma_2)$ and $(\mu_0, \beta_1, \lambda)$ and show that:

$$v_t = e_t - (1-\lambda)e_{t-1}.$$

Hint: **If you want you may use the following shorthand notation in your derivations: $y_t = infl_t$, $y_t^e = infl_t^e$ and $x_t = unem_t$.**

(7 marks)

(c) **Discuss what assumptions you would like to make about e_t (the supply shock) that will guarantee that the OLS estimator for the parameters in (9.1) is consistent.** ***Hint:*** **You may want to give the assumptions you need to make about v_t (composite error term) first.**

(3 marks)

(d) **Show how you can obtain a consistent estimator for λ using your consistent estimates for $(\gamma_0, \gamma_1, \gamma_2)$. Provide a proof of its consistency. [*Note:* If you did not manage to get an explicit relation between λ and $(\gamma_0, \gamma_1, \gamma_2)$, consider $\lambda = g(\gamma_0, \gamma_1, \gamma_2)$ where $g(\cdot)$ is some continuous function].**

(3 marks)

(e) **One of the assumptions provided in (c) is rather unreasonable (*Hint:* future unemployment may be related to current supply shocks). Discuss how you could use IV (TSLS) to obtain a consistent estimator for the parameters in (9.1). Discuss what conditions your instruments need to satisfy and propose a suitable instrument in this setting.**

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections R.14, 9.3, 11.3, 11.4 and 11.5.

Dougherty, C. Subject guide (2016): Chapters 8 and 11.

Approaching the question

(a) This is an adaptive expectations model. The parameter λ, which should lie between 0 and 1, indicates the speed with which expectations are adjusted in response to deviations between the actual and the expected rate of inflation in the previous period.

(b) To obtain this it is useful to realise that the equation given in part (a) can be rewritten as:

$$y_t^e = (1-\lambda)y_{t-1}^e + \lambda y_{t-1}. \qquad (*)$$

This shows that if we subtract $(1-\lambda)$ times the lagged expectation augmented Phillips curve, we can get rid of the inflation expectation variable. Let us substract from the first equation given in the question, $1-\lambda$ times the equation one period lagged:

$$y_t - y_t^e - (1-\lambda)[y_{t-1} - y_{t-1}^e] = \beta_1(x_t - \mu_0) + e_t - (1-\lambda)[\beta_1(x_{t-1} - \mu_0) + e_{t-1}].$$

Rearranging yields:

$$y_t - (1-\lambda)y_{t-1} - (y_t^e - (1-\lambda)y_{t-1}^e) = \beta_1 x_t - (1-\lambda)\beta_1 x_{t-1} - \lambda\beta_1\mu_0 + e_t - (1-\lambda)e_{t-1}.$$

Using (*) we get:

$$y_t - (1-\lambda)y_{t-1} - \lambda y_{t-1} = \beta_1 x_t - (1-\lambda)\beta_1 x_{t-1} - \lambda\beta_1\mu_0 + v_t$$

which yields:

$$\Delta y_t = -\beta_1\lambda\mu_0 + \beta_1 x_t - (1-\lambda)\beta_1 x_{t-1} + v_t$$

which gives rise to the following relationship between the γs and original parameters:

$$\gamma_0 = -\beta_1\lambda\mu_0, \quad \gamma_1 = \beta_1 \quad \text{and} \quad \gamma_2 = -(1-\lambda)\beta_1.$$

(c) Consistency requires the error term in the new model v_t and both regressors, x_t and x_{t-1} to be uncorrelated. Since e_t has a zero mean this means:

$$\mathrm{E}(x_t(e_t - (1-\lambda)e_{t-1})) = 0 \quad \text{and} \quad \mathrm{E}(x_{t-1}(e_t - (1-\lambda)e_{t-1})) = 0.$$

A sufficient condition for this would be that $\mathrm{E}(e_t x_s) = 0$ for $s \in \{t-1, t, t+1\}$ or uncorrelatedness between x_t and e_s $\forall$ s, t.

(d) The above relationships can be solved for λ as:

$$\lambda = \frac{\gamma_2}{\gamma_1} + 1.$$

Using the estimators $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$, an estimator $\widehat{\lambda}$ can be constructed accordingly. For consistent estimators $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$, the estimator $\widehat{\lambda}$ is also consistent by the Slutsky theorem:

$$\mathrm{plim}(\widehat{\lambda}) = \mathrm{plim}\left(\frac{\widehat{\gamma}_2}{\widehat{\gamma}_1} + 1\right) = \frac{\mathrm{plim}(\widehat{\gamma}_2)}{\mathrm{plim}(\widehat{\gamma}_1)} + 1 = \frac{\gamma_2}{\gamma_1} + 1 = \lambda.$$

Slutsky's theorem (probability limit rules) guarantees the second equality since these are continuous functions.

(e) If future unemployment is related to current supply shocks, then $\mathrm{Cov}(x_t, e_{t-1}) \neq 0$ and we get inconsistency as that would mean $\mathrm{Cov}(x_t, v_t) \neq 0$.

As long as unemployment is only related to past supply shocks, but not current or future ones x_{t-1} is still uncorrelated with v_t, $\mathrm{Cov}(x_{t-1}, v_t) = 0$.

We should proceed using an *instrumental variable approach* (2SLS), where we need to look for an instrument to deal with the endogeneity of x_t (this could for instance be a demand shock which is unrelated with supply shocks). Call this instrument z_t.

The instrument must satisfy instrument relevance and instrument exogeneity. Instrument relevance means that the instrument must be correlated with the original regressor, that is $\mathrm{Cov}(x_t, z_t) \neq 0$. Instrument exogeneity requires the instrument to be uncorrelated with the error term; that is, $\mathrm{Cov}(z_t, v_t) = 0$. Since v_t is a function of both e_t and e_{t-1}, the latter

assumption is satisfied when the instrument is unrelated with any current and past supply shocks.

In this particular case, due to the fact that v_t has an MA(1) error structure, we can also use x_{t-2} as an instrument. We expect there to be correlation between x_t and x_{t-2} and x_{t-2} (the past) is uncorrelated with e_t and e_{t-1} which make up v_t (we cannot use x_{t-1} because that variable is already included in the model).

A TSLS procedure would proceed by running a regression of x_t on the instrument and the other regressors to obtain fitted values $\widehat{x}_t$. In the second stage, the outcome variable Δy_t would be regressed on an intercept, the predicted values $\widehat{x}_t$ and x_{t-1}. Alternatively, we can apply IV in the second stage where we use $\widehat{x}_t$ as our instrument for x_t. This would give rise to consistent TSLS estimators $(\gamma_0^{TSLS}, \gamma_1^{TSLS}, \gamma_2^{TSLS})$.

Question 10

This question is based on 'Capital Accumulation and Growth: A New Look at the Empirical Evidence', by Bond et al. (*Journal of Applied Econometrics*, 2010). In this article, the authors are interested in a regression model for the (logarithm of) output-per-capita y_t in a given country and time period t similar to:

$$y_t = \alpha + \rho y_{t-1} + \beta x_t + \gamma_t + \varepsilon_t, \quad |\rho| < 1 \tag{10.1}$$

where x_t is (the logarithm of) investment-per output. Imagine that investment rates are also affected by current output-per-capita, so that:

$$x_t = \varphi_0 + \varphi_1 y_t + u_t. \tag{10.2}$$

Equations (10.1) and (10.2) then form a simultaneous equation model. Both errors ε_t and u_t have zero mean.

For (a) and (b), we start by assuming that ε_t and u_t are not serially correlated.

(a) Obtain the reduced form equations for y_t and x_t. (Note: the variables t and y_{t-1} are exogenous.) Are equations (10.1) and (10.2) identified? Discuss.

(5 marks)

(b) Assume that $x_t = \varphi_0 + u_t$ (that is $\varphi_1 = 0$). Under what conditions is the OLS estimator for the parameters in the structural equation for output-per-capita (i.e., equation (10.1)) consistent? Discuss.

(5 marks)

(c) Assume that $x_t = \varphi_0 + u_t$ (that is $\varphi_1 = 0$), u_t is not serially correlated, and $\gamma = 0$. How would you test whether ε_t is serially correlated?

(5 marks)

(d) Assume that $x_t = \varphi_0 + u_t$ (that is $\varphi_1 = 0$), $\rho = 0$ and $\gamma = 0$. Using the Dickey–Fuller test, the authors fail to reject the hypothesis of a unit root in y_t and x_t for most of the countries in their sample at usual significance levels (5%). (Note: we are no longer assuming ε_t or u_t to not be serially correlated). Explaining your answers, what can you say about the OLS estimator for α and β applied to equation (10.1):

i. if ε_t is stationary and weakly dependent?

ii. if ε_t is non-stationary and strongly dependent?

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 9.1, 12.2, 13.2 and 13.6.

Dougherty, C. Subject guide (2016): Chapters 9 and 13.

Approaching the question

(a) Substituting x_t and y_t in appropriately gives the following reduced forms:

$$y_t = \frac{\alpha + \beta\varphi_0}{1 - \beta\varphi_1} + \frac{\rho}{1 - \beta\varphi_1} y_{t-1} + \frac{\gamma}{1 - \beta\varphi_1} t + \frac{\beta u_t + \varepsilon_t}{1 - \beta\varphi_1}$$

$$x_t = \frac{\varphi_0 + \varphi_1\alpha}{1 - \beta\varphi_1} + \frac{\rho\varphi_1}{1 - \beta\varphi_1} y_{t-1} + \frac{\varphi_1\gamma}{1 - \beta\varphi_1} t + \frac{\varphi_1\varepsilon_t + u_t}{1 - \beta\varphi_1}.$$

Equation (10.2) is identified since the exogenous variables t and y_{t-1} can be used as instruments for the single endogenous variable y_t in this equation. Both exogenous variables are excluded from (10.2). This equation is overidentified. Equation (10.1), though, is not identified since there is an instrument we can use for the single endogenous variable x_t in this equation. There are no exclusions in (10.1), hence this equation is underidentified.

(b) Consistency of the OLS estimator of (10.1) requires that the error ε_t is uncorrelated with the regressors y_{t-1}, x_t, t. We do not need to worry about correlation between ε_t and fixed regressors such at t. If $x_t = \varphi_0 + u_t$, uncorrelatedness between x_t and ε_t requires uncorrelatedness between ε_t and u_t which is given in the question. To ensure that y_{t-1} is uncorrelated with ε_t finally we require the absence of autocorrelation in ε_t which is true by assumption.

(c) We are asked to test for serial correlation in the error ε_t, from the following regression model:

$$y_t = \alpha + \rho y_{t-1} + \beta x_t + \varepsilon_t, \quad |\rho| < 1.$$

Say we postulate an AR(1) process for ε_t we test H_0: no autocorrelation against H_1: AR(1) presence of autocorrelation. We should apply the Breusch–Godfrey test (the Durbin–Watson test cannot be applied because of the presence of the lagged endogenous variable).

Using the OLS residuals, we run the following regression:

$$\widehat{\varepsilon}_t = \gamma_0 + \gamma_1 y_{t-1} + \gamma_2 x_t + \delta\widehat{\varepsilon}_{t-1} + v_t.$$

The test statistic is given by nR^2 of this regression, where n is the sample size. Under the null hypothesis its asymptotic distribution is χ^2_1. We reject H_0 if the realisation exceeds the critical value given by this distribution for a given significance level.

(d) We are asked to consider the following model:

$$y_t = \alpha + \beta x_t + \varepsilon_t$$

in the setting where both y_t and x_t are expected to be integrated of order one.

i. If ε_t is non-stationary and strongly dependent, y_t and x_t are unrelated. Nevertheless, running the regression (10.1) may still give significant parameter estimates since we face the case of spurious regression, wrongly suggesting that there is a relationship between y_t and x_t. This is because the estimates are inconsistent. This is the spurious regression problem.

ii. If ε_t is stationary and weakly dependent, then the two unit root processes y_t and x_t must be cointegrated. This means the parameters α and β can be consistently estimated by OLS.

Examiners' commentaries 2019

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2018–19. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone B

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Section A

Answer all questions from this section.

Question 1

Consider the simple linear regression model:

$$Y_i = \beta X_i + u_i, \quad i = 1, \ldots, n.$$

We assume that the errors $\{u_i\}_{i=1}^n$ are independent normal random variables with zero mean. The regressor $\{X_i\}_{i=1}^n$ is non-stochastic (fixed under repeated sampling). You suspect that the errors exhibit heteroskedasticity.

(a) **Explain what we mean by the concept of heteroskedasticity. Enhance your answer with the help of a graphical illustration.**

(3 marks)

(b) **Derive the variance of the OLS estimator for $\widehat{\beta}$ in the presence of heteroskedastity and explain how a robust (White) standard error for $\widehat{\beta}$ can be obtained.**

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 7.1 and 7.3.

Dougherty, C. Subject guide (2016): Chapter 7.

Approaching the question

(a) Heteroskedasticity means that the error variance depends on the regressors, it violates the assumption that $\mathrm{E}(u_i^2) = \sigma^2$.

A graphical illustration should show data points and some depiction of the variance – for example, drawing the distribution for the errors at different levels of X. The graph should be clearly labelled.

(b) The variance of the estimator $\widehat{\beta}$ depends on X_i and, given the unbiasedness, is given by:

$$\mathrm{Var}(\widehat{\beta}) = \mathrm{E}\left((\widehat{\beta} - \beta)^2\right) = \mathrm{E}\left(\frac{\left(\sum_i X_i u_i\right)^2}{\left(\sum_i X_i^2\right)^2}\right) = \frac{\sum_i X_i^2 \mathrm{E}(u_i^2)}{\left(\sum_i X_i^2\right)^2} = \frac{\sum_i X_i^2 \sigma_i^2}{\left(\sum_i X_i^2\right)^2}$$

where σ_i^2 is the variance of the error term which depends on i and the second equality follows from fixed X in repeated samples and independence of u_i across observations giving $\mathrm{E}(u_i u_j) = 0 \ \forall \ i \neq j$. Many candidates struggled to derive the variance.

The White robust standard error for $\widehat{\beta}$ is the square root of the estimate of the above variance. The estimate would be constructed using the residuals from an OLS regression on the original model:

$$\widehat{\mathrm{Var}}(\widehat{\beta}) = \frac{\sum_i X_i^2 \widehat{u}_i^2}{\left(\sum_i X_i^2\right)^2}.$$

Question 2

Consider the human capital earnings function given by:

$$earnings_i = \beta_0 + \beta_1 educ_i + \beta_2 exper_i + \beta_3 exper_i^2 + u_i, \quad i = 1, \ldots, n$$

where *earnings* denotes the hourly earnings of an individual and *educ* and *exper* denote the years of schooling and experience, respectively. We assume we have obtained a random sample $\{(earnings_i, educ_i, exper_i)\}_{i=1}^n$ from the population. The errors $\{u_i\}_{i=1}^n$ are i.i.d. normal random variables with zero mean and variance σ^2. We assume independence between the errors and regressors (i.e., we ignore the usual ability bias problem).

(a) Discuss the rationale for including both *exper* and $exper^2$ in this model. In your answer be explicit about the expected signs for β_2 and β_3.

(2 marks)

(b) Discuss briefly how would you test that *exper* has a significant effect on *earnings*?

(2 marks)

(c) Provide a clear interpretation of the parameter β_1. You are told that you can obtain the estimator for β_1 by running the regression:

$$earnings_i^* = \beta_1 educ_i^* + e_i, \quad i = 1, \ldots, n$$

where $earnings_i^*$ and $educ_i^*$ are obtained from running a regression of *earnings* (and *educ*) on an intercept, *exper* and $exper^2$. Explain this statement.

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapter 3.2 and 3.5.

Dougherty, C. Subject guide (2016): Chapter 3.

Approaching the question

(a) The rationale is to allow for diminishing returns to experience, β_2 positive and β_3 negative. For other values of the βs a different explanation should be given.

(b) Since both β_2 and β_3 capture experience, such a test needs to be a joint test with the hypotheses:

$$\mathrm{H}_0 : \beta_2 = \beta_3 = 0 \quad \text{vs.} \quad \mathrm{H}_1 : \beta_2 \neq 0 \text{ and/or } \beta_3 \neq 0.$$

The F test can be used for such a joint hypothesis with the critical values on the $F_{2,\,n-4}$ distribution.

Many candidates failed to recognise that this needs to be a joint test since we have two different regressors which capture experience.

(c) The parameter β_1 is the marginal effect of education on earnings; that is, it measures by how many (monetary) units hourly earnings would increase as a response to someone going to school for one more year, *ceteris paribus*.

The parameter estimate in the original model can also be obtained by the given equation where the starred variables denote the residuals from running a regression of the respective variable on the left-out intercept and regressors *exper* and *exper*^{2}. For both *earnings* and *educ*, this would isolate the remaining variation in the variables after controlling for the intercept, *exper* and *exper*^{2}. Hence this procedure in two steps is the same as controlling for the three left-out variables in the original regression equation. The is the Frisch–Waugh–Lovell result (regression anatomy).

Question 3

Consider the OLS estimator for β in the linear regression model:

$$Y_i = \beta X_i + \varepsilon_i, \quad i = 1, \ldots, n$$

where $\{(Y_i, X_i)\}_{i=1}^n$ form an i.i.d. sample from a population and the errors are drawn from an <u>unknown</u> distribution with mean zero and variance σ^2. You are told that X_i and ε_i are <u>uncorrelated</u> (not necessarily independent therefore!).

(a) Discuss the importance of convergence in probability.

(4 marks)

(b) Discuss the importance of convergence in distribution.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections R.14 and R.15.

Dougherty, C. Subject guide (2016): Chapter 8.

Approaching the question

(a) Convergence in probability plays an important role when proving *consistency* of estimators. A consistent estimator converges in probability to the true parameter and we either write $\text{plim}(\widehat{\beta}) = \beta$ or $\widehat{\beta} \xrightarrow{p} \beta$.

Proofs of consistency (a large sample property) are particularly important in settings where we cannot show that our estimator is unbiased (a finite sample property). For the OLS estimator:

$$\text{plim}(\widehat{\beta}) = \text{plim}(\beta) + \frac{\text{plim}(\frac{1}{n}\sum X_i\varepsilon_i)}{\text{plim}(\frac{1}{n}\sum X_i^2)} = \beta + \frac{\text{E}(X_i\varepsilon_i)}{\text{E}(X_i^2)} = \beta$$

where we use the properties of the plim operator and the Law of Large Numbers to argue that the sample averages converge in probability to their population equivalents.

(b) Convergence in distribution plays an important role in settings where we do not know the exact sampling distribution of our estimator, such as in the linear regression setting when the distribution of u is unknown. We need a distributional result in order for us to conduct hypothesis testing.

The central limit theorem allows us to obtain an approximate (asymptotic) distribution for $\widehat{\beta}$ even without knowing the true distribution of u. The asymptotic distribution provides a good approximation of the true sampling distribution as long as the sample is sufficiently large. This powerful result enables us to conduct hypothesis testing which is valid assuming our sample is large.

Question 4

Consider the following regression model:

$$\boldsymbol{Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t, \quad t = 1, \ldots, T.}$$

Both $\{Y_t\}_{t=1}^T$ and $\{X_t\}_{t=1}^T$ are difference stationary processes.

(a) Discuss the concept 'difference stationarity' and contrast it to the concept 'trend stationarity'. In your answer make sure you also explain what stationarity means.

(4 marks)

(b) How can you test whether the above relation is spurious or cointegrating? In your answer you are expected to explain the difference between a spurious and cointegrating relationship.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 13.1 and 13.4.

Dougherty, C. Subject guide (2016): Chapter 13.

Approaching the question

(a) *Trend stationarity* relates to a process which has a deterministic trend ($y_t = \alpha + \beta t + \varepsilon_t$ where ε_t is weakly dependent). It is a type of non-stationarity which can be removed by detrending. Once we include a trend in the regression, we can use standard statistical inference methods: the errors (and, therefore, Y_t) will be weakly dependent and consistent estimates can be obtained.

Difference stationarity relates to a process which can be made stationary by taking first differences: ΔY_t, examples are random walk and random walk with drift. Here Y_t is persistent, strongly dependent, and estimation of the process Y_t on Y_{t-1} would fail to produce consistent parameter estimates.

By stationarity we mean *covariance stationarity*, which requires constant (finite) means and variances, and covariances which only depend on distance in time, i.e. $\text{E}(Y_t, Y_{t-s})$ only depends on s, not on location t.

(b) If Y_t and X_t are cointegrated, the error term in this model will be covariance stationary and consistent estimates of β_0 and β_1 can be obtained. The existence of a cointegrating relationship indicates the presence of a long-run relationship (existence of an ECM). If the error term has a unit root, however, the problem of a spurious relation will arise (the relationship is meaningless). Even if Y_t and X_t are independent, estimates of β_0 and β_1 may show up significant and suggest there would be a relationship. This is because a unit root error term implies that the parameters cannot consistently be estimated and we cannot use the usual distribution to obtain critical values.

A test which allows us to distinguish between the two cases would be a test whether the error term ε_t has a unit root. We can conduct this by performing a *Dickey–Fuller test* on the differenced residuals $\Delta\widehat{\varepsilon}_t$.

Question 5

Consider the simple linear regression model:

$$\boldsymbol{Y_i = \beta_0 + \beta_1 X_i + u_i, \quad i = 1, \ldots, n.}$$

We assume that the errors $\{u_i\}_{i=1}^n$ are independent random variables with zero mean. The regressor $\{X_i\}_{i=1}^n$ is non-stochastic (fixed under repeated sampling). Under these conditions, the OLS estimator for β_1, $\widehat{\beta}_1$, is unbiased. (You are not asked to derive $\widehat{\beta}_1$).

(a) Explain the concept of unbiasedness of an estimator.

(2 marks)

(b) Let us consider two other estimators for the slope β_1:

$$\widehat{\beta}_1^{\circ} = \frac{\sum_{i=1}^{n}(Z_i - \bar{Z})Y_i}{\sum_{i=1}^{n}(Z_i - \bar{Z})X_i} \quad \text{and} \quad \widehat{\beta}_1^{*} = \frac{\sum_{i=1}^{n}(Z_i - \bar{Z})Y_i}{\sum_{i=1}^{n}(Z_i - \bar{Z})Z_i}$$

where $Z_i = \sqrt{X_i}$ for all i and $\bar{Z} = \frac{1}{n}\sum_{i=1}^{n} Z_i$. Please indicate whether $\widehat{\beta}_1^{\circ}$ and $\widehat{\beta}_1^{*}$ are unbiased estimators for β_1. Clearly show your derivations.

(4 marks)

(c) Briefly indicate how you would choose between the three estimators, $\widehat{\beta}_1$, $\widehat{\beta}_1^{\circ}$ and $\widehat{\beta}_1^{*}$.

(2 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 2.3 and 2.5.

Dougherty, C. Subject guide (2016): Chapter 2.

Approaching the question

(a) Unbiasedness means $\mathrm{E}(\widehat{\beta}_1) = \beta_1$. It means that the expected value of the estimator is the true parameter; that is, we are correct on average in repeated samples. This ensures that we will not make systematic errors when estimating β.

(b) In answering this question, it is easiest to derive results using Z_i rather than working with $Z_i = \sqrt{X_i}$. It is important to note furthermore that $\sum(Z_i - \bar{Z}) = 0$.

When plugging in the true model ($Y_i = \beta_0 + \beta_1 X_i + u_i$), we can then write:

$$\widehat{\beta}_1^{\circ} = \beta_1 + \frac{\sum(Z_i - \bar{Z})u_i}{\sum(Z_i - \bar{Z})X_i}$$

and:

$$\mathrm{E}(\widehat{\beta}_1^{\circ}) = \beta_1 + \frac{\sum(Z_i - \bar{Z})\mathrm{E}(u_i)}{\sum(Z_i - \bar{Z})X_i} = \beta_1$$

(X and Z are assumed to be non-stochastic). Unbiasedness follows as $\mathrm{E}(u_i) = 0$.

For the second estimator, when plugging in the true model, we can write:

$$\widehat{\beta}_1^{*} = \beta_1 + \frac{\sum(Z_i - \bar{Z})X_i}{\sum(Z_i - \bar{Z})Z_i} + \frac{\sum(Z_i - \bar{Z})u_i}{\sum(Z_i - \bar{Z})Z_i}$$

and:

$$\mathrm{E}(\widehat{\beta}_1^{*}) = \beta_1 + \frac{\sum(Z_i - \bar{Z})Z_i^2}{\sum(Z_i - \bar{Z})Z_i} \neq \beta_1$$

using again the fact that Z is non-stochastic and $\mathrm{E}(u_i) = 0$, hence $\widehat{\beta}_1^{*}$ is biased.

Side notes: (i) $\widehat{\beta}_1^{\circ}$ essentially is an IV estimator where Z is used as an instrument for X. Both Z and X are exogenous here. (ii) $\widehat{\beta}_1^{*}$, on the other hand, is the slope parameter of a regression of Z instead of X on Y, and hence should not give us an unbiased estimator of β_1.

Many candidates failed to realise that $\sum(Z_i - \bar{Z}) = 0$ and wrongly concluded that the first estimator $\widehat{\beta}_1^{\circ}$ is biased.

(c) When choosing between $\widehat{\beta}_1$ and $\widehat{\beta}_1^{\circ}$ (both unbiased), efficiency considerations matter. Since the model satisfies the Gauss–Markov assumptions, we know that by the Gauss–Markov theorem, $\widehat{\beta}_1$ is BLUE and should, therefore, be chosen over $\widehat{\beta}_1^{\circ}$. Nevertheless, there may be a trade-off between bias and variance, and the mean squared error of $\widehat{\beta}_1^{*}$ could be smaller than that of $\widehat{\beta}_1$ (which is unbiased).

The bias–variance trade-off was mentioned only by a few candidates.

Section B

Answer three questions from this section.

Question 6

Let us consider the expectations augmented Phillips curve (see also Mankiw, 1994):

$$\boldsymbol{infl_t - infl_t^e = \beta_1(unem_t - \mu_0) + e_t}$$

where μ_0 is the natural rate of unemployment (assumed to be constant over time) and $infl_t^e$ is the expected rate of inflation formed in $t-1$.

This model suggests that there is a trade-off between unanticipated inflation ($infl_t - infl_t^e$) and cyclical unemployment (difference between actual unemployment and the natural rate of unemployment). We assume that e_t (also called supply shock) is an i.i.d. random variable with zero mean.

(a) You are told that expectations are formed as follows:

$$\boldsymbol{infl_t^e - infl_{t-1}^e = \lambda(infl_{t-1} - infl_{t-1}^e).}$$

What name do we give such a process and how should we interpret λ?

(2 marks)

(b) Show that you can rewrite the model as:

$$\boldsymbol{\Delta infl_t = \gamma_0 + \gamma_1 unem_t + \gamma_2 unem_{t-1} + v_t} \tag{6.1}$$

where $\boldsymbol{\Delta infl_t = infl_t - infl_{t-1}}$. Clearly indicate the relation between $\boldsymbol{(\gamma_0, \gamma_1, \gamma_2)}$ and $\boldsymbol{(\mu_0, \beta_1, \lambda)}$ and show that:

$$\boldsymbol{v_t = e_t - (1 - \lambda)e_{t-1}.}$$

***Hint:* If you want you may use the following shorthand notation in your derivations: $\boldsymbol{y_t = infl_t}$, $\boldsymbol{y_t^e = infl_t^e}$ and $\boldsymbol{x_t = unem_t}$.**

(7 marks)

(c) Discuss what assumptions you would like to make about $\boldsymbol{e_t}$ (the supply shock) that will guarantee that the OLS estimator for the parameters in (6.1) is consistent. *Hint:* You may want to give the assumptions you need to make about $\boldsymbol{v_t}$ (composite error term) first.

(3 marks)

(d) Show how you can obtain a consistent estimator for $\boldsymbol{\lambda}$ using your consistent estimates for $\boldsymbol{(\gamma_0, \gamma_1, \gamma_2)}$. Provide a proof of its consistency. [*Note:* If you did not manage to get an explicit relation between $\boldsymbol{\lambda}$ and $\boldsymbol{(\gamma_0, \gamma_1, \gamma_2)}$, consider $\boldsymbol{\lambda = g(\gamma_0, \gamma_1, \gamma_2)}$ where $\boldsymbol{g(\cdot)}$ is some continuous function].

(3 marks)

(e) One of the assumptions provided in (c) is rather unreasonable (*Hint:* future unemployment may be related to current supply shocks). Discuss how you could use IV (TSLS) to obtain a consistent estimator for the parameters in (6.1). Discuss what conditions your instruments need to satisfy and propose a suitable instrument.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections R.14, 9.3, 11.3, 11.4 and 11.5.

Dougherty, C. Subject guide (2016): Chapters 8 and 11.

Approaching the question

(a) This is an adaptive expectations model. The parameter λ, which should lie between 0 and 1, indicates the speed with which expectations are adjusted in response to deviations between the actual and the expected rate of inflation in the previous period.

(b) To obtain this it is useful to realise that the equation given in part (a) can be rewritten as:

$$y_t^e = (1 - \lambda)y_{t-1}^e + \lambda y_{t-1}. \tag{*}$$

This shows that if we subtract $(1 - \lambda)$ times the lagged expectation augmented Phillips curve, we can get rid of the inflation expectation variable. Let us substract from the first equation given in the question, $1 - \lambda$ times the equation one period lagged:

$$y_t - y_t^e - (1 - \lambda)[y_{t-1} - y_{t-1}^e] = \beta_1(x_t - \mu_0) + e_t - (1 - \lambda)[\beta_1(x_{t-1} - \mu_0) + e_{t-1}].$$

Rearranging yields:

$$y_t - (1 - \lambda)y_{t-1} - (y_t^e - (1 - \lambda)y_{t-1}^e) = \beta_1 x_t - (1 - \lambda)\beta_1 x_{t-1} - \lambda\beta_1\mu_0 + e_t - (1 - \lambda)e_{t-1}.$$

Using (*) we get:

$$y_t - (1 - \lambda)y_{t-1} - \lambda y_{t-1} = \beta_1 x_t - (1 - \lambda)\beta_1 x_{t-1} - \lambda\beta_1\mu_0 + v_t$$

which yields:

$$\Delta y_t = -\beta_1 \lambda \mu_0 + \beta_1 x_t - (1-\lambda)\beta_1 x_{t-1} + v_t$$

which gives rise to the following relationship between the γs and original parameters:

$$\gamma_0 = -\beta_1 \lambda \mu_0, \quad \gamma_1 = \beta_1 \quad \text{and} \quad \gamma_2 = -(1-\lambda)\beta_1.$$

(c) Consistency requires the error term in the new model v_t and both regressors, x_t and x_{t-1} to be uncorrelated. Since e_t has a zero mean this means:

$$\mathrm{E}(x_t(e_t - (1-\lambda)e_{t-1})) = 0 \quad \text{and} \quad \mathrm{E}(x_{t-1}(e_t - (1-\lambda)e_{t-1})) = 0.$$

A sufficient condition for this would be that $\mathrm{E}(e_t x_s) = 0$ for $s \in \{t-1, t, t+1\}$ or uncorrelatedness between x_t and e_s $\forall$ s, t.

(d) The above relationships can be solved for λ as:

$$\lambda = \frac{\gamma_2}{\gamma_1} + 1.$$

Using the estimators $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$, an estimator $\widehat{\lambda}$ can be constructed accordingly. For consistent estimators $\widehat{\gamma}_1$ and $\widehat{\gamma}_2$, the estimator $\widehat{\lambda}$ is also consistent by the Slutsky theorem:

$$\mathrm{plim}(\widehat{\lambda}) = \mathrm{plim}\left(\frac{\widehat{\gamma}_2}{\widehat{\gamma}_1} + 1\right) = \frac{\mathrm{plim}(\widehat{\gamma}_2)}{\mathrm{plim}(\widehat{\gamma}_1)} + 1 = \frac{\gamma_2}{\gamma_1} + 1 = \lambda.$$

Slutsky's theorem (probability limit rules) guarantees the second equality since these are continuous functions.

(e) If future unemployment is related to current supply shocks, then $\mathrm{Cov}(x_t, e_{t-1}) \neq 0$ and we get inconsistency as that would mean $\mathrm{Cov}(x_t, v_t) \neq 0$.

As long as unemployment is only related to past supply shocks, but not current or future ones x_{t-1} is still uncorrelated with v_t, $\mathrm{Cov}(x_{t-1}, v_t) = 0$.

We should proceed using an *instrumental variable approach* (2SLS), where we need to look for an instrument to deal with the endogeneity of x_t (this could for instance be a demand shock which is unrelated with supply shocks). Call this instrument z_t.

The instrument must satisfy instrument relevance and instrument exogeneity. Instrument relevance means that the instrument must be correlated with the original regressor, that is $\mathrm{Cov}(x_t, z_t) \neq 0$. Instrument exogeneity requires the instrument to be uncorrelated with the error term; that is, $\mathrm{Cov}(z_t, v_t) = 0$. Since v_t is a function of both e_t and e_{t-1}, the latter assumption is satisfied when the instrument is unrelated with any current and past supply shocks.

In this particular case, due to the fact that v_t has an MA(1) error structure, we can also use x_{t-2} as an instrument. We expect there to be correlation between x_t and x_{t-2} and x_{t-2} (the past) is uncorrelated with e_t and e_{t-1} which make up v_t (we cannot use x_{t-1} because that variable is already included in the model).

A TSLS procedure would proceed by running a regression of x_t on the instrument and the other regressors to obtain fitted values $\widehat{x}_t$. In the second stage, the outcome variable Δy_t would be regressed on an intercept, the predicted values $\widehat{x}_t$ and x_{t-1}. Alternatively, we can apply IV in the second stage where we use $\widehat{x}_t$ as our instrument for x_t. This would give rise to consistent TSLS estimators $(\gamma_0^{TSLS}, \gamma_1^{TSLS}, \gamma_2^{TSLS})$.

Question 7

This question is based on 'Openness and Inflation: Theory and Evidence', by Romer (*Quarterly Journal of Economics*, 1993). Romer proposed theoretical models of inflation that imply that more 'open' countries should have lower inflation. A simple macroeconomic model is:

$$infl_i = \beta_0 + \beta_1 open_i + \beta_2 \ln(pcinc_i) + u_{1i} \tag{7.1}$$

$$open_i = \alpha_0 + \alpha_1 infl_i + \alpha_2 \ln(pcinc_i) + \alpha_3 \ln(land_i) + u_{2i}, \quad i = 1, \ldots, n \tag{7.2}$$

where *infl* is the average annual inflation rate (since 1973), *open* is the average share of imports in gross domestic (or national) product since 1973, $\ln(pcinc)$ is the log of per capita income in US dollars, and $\ln(land)$ is the log of land area in square miles. The variables $\ln(pcinc)$ and $\ln(land)$ are treated as exogenous variables; u_{1i} and u_{2i} are serially uncorrelated disturbances with zero mean, variances σ_1^2 and σ_2^2 and covariance σ_{12}.

(a) Explain the concepts of endogenous versus exogenous explanatory variables and show that *open* is an endogenous variable in (7.1).

(5 marks)

(b) Discuss the identification of each structural form equation and explain what it means to say that an equation is exact identified.

(5 marks)

The table below shows the OLS and IV estimation results for equation (7.1). Standard errors are reported in parenthesis.

	OLS	IV
constant	25.109	26.90
	(15.205)	(15.40)
open	−0.215	−0.337
	(0.095)	(0.144)
$\ln(pcinc)$	0.018	0.376
	1.975	(2.015)

The IV estimator is obtained by solving the following three conditions (no need to show!):

$$\sum_{i=1}^{n} \left(infl_i - \widehat{\beta}_0^{IV} - \widehat{\beta}_1^{IV} open_i - \widehat{\beta}_1^{IV} \ln(pcinc_i) \right) = 0$$

$$\sum_{i=1}^{n} \left(infl_i - \widehat{\beta}_0^{IV} - \widehat{\beta}_1^{IV} open_i - \widehat{\beta}_1^{IV} \ln(pcinc_i) \right) \ln(pcinc_i) = 0$$

$$\sum_{i=1}^{n} \left(infl_i - \widehat{\beta}_0^{IV} - \widehat{\beta}_1^{IV} open_i - \widehat{\beta}_1^{IV} \ln(pcinc_i) \right) \ln(land_i) = 0.$$

(c) Describe in detail how you would estimate the parameters of (7.1) using Two Stage Least Squares (TSLS). How would these estimates compare with the reported IV estimates?

(5 marks)

(d) Conduct a test to see whether you can find evidence that more 'open' countries have lower inflation. Clearly indicate the null and alternative hypothesis, test statistic and its (asymptotic) distribution under the null. Briefly indicate how you could test whether the OLS and IV results are significantly different (so that it matters whether you use the OLS or IV parameter estimates).

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 2.6, 9.1 and 9.3.

Dougherty, C. Subject guide (2016): Chapters 8 and 9.

Approaching the question

(a) Exogenous variables are uncorrelated with the error term ('outside of the model') while endogenous variables are correlated with the error term ('inside the model'). This has implications for the estimates of the coefficients of these variables. While we can obtain consistent estimates under weak exogeneity and unbiased estimates under strong exogeneity, endogeneity implies that the estimates are generally inconsistent and biased. Since we face a simultaneity problem between infl_i and open_i in this problem, both will be endogenous. To show that open_i is an endogenous variable in (7.1), i.e. that $\text{Cov}(\mathit{open}_i, u_{1i}) \neq 0$, candidates should first derive the reduced form for open_i, which is:

$$\begin{aligned}
\mathit{open}_i &= \alpha_0 + \alpha_1(\beta_0 + \beta_1 \mathit{open}_i + \beta_2 \ln(\mathit{pcinc}_i) + u_{1i}) + \alpha_2 \ln(\mathit{pcinc}_i) \\
&\quad + \alpha_3 \ln(\mathit{land}_i) + u_{2i} \\
(1 - \alpha_1\beta_1)\mathit{open}_i &= (\alpha_0 + \alpha_1\beta_0) + (\alpha_1\beta_2 + \alpha_2)\ln(\mathit{pcinc}_i) + \alpha_3 \ln(\mathit{land}_i) + \alpha_1 u_{1i} + u_{2i} \\
\mathit{open}_i &= \frac{\alpha_0 + \alpha_1\beta_0}{1 - \alpha_1\beta_1} + \frac{\alpha_1\beta_2 + \alpha_2}{1 - \alpha_1\beta_1}\ln(\mathit{pcinc}_i) + \frac{\alpha_3}{1 - \alpha_1\beta_1}\ln(\mathit{land}_i) + \frac{\alpha_1 u_{1i} + u_{2i}}{1 - \alpha_1\beta_1}.
\end{aligned}$$

Using the exogeneity of $\ln(\mathit{pcinc}_i)$ and $\ln(\mathit{land}_i)$ we can then show:

$$\text{Cov}(\mathit{open}_i, u_{1i}) = \text{Cov}\left(\frac{\alpha_1 u_{1i} + u_{2i}}{1 - \alpha_1\beta_1}, u_{1i}\right) = \frac{\alpha_1\sigma_1^2 + \sigma_{12}}{1 - \alpha_1\beta_1} \neq 0 \quad \text{generally.}$$

(b) Identification of equation (7.1) requires that we have at least one variable which we can use for the single endogenous variable in this equation open_i $(G-1)$. Here we have exactly one such variable $\ln(\mathit{land}_i)$ (R) which gives us exact identification $(G - 1 = R)$. The reduced form reveals the relevance of this instrument, the validity of the instrument is by assumption, and finally the equation is excluded from (7.1) itself. Hence (7.1) is exactly identified. In equation (7.2) there is also a single $(G-1)$ endogenous variable, infl_i, for which at least one instrument is required. Unfortunately, there is no exogenous variable excluded from (7.2) which we can use. Hence (7.2) is underidentified $(R < G - 1)$.

(c) Using land_i as an instrument for open_i, we would proceed in two stages. In the first stage, we regress open_i on all remaining exogenous regressors in the original specification and the instrument to obtain fitted values of open_i:

$$\widehat{\mathit{open}}_i = \widehat{\pi}_0 + \widehat{\pi}_1 \ln(\mathit{pcinc}_i) + \widehat{\pi}_2 \ln(\mathit{land}_i).$$

For the second stage, we regress the outcome variable infl_i on the original exogenous regressors and the fitted values of open_i:

$$\mathit{infl}_i = \beta_0 + \beta_1 \widehat{\mathit{open}}_i + \beta_2 \ln(\mathit{pcinc}_i) + u_{1i}.$$

Alternatively, the second stage applies IV to the original equation using the fitted values of open_i as an instrument for open_i.

Under exact identification, the two-stage least squares estimator coincides with the IV estimator which uses the moment conditions, $\widehat{\beta}_1^{TSLS} = \widehat{\beta}_1^{IV}$.

(d) This is a one-sided test on whether the IV estimator of β_1 is significantly smaller than zero or not, that is:

$$\text{H}_0 : \beta_1 = 0 \quad \text{vs.} \quad \text{H}_1 : \beta_1 < 0.$$

The test uses the test statistic:

$$T = \frac{\widehat{\beta}_1^{IV}}{\text{s.e.}(\widehat{\beta}_1^{IV})} \xrightarrow{d} N(0,\, 1) \text{ under } \text{H}_0.$$

We would evaluate the test statistic using the estimates above to obtain:

$$t = \frac{-0.337}{0.144} = -2.34$$

and we should reject H_0 if it takes values lower than the critical value of the standard normal distribution, -1.645 at the 5% significance level (one-sided). We reject H_0 if $t < z$ and fail to reject otherwise. Here, we would reject H_0. This test only works (approximately) for large samples since we are using the asymptotic distribution of the test statistic (IV estimators only have nice large sample properties).

We should conduct a Hausman test to detect whether the OLS and IV estimates differ, $H_0 : \text{Cov}(open_i, u_{1i}) = 0$ against $H_1 : \text{Cov}(open_i, u_{1i}) \neq 0$. Under the null hypothesis both IV and OLS are consistent, whereas under the alternative only IV is consistent. The test is:

$$\left(\widehat{\beta}^{OLS} - \widehat{\beta}^{IV}\right)' V \left(\widehat{\beta}^{OLS} - \widehat{\beta}^{IV}\right)^{-1} \left(\widehat{\beta}^{OLS} - \widehat{\beta}^{IV}\right) \xrightarrow{d} \chi^2_3 \text{ under } H_0.$$

For a given significance level we reject H_0 if the test statistic takes values which are too large relative to the critical values given by the χ^2_3 distribution.

Question 8

Let us consider how workplace smoking bans affect the incidence of smoking. Below, we use data on 10,000 US indoor workers from 1991 to 1993 taken from 'Do Workplace Smoking Bans Reduce Smoking', by Evans et al. (*American Economic Review*, 1999).

Let *smoker* be a dummy variable indicating whether a worker smokes (1 = yes, 0 = no) and *smkban* a dummy variable indicating whether there is a ban on smoking in the workplace (1 = yes, 0 = no).

(a) The following OLS regression results were obtained:

$$\widehat{smoker} = \underset{(0.007)}{0.290} - \underset{(0.009)}{0.078}\, smkban \tag{8.1}$$

$$n = 10000,\ R^2 = 0.0078,\ RSS = 1821.59$$

The standard errors (SEs) are in parentheses. Interpret the parameter estimates of the coefficient on *smkban*. Provide the (approximate) 95% confidence interval for the coefficient on *smkban*. How can we use this confidence interval to test the hypothesis that $\beta_{smkban} = 0$?

(6 marks)

A further specification was considered that included other characteristics of the worker: the age (in years), gender (male/female), ethnicity (black/hispanic/white), and level of education ($E1$ = highschool dropout, $E2$ = highschool graduate, $E3$ = some college, $E4$ = college graduate, $E5$ = Master degree or above). The following OLS regression results were obtained for this multiple linear regression model:

$$\widehat{smoker} = \underset{(0.019)}{0.201} - \underset{(0.009)}{0.045}\, smkban - \underset{(.009)}{0.033} female - \underset{(.0003)}{0.001}\, age - \underset{(.016)}{0.027} black \tag{8.2}$$

$$-\underset{(.014)}{0.104} hispanic + \underset{(.019)}{0.310} E1 + \underset{(.012)}{0.224} E2 + \underset{(.012)}{0.156} E3 + \underset{(.012)}{0.042} E4$$

$$n = 10000,\ R^2 = 0.0526,\ RSS = 1736.81$$

The SEs are in parentheses.

(b) Compare the coefficient estimates on *smkban* from the simple and multiple regression model in (8.1) and (8.2) and explain why the estimates differ.

(3 marks)

(c) Interpret the estimated parameter on $E2$ (highschool graduate) in (8.2) and indicate how you can obtain its p-value and what information the p-value provides.

(5 marks)

(d) Both the simple and multiple regression model suffer from heteroskedasticity. Explain why. What are the implications of heteroskedasticy for the parameter estimates and the standard errors in (8.1) and (8.2)? What can you do to resolve this problem? Explain your answer.

(6 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 2.6, 3.2, 6.2, 7.3 and 10.1.

Dougherty, C. Subject guide (2016): Chapters 5, 6 and 10.

Approaching the question

(a) This is the linear probability model, where:

$$\mathrm{E}(smoker = 1 \,|\, X) = P(smoker = 1 \,|\, X) = \beta_0 + \beta_1 smkban.$$

If the smoking ban is introduced, the probability of a worker smoking decreases by 7.8 percentage points (not 7.8%).

An (approximate) 95% confidence interval is:

$$\begin{aligned}&(\widehat{\beta}_1 - z_{crit,\,0.025} \times \text{s.e.}(\widehat{\beta}_1),\ \widehat{\beta}_1 + z_{crit,\,0.025} \times \text{s.e.}(\widehat{\beta}_1)\\ &= (-0.078 - 1.96 \times 0.009,\ -0.078 + 1.96 \times 0.009) \approx (-0.096,\ -0.060)\end{aligned}$$

where z relates to the standard normal distribution (for large degrees of freedom the t distribution converges to the standard normal).

We can use confidence intervals for testing. For a two-sided $\mathrm{H}_0 : \beta_1 = 0$ test, we reject H_0 at the 5% significance level if zero does not lie in the 95% confidence interval.

(b) The reason why the coefficient estimates on *smkban* differ in (8.1) and (8.2) is due to the *omitted variable bias problem*. When we omit relevant variables (for example, education and gender) which are related to the included variable *smkban*), then the parameter estimates on *smkban* will estimate not only the direct effect that *smkban* has on smoking (the parameter of interest) but the indirect effect of these omitted variables as well. We also call these omitted variables confounders.

The more negative effect found in (a) signals that once the effect of education on smoking (gender, age and ethnicity) is controlled for, the true effect of a smoking ban on the probability of smoking is smaller.

(c) The coefficient on $E2$ means that highschool graduates are 22.4 percentage points more likely to smoke relative to people with a Master's degree or above since $E5$ *is the left out base category*. (Many candidates failed to recognise that the left out category was required to interpret the parameter clearly). The t statistic for this effect is:

$$t = \frac{\widehat{\beta}}{\text{s.e.}(\widehat{\beta})} = \frac{0.224}{0.012} = 18.667$$

which has a p-value smaller than 0.05 (we would reject H_0 at the 5% significance level). The p-value is the lowest level of significance at which we can reject the null hypothesis. Since the sample size is large, we can use the standard normal distribution to obtain the p-value (candidates may provide a graphical discussion instead). For a one-sided test, $p = 1 - \Phi(t)$ and for a two-sided test, $p = 2 \times (1 - \Phi(t))$, assuming $t \geq 0$.

(d) In the linear probability model (LPM) we have to deal with the problem of heteroskedasticity. Conditional on X, $\mathrm{Var}(y \mid X) = p(X)(1 - p(X)) \equiv \mathrm{Var}(u \mid X)$. The implication of heteroskedasticity for the estimated standard errors is that they are wrong (they rely on homoskedasticity) and inference based on them would be invalid. The parameter estimates themselves remain unbiased and consistent. For inference, we should use heteroskedasticity-robust standard errors or apply WLS instead.

Question 9

In question 8 we considered the linear regression model to study how workplace smoking bans affect the incidence of smoking. Here we consider the results from applying a probit regression of *smoker* (1 = yes, 0 = no) on *smkban* (1 = yes, 0 = no) and the other explanatory variables:

```
. probit smoker smkban female age black hispanic E1 E2 E3 E4

Iteration 0:   log likelihood = -5537.1662
Iteration 1:   log likelihood = -5255.1526
Iteration 2:   log likelihood =  -5252.349
Iteration 3:   log likelihood = -5252.3489

Probit regression                               Number of obs     =     10,000
                                                LR chi2(9)        =     569.63
                                                Prob > chi2       =     0.0000
Log likelihood = -5252.3489                     Pseudo R2         =     0.0514
```

smoker	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
smkban	-.1517626	.0289268	-5.25	0.000	-.208458	-.0950671
female	-.1106249	.0287785	-3.84	0.000	-.1670298	-.05422
age	-.0042031	.0011748	-3.58	0.000	-.0065057	-.0019006
black	-.07969	.0525369	-1.52	0.129	-.1826604	.0232804
hispanic	-.3327039	.0476677	-6.98	0.000	-.4261308	-.2392769
E1	1.094231	.0714121	15.32	0.000	.9542663	1.234197
E2	.8518588	.0594747	14.32	0.000	.7352906	.9684271
E3	.6492566	.0606989	10.70	0.000	.530289	.7682241
E4	.2224747	.0649939	3.42	0.001	.0950891	.3498603
_cons	-.9842425	.0756055	-13.02	0.000	-1.132427	-.8360584

(a) **It is argued that using the probit regression model is better than using the linear probability model when explaining the binary variable *smoker*. Discuss the benefits/drawbacks of using the Probit model when trying to explain a binary variable.**

(5 marks)

(b) **Explain briefly how the Probit estimates are obtained and discuss the properties of the parameter estimates.**

Hint: **You may recall that for the Probit model, we will specify:**

$$\Pr(smoker = 1) = \Phi(\beta_0 + \beta_1 smkban + \beta_2 female + \cdots + \beta_8 E3 + \beta_9 E4)$$

where Φ is the standard normal CDF (cumulative distribution function).

(5 marks)

(c) **Explain how you can estimate the effect of the smoking ban on the probability of smoking for a 50-year old white, college graduated man. You are not expected to use your calculator, clarity of the computations required is enough.**

(5 marks)

(d) Discuss how you could test the joint significance of the worker's characteristics (gender, age, ethnicity and level of education) using the likelihood ratio test. Clearly indicate the test statistic, its distribution, the rejection rule and the additional information you would need to implement it.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 10.3 and 10.6.

Dougherty, C. Subject guide (2016): Chapter 10.

Approaching the question

(a) The probit model has three main advantages over the linear probability model (LPM): (i) predicted probabilities are restricted to lie in $[0, 1]$, (ii) maximum likelihood estimators are (asymptotically) efficient whereas OLS (LPM) estimators will be inefficient, and (iii) maximum likelihood estimators automatically deal with heteroskedasticity.

The main drawbacks of the probit model relative to the LPM are that (i) the coefficients cannot be directly interpreted as the marginal effects of the regressor(s) of interest, so we need to compute predicted probabilities using the probit specification, and (ii) it is also computationally more complicated.

(b) The parameters are estimated by maximum likelihood estimation, where the log-likelihood function is given by:

$$\begin{aligned}\log L(\beta) &= \sum_{i=1}^{n}\{smoker_i \log(P(smoker_i = 1 \mid X) + (1 - smoker_i)\log(P(smoker_i = 0 \mid X)\} \\ &= \sum_{i=1}^{n}\{smoker_i \log(\Phi(\beta_0 + \beta_1 smkban_i + \beta_2 female_i + \cdots + \beta_9 E4_i) + \\ &\quad (1 - smoker_i)\log(1 - \Phi(\beta_0 + \beta_1 smkban_i + \beta_2 female_i + \cdots + \beta_9 E4_i))\}.\end{aligned}$$

To obtain the parameter estimates the first-order conditions are solved (numerically as no explicit formulae exist).

Under suitable regularity conditions, the estimates are consistent, asymptotically normally, and asymptotically efficient.

(c) The effect is the difference in predicted probabilities between the man with the characteristics given, with $smkban = 1$ versus $smkban = 0$:

$$\begin{aligned}&\widehat{P}(y_i = 1 \mid X, smkban = 1) - \widehat{P}(y_i = 1 \mid X, smkban = 0) \\ &= \Phi\left(\sum_k x_{ki}\widehat{\beta}_k + \widehat{\beta}_1\right) - \Phi\left(\sum_k x_{ki}\widehat{\beta}_k\right) \\ &= \Phi(-0.984 + 0.222 - 0.004 \times 50 - 0.152) - \Phi(-0.984 + 0.222 - 0.004 \times 50).\end{aligned}$$

A discussion of marginal effects (ignoring the fact that *smkban* is a dummy variable) is acceptable as well.

(d) We would like to use the LR test, to test:

$$\begin{aligned}&\mathrm{H}_0 : \beta_j = 0 \ \forall\ j \in \{\text{worker characteristics}\} \\ &\mathrm{H}_1 : \beta_j \neq 0 \text{ for at least one } j \in \{\text{worker characteristics}\}.\end{aligned}$$

Estimate the restricted (R) and the unrestricted probit model (U). The restricted model imposes H_0. The unrestricted model is just the originally estimated model above. The LR

test makes use of the difference between the two estimated log-likelihood functions (the log ratio of the likelihood functions). The test statistic is:

$$LR = 2(\log L^U - \log L^R) \sim \chi^2_J$$

where J denotes the degrees of freedom which is equal to the number of restrictions; that is, $J = 8$ here. For a given significance level we reject H_0 if its realisation exceeds the critical value given by the χ^2_8 distribution, which equals 15.51 for the 5% significance level.

Question 10

This question is based on 'An Empirical Comparison of Alternative Models of the Short-Term Interest Rate', by Chan et al. (*The Journal of Finance*, 1992). In this article, the authors are interested in a regression model for the short-term interest rate r_t given by:

$$r_t - r_{t-1} = \beta_0 + \beta_1 r_{t-1} + \varepsilon_t. \tag{10.1}$$

Assume that the errors ε_t are not serially correlated.

(a) Provide sufficient conditions (on β_1) for r_t to be stationary and weakly dependent. Will the OLS estimator for a linear regression of r_t on r_{t-1} in this case be unbiased and consistent? Explain your answer.

(5 marks)

(b) A few well-known models in finance hypothesise that $\beta_1 = 0$. Explain how you would test this hypothesis. What name do we give the process r_t when $\beta_1 = 0$?

(5 marks)

Suppose you are interested in the following relation between long-term interest rates R_t and short-term interest rates r_t:

$$R_t - r_t = \alpha_0 + \alpha_1 r_t + u_t \tag{10.2}$$

where u_t is an error term that has zero mean.

(c) Let us assume $r_t = \beta_0 + \varepsilon_t$ (that is $\beta_1 = -1$). How would you test whether u_t is serially correlated? Explain your answer.

(5 marks)

(d) Let us assume $\beta_1 = 0$. Explaining your answers, what can you say about the OLS estimator for α_0 and α_1 applied to equation (10.2):

i. if u_t is stationary and weakly dependent?

ii. if u_t is non-stationary and strongly dependent?

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 12.2, 13.1, 13.2, 13.4 and 13.6.

Dougherty, C. Subject guide (2016): Chapter 13.

Approaching the question

(a) The model can be rewritten in the classic AR(1) structure:

$$r_t = \beta_0 + (1 + \beta_1) r_{t-1} + \varepsilon_t.$$

Stationarity in an AR(1) process requires the AR coefficient to be smaller than 1 implying $\beta_1 < 0$. We have weak dependence here. Under this assumption and given the absence of

serial correlation in ε_t, there is no correlation between the error and the regressor r_{t-1}, hence we can obtain consistent parameter estimates for β_1. Nevertheless, we will not get unbiasedness as ε_t will be correlated with future r_{t+j} $\forall\ j \geq 0$ due to the presence of lagged endogenous variables.

(b) Under $\beta_1 = 0$, the AR coefficient is 1 which indicates that r_t is a random walk with drift, and r_t is difference stationary, that is $\Delta r_t := r_t - r_{t-1}$ is stationary. Here we get strong dependence. To test whether we have a unit root, we can carry out the Dickey–Fuller test. Under $\mathrm{H}_0 : \beta_1 = 0$, r_t has a unit root, whereas under $\mathrm{H}_1 : \beta_1 < 0$, r_t is stationary (and weakly dependent). Using the Dickey–Fuller statistical tables, for a given significance level, we reject H_0 if the test statistic $\widehat{\beta}_1/\text{s.e.}(\widehat{\beta}_1)$ exceeds the appropriate critical value.

(c) If $r_t = \beta_0 + \varepsilon_t$, it is just a random error plus a constant. The proposed equation (10.2) can be estimated using OLS and yields consistent parameter estimates assuming u_t is uncorrelated with r_t.

Say we postulate an AR(1) processs for ε_t. We then test H_0: no autocorrelation against H_1: AR(1) presence of autocorrelation. We should apply the Breusch–Godfrey test (the Durbin–Watson test cannot be applied because of the presence of stochastic regressors).

Using the OLS residuals, we run the following regression:

$$\widehat{u}_t = \gamma_0 + \gamma_1 r_t + \delta \widehat{u}_{t-1} + v_t.$$

The test statistic is given by nR^2 of this regression, where n is the sample size. Under the null hypothesis its asymptotic distribution is χ^2_1. We reject H_0 if the realisation exceeds the critical value given by this distribution for a given significance level.

(d) Under $\beta_1 = 0$, r_t is integrated of order 1 and by implication so will be R_t.

i. If u_t is stationary and weakly dependent, then the two unit root processes R_t and r_t must be cointegrated. This means the parameters α and β can be consistently estimated by OLS.

ii. If u_t is non-stationary and strongly dependent, R_t and r_t are unrelated. Nevertheless, running the regression (10.1) may still give significant parameter estimates since we face the case of spurious regressions, wrongly suggesting that there is a relationship between R_t and r_t. This is because the estimates are inconsistent. This is the spurious regression problem.

THIS PAPER IS NOT TO BE REMOVED FROM THE EXAMINATION HALL

UNIVERSITY OF LONDON

EC2020 ZB

BSc DEGREES AND GRADUATE DIPLOMAS IN ECONOMICS, MANAGEMENT, FINANCE AND THE SOCIAL SCIENCES, THE DIPLOMA IN ECONOMICS AND SOCIAL SCIENCES AND THE CERTIFICATE IN EDUCATION IN SOCIAL SCIENCES

Elements of Econometrics

Thursday 23 May 2019 : 10.00 – 13.00

Time allowed: 3 hours

DO NOT TURN OVER UNTIL TOLD TO BEGIN

Candidates should answer **EIGHT** of the following **TEN** questions: **ALL** of the questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). **Candidates are strongly advised to divide their time accordingly.**

Extracts from statistical tables are given after the final question on this paper.

Graph paper is provided at the end of this question paper. If used, it must be detached and fastened securely inside the answer book.

A handheld calculator may be used when answering questions on this paper and it must comply in all respects with the specification given with your Admission Notice. The make and type of machine must be clearly stated on the front cover of the answer book.

If more questions are answered than requested, only the first answers attempted will be counted.

© University of London 2019

UL19/0000

SECTION A

Answer all questions from this section

1. Consider the simple linear regression model

$$Y_i = \beta X_i + u_i, \qquad i = 1, ..., n.$$

We assume that the errors $\{u_i\}_{i=1}^n$ are independent normal random variables with zero mean. The regressor $\{X_i\}_{i=1}^n$ is non-stochastic (fixed under repeated sampling). You suspect that the errors exhibit heteroskedasticity.

(a) **(3 marks)** Explain what we mean by the concept of heteroskedasticity. Enhance your answer with the help of a graphical illustration.

(b) **(5 marks)** Derive the variance of the OLS estimator for $\hat{\beta}$ in the presence of heteroskedastity and explain how a robust (White) standard error for $\hat{\beta}$ can be obtained.

2. Consider the human capital earnings function given by

$$\textit{earnings}_i = \beta_0 + \beta_1 \textit{educ}_i + \beta_2 \textit{exper}_i + \beta_3 \textit{exper}_i^2 + u_i, \qquad i = 1, ..., n,$$

where *earnings* denotes the hourly earnings of an individual and *educ* and *exper* denote the years of schooling and experience, respectively. We assume we have obtained a random sample $\{(\textit{earnings}_i, \textit{educ}_i, \textit{exper}_i)\}_{i=1}^n$ from the population. The errors $\{u_i\}_{i=1}^n$ are i.i.d. normal random variables with zero mean and variance σ^2. We assume independence between the errors and regressors (i.e., we ignore the usual ability bias problem).

(a) **(2 marks)** Discuss the rationale for including both *exper* and *exper*^{2} in this model. In your answer be explicit about the expected signs for β_2 and β_3.

(b) **(2 marks)** Discuss briefly how would you test that *exper* has a significant effect on *earnings?*

(c) **(4 marks)** Provide a clear interpretation of the parameter β_1. You are told that you can obtain the estimator for β_1 by running the regression

$$\textit{earnings}_i^* = \beta_1 \textit{educ}_i^* + e_i, \qquad i = 1, ..., n,$$

where $\textit{earnings}_i^*$ and $\textit{educ}_i^*$ are obtained from running a regression of *earnings* (and *educ)* on an intercept, *exper* and *exper*^{2}. Explain this statement.

3. Consider the OLS estimator for β in the linear regression model

$$Y_i = \beta X_i + \varepsilon_i, \qquad i = 1, ..., n,$$

where $\{(Y_i, X_i)\}_{i=1}^n$ form an i.i.d. sample from a population and the errors are drawn from an <u>unknown</u> distribution with mean zero and variance σ^2. You are told that X_i and ε_i are <u>uncorrelated</u> (not necessarily independent therefore!).

(a) **(4 marks)** Discuss the importance of convergence in probability.

(b) **(4 marks)** Discuss the importance of convergence in distribution.

4. Consider the following regression model

$$Y_t = \beta_0 + \beta_1 X_t + \varepsilon_t, \qquad t = 1, ..., T.$$

Both $\{Y_t\}_{t=1}^T$ and $\{X_t\}_{t=1}^T$ are difference stationary processes.

(a) **(4 marks)** Discuss the concept "difference stationarity" and contrast it to the concept "trend stationarity". In your answer make sure you also explain what stationarity means.

(b) **(4 marks)** How can you test whether the above relation is spurious or cointegrating? In your answer you are expected to explain the difference between a spurious and cointegrating relationship.

5. Consider the simple linear regression model

$$Y_i = \beta_0 + \beta_1 X_i + u_i, \qquad i = 1, ..., n.$$

We assume that the errors $\{u_i\}_{i=1}^n$ are independent random variables with zero mean. The regressor $\{X_i\}_{i=1}^n$ is non-stochastic (fixed under repeated sampling). Under these conditions, the OLS estimator for β_1, $\hat{\beta}_1$, is unbiased. (You are not asked to derive $\hat{\beta}_1$).

(a) **(2 marks)** Explain the concept of unbiasedness of an estimator.

(b) **(4 marks)** Let us consider two other estimators for the slope β_1 :

$$\hat{\beta}_1^\circ = \frac{\sum_{i=1}^n (Z_i - \bar{Z}) Y_i}{\sum_{i=1}^n (Z_i - \bar{Z}) X_i}; \quad \hat{\beta}_1^* = \frac{\sum_{i=1}^n (Z_i - \bar{Z}) Y_i}{\sum_{i=1}^n (Z_i - \bar{Z}) Z_i}$$

where $Z_i = \sqrt{X_i}$ for all i and $\bar{Z} = \frac{1}{n}\sum_{i=1}^n Z_i$. Please indicate whether $\hat{\beta}_1^\circ$ and $\hat{\beta}_1^*$ are unbiased estimators for β_1. Clearly show your derivations.

(c) **(2 marks)** Briefly indicate how you would choose between the three estimators, $\hat{\beta}_1$, $\hat{\beta}_1^\circ$ and $\hat{\beta}_1^*$.

SECTION B

Answer three questions from this section.

6. Let us consider the expectations augmented Phillips curve (see also Mankiw, 1994):

$$\textit{infl}_t - \textit{infl}_t^e = \beta_1 \left(\textit{unem}_t - \mu_0\right) + e_t,$$

where μ_0 is the natural rate of unemployment (assumed to be constant over time) and $\textit{infl}_t^e$ is the expected rate of inflation formed in $t-1$.

This model suggests that there is a tradeoff between unanticipated inflation ($\textit{infl}_t$-$\textit{infl}_t^e$) and cyclical unemployment (difference between actual unemployment and the natural rate of unemployment). We assume that e_t (also called supply shock) is an i.i.d. random variable with zero mean.

(a) **(2 marks)** You are told that expectations are formed as follows

$$\textit{infl}_t^e - \textit{infl}_{t-1}^e = \lambda \left(\textit{infl}_{t-1} - \textit{infl}_{t-1}^e\right)$$

What name do we give such a process and how should we interpret λ?

(b) **(7 marks)** Show that you can rewrite the model as

$$\Delta\textit{infl}_t = \gamma_0 + \gamma_1 \textit{unem}_t + \gamma_2 \textit{unem}_{t-1} + v_t, \tag{6.1}$$

where $\Delta\textit{infl}_t = \textit{infl}_t - \textit{infl}_{t-1}$. Clearly indicate the relation between $(\gamma_0, \gamma_1, \gamma_2)$ and $(\mu_0, \beta_1, \lambda)$ and show that

$$v_t = e_t - (1-\lambda)\, e_{t-1}.$$

Hint: If you want you may use the following shorthand notation in your derivations: $y_t = \textit{infl}_t$, $y_t^e = \textit{infl}_t^e$ and $x_t = \textit{unem}_t$.

(c) **(3 marks)** Discuss what assumptions you would like to make about e_t (the supply shock) that will guarantee consistency of the OLS estimator for (6.1). *Hint*: You may want to give the assumptions you need to make about v_t (composite error term) first.

(d) **(3 marks)** Show how you can obtain a consistent estimator for λ using your consistent estimates for $(\gamma_0, \gamma_1, \gamma_2)$? Provide a proof of its consistency. [*Note:* If you did not manage to get an explicit relation between λ and $(\gamma_0, \gamma_1, \gamma_2)$, consider $\lambda = g(\gamma_0, \gamma_1, \gamma_2)$ where $g(\cdot)$ is some continuous function.]

(e) **(5 marks)** One of the assumptions provided in (c) is rather unreasonable (*Hint*: future unemployment may be related to current supply shocks). Discuss how you could use IV (TSLS) to obtain a consistent estimator for the parameters in (6.1). Discuss what conditions you instruments need to satisfy and propose a suitable instrument.

7. This question is based on "Openness and Inflation: Theory and Evidence", by Romer (*Quarterly Journal of Economics,* 1993). Romer proposed theoretical models of inflation that imply that more "open" countries should have lower inflation. A simple macroeconomic model is

$$\textit{infl}_i = \beta_0 + \beta_1 \textit{open}_i + \beta_2 \ln(\textit{pcinc}_i) + u_{1i} \tag{7.1}$$

$$\textit{open}_i = \alpha_0 + \alpha_1 \textit{infl}_i + \alpha_2 \ln(\textit{pcinc}_i) + \alpha_3 \ln(\textit{land}_i) + u_{2i}, \qquad i = 1, ..., n \tag{7.2}$$

where *infl* is the average annual inflation rate (since 1973), *open* is the average share of imports in gross domestic (or national) product since 1973, $\ln(\textit{pcinc})$ is the log of per capita income in US dollars, and $\ln(\textit{land})$ is the log of land area in square miles. The variables $\ln(\textit{pcinc})$ and $\ln(\textit{land})$ are treated as exogenous variables; u_{1i} and u_{2i} are serially uncorrelated disturbances with zero mean, variances σ_1^2 and σ_2^2 and covariance σ_{12}.

(a) **(5 marks)** Explain the concepts of endogenous versus exogenous explanatory variables and show that $open$ is an endogenous variable in (7.1).

(b) **(5 marks)** Discuss the identification of each structural form equation and explain what it means to say that an equation is exact identified.

The table below shows the OLS and IV estimation results for equation (7.1). Standard errors are reported in parenthesis.

	OLS	IV
constant	25.109	26.90
	(15.205)	(15.40)
open	−0.215	−0.337
	(0.095)	(0.144)
ln(*pcinc*)	0.018	0.376
	(1.975)	(2.015)

The IV estimator is obtained by solving the following three conditions (no need to show!):

$$\sum_{i=1}^n \left(\textit{infl}_i - \hat{\beta}_0^{IV} - \hat{\beta}_1^{IV} \textit{open}_i - \hat{\beta}_2^{IV} \ln(\textit{pcinc}_i)\right) = 0$$

$$\sum_{i=1}^n \left(\textit{infl}_i - \hat{\beta}_0^{IV} - \hat{\beta}_1^{IV} \textit{open}_i - \hat{\beta}_2^{IV} \ln(\textit{pcinc}_i)\right) \ln(\textit{pcinc}_i) = 0$$

$$\sum_{i=1}^n \left(\textit{infl}_i - \hat{\beta}_0^{IV} - \hat{\beta}_1^{IV} \textit{open}_i - \hat{\beta}_2^{IV} \ln(\textit{pcinc}_i)\right) \ln(\textit{land}_i) = 0$$

(c) **(5 marks)** Describe in detail how you would estimate the parameters of (7.1) using Two Stage Least Squares (TSLS). How would these estimates compare with the reported IV estimates?

(d) **(5 marks)** Conduct a test to see whether you can find evidence that more "open" countries have lower inflation. Clearly indicate the null and alternative hypothesis, test statistic and its (asymptotic) distribution under the null. Briefly indicate how you could test whether the OLS and IV results are significantly different (so that it matters whether you use the OLS or IV parameter estimates).

8. Let us consider how workplace smoking bans affect the incidence of smoking. Below, we use data on 10,000 US indoor workers from 1991 to 1993 taken from "Do Workplace Smoking Bans Reduce Smoking", by Evans et al. (*American Economic Review*, 1999). Let *smoker* be a dummy variable indicating whether a worker smokes (1=yes, 0=no) and *smkban* a dummy variable indicating whether there is a ban on smoking in the workplace (1=yes, 0=no).

   (a) **(5 marks)** OLS regression results were obtained for the simple linear regression model

$$\begin{aligned} \widehat{\textit{smoker}} &= \underset{(0.007)}{0.290} - \underset{(0.009)}{0.078}\textit{smkban} \\ n &= 10000,\ R^2 = 0.0078,\ RSS = 1821.59 \end{aligned} \tag{8.1}$$

   The standard errors are in parentheses. Interpret the parameter estimates of the intercept and the coefficient on *smkban*. Provide the (approximate) 95% confidence interval for the coefficient on *smkban.* How can we use this confidence interval to test the hypothesis that $\beta_{\textit{smkban}} = 0$?

A further specification was considered that included other characteristics of the worker: the age (in years), gender (male/female), ethnicity (black/hispanic/white), and level of education (*E1*=highschool dropout, *E2*=highschool graduate, *E3*=some college, *E4*=college graduate, E5 Master degree or above). The following OLS regression results were obtained for this multiple linear regression model

$$\begin{aligned} \widehat{\textit{smoker}} &= \underset{(0.019)}{0.201} - \underset{(0.009)}{0.045}\textit{smkban} - \underset{(.009)}{0.033}\textit{female} - \underset{(.0003)}{0.001}\textit{age} - \underset{(.016)}{0.027}\textit{black} \\ &\quad -\underset{(.014)}{0.104}\textit{hispanic} + \underset{(.019)}{0.310}\textit{E1} + \underset{(.012)}{0.224}\textit{E2} + \underset{(.012)}{0.156}\textit{E3} + \underset{(.012)}{0.042}\textit{E4} \\ n &= 10000,\ R^2 = 0.0526,\ RSS = 1736.81 \end{aligned} \tag{8.2}$$

The standard errors are in parentheses.

   (b) **(3 marks)** Compare the estimates of the coefficient on *smkban* from the simple and multiple regression model in (8.1) and (8.2) and explain why the estimates differ.

   (c) **(4 marks)** Interpret the estimated parameter on *E2* (highschool graduate) in (8.2) and indicate how you can obtain its p-value and what information the p-value provides.

   (d) **(4 marks)** Both the simple and multiple regression model suffer from heteroskedasticity. Explain why. What are the implications of heteroskedasticy for the parameter estimates and the standard errors in (8.1) and (8.2)? What can you do to resolve this problem? Explain your answer.

9. In question 8 we considered the linear regression model to study how workplace smoking bans affect the incidence of smoking. Here we consider the results from applying a probit regression of *smoker* (1=yes, 0=no) on *smkban* (1=yes, 0=no) and the other explanatory variables

```
. probit smoker smkban female age black hispanic E1 E2 E3 E4

Iteration 0:   log likelihood = -5537.1662
Iteration 1:   log likelihood = -5255.1526
Iteration 2:   log likelihood =  -5252.349
Iteration 3:   log likelihood = -5252.3489

Probit regression                                 Number of obs     =     10,000
                                                  LR chi2(9)        =     569.63
                                                  Prob > chi2       =     0.0000
Log likelihood = -5252.3489                       Pseudo R2         =     0.0514
```

smoker	Coef.	Std. Err.	z	P>|z|	[95% Conf.	Interval]
smkban	-.1517626	.0289268	-5.25	0.000	-.208458	-.0950671
female	-.1106249	.0287785	-3.84	0.000	-.1670298	-.05422
age	-.0042031	.0011748	-3.58	0.000	-.0065057	-.0019006
black	-.07969	.0525369	-1.52	0.129	-.1826604	.0232804
hispanic	-.3327039	.0476677	-6.98	0.000	-.4261308	-.2392769
E1	1.094231	.0714121	15.32	0.000	.9542663	1.234197
E2	.8518588	.0594747	14.32	0.000	.7352906	.9684271
E3	.6492566	.0606989	10.70	0.000	.530289	.7682241
E4	.2224747	.0649939	3.42	0.001	.0950891	.3498603
_cons	-.9842425	.0756055	-13.02	0.000	-1.132427	-.8360584

(a) **(5 marks)** It is argued that using the probit regression model is better than using the linear probability model when explaining the binary variable *smoker*. Discuss the benefits/drawback of using the Probit model when trying to explain a binary variable.

(b) **(5 marks)** Explain briefly how the Probit estimates are obtained and discuss the properties of the parameter estimates.

Hint: You may recall that for the Probit model, we will specify

$$\Pr(\textit{smoker} = 1) = \Phi\left(\beta_0 + \beta_1 \textit{smkban} + \beta_2 \textit{female} + ... + \beta_8 E3 + \beta_9 E4\right),$$

where Φ is the standard normal CDF (cumulative distribution function).

(c) **(5 marks)** Explain how you can estimate the effect of the smoking ban on the probability of smoking for a 50-year old white, college graduated man. You are not expected to use your calculator, clarity of the computations required is enough.

(d) **(5 marks)** Discuss how you could test the joint significance of the worker's characteristics (age, gender, ethnicity, and level of education) using the likelihood ratio test. Clearly indicate the test statistic, its distribution, the rejection rule and the additional information you would need to implement it.

10. This question is based on "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate", by Chan et al. (*The Journal of Finance,* 1992). In this article, the authors are interested in a regression model for the short-term interest rate r_t given by

$$r_t - r_{t-1} = \beta_0 + \beta_1 r_{t-1} + \varepsilon_t. \tag{10.1}$$

Assume that the errors ε_t are not serially correlated.

(a) **(5 marks)** Provide sufficient conditions (on β_1) for r_t to be stationary and weakly dependent. Will the OLS estimator for a linear regression of r_t on r_{t-1} in this case be unbiased and consistent? Explain your answer.

(b) **(5 marks)** A few well-known models in finance hypothesize that $\beta_1 = 0$. Explain how you would test this hypothesis. What name do we give the process r_t when $\beta_1 = 0$.

Suppose you are interested in the following relation between long-term interest rates R_t and short-term interest rates r_t:

$$R_t - r_t = \alpha_0 + \alpha_1 r_t + u_t, \tag{10.2}$$

where u_t is an error term that has zero mean.

(c) **(5 marks)** Let us assume $r_t = \beta_0 + \varepsilon_t$ (that is $\beta_1 = -1$). How would you test whether u_t is serially correlated? Explain your answer.

(d) **(5 marks)** Let us assume $\beta_1 = 0$. Explaining your answers, what can you say about the OLS estimator for α_0 and α_1 applied to equation (10.2)

i. if u_t is stationary and weakly dependent?

i. if u_t is non-stationary and strongly dependent?

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

© University of London 2019

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	5·0	1·6449	3·0	1·8808	2·0	2·0537	1·0	2·3263	0·10	3·0902
45	0·1257	4·8	1·6646	2·9	1·8957	1·9	2·0749	0·9	2·3656	0·09	3·1214
40	0·2533	4·6	1·6849	2·8	1·9110	1·8	2·0969	0·8	2·4089	0·08	3·1559
35	0·3853	4·4	1·7060	2·7	1·9268	1·7	2·1201	0·7	2·4573	0·07	3·1947
30	0·5244	4·2	1·7279	2·6	1·9431	1·6	2·1444	0·6	2·5121	0·06	3·2389
25	0·6745	4·0	1·7507	2·5	1·9600	1·5	2·1701	0·5	2·5758	0·05	3·2905
20	0·8416	3·8	1·7744	2·4	1·9774	1·4	2·1973	0·4	2·6521	0·01	3·7190
15	1·0364	3·6	1·7991	2·3	1·9954	1·3	2·2262	0·3	2·7478	0·005	3·8906
10	1·2816	3·4	1·8250	2·2	2·0141	1·2	2·2571	0·2	2·8782	0·001	4·2649
5	1·6449	3·2	1·8522	2·1	2·0335	1·1	2·2904	0·1	3·0902	0·0005	4·4172

© University of London 2019

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

© University of London 2019

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

© University of London 2019

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

© University of London 2019

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

© University of London 2019

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

© University of London 2019

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

© University of London 2019

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

END OF PAPER

UNIVERSITY OF LONDON

Summer 2020 online assessment guidance

EC2020 Elements of Econometrics

The assessment will be an **open-book take-home online assessment within a 24-hour window**. The requirements for this assessment remain the same as the originally planned closed-book exam, with an expected time/effort of 3 hours and 15 minutes (inclusive of fifteen minutes reading time).

Candidates should answer **EIGHT** of the following **TEN** questions: all **FIVE** questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly. **If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.**

Extracts from statistical tables are given after the final question on this paper.

You should complete your assessment using **pen and paper**. Please use **BLACK ink only**.

Handwritten work then needs to be scanned, converted to PDF and then uploaded to the VLE as **ONE individual file** including the coversheet. Each scanned sheet should have your **candidate number** written clearly at the top. Please **do not write your name anywhere** on any sheet.

The paper will be available at **12.00 pm (BST) on Monday 6 July 2020**.

You have **until 12.00 pm (BST) on Tuesday 7 July 2020** to upload your file into the VLE submission portal. However, you are advised not to leave your submission to the last minute. *We will deduct 5 marks if your submission is up to one hour late, 10 marks if your submission is more than one hour late but less than two hours late (etc.).*

The assessment has been designed with a duration of 24 hours to provide a more flexible window in which to complete the assessment and to appropriately test the course learning outcomes. As an open-book assessment, the expected amount of effort required to complete all questions and upload your answers during this window is no more than 3 hours and 15 minutes. Organise your time well and avoid working all night.

You may use *any* calculator for any appropriate calculations, but you may not use any computer software to obtain solutions. Credit will only be given if all workings are shown.

UL20/0391

You are assured that there will be no benefit in you going beyond the expected 3 hours and 15 minutes of effort. Your assessment has been carefully designed to help you show what you have learned in the hours allocated.

This is an open book assessment and as such you may have access to additional materials including but not limited to subject guides and any recommended reading. But the work you submit is expected to be 100% your own. Therefore, unless instructed otherwise, you must not collaborate or confer with anyone during the assessment. The University of London will carry out checks to ensure the academic integrity of your work. Many students that break the University of London's assessment regulations did not intend to cheat but did not properly understand the University of London's regulations on referencing and plagiarism. The University of London considers all forms of plagiarism, whether deliberate or otherwise, a very serious matter and can apply severe penalties that might impact on your award. The University of London 2019-20 Procedure for the Consideration of Allegations of Assessment offences is available online at:

https://london.ac.uk/sites/default/files/governance/assessment-offence-procedure-year-2019-2020.pdf

The University of London's Rules for Taking Online Timed Assessments have been included in an update to the University of London General Regulations and are available at:

https://london.ac.uk/sites/default/files/regulations/progregs-general-2019-2020.pdf

© University of London 2020

SECTION A

Answer all questions from this section

1. Consider the following regression model

$$Y_i = \beta X_i + u_i, \qquad i = 1, ..., n.$$

The error term has a zero mean, variance equal to σ^2/X_i^2, and $E(u_i u_j) = 0$ for $i \neq j$. You are given a sample of observations $\{(Y_i, X_i)\}_{i=1}^n$. You may treat X_i as being non-stochastic. Clearly annotating your answers:

(a) **(5 marks)** Derive the OLS estimator of β. In the presence of heteroskedasticity, the OLS estimator remains unbiased (you are not asked to show this). Derive the variance of the OLS estimator of β.

(b) **(3 marks)** Discuss how you can obtain the Best Linear Unbiased Estimator (BLUE) of β given the heteroskedasticity.

2. Consider the simple linear regression model

$$Y_i = \alpha + \beta X_i + u_i, \qquad i = 1, ..., n$$

in the presence of correlation between the error and regressor. The regressor exhibits variability in the sample, i.e., $\sum_{i=1}^n (X_i - \bar{X})^2 \neq 0$. Under assumptions of homoskedasticity and the absence of autocorrelation, the IV estimator of β that uses the instrument Z has the following (asymptotic) variance (no need to prove this statement)

$$Var\left(\hat{\beta}_{IV}\right) = \frac{\sigma_u^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \times \frac{1}{r_{XZ}^2},$$

where $r_{XZ} \neq 0$ is the sample correlation between X and Z and σ_u^2 is the variance of the disturbance term u.

(a) **(1 marks)** Give the formula for $\hat{\beta}_{IV}$ (you are not asked to derive it).

(b) **(4 marks)** Provide at least three factors that will help obtain more precise IV parameter estimates for β. In your answer explain why the precision of parameter estimates is important.

(c) **(3 marks)** Discuss the following statement: "If X is not correlated with u, the best choice of instrument is using the regressor itself".

3. Consider a linear regression model

$$y_t = \alpha + \beta x_t + \varepsilon_t, \qquad t = 1, ..., T$$

where the zero mean error ε_t exhibits autocorrelation of an unknown form. We assume our processes are covariance stationary and exhibit weak dependence. The regressor and error may be assumed to be independent.

(a) **(3 marks)** Explain what it means to say that $\{\varepsilon_t\}_{t=1}^T$ is covariance stationary. Provide an intuitive discussion of the requirements and indicate why these requirements are desirable.

(b) **(5 marks)** Recognizing that the errors exhibit autocorrelation, discuss how we can conduct statistical inference on β using the OLS estimator. Specifically, discuss how you can test the hypothesis $H_0 : \beta = 0.7$ against the alternative $H_1 : \beta < 0.7$ using the OLS estimator.

4. Consider the following time series model for $\{y_t\}_{t=1}^T$:

$$\begin{aligned} y_t &= \alpha + \beta t + u_t, \qquad t = 1, ..., T \\ &\quad \text{with } u_t = \rho u_{t-1} + \varepsilon_t \text{ and } |\rho| \leq 1, \end{aligned}$$

where ε_t is an i.i.d. $\left(0, \sigma_\varepsilon^2\right)$ error that is uncorrelated with anything in the past (white noise).

(a) **(3 marks)** Show that y_t is trend stationary when $|\rho| < 1$.

(b) **(3 marks)** Show that y_t is difference stationary when $\rho = 1$.

(c) **(2 marks)** Discuss the importance of distinguishing between trend stationary and difference stationary processes.

5. (a) **(2 marks)** Define the term “multicollinearity” and provide a real life example where this problem is likely to occur.

(b) Examine whether the following statements are true or false. Give an explanation.

i) **(3 marks)** In multiple regression, multicollinearity implies that the least squares estimators of the coefficients are biased and standard errors invalid.

ii) **(3 marks)** If the coefficient estimates in an equation have high standard errors, this is evidence of high multicollinearity.

SECTION B

Answer three questions from this section.

6. We are interested in explaining the willingness of households to buy ecologically produced apples. We use data where each family was presented with a description of ecologically friendly apples, along with prices (in $) of regular apples (*regprc*) and prices of the hypothetical ecolabeled apple (*ecoprc*).

   The variable we want to explain is the dummy variable, *ecobuy* which equals 1 if the household wants to buy ecologically friendly apples and 0 otherwise. Additional household variables we have are, family income in $1000s, *faminc,* household size, *hhsize*, years of schooling, *educ*, and *age.*

   Using a sample of 660 households, the following results were obtained

	OLS A	Probit A	Probit B
constant	0.890 [.068]	1.088 (.206)	−0.244 (.474)
regprc	0.735 [.132]	2.029 (.378)	2.030 (.381)
ecoprc	−0.845 [.106]	−2.344 (.318)	−2.267 (.321)
faminc			0.0014 (.0015)
hhsize			0.069 (.037)
educ			0.071 (.024)
age			−0.001 (.004)
R^2	.086		
$\log L$		−407.60	−399.04

   The heteroskedasticity robust standard errors are reported in square brackets and the (asymptotic) standard errors are reported in parentheses.

   (a) **(5 marks)** Carefully interpret the estimated coefficient of the price of ecologically friendly apples reported in column "OLS A" and discuss whether the effect is statistically significant. In your answer explain why it is important to use robust standard errors.

(b) **(5 marks)** It is argued that using the Probit model is better than using the linear probability model when explaining the binary variable *ecobuy*. Discuss the benefits/drawbacks of using the Probit model when trying to explain a binary variable. In your answer explain what the linear probability model refers to.

(c) **(5 marks)** The Probit model B postulates that

$$\Pr(\textit{ecobuy} = 1|x) =$$
$$\Phi\left(\beta_0 + \beta_1 \textit{regprc} + \beta_2 \textit{ecoprc} + \beta_3 \textit{faminc} + \beta_4 \textit{hhsize} + \beta_5 \textit{educ} + \beta_6 \textit{age}\right),$$

where $\Phi(z)$ is the standard normal cumulative distribution function.

Use the likelihood ratio test, to test the joint significance of the <u>nonprice</u> variables. Clearly indicate the null and alternative hypothesis, the test statistic and the rejection rule.

(d) In this question we are interested in the marginal effect of the price of ecologically friendly apples using Probit model A holding constant the price of regular apples.

i. **(2 marks)** Indicate how you can obtain the marginal effect of *ecoprc* using the probit model.

ii. **(3 marks)** Unlike in the LPM this marginal effect will not be constant. Discuss what computations you would carry out to obtain the marginal effect of a 0.10$ reduction in *ecoprc* when evaluated at the mean of our explanatory variables (*regprc* mean equals 0.884$ and *ecoprc* mean equals 1.082$). You are not expected to use your calculator. Clarity of the computations required is enough.

7. Let us consider the following ADL(1,1) model

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + u_t, \quad |\beta_1| < 1$$

where X_t and Y_t are $I(0)$ variables and u_t is an i.i.d. $\left(0, \sigma_u^2\right)$ error term that is uncorrelated with $Y_{t-1}, Y_{t-2}, Y_{t-3}$..... You may assume that u_t is independent of $X_{t'}$ for all t, t'.

(a) **(3 marks)** Provide the short-run and long-run effect of X on Y. Explain the difference between these effects.

(b) **(7 marks)** Discuss what properties your OLS estimators for the ADL(1,1) parameters will have in the presence of the lagged dependent variable. In particular, (i) are the estimators unbiased and consistent, and (ii) should we use robust standard errors? Provide supportive arguments for your answers.

(c) **(5 marks)** Show that when you omit the relevant variable Y_{t-1} in the above model, you will get evidence of autocorrelation. Explain the result.
Hint: You are expected to reformulate your model as

$$Y_t = \beta_0 + \beta_2 X_t + \beta_3 X_{t-1} + v_t,$$

where $v_t = \beta_1 Y_{t-1} + u_t$, and show that, e.g., $Cov(v_{t+1}, v_t) \neq 0$.

(d) **(5 marks)** Discuss how you would proceed to test for the presence of autocorrelation in the model in (b) using the Breusch-Godfrey test. You may assume that under the alternative, u_t is (or, can be suitably well approximated by) a stationary AR(1).

8. Let us consider the following ADL(1,1) model

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + u_t$$

where X_t and Y_t are $I(1)$ variables and $|\beta_1| < 1$. Let us assume that the error term has zero mean and that the error is uncorrelated with Y_{t-1}, X_t and X_{t-1}.

(a) **(5 marks)** It is important to distinguish whether the above model is spurious or cointegrating. Explain these concepts clearly and highlight their differences.

(b) **(5 marks)** Discuss how you can test whether the above specification is spurious (null hypothesis) or cointegrating (alternative) using a standard Augmented Dickey Fuller test. Clearly indicate the test equation, test statistic and the rejection rule. The critical value provided by Dickey and Fuller for this test at the 5% level of significance is equal to -2.86.

(c) **(5 marks)** You are told that there exists an Error Correction Model (ECM) that describes the short-run and long-run dynamics between Y_t and X_t. What does this tell you as it relates to your answers in (a) and (b)? Provide the Error Correction Model and interpret the various components of the ECM.

(d) **(5 marks)** Discuss how the ECM can be estimated using Ordinary Least Squares.

9. It is postulated that a reasonable demand-supply model for the wine industry in Australia, under market clearing assumption, would be given by

$$Q_t = \alpha_0 + \alpha_1 P_t^w + \alpha_2 P_t^b + \alpha_3 Y_t + \alpha_4 A_t + u_t \quad \text{demand}$$
$$Q_t = \beta_0 + \beta_1 P_t^w + \beta_2 S_t + v_t \quad \text{supply}$$

where $Q_t =$ real per capita consumption of wine, $P_t^w =$ price of wine relative to CPI, $P_t^b =$ price of beer relative to CPI, $Y_t =$ real per capita disposable income, $A_t =$ real per capital advertising expenditure, and $S_t =$ storage cost. CPI is the Consumer Price Index.

The endogenous variables in this model are Q and P^w, and the exogenous variables are P^b, Y, A and S.

The variance of u_t and v_t are, respectively σ_u^2, and σ_v^2, and $Cov(u_t, v_t) = \sigma_{uv} \neq 0$. The errors do not exhibit any correlation over time.

(a) **(5 marks)** Provide the reduced form for P_t^w.

(b) **(5 marks)** The OLS estimation of the demand function, based on annual data from 1955-1975 ($T = 20$), gave the following results (all variables are in logs and figures in parentheses are t-ratios).

$$\hat{Q}_t = \underset{(-6.04)}{-23.651} + \underset{(4.0)}{1.158} P_t^w \underset{(-0.45)}{-0.275} P_t^b \underset{(4.5)}{+3.212} Y_t \underset{(-1.3)}{-0.603} A_t$$

All the coefficients except that of Y have the wrong signs. The coefficient of P^w (price elasticity of demand, α_1) not only has the wrong sign but also appears significant.

Explain why the OLS parameter estimator may give rise to these counter-intuitive results. You are expected to use your results in answer (a) to support your answer.

(c) **(5 marks)** The supply equation is overidentified. Clearly explain this terminology. What distinguishes overidentification from exact identification and underidentification? Provide one set of assumptions that would render the supply equation exactly identified.

(d) **(5 marks)** Discuss how you should estimate the supply equation in light of the overidentification.

10. To investigate the relationship between the price of wine and consumption of wine, an economist estimates the following regression using a sample of 32 individuals for one week in 2013:

$$\widehat{\log(\textit{wine})} = \underset{(0.8911)}{4.2514} - \underset{(0.0031)}{0.8328} \log(\textit{price}),$$
$$n = 32,\ R^2 = 0.89.$$

wine denotes the amount of wine consumed per week in milliliters (a medium glass contains 175ml), and *price* denotes the average price of a medium glass of wine of a selection of wines during the week in GBP (£). The numbers in parentheses are the standard errors.

(a) **(3 marks)** Discuss what would happen to the parameter estimate of the slope coefficient if we had measured the amount of wine consumed per week in number of medium glasses instead of milliliters. Explain your answer.

(b) **(5 marks)** You are asked to test the hypothesis that the demand for wine has an elasticity equal to -1 against a two-sided alternative, using a 5% level of significance. Clearly stating any assumptions you may need, carry out this test.

(c) **(5 marks)** Construct a 95% confidence interval for the price elasticity of demand and discuss how this interval can be used to carry out the test in (b).

(d) A famous TV chef suggests in a talk show that the demand for wine is less elastic (i.e., less negative) for people who have eaten at a restaurant during the week, arguing that eating in a restaurant encourages people to drink wine regardless of the price. To test this theory, the economist defines a dummy variable D_i that takes the value 1 if individual i ate at a restaurant during the week, and 0 otherwise. She obtains the following regression result:

$$\widehat{\log(\textit{wine})} = \underset{(0.8911)}{4.2133} - \underset{(0.0028)}{0.8218} \log(\textit{price}) + \underset{(0.0011)}{0.0889} D \times \log(\textit{price}),$$
$$n = 32,\ R^2 = 0.92.$$

i. **(3 marks)** How does this regression help in assessing the TV chef's claim?

ii. **(4 marks)** Conduct a test that may offer support for the TV chef's claim. Clearly specify the null and the alternative hypothesis. Explain clearly why your test is effective in answering the question of interest.

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

When $x > 3·3$ the formula $1-\Phi(x) \doteqdot \frac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\frac{1}{x^2}+\frac{3}{x^4}-\frac{15}{x^6}+\frac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	5·0	1·6449	3·0	1·8808	2·0	2·0537	1·0	2·3263	0·10	3·0902
45	0·1257	4·8	1·6646	2·9	1·8957	1·9	2·0749	0·9	2·3656	0·09	3·1214
40	0·2533	4·6	1·6849	2·8	1·9110	1·8	2·0969	0·8	2·4089	0·08	3·1559
35	0·3853	4·4	1·7060	2·7	1·9268	1·7	2·1201	0·7	2·4573	0·07	3·1947
30	0·5244	4·2	1·7279	2·6	1·9431	1·6	2·1444	0·6	2·5121	0·06	3·2389
25	0·6745	4·0	1·7507	2·5	1·9600	1·5	2·1701	0·5	2·5758	0·05	3·2905
20	0·8416	3·8	1·7744	2·4	1·9774	1·4	2·1973	0·4	2·6521	0·01	3·7190
15	1·0364	3·6	1·7991	2·3	1·9954	1·3	2·2262	0·3	2·7478	0·005	3·8906
10	1·2816	3·4	1·8250	2·2	2·0141	1·2	2·2571	0·2	2·8782	0·001	4·2649
5	1·6449	3·2	1·8522	2·1	2·0335	1·1	2·2904	0·1	3·0902	0·0005	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10$, 5, 2·5, 1, 0·5 and 0·1. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

Examiners' commentaries 2020

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2019–20. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

General remarks

Learning outcomes

At the end of the course, and having completed the Essential reading and activities, you should be able to:

- describe and apply the classical regression model and its application to cross-section data
- describe and apply the:
  - Gauss–Markov conditions and other assumptions required in the application of the classical regression model
  - reasons for expecting violations of these assumptions in certain circumstances
  - tests for violations
  - potential remedial measures, including, where appropriate, the use of instrumental variables
- recognise and apply the advantages of logit, probit and similar models over regression analysis when fitting binary choice models
- competently use regression, logit and probit analysis to quantify economic relationships using standard regression programmes (Stata and EViews) in simple applications
- describe and explain the principles underlying the use of maximum likelihood estimation
- apply regression analysis to time-series models using stationary time series, with awareness of some of the econometric problems specific to time series applications (for example, autocorrelation) and remedial measures
- recognise the difficulties that arise in the application of regression analysis to nonstationary time series, know how to test for unit roots, and know what is meant by cointegration.

Common mistakes committed by candidates

Due to the impact of the Covid-19 pandemic, there was a shift from in-person examinations to open-book examinations completed within a 24-hour window. Despite the challenges, many candidates performed extremely well. We discuss below common mistakes made by candidates, some of them are repeated year after year.

Candidates should realise that $\frac{1}{n}\sum_{i=1}^{n} u_i$ is not the same as $\mathrm{E}(u_i)$. So, while we typically assume $\mathrm{E}(u_i) = 0$, this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} u_i = 0$. Also, while we may be happy to assume $\mathrm{E}(x_i u_i) = 0$ (uncorrelatedness between the errors and regressors), this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} x_i u_i = 0$. Note that:

- both $\frac{1}{n}\sum_{i=1}^{n} u_i$ and $\frac{1}{n}\sum_{i=1}^{n} x_i u_i$ are random variables, which take the value 0 with probability 0 (continuous random variables)!
- $\mathrm{E}(u_i) = 0$ and $\mathrm{E}(x_i u_i) = 0$ are fixed, not stochastic!

The differences between sample and population moments need to come across clearly when looking at unbiasedness and making consistency arguments. In both cases, we first simplify our estimator (plug in the true model) to obtain:

$$\widehat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \beta + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \quad \text{with } x_i = X_i - \bar{X}.$$

- For *unbiasedness*, clearly indicate that you want to show that $\mathrm{E}(\widehat{\beta}) = \beta$. Unbiasedness does not follow from $\sum_{i=1}^{n} x_i u_i = 0$, instead it follows from $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) = 0$.

  If we treat x_i as fixed:

  $$\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) \equiv \mathrm{E}\left(\sum_{i=1}^{n} d_i u_i\right) = \sum_{i=1}^{n} d_i \mathrm{E}(u_i)$$

  and then unbiasedness follows as $\mathrm{E}(u_i) = 0$.

- For *consistency*, clearly indicate that you want to show that $\mathrm{plim}(\widehat{\beta}) = \beta$. Using the plim properties, we show:

  $$\begin{aligned}\mathrm{plim}\ \widehat{\beta} = \beta + \mathrm{plim}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) &= \beta + \frac{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i u_i\right)}{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)} \\ &\equiv \beta + \frac{\mathrm{plim}(\text{Sample Cov}(x,u))}{\mathrm{plim}(\text{Sample Var}(x))} \\ &= \beta + \frac{\mathrm{Cov}(x,u)}{\mathrm{Var}(x)} \quad \text{using the law of large numbers}\end{aligned}$$

  where $\mathrm{Cov}(x,u) = 0$ and $\mathrm{Var}(x) > 0$, ensuring we get consistency.

- Remember, the law of large numbers ensures that sample averages converge to their population analogues.

A large number of candidates are not able to clearly distinguish between sample variance and covariance, and population variance and covariance.

The use of $\text{Cov}(X, Y)$ and $\text{Var}(X)$ should be restricted to describing the population covariance and variances, respectively, with definitions:

$$\text{Cov}(X, Y) = \text{E}((X - \text{E}(X))(Y - \text{E}(Y))) = \text{E}(XY) - \text{E}(X)\,\text{E}(Y)$$

and:

$$\text{Var}(X) = \text{E}((X - \text{E}(X))^2) = \text{E}(X^2) - (\text{E}(X))^2$$

(you also may denote $\text{Cov}(X, Y) = \sigma_{XY}$ and $\text{Var}(X) = \sigma_X^2$). They are typically unknown, but fixed, quantities.

The sample covariance and variance are estimators of the population covariance and variance, respectively. They are defined as:

$$\text{Sample Cov}(X, Y) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

and:

$$\text{Sample Var}(X) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

(you also may use $\widehat{\sigma}_{XY}$ and $\widehat{\sigma}_X^2$). You can compute them given the data.

With a slight abuse of notation, we often divide by n instead, which is irrelevant if we let n be large. The division by $n - 1$ is a finite sample issue only (unbiasedness).

The sample covariance and variance show up in our definition of the OLS estimator of the slope in the simple linear regression model, not the population covariance and variance, as:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{\text{Sample Cov}(X, Y)}{\text{Sample Var}(X)} \neq \frac{\text{Cov}(X, Y)}{\text{Var}(X)}.$$

Treating them as being the same results in incorrect analyses and candidates lose significant marks. Similarly, a handful of candidates lost marks in Question 1 (b) for using the population variances instead of sample variances in the IV estimate formula:

$$\widehat{\beta}_{IV} = \frac{\sum_{i=1}^{n}(Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(Z_i - \bar{Z})(X_i - \bar{X})} = \frac{\text{Sample Cov}(Z_i, Y_i)}{\text{Sample Cov}(Z_i, X_i)} \neq \frac{\text{Cov}(Z_i, Y_i)}{\text{Cov}(Z_i, X_i)}.$$

Candidates often made mistakes in the last steps to correctly derive the variance of the estimator β in Question 1 (a). In particular, candidates often mistook $\sum_{i=1}^{n}\sigma^2$ as simply σ^2; whereas we must have:

$$\text{Var}(\widehat{\beta}) = \frac{\sum_{i=1}^{n}\sigma^2}{\left(\sum_{i=1}^{n}X_i^2\right)^2} = \frac{\sigma^2 + \sigma^2 + \cdots + \sigma^2}{\left(\sum_{i=1}^{n}X_i^2\right)^2} = \frac{n\sigma^2}{\left(\sum_{i=1}^{n}X_i^2\right)^2}.$$

A number of candidates seemed to be confused between the conditions for a valid instrument (exclusion restriction and relevance) and factors related to obtaining a more precise IV estimate (a higher sample size, a higher sample variability of X_i, a lower variance of the error term, and a

stronger correlation between X_i and Z_i). Candidates describing properties required of instrumental variables, only received 1 mark for discussing the stronger correlation point (analogous to the relevance of an instrument).

When discussing the importance of distinguishing trend stationary and difference stationary processes, two points are expected: discussing the implication for the long-run behaviour of the processes, and discussing statistical implications (the consistency of the estimators). Most candidates discussed the former very well, only a few touched on the latter.

Many candidates simply described multicollinearity as independent variables being correlated. Although we accept answers discussing high correlation among the independent variables as an alternative to a close linear relation among the regressors, simply taking correlation among regressors as the definition of multicollinearity is incorrect. In attempting True/False questions with a specific instruction to explain your answers, several candidates failed to demonstrate adequate understanding of the course material and simply reproduced the questions at hand.

Candidates struggled to give competent answers to the interpretation of empirical results. When interpreting an empirical result you should discuss the significance of the coefficients, magnitude and sign of the coefficients. Relatedly, *ceteris paribus* (or holding other factors constant) is needed to ensure that the estimate can be interpreted as the marginal change in the dependent variable given a change in the independent variable of interest. Many candidates lost marks from these mistakes.

Candidates particularly struggled with the difference between percentage points and % changes when it comes to the interpretation of coefficients in binary choice models. A coefficient in such a model gives the percentage point difference in the outcome variable when the regressor changes by 1 unit. The % change is a *relative* change that depends on the baseline value of the outcome variable (for example, the realisation when a regressor equals zero in the case of a dummy variable regressor). For example, if we have initially 50% of a sample buying ecological apples, a 10% increase will mean now $(1 + 0.1) \times 50 = 55\%$ of the sample buys ecological apples. A 10 percentage point increase instead will result in $50 + 10 = 60\%$ of the sample choosing to buy ecological apples.

It was very common to misunderstand the Central Limit Theorem. While it provides the distribution of the *estimator* $\widehat{\beta}$, no matter what the true distribution of the errors is, it does *not* say that the distribution of the error term is approximately normal in large samples.

When discussing the marginal effect after a Probit estimation, candidates often did not specify the term z in the pdf $\phi(z)$.

A number of candidates described the Dickey–Fuller test when an Augmented Dickey–Fuller (ADF) test was asked. The difference between the two tests is in the test equation. In the augmented specification, we add sufficient lagged differences of the residuals to ensure that there is no autocorrelation left in the test equation regression.

Exogeneity and endogeneity are often very simply explained as variables 'outside of the model' or 'inside the model', respectively. While this approach is not wrong; it is not a particularly clear or helpful definition in general contexts that are not related to simultaneous equations models. Defining these concepts in terms of the relationship between regressor and error term is much clearer and naturally leads to a way to show exogeneity or endogeneity by considering the covariance between regressors and error terms.

Candidates often missed that all exogenous regressors and instruments have to be included in the first stage when performing a TSLS estimation. Some candidates used the fitted values of the endogenous variable obtained from the first stage as the dependent variable in the second stage. This mistake is a serious misunderstanding. Other candidates regressed the reduced forms of the demand and supply equations in each stage as a TSLS estimation. This method is not helpful as we are interested in the relationship between the quantity and the price (between the two endogenous variables), not between each endogenous variable and the exogenous variables.

When asked about identification of structural form equations, candidates often compared the number of excluded regressors to the number of endogenous variables in general, without pointing

out which variables are the excluded regressors and could therefore be used as an instrument. Candidates were also often unclear of the definitions of over-identification, under-identification and exact-identification. For example, some candidates discussed the concepts by comparing the number of variables and the number of numerical values of the parameters that can be obtained from the structural equations. This point is technically correct in the sense that we do not have enough data on the instruments to identify the parameters in the under-identification case. Candidates, however, often did not make that discussion clear.

When deriving the reduced forms, candidates should check conditions for their existence, in particular, whether the denominator of the expression is different from zero. Only a few candidates made this check.

Gauss–Markov conditions were used in many places in the examination, for example, when discussing the BLUE property of an estimator or when conducting hypothesis tests. The Gauss–Markov conditions have to be explicitly specified. Only writing that the Gauss–Markov conditions hold is not sufficient. A good practice that many candidates have adopted is: begin the script by explicitly providing the Gauss–Markov conditions and refer back to them thereafter. Many candidates did not mention either the Gauss–Markov conditions or the normality assumption when conducting t tests.

When conducting hypothesis testing, ensure that you clearly indicate the null and alternative hypotheses (in terms of the true parameters, say β_1), the test statistic (in terms of the parameter estimates, here $\widehat{\beta}_1$), its distribution (with degrees of freedom), the rejection rule (one-sided or two-sided) for a given significance level (typically 5%) with suitable critical values, and provide an interpretation of your result. Candidates often lost marks for missing parts of these steps. Choosing the critical value appropriate to the sample size is also important. Several candidates used 1.96 (a critical value for a very large number of degrees of freedom) in a small sample (30) instead of the appropriate one (2.042). This mistake also cost marks.

When using confidence intervals in a hypothesis test, candidates should decide if the hypothesised value for the population parameter of interest β lies within the confidence interval. Many candidates used the default value of zero when discussing their confidence interval regardless of the population parameter in question.

Another common mistake was to conduct a two-sided test when the statement was worded with a clear direction of the relationship. In particular, in Question 10 (d), the chef's claim had such a hint of the direction ('less elastic') so that a one-sided test was appropriate. Many candidates conducted a two-sided test and lost marks.

Just as last year, many candidates did not answer all parts of a question. Make sure you read the questions properly and provide all details that are requested. Not answering a question will automatically earn you zero marks for that question.

Key steps to improvement

Essential reading for this course includes the subject guide and the following:

- Dougherty, C. *Introduction to econometrics.* (Oxford: Oxford University Press, 2016) 5th edition [ISBN 9780199676828]; http://oxfordtextbooks.co.uk/orc/dougherty5e/

Apart from the Essential readings you should do some supplementary reading. One very good book at the same level is:

- Gujarati, D.N. and D.C. Porter *Basic econometrics.* (McGraw–Hill, 2009, International edition) 5th edition [ISBN 9780071276252].

To understand the subject clearly it is important to supplement Dougherty's *Introduction to econometrics* (fifth edition) with the subject guide **EC2020 Elements of econometrics** (2016), especially Chapter 10 which covers maximum likelihood estimation. It is very important to carefully

go through the subject guide. The subject guide contains solutions to the questions given in the main textbook and also some additional questions and solutions. Working through these will improve your understanding of the subject.

The chapter in the subject guide on maximum likelihood (Chapter 10) includes some additional theory which has not been covered in the main textbook. It is important to read the additional theory given in the subject guide to have a better understanding of the principles of maximum likelihood and tests based on the likelihood function.

Please check the VLE course page for resources for this subject such as a downloadable copy of the subject guide **EC2020 Elements of econometrics** (2016), PowerPoint slideshows that provide a graphical treatment of the topics covered in the textbook, datasets and statistical tables. Candidates should utilise datasets using standard regression programmes (STATA or EViews). This will help in the understanding of the subject.

Examination revision strategy

Many candidates are disappointed to find that their examination performance is poorer than they expected. This may be due to a number of reasons, but one particular failing is '**question spotting**', that is, confining your examination preparation to a few questions and/or topics which have come up in past papers for the course. This can have serious consequences.

We recognise that candidates might not cover all topics in the syllabus in the same depth, but you need to be aware that examiners are free to set questions on **any aspect** of the syllabus. This means that you need to study enough of the syllabus to enable you to answer the required number of examination questions.

The syllabus can be found in the Course information sheet available on the VLE. You should read the syllabus carefully and ensure that you cover sufficient material in preparation for the examination. Examiners will vary the topics and questions from year to year and may well set questions that have not appeared in past papers. Examination papers may legitimately include questions on any topic in the syllabus. So, although past papers can be helpful during your revision, you cannot assume that topics or specific questions that have come up in past examinations will occur again.

If you rely on a question-spotting strategy, it is likely you will find yourself in difficulties when you sit the examination. We strongly advise you not to adopt this strategy.

Examiners' commentaries 2020

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2019–20. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions

Candidates should answer **EIGHT** of the following **TEN** questions: all **FIVE** questions in Section A (8 marks each) and **THREE** questions from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly. **If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.**

Section A

Answer all questions from this section.

Question 1

Consider the following regression model:

$$Y_i = \beta X_i + u_i, \quad i = 1, \ldots, n.$$

The error term has a zero mean, variance equal to σ^2/X_i^2, and $\mathrm{E}(u_i u_j) = 0$ for $i \neq j$. You are given a sample of observations $\{(Y_i, X_i)\}_{i=1}^n$. You may treat X_i as being non-stochastic. Clearly annotating your answers:

(a) **The OLS estimator of β can be shown to be unbiased in the presence of heteroskedasticity (you are not asked to show this). Derive the variance of the OLS estimator of β. You are expected to clearly define the OLS estimator of β.**

(5 marks)

(b) **Discuss how you can obtain the Best Linear Unbiased Estimator (BLUE) of β given the heteroskedasticity.**

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 2.3, 2.4 and 7.3.

Dougherty, C. Subject guide (2016): Chapters 2 and 7, and Section 15.15.

Approaching the question

(a) The OLS estimator of β is:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} X_i^2} = \beta + \sum_{i=1}^{n} d_i u_i$$

where $d_i = X_i / \sum_{i=1}^{n} X_i^2$ are constants for $i = 1, \ldots, n$. We have:

$$\begin{aligned}
\mathrm{Var}(\widehat{\beta}) &= \mathrm{Var}\left(\beta + \sum_{i=1}^{n} d_i u_i\right) \\
&= \mathrm{Var}\left(\sum_{i=1}^{n} d_i u_i\right) && \text{(additive constant does not carry any variability)} \\
&= \sum_{i=1}^{n} d_i^2 \mathrm{Var}(u_i) + \sum_{i \neq j}^{n} d_i d_j \mathrm{Cov}(u_i, u_j) && \text{(because } d_i \text{ is non-stochastic)} \\
&= \sum_{i=1}^{n} d_i^2 \left(\frac{\sigma^2}{X_i^2}\right) + 0 && \text{(uses properties of } u_i \text{ (heteroskedastic, no autocorrelation))} \\
&= \sum_{i=1}^{n} \left(\frac{X_i}{\sum_{i=1}^{n} X_i^2}\right)^2 \left(\frac{\sigma^2}{X_i^2}\right) \\
&= \frac{\sum_{i=1}^{n} \sigma^2}{\left(\sum_{i=1}^{n} X_i^2\right)^2} \\
&= \frac{n\sigma^2}{\left(\sum_{i=1}^{n} X_i^2\right)^2}.
\end{aligned}$$

(b) Because the variances depend on X_i, we have heteroskedasticity in the regression model. An OLS on the above model will not be the BLUE of β. To get the BLUE estimator of β, we need to transform the model so that it satisfies the Gauss–Markov conditions (which have to be specified somewhere in the script). Specifically, we should consider:

$$Y_i X_i = \beta X_i^2 + u_i X_i \quad \text{for } i = 1, \ldots, n.$$

Because $\mathrm{E}(u_i X_i) = 0$ and $\mathrm{Var}(u_i X_i) = \sigma^2$, there is no autocorrelation in this transformed model. An OLS on this model now gives us the BLUE. The estimator of β is:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} X_i^3 Y_i}{\sum_{i=1}^{n} X_i^4}.$$

Comments:

- In part (a), several candidates suggested the use of GLS, which was wrong and given zero marks. Many candidates failed to correctly derive the final answer.
- In part (b), many candidates correctly identified the need to use a weighted least squares estimator (WLS) and were well-rewarded. Some suggested the use of GLS to correct for heteroskedasticity and obtain GLS. Both estimators are equivalent in the presence of heteroskedasticity.

Question 2

Consider the simple linear regression model:

$$Y_i = \alpha + \beta X_i + u_i, \quad i = 1, \ldots, n$$

in the presence of correlation between the errors and regressors. The regressors exhibit variability in the sample, i.e. $\sum_{i=1}^{n}(X_i - \bar{X})^2 \neq 0$. Under assumptions of homoskedasticity and the absence of autocorrelation, the IV estimator for β that uses the instrument Z has the following (asymptotic) variance (no need to prove this statement):

$$\text{Var}(\widehat{\beta}_{IV}) = \frac{\sigma_u^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \times \frac{1}{r_{XZ}^2}$$

where $r_{XZ} \neq 0$ is the sample correlation between X and Z and σ_u^2 is the variance of the disturbance term.

(a) Give the formula for $\widehat{\beta}_{IV}$ (you are not asked to derive it).

(1 mark)

(b) Provide at least three factors that will help obtain more <u>precise</u> IV parameter estimates for β. In your answer explain why the precision of parameter estimates is important.

(4 marks)

(c) Discuss the following statement: 'If X is not correlated with u, the best choice of instrument is using the regressor itself.'

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 2.5, 8.2, and 8.5.

Dougherty, C. Subject guide (2016): Chapter 8 and Section 15.16.

Approaching the question

(a) There are two ways to write the IV estimate formula:

$$\widehat{\beta}_{IV} = \frac{\sum_{i=1}^{n}(Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(Z_i - \bar{Z})(X_i - \bar{X})} = \frac{\text{Sample Cov}(Z_i, Y_i)}{\text{Sample Cov}(Z_i, X_i)}.$$

Note that sample variances have to be used instead of their population counterparts.

(b) We can rewrite the definition of the variance as follows:

$$\text{Var}(\widehat{\beta}_{IV}) = \frac{\sigma^2}{n \times \text{Sample Var}(X_i)} \times \frac{1}{r_{XZ}^2}$$

which shows that we can get more precise estimates by:

- increasing the sample size
- increasing the sample variability of X_i
- reducing the variance of the error term σ^2
- increasing the correlation between X_i and Z_i.

The importance of the precision is that it is useful when conducting hypothesis testing. It makes it easier to detect when the null hypothesis is false. Precision is good when the dispersion of our parameter estimates over different samples is quite small, giving more credence to a particular point estimate.

(c) If X_i is not correlated with u_i, the instrument that has the highest correlation with X_i is X_i itself (perfect correlation) – hence the highest precision as $r_{XZ}^2 = 1$ (and r_{XZ}^2 cannot be larger than 1). If we use X as an instrument for X, we get the OLS estimator of β. OLS is the BLUE estimator in this setting.

Comments:

- In part (b), candidates who discussed the conditions for a valid instrument often received one mark for highlighting the need for a strong correlation between X_i and Z_i. Many candidates did not discuss the importance of the IV estimate precision asked by the question. Some used the IV formula to discuss the conditions and were well-rewarded.
- In part (c), some candidates correctly realised that since there is no correlation between the error term and the regressor, there is no endogeneity present in the model and we could use the OLS as the BLUE. To get the full three marks using this approach, candidates should at least discuss the Gauss–Markov conditions or relate to how the OLS estimate is more efficient in this case.

Question 3

Consider a linear regression model:

$$y_t = \alpha + \beta x_t + \varepsilon_t, \quad t = 1, \ldots, T$$

where the zero mean error ε_t exhibits autocorrelation of an unknown form. We assume our processes are covariance stationary and exhibit weak dependence. The regressor and error may be assumed to be independent.

(a) Explain what it means to say that $\{\varepsilon_t\}_{t=1}^T$ is covariance stationary. Provide an intuitive discussion of the requirements and indicate why these requirements are desirable.

(3 marks)

(b) Recognising that the errors exhibit autocorrelation, discuss how we can conduct statistical inference on β using the OLS estimator. Specifically, discuss how you can test the hypothesis $\mathrm{H}_0 : \beta = 0.7$ against the alternative $\mathrm{H}_1 : \beta < 0.7$ using the OLS estimator.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 7.1, 7.2, 12.1 and 12.3.

Dougherty, C. Subject guide (2016): Chapters 7 and 12.

Approaching the question

(a) Candidates were expected to discuss the three following conditions, without a need to specifically mention the finite requirement.

- The expected value of ε_t does not change over time (and is finite), i.e. $\mathrm{E}(\varepsilon_t) = \mu$ for all t.
- The variance of ε_t exists and does not change over time (and is finite), i.e. $\mathrm{Var}(\varepsilon_t) = \sigma^2$ for all t.
- The covariance of ε_t and ε_{t+h} only depends on the distance of time h, i.e. $\mathrm{Cov}(\varepsilon_t, \varepsilon_{t+h}) = \mathrm{Cov}(\varepsilon_s, \varepsilon_{s+h})$ for all s and t.

For the intuition, candidates were expected to explain that non-stationarity is a problem that may need to be dealt with (due to consistency/spurious regression). One of the following discussions works well.

- These assumptions replace the assumption typically made in a cross-sectional setting in which we assume the errors to be i.i.d. (independent and identically distributed). These assumptions are more reasonable in time series, where generally there is dependence.
- The fact that the mean and the variance do not change over time ensures that the process (or relationship) is stable over time. The second assumption provides a homoskedastic setting, which ensures the standard errors are valid for testing.

(b) Candidates are expected to discuss the use of Newey–West standard errors (or heteroskedasticity and autocorrelation consistent (HAC) standard errors) to obtain the valid standard error for statistical tests. A complete testing procedure with usual components is also expected. One way to approach the question is the following.

In the presence of autocorrelation, we cannot rely on the usual standard errors of our OLS estimators. Instead, we need to use the Newey–West (HAC) robust standard errors for hypothesis testing. Using the robust standard errors, we obtain our t test statistic:

$$\frac{\widehat{\beta} - 0.7}{\text{Robust SE}(\widehat{\beta})}.$$

Using a 5% significance level we will reject H_0 if:

$$\frac{\widehat{\beta} - 0.7}{\text{Robust SE}(\widehat{\beta})} < -1.645.$$

Candidates should also note that the test is an asymptotic t test (or z test) with a test distribution of $N(0,1)$ (a t_{n-2} distribution is also accepted).

Comments:

- In part (a), candidates often missed one of the requirements (particularly the one regarding the covariance).
- In part (b), many candidates lost marks for an incorrect test statistic or not providing the test statistic distribution. Several suggested to use GLS instead, which was not correct as the question specifically asked for a hypothesis test using the OLS estimator.

Question 4

Consider the following time series model for $\{y_t\}_{t=1}^{T}$:

$$y_t = \alpha + \beta t + u_t, \quad t = 1, \ldots, T$$

$$\text{with} \quad u_t = \rho u_{t-1} + \varepsilon_t \quad \text{and} \quad |\rho| \leq 1$$

where ε_t is an i.i.d. $(0, \sigma_\varepsilon^2)$ error that is uncorrelated with anything in the past (white noise).

(a) Show that y_t is trend stationary when $|\rho| < 1$.

(3 marks)

(b) Show that y_t is difference stationary when $\rho = 1$.

(3 marks)

(c) Discuss the importance of distinguishing between trend stationary and difference stationary processes.

(2 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 13.1 and 13.2.

Dougherty, C. Subject guide (2016): Chapter 13.

Approaching the question

(a) Candidates were expected to discuss the following points.

y_t is trend stationary if by detrending we can make y_t stationary. By detrending y_t, we have:

$$y_t - \beta t = \alpha + u_t.$$

This would be the case provided u_t is stationary. If $|\rho| < 1$, we know that AR(1) is a stationary process, hence y_t is trend stationary under this condition.

(b) Candidates were expected to discuss the following points.

y_t is difference stationary if by differencing we can make y_t stationary. Taking the difference of $y_t - y_{t-1}$, we have:

$$\begin{array}{ll} y_t & = \alpha + \beta t + u_t \\ y_{t-1} & = \alpha + \beta(t-1) + u_{t-1} \\ \hline y_t - y_{t-1} & = \beta + u_t - u_{t-1} \\ \Delta y_t & = \beta + \varepsilon_t \end{array}$$

where the latter line recognises that if $\rho = 1$ then $u_t - u_{t-1} = \varepsilon_t$. Since ε_t is i.i.d., we know ε_t is stationary and so $\beta + \varepsilon_t$ also is, hence y_t is difference stationary.

(c) Candidates are expected to discuss the two following points:

- Implication for the long-run behaviour of a process: Time series that are trend stationary always revert to the trend in the long run (the effects of shocks are eventually eliminated). That is, they exhibit weak dependence. Time series which are difference stationary never recover from shocks to the system (the effects of shocks are permanent). That is, they exhibit strong dependence or persistence of the effect.
- There are also statistical implications regarding consistency of the estimators if the assumptions are not satisfied.

Comments:

- Candidates should explicitly discuss the definitions of trend stationary and difference stationary somewhere in the answers.
- In part (b), many candidates lost marks for stating explicitly that the reason Δy_t is stationary follows from the fact that ε_t is i.i.d. (independent and identically distributed).
- In part (c), many candidates focused on the long-run behaviour but did not discuss the statistical implications when the requirements were not met.

Question 5

(a) Define the term 'multicollinearity' and provide a real life example where this problem is likely to occur.

(2 marks)

(b) Examine whether the following statements are true or false. Give an explanation.

i. In multiple regression, multicollinearity implies that the least squares estimators of the coefficients are biased and standard errors invalid.

(3 marks)

ii. If the coefficient estimates in an equation have high standard errors, this is evidence of high multicollinearity.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Section 3.4.

Dougherty, C. Subject guide (2016): Chapter 3.

Approaching the question

(a) Perfect multicollinearity means that some of the regressors can be written as a linear combination of other regressors. For example, $X_i = \sum_{j \neq i}^{n} \lambda_j X_j$, where the X_js are regressors.

Multicollinearity (often referred to as near multicollinearity) indicates that there is a close linear relation between the regressors. Answers stating that regressors exhibit high correlations were also accepted.

A real-life example of multicollinearity is in a setting where we try to explain the determinants of educational attainment and we use various measures of cognitive skills, which are highly correlated.

(b) Candidates were expected to give clear explanations to their answers.

i. Presence of near multicollinearity is not a violation of the Gauss–Markov assumptions. In the presence of multicollinearity, the OLS estimators remain unbiased and their standard errors remain valid as long as all the Gauss–Markov assumptions are satisfied. The statement is false.

ii. It is true that the presence of near multicollinearity gives rise to high standard errors, but having high standard errors could also be due to the fact that we have a very small sample or high variance of the error term. The statement is false.

Comments:

- In part (a), candidates often lost marks for describing multicollinearity as the issue of regressors being simply *correlated*. Some provided a real-life example of the dummy variable trap, which is a special case of perfect multicollinearity. Many candidates did not provide specific examples with a clear regression. They simply discussed two highly correlated (potential) regressors without pointing to the regression at hand.
- In part (b), many candidates lost marks for not explaining their answers but rather repeating the statements.

Section B

Answer three questions from this section.

Question 6

We are interested in explaining the willingness of households to buy ecologically produced apples. We use data where each family was presented with a description of ecologically friendly apples, along with prices (in \$) of regular apples (*regprc*) and prices of the hypothetical ecolabeled apple (*ecoprc*).

The variable we want to explain is the dummy variable, *ecobuy* which equals 1 if the household wants to buy ecologically friendly apples and 0 otherwise. Additional household variables we have are family income in \$1000s, *faminc*, household size, *hhsize*, years of schooling, *educ*, and *age*.

Using a sample of 660 households, the following results were obtained:

	OLS A	Probit A	Probit B
constant	0.890	1.088	−0.244
	[.068]	(.206)	(.474)
regprc	0.735	2.029	2.030
	[.132]	(.378)	(.381)
ecoprc	−0.845	−2.344	−2.267
	[.106]	(.318)	(.321)
faminc			0.0014
			(.0015)
hhsize			0.069
			(.037)
educ			0.071
			(.024)
age			−0.001
			(.004)
R^2	.086		
$\log L$		−407.60	−399.04

The heteroskedasticity robust standard errors are reported in squared brackets and the (asymptotic) standard errors are reported in parentheses.

(a) Carefully interpret the estimated coefficient on the price of ecologically friendly apples reported in 'OLS A' and discuss whether the effect is statistically significant. In your answer explain why it is important to use robust standard errors.

(5 marks)

(b) It is argued that using the Probit model is better than using the linear probability model when explaining the binary variable *ecobuy*. Discuss the benefits/drawbacks of using the Probit model when trying to explain a binary variable. In your answer explain what the linear probability model refers to.

(5 marks)

(c) The Probit model B postulates that:

$$\Pr(ecobuy = 1 \mid x)$$
$$= \Phi(\beta_0 + \beta_1 regprc + \beta_2 ecoprc + \beta_3 faminc + \beta_4 hhsize + \beta_5 educ + \beta_6 age)$$

where $\Phi(z)$ is the standard normal cumulative distribution function.

Use the likelihood ratio test, to test the joint significance of the nonprice variables. Clearly indicate the null and alternative hypothesis, the test statistic and the rejection rule.

(5 marks)

(d) In this question we are interested in the marginal effect of the price of ecologically friendly apples using Probit model A holding constant the price of regular apples.

i. Indicate how you can obtain the marginal effect of *ecoprc* using the Probit model.

(2 marks)

ii. Unlike in the LPM this marginal effect will not be constant. Discuss what computations you would carry out to obtain the marginal effect of a \$0.10 reduction in *ecoprc* when evaluated at the mean of our explanatory variables (*regprc* mean equals \$0.884 and *ecoprc* mean equals \$1.082). You are not expected to use your calculator. Clarity of the computations required is enough.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 10.1, 10.3 and 10.6.

Dougherty, C. Subject guide (2016): Chapter 10.

Approaching the question

(a) Candidates are expected to correctly interpret the coefficient of ecologically produced apple price before conducting a two-sided test for its statistical significance. An excellent answer would have the following points.

The coefficient is -0.845, hence when the price of ecologically produced apples increases one dollar, the probability that households buy ecologically produced apples decreases by 84.5% (or a drop in probability of 0.845), *ceteris paribus*.

Conduct a t test with the t statistic of:

$$\frac{\widehat{\beta}_{ecoprc}}{\mathrm{SE}(\widehat{\beta}_{ecoprc})}.$$

We can see that the null hypothesis is rejected, suggesting that the estimate is significant. The null hypothesis, the test statistic, the test statistic distribution, the critical value, and the rejection rule were expected.

The key reason to use robust standard errors is to deal with the heteroskedasticity in the error term inherent in the linear probability model (LPM).

(b) Candidates should explain the difference in terms of linearity between OLS and LPM estimators somewhere in the answer before discussing the advantages and drawbacks of the Probit estimation.

- The LPM is a special case of a binary regression model. When the dependent variable is binary $(0, 1)$, we have $\mathrm{E}(y \mid x) = \Pr(y = 1 \mid x)$. If we assume the probability is linear in the parameters, the model can be estimated using least squares.
- The Probit model has various advantages over the LPM: Predicted probabilities are restricted to lie in $[0, 1]$, maximum likelihood estimation (MLE) is (asymptotically) efficient whereas least squares estimators such as OLS (LPM) will be inefficient. Second, MLE automatically deals with heteroskedasticity inherent in the LPM, the marginal effect obtained from Probit also is allowed to depend on individual characteristics.
- The main drawback(s) of the Probit model relative to the linear probability model is that the coefficients cannot be directly interpreted as the marginal effects of the regressor(s) of interest. Instead, we need to compute predicted probabilities using the Probit specification, and it is also computationally more complicated.

(c) Candidates are expected to only test the joint significance of β_3, β_4, β_5 and β_6. This joint significance test should be clearly notated. An excellent answer should cover the following points:

We are asked to test:

$$\mathrm{H}_0 : \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0 \quad \text{vs.} \quad \mathrm{H}_1 : \exists\, i \in [3,4,5,6], \beta_i \neq 0.$$

We need to use the likelihood ratio (LR) test for hypothesis testing. This requires us to estimate the restricted Probit (where only *regprc* and *ecoprc* are used, i.e. Probit A) and unrestricted Probit (where all variables are included, i.e. Probit B). The test statistic is given by:

$$LR = 2(\log L^U - \log L^R) \overset{a}{\sim} \chi^2_4 \quad \text{under } \mathrm{H}_0.$$

The χ^2 distribution has four degrees of freedom because we are testing four restrictions. We should reject at the 5% significance level if $LR > \chi^2_{0.05,\,4} = 9.488$. We note that our test statistic value is:

$$2 \times (-399.04 - (-407.60)) = 17.12$$

which exceeds the critical value. We reject the joint significance of non-price variables. (The test uses the property that our maximum likelihood estimators are asymptotically normally distributed.)

(d) Candidates are expected to note down the marginal effect formula for *ecoprc* and also explain the formula somewhere in the answer.

i. To get the marginal effect of *ecoprc* we need:

$$\phi(z)\widehat{\beta}_{ecoprc} \quad \text{with } z = \widehat{\beta}_0 + \widehat{\beta}_1 regprc + \widehat{\beta}_2 ecoprc$$

and $\phi(\cdot)$ is the pdf of $N(0,1)$. Specifically:

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right).$$

ii. In this setting, we have:

$$z = 1.088 + 2.029 \times 0.884 - 2.344 \times 1.082.$$

Now, we have two ways to obtain the marginal effect. First, we can use the fact that the marginal effect of a one-unit change is given by:

$$\phi(1.088 + 2.029 \times 0.884 - 2.344 \times 1.082) \times -2.344.$$

To get the effect of a \$0.10 reduction, we will then need to multiply the above by -0.10.

Second, candidates could opt for computing:

$$\Phi(1.088 + 2.029 \times 0.884 - 2.344 \times (1.082 - 0.10)) - \Phi(1.088 + 2.029 \times 0.884 - 2.344 \times 1.082).$$

Comments:

- Most candidates lost marks in part (a) for either forgetting the *ceteris paribus* condition or missing several parts required for hypothesis testing.
- Several candidates lost marks in part (c) for choosing the wrong degrees of freedom or misspecifying the null hypothesis.
- Some candidates found part (d) particularly challenging. Note that the question did not require the use of a calculator but some very good candidates went beyond and computed the exact marginal effect.

Question 7

Let us consider the following ADL(1, 1) model:

$$\boldsymbol{Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + u_t, \quad |\beta_1| < 1}$$

where X_t and Y_t are $I(0)$ variables and u_t is an i.i.d. $(0, \sigma_u^2)$ error term that is uncorrelated with $Y_{t-1}, Y_{t-2}, Y_{t-3}, \ldots$. You may assume that u_t is independent of $X_{t'}$ for all t, t'.

(a) **Provide the short-run and long-run effect of X on Y. Explain the difference between these effects.**

(3 marks)

(b) **Discuss what properties your OLS estimators for the ADL(1, 1) parameters will have in the presence of the lagged dependent variable. In particular, (i) are the estimators unbiased and consistent, and (ii) should we use robust standard errors? Provide supportive arguments for your answers.**

(7 marks)

(c) **Show that when you omit the relevant variable Y_{t-1} in the above model, you will get evidence of autocorrelation. Explain the result.**

Hint: **You are expected to reformulate your model as:**

$$\boldsymbol{Y_t = \beta_0 + \beta_2 X_t + \beta_3 X_{t-1} + v_t}$$

where $v_t = \beta_1 Y_{t-1} + u_t$, and show that, for example, $\mathrm{Cov}(v_{t+1}, v_t) \neq 0$.

(5 marks)

(d) **Discuss how you would proceed to test for the presence of autocorrelation in the model in (b) using the Breusch–Godfrey test. You may assume that under the alternative, u_t is (or, can be suitably well approximated by) a stationary AR(1).**

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 11.3, 11.4, 11.5 and 12.3.

Dougherty, C. Subject guide (2016): Chapters 11 and 12.

Approaching the question

(a) Candidates were expected to clearly indicate the estimates for the short-run and long-run effects of X on Y before discussing their differences as follows.

The short-run effect is β_2. The long-run effect is:

$$\frac{\beta_2 + \beta_3}{1 - \beta_1}.$$

We obtain this by solving the model at equilibrium:

$$Y^e = \beta_0 + \beta_1 Y^e + \beta_2 X^e + \beta_3 X^e \quad \Rightarrow \quad Y^e = \frac{\beta_0}{1 - \beta_1} + \underbrace{\frac{\beta_2 + \beta_3}{1 - \beta_1}}_{\text{long-run effect}} X^e.$$

The short-run effect gives the immediate effect that a change (potentially just temporary) in X has on Y (temporary change); whereas the long-run effect indicates what in equilibrium the effect would be of a permanent change in X.

(b) Candidates were expected to clearly indicate the biasedness of the estimator due to the presence of Y_{t-1} and further discuss the assumptions for consistency.

Because of the presence of Y_{t-1} in the above model, the OLS estimator is biased. The assumptions for consistency are satisfied because for this we only need to ensure that u_t is uncorrelated with the regressors Y_{t-1}, X_t and X_{t-1}. Because the X variables are independent of the errors by assumption, they are uncorrelated. As long as there is no autocorrelation, we will not have any correlation between Y_{t-1} and u_t. The requirement for unbiasedness is stronger than before. We do not need to use robust standard errors because there is no heteroskedasticity or autocorrelation.

(c) Candidates were expected to show that $\text{Cov}(v_{t+1}, v_t) \neq 0$. Although some candidates chose to discuss the intuitions underlying these derivations, it is clearer to combine equations and the intuitions as follows.

Using our definition of v_t and v_{t+1}, we note:

$$\begin{aligned}\text{Cov}(v_{t+1}, v_t) &= \text{Cov}(\beta_1 Y_t + u_{t+1}, \beta_1 Y_{t-1} + u_t) \\ &= \beta_1^2 \text{Cov}(Y_t, Y_{t-1}) + \beta_1 \text{Cov}(Y_t, u_t) + \beta_1 \text{Cov}(u_{t+1}, Y_{t-1}) + \text{Cov}(u_{t+1}, u_t).\end{aligned}$$

Because u_t is white noise (i.i.d.), the errors are unrelated to anything in the past. As a result, $\text{Cov}(u_{t+1}, u_t) = 0$ and $\text{Cov}(u_{t+1}, Y_{t-1}) = 0$.

To show that $\text{Cov}(v_{t+1}, v_t) \neq 0$, it is now sufficient to point out that $\text{Cov}(Y_t, u_t) \neq 0$. This follows quickly when noting that as our model is given by:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + u_t, \quad |\beta_1| < 1$$

this expression is a direct relationship between Y_t and u_t. We immediately have that $\text{Cov}(Y_t, u_t) \neq 0$ because $\text{Cov}(Y_t, Y_{t-1}) \neq 0$ as $\beta_1 \neq 0$.

(d) Candidates were expected to perform the following steps to get full marks, in particular, the test equation. A standard testing procedure was expected as follows.

We should test $\text{H}_0 : \rho = 0$ against $\text{H}_1 : \rho \neq 0$ with $u_t = \rho u_{t-1} + e_t$, where e_t is white noise. Run OLS on the original equation to obtain the residuals $\widehat{u}_t$:

$$\widehat{u}_t = Y_t - \widehat{\beta}_0 - \widehat{\beta}_1 Y_{t-1} - \widehat{\beta}_2 X_t - \widehat{\beta}_3 X_{t-1}.$$

By the assumption that the alternative is a stationary AR(1) process, the test equation is:

$$\widehat{u}_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 X_t + \alpha_3 X_{t-1} + \rho \widehat{u}_{t-1} + e_t.$$

Using the sample size of the latter regression, n, and its R^2 (goodness-of-fit measure), we obtain our test statistic $nR^2 \overset{a}{\sim} \chi_1^2$. We should reject the null hypothesis if nR^2 exceeds the critical value of the χ_1^2 distribution.

Comments:

- In part (a), several candidates failed to recognise β_2 as the short-run effect; others did not clearly discuss the fact that the long-run effect pertains to the permanent change in X.
- In part (b), only a few candidates discussed the fact that the estimator remains consistent because u_t is uncorrelated with the regressors.
- In part (c), only a few candidates did well in this subquestion. Several did not follow the hint and relied on an incorrect covariance: $\text{Cov}(Y_{t-1}, u_t)$ or $\text{Cov}(Y_t, u_t)$.
- In part (d), like in other questions on hypothesis testing, candidates lost marks for missing one or more parts of the testing process. Several candidates did not include $\rho \widehat{u}_{t-1}$ in the testing equation or failed to indicate the correct degrees of freedom.

Question 8

Let us consider the following ADL(1, 1) model:

$$\boldsymbol{Y_t = \beta_0 + \beta_1 Y_{t-1} + \beta_2 X_t + \beta_3 X_{t-1} + u_t}$$

where X_t and Y_t are $I(1)$ variables and $|\beta_1| < 1$. Let us assume that the error term has zero mean and that the error is uncorrelated with Y_{t-1}, X_t and X_{t-1}.

(a) It is important to distinguish whether the above model is spurious or co-integrating. Explain these concepts clearly and highlight their differences.

(5 marks)

(b) Discuss how you can test whether the above specification is spurious (null hypothesis) or co-integrating (alternative) using a standard Augmented Dickey–Fuller test. Clearly indicate the test equation, test statistic and the rejection rule. The critical value provided by Dickey and Fuller for this test at the 5% level of significance is equal to −2.86.

(5 marks)

(c) You are told that there exists an Error Correction Model (ECM) that describes the short-run and long-run dynamics between Y_t and X_t. What does this tell you as it relates to your answers in (a) and (b)? Provide the Error Correction Model and interpret the various components of the ECM.

(5 marks)

(d) Discuss how the ECM can be estimated using Ordinary Least Squares.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 13.1, 13.2, 13.4, and 13.6.

Dougherty, C. Subject guide (2016): Chapter 13.

Approaching the question

(a) Candidates were expected to discuss the following points.

Spurious regression refers to the scenario where *two unrelated series* give the impression that they are related, purely because of the fact that both are trending. A classic example is where Y_t and X_t are both unrelated $I(1)$ series (random walk or random walk with drift). Running a regression of Y_t on X_t (a spurious relationship) has a large R^2 and a seemingly significant slope parameter. The problem is that here the error term is $I(1)$ as well and statistical inference is invalid. The regression is meaningless!

Another example of spurious regression is where two completely unrelated trend stationary series may give the impression that they are related purely because of the omission of the 'time' variable (OVB). Once we account for the deterministic trend, the regression will be able to detect that there is no relationship between X and Y. By omitting the time trend in the regression, the parameter on X will capture their joint trending behaviour.

If, on the other hand, we have co-integration then we do have a long-run relationship between our $I(1)$ variables X_t and Y_t. We have a linear combination of $I(1)$ variables that becomes stationary $I(0)$. For example, if X_t and Y_t are $I(1)$ processes, but $Y_t - \beta X_t$ is stationary, we can say X and Y are co-integrated, meaning there is a long-run relationship between these two. In that case, there will be an ECM representation – see part (c).

If we have a spurious relationship, the OLS parameter estimates will not be consistent, the R^2 will be high and the DW test statistic small (as evidence of strong dependence in the errors). For a co-integrating relationship, the OLS parameter estimates will be consistent.

(b) Candidates were expected to perform the augmented version of the Dickey–Fuller test (the ADF test). To test for co-integration, candidates were expected to evaluate whether the disturbance term is a stationary process (co-integrating) or contains unit roots (spurious). With this intuition in mind, candidates should be able to perform an ADF test on any type of model equations. The following steps were expected.

We run the above equation to obtain the residuals:

$$\widehat{u}_t = Y_t - \widehat{\beta}_0 - \widehat{\beta}_1 Y_{t-1} - \widehat{\beta}_2 X_t - \widehat{\beta}_3 X_{t-1}.$$

The ADF specification we can use to test for a spurious relationship is given by:

$$\Delta\widehat{u}_t = \rho\widehat{u}_{t-1} + \delta\Delta\widehat{u}_{t-1} + e_t$$

where by having added sufficient lagged differences we ensure that e_t does not exhibit autocorrelation.

We now test $\mathrm{H}_0 : \rho = 0$ (spurious, nonstationary) against $\mathrm{H}_1 : \rho < 0$ (co-integrating, stationary). The test statistic is:

$$\frac{\widehat{\rho}}{\mathrm{se}(\widehat{\rho})}$$

and we reject at the 5% level of significance if this value is smaller than -2.86.

(c) Candidates were expected to discuss the implication of the existence of an ECM in the model before discussing the error correction model.

If there is an ECM model for X and Y, then X and Y must be co-integrated and the test conducted in (b) should have rejected the null hypothesis. The error correction model can be written as:

$$\Delta Y_t = (\beta_1 - 1)\left(Y_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta_1}X_{t-1}\right) + \beta_2\Delta X_t + u_t$$

where:

- the error correction term is:

$$Y_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta_1}X_{t-1}$$

  measuring the disequilibrium at time $t-1$
- $(\beta_1 - 1)$ indicates how fast the two processes are coming back to the steady state
- β_2 is the direct effect of a change in X on Y.

(d) Candidates were expected to discuss the following Engel–Granger two-step estimation.

Firstly, we estimate the co-integrating relationship between X and Y by OLS to obtain residuals:

$$\widehat{diseq}_{t-1} = Y_{t-1} - \widehat{\gamma}_1 - \widehat{\gamma}X_{t-1}.$$

Secondly, using $\widehat{diseq}_{t-1}$ we run OLS on the error correction equation:

$$\Delta Y_t = (\beta_1 - 1)\widehat{diseq}_{t-1} + \beta_2\Delta X_t + u_t.$$

Comments:

- In part (a), only a few candidates clearly distinguished the two concepts and their implications for an OLS estimator (particularly the consistency).
- In part (b), several candidates used the Dickey–Fuller version of the test (without the lagged differences of the residuals) instead of the augmented version. Some candidates used Y_t instead of the residuals $\widehat{u}_t$.
- Candidates did very well in parts (c) and (d). A few made mistakes when rewriting the ECM model and failed to interpret the error correction term.

Question 9

It is postulated that a reasonable demand–supply model for the wine industry in Australia would be given by:

$$Q_t = \alpha_0 + \alpha_1 P_t^w + \alpha_2 P_t^b + \alpha_3 Y_t + \alpha_4 A_t + u_t \quad \text{demand}$$

$$Q_t = \beta_0 + \beta_1 P_t^w + \beta_3 S_t + v_t \quad \text{supply}$$

where Q_t = real per capita consumption of wine, P_t^w = price of wine relative to CPI, P_t^b = price of beer relative to CPI, Y_i = real per capita disposable income, A_t = real per capital advertising expenditure, and S_t = storage cost. CPI is the Consumer Price Index.

The endogenous variables in this model are Q and P^w, and the exogenous variables are P^b, Y, A and S.

The variance of u_t and v_t are, respectively, σ_u^2 and σ_v^2, and $\text{Cov}(u_t, v_t) = \sigma_{uv} \neq 0$. The errors do not exhibit any correlation over time.

(a) Provide the reduced form for P_t^w.

(5 marks)

(b) The OLS estimation of the demand function, based on annual data from 1955–1975 ($T = 20$), gave the following results (all variables are in logs and figures in parentheses are t-ratios).

$$\widehat{Q}_t = \underset{(-6.04)}{-23.651} + \underset{(4.0)}{1.158} P_t^w \underset{(-0.45)}{-0.275} P_t^b \underset{(4.5)}{+3.212} Y_t \underset{(-1.3)}{-0.603} A_t.$$

All the coefficients except that of Y have the wrong signs. The coefficient of P^w (price elasticity of demand, α_1) not only has the wrong sign but also appears significant.

Explain why the OLS parameter estimator may give rise to these counter-intuitive results. You are expected to use your results in part (a) to support your answer.

(5 marks)

(c) The supply equation is overidentified. Clearly explain this terminology. What distinguishes overidentification from exact identification and underidentification? Provide one set of assumptions that would render the supply equation exactly identified.

(5 marks)

(d) Discuss how you should estimate the supply equation in light of the overidentification.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections 2.6, 9.1, and 9.3.

Dougherty, C. Subject guide (2016): Chapters 8 and 9, and Section 15.15.

Approaching the question

(a) To derive the reduced form for P_t^w, candidates were expected to explicitly combine the demand curve and supply curve and set the quantity of demand equal to the quantity of supply as follows:

$$\alpha_0 + \alpha_1 P_t^w + \alpha_2 P_t^b + \alpha_3 Y_t + \alpha_4 A_t + u_t = \beta_0 + \beta_1 P_t^w + \beta_3 S_t + v_t.$$

We then rewrite to get an expression of P_t^w in terms of the exogenous regressors and errors only:

$$\begin{aligned} P_t^w &= \frac{\beta_0 + \beta_3 S_t + v_t - \alpha_0 - \alpha_2 P_t^b - \alpha_3 Y_t - \alpha_4 A_t - u_t}{\alpha_1 - \beta_1} \\ &= \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \frac{\beta_3}{\alpha_1 - \beta_1} S_t - \frac{\alpha_2}{\alpha_1 - \beta_1} P_t^b - \frac{\alpha_3}{\alpha_1 - \beta_1} Y_t - \frac{\alpha_4}{\alpha_1 - \beta_1} A_t + \frac{v_t - u_t}{\alpha_1 - \beta_1}. \end{aligned}$$

Candidates were expected to discuss the fact that $\alpha_1 \neq \beta_1$, which allows us to derive the last step (demand/supply slope different). Therefore, the reduced form equation equals:

$$P_t^w = \pi_0 + \pi_1 S_t + \pi_2 P_t^b + \pi_3 Y_t + \pi_4 A_t + V_t$$

where π_j are our reduced-form parameters.

(b) Candidates were expected to discuss the endogeneity present in the structural equations and relate the problems with the estimation model. Points for discussion are the following.

The key reason that the OLS gives rise to these counterintuitive results is the endogeneity of P_t^w, which renders our OLS inconsistent and biased. Q_t and P_t^w are jointly determined in the system of simultaneous equations, and this joint determination results in the endogeneity problem as $\text{Cov}(P_t^w, u_t) \neq 0$.

Our parameter estimates above therefore suffer from simultaneity bias which may indeed even lead to parameter estimates of the wrong sign. Using the result from (a), we have:

$$\text{Cov}(P_t^w, u_t) = \text{Cov}(\pi_0 + \pi_1 S_t + \pi_2 P_t^b + \pi_3 Y_t + \pi_4 A_t + V_t, u_t)$$

where:

$$V_t = \frac{v_t - u_t}{\alpha_1 - \beta_1}.$$

As we are told that S_t, P_t^b, Y_t, and A_t are exogenous, we get:

$$\text{Cov}(P_t^w, u_t) = \text{Cov}\left(\frac{v_t - u_t}{\alpha_1 - \beta_1}, u_t\right) = \frac{1}{\alpha_1 - \beta_1}(\sigma_{vu} - \sigma_u^2) \neq 0.$$

This result demonstrates the correlation between the regressor and the error term in the structural equations.

(c) Candidates were expected to give clear definitions of over-identification and discuss how it differs from under-identification and exact-identification. Excellent answers should specifically recognise that two parameters in the demand equation could be made zero for the supply equation to become exactly identified.

The supply equation indeed is overidentified because we have three instrumental variables P_t^b, Y_t and A_t ($k = 3$) and only one endogenous variable P_t^w ($G - 1 = 1$). The scenario that the number of instrumental variables is larger than the number of endogenous variables is regarded as overidentification.

If in the demand equation we have, say, $\alpha_3 = \alpha_4 = 0$, the supply equation becomes exactly identified, because then we would only have one instrument P_t^b ($k = 1$) available for the endogenous variable P_t^w ($G - 1 = 1$). Other assumptions may be offered as well.

Set the number of instrumental variables as K and the number of endogenous variables as E. We can distinguish the three cases by the following criteria:

(1) $K > E$ means overidentification.

(2) $K = E$ means exact identification.

(3) $K < E$ means underidentification.

(d) Candidates were expected to perform a Two Stage Least Squares estimation using all of the available instruments as follows.

First, we estimate the reduced form of P_t^w and get the fitted value $\widehat{P}_t^w$:

$$\widehat{P}_t^w = \widehat{\pi}_0 + \widehat{\pi}_1 S_t + \widehat{\pi}_2 P_t^b + \widehat{\pi}_3 Y_t + \widehat{\pi}_4 A_t.$$

Second, we plug $\widehat{P}_t^w$ instead of P_t^w into the supply equation and run OLS to estimate the β parameters, i.e. we have:

$$Q_t = \beta_0 + \beta_1 \widehat{P}_t^w + \beta_2 S_t + \varepsilon_t.$$

Alternatively, the second step would involve performing an IV estimation on the supply equation where we use $\widehat{P}_t^w$ as instrument for P_t^w. Candidates could use the IV estimate formula at this step. Notice that since the second stage equation also includes S_t, candidates should not use the simple linear regression formula, which is correct only in the bivariate case.

Comments:

- In part (a), only a few candidates received full marks for this question. Most candidates failed to explicitly state that $\alpha_1 \neq \beta_1$ so that the reduced form equation exists. Others failed to clearly notate their derivations, in particular explicitly stating the first step of equating the demand and supply equations.
- In part (b), some candidates did not explicitly mention the bias caused by the endogeneity (note a very small sample in this question) and discuss the consistency of the estimator instead. This discussion still received marks. Most candidates failed to use the result in (a) to support their discussion of the correlation between the error term and the regressor.
- In part (c), some candidates discussed the assumptions for a valid instrument but failed to explicitly explain when the supply equation can be exactly identified. Some candidates mistook the number of exogenous variables and endogenous variables available in the model even though they correctly identified that the model was over-identified.
- In part (d), most candidates did very well on this question. There were three common mistakes. First, some candidates omitted some variables from the first-stage regression. Second, a handful of candidates estimated the demand and supply equations instead. Third, instead of running the first stage as described above, some candidates constructed a new instrument (Z_t) as the linear combination of the instruments ($Z_t = \alpha S_t + \beta P_t^b + \gamma Y_t$) to maximise the correlation between this new instrument and the endogenous variable, i.e. max $\text{corr}(Z_t, P_t^w)$. The candidates then used this instrument Z_i in a simple linear regression formula with Q_t as the second stage regression. While the idea underlying this method provides a consistent estimator, the candidates failed to discuss how such maximisation could be done. Furthermore, this is not the TSLS procedure.

Question 10

To investigate the relationship between the price of wine and consumption of wine, an economist runs the following regression on a sample of 32 individuals for one week in 2013:

$$\widehat{\log(wine)} = \underset{(0.8911)}{4.2514} - \underset{(0.0031)}{0.8328} \log(price)$$

$$n = 32, \ R^2 = 0.89.$$

wine **denotes the amount of wine consumed per week in millilitres (a medium glass contains 175ml) and** ***price*** **denotes the average price of a selection of wines during the week in GBP (£). The numbers in parentheses are the standard errors.**

(a) Discuss what would happen to the parameter estimate of the slope coefficient if we had measured the amount of wine consumed per week in number of medium glasses instead of millilitres. Explain your answer.

(3 marks)

(b) You are asked to test the hypothesis that the demand for wine has an elasticity equal to -1 against a two-sided alternative, using a 5% level of significance. Clearly stating any assumptions you may need, carry out this test.

(5 marks)

(c) Construct a 95% confidence interval for the price elasticity of demand and discuss how this interval can be used to carry out the test in (b).

(5 marks)

(d) A famous TV chef suggests in a talk show that the demand for wine is less elastic (i.e. less negative) for people who have eaten at a restaurant during the week, arguing that eating in a restaurant encourages people to drink wine regardless of the price. To test this theory, the economist defines a dummy variable D_i that takes the value 1 if individual i ate at a restaurant during the week, and 0 otherwise. She obtains the following regression result:

$$\widehat{\log(wine)} = \underset{(0.8911)}{4.2133} - \underset{(0.0028)}{0.8218}\log(price) + \underset{(0.0011)}{0.0889}D \times \log(price)$$

$$n = 32, \;\; R^2 = 0.92.$$

i. How does this regression help in assessing the TV chef's claim?

(3 marks)

ii. Conduct a test that may offer support for the TV chef's claim. Clearly specify the null and the alternative hypothesis. Explain clearly why your test is effective in answering the question of interest.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Sections R12, R13, 1.4, 2.6, 4.2 and 5.2.

Dougherty, C. Subject guide (2016): Chapters 2, 4 and 5.

Approaching the question

(a) Candidates were expected to realise that the slope coefficient can be interpreted as the price elasticity of wine consumption and hence changes in measurement will not affect its interpretation. An excellent discussion could follow the points below.

If we measure the amount of wine consumed per week in medium glasses instead of millilitres, nothing will happen to the parameter estimate of the slope coefficient since this coefficient refers to the elasticity of wine consumption with respect to price. A 1.0% increase in price will result in a 0.8328% decline in amount of wine consumed. This elasticity does not vary with the way we measure the wine consumption (or prices for that matter).

(b) Candidates were expected to perform a two-sided test with all necessary components of a testing procedure. An excellent answer should cover the following points.

Suppose the true model is:

$$\log(wine) = \alpha + \beta\log(price) + u.$$

We perform a two-sided test as follows. We formulate the null hypothesis as $\mathrm{H}_0 : \beta = -1$ against $\mathrm{H}_1 : \beta \neq -1$. Under the Gauss–Markov assumptions, whose details must be provided

somewhere in the script, and under normality of the errors assumption, we can use the t test. The test statistic is:

$$\frac{\widehat{\beta}+1}{\mathrm{SE}(\widehat{\beta})} \sim t_{30}$$

under H_0. Decision rule: we should reject at the 5% level of significance if:

$$\left|\frac{\widehat{\beta}+1}{\mathrm{SE}(\widehat{\beta})}\right| > 2.042.$$

Because our t statistic equals $(-0.8328+1)/0.0031 = 53.9355$, which is far larger than any reasonable critical value, we reject the null hypothesis, in favour of the alternative hypothesis.

(c) Candidates were expected to compute the confidence interval with the appropriate critical value, then decide whether we could reject the null hypothesis specified in part (b). An excellent answer should cover the following points.

Using the critical value of 2.042 for $n = 32$, the 95% confidence interval is given by:

$$\left[\widehat{\beta} - 2.042 \times \mathrm{SE}(\widehat{\beta}), \widehat{\beta} + 2.042 \times \mathrm{SE}(\widehat{\beta})\right].$$

Plugging in the numbers, we have:

- the lower bound: $-0.8328 - 2.042 \times 0.0031 = -0.839$
- the upper bound: $-0.8328 + 2.042 \times 0.0031 = -0.826$.

Therefore, the 95% confidence interval equals $[-0.839, -0.826]$. Note that we are using the critical value of 2.042 for the sample size of 30 instead of 1.96, which is appropriate for a large (infinite) sample size.

As -1 does not lie in this interval, we should reject H_0 at the 5% level of significance.

(d) Candidates were expected to discuss the interaction term in part i. and perform a one-sided t test in part ii.

i. Rewrite the regression equation in the following way:

$$\log(wine_i) = \alpha + \beta_1 \log(price_i) + \beta_2 D_i \log(price_i) + \varepsilon_i.$$

β_2 allows us to detect whether there are differences in the elasticity for individuals who went to the restaurant and those who did not. Candidates could provide the following expressions to support the discussion:

$$\mathrm{E}(\log(wine) \mid D = 1, \log(price)) = \alpha + (\beta_1 + \beta_2)\log(price) + \varepsilon$$

$$\mathrm{E}(\log(wine) \mid D = 0, \log(price)) = \alpha + \beta_1 \log(price) + \varepsilon.$$

ii. Notice that the TV chef clearly indicated the direction of the elasticity in their claim (less negative). Hence candidates were expected to perform a one-sided test to evaluate the claim of the TV chef instead of a two-sided test. Candidates were also expected to describe all components of the testing procedure.

We should test the null hypotheses: $\mathrm{H}_0 : \beta_2 = 0$, against the one-sided alternative $\mathrm{H}_1 : \beta_2 > 0$. Candidates lost marks for an alternative hypothesis for a two-sided test.

Under the Gauss–Markov assumptions and the normality assumption, we can use the t test. The t statistic and the test statistic distribution are:

$$\frac{\widehat{\beta}_2}{\mathrm{SE}(\widehat{\beta}_2)} \sim t_{n-3}.$$

We should reject the null hypothesis at the 5% level of significance if:

$$\frac{\widehat{\beta}_2}{\mathrm{SE}(\widehat{\beta}_2)} > 1.699.$$

Interpretation: Because:

$$\frac{\widehat{\beta}_2}{\mathrm{SE}(\widehat{\beta}_2)} = \frac{0.0889}{0.0011} = 80.818 > 1.699$$

we reject the null hypothesis, in favour of the alternative hypothesis. We found evidence that the TV chef's claim is true in our sample.

Comments:

- In part (a), candidates could provide either a discussion on the elasticity interpretation of the estimate or a formal derivation of the new regression equation. If the latter method is chosen, careful steps should be taken to derive the new intercept correctly, which should go up as a result of the transformation. Some candidates derived the new regression equation but mistakenly interpreted the new intercept and had it gone down.
- In part (b), many candidates lost marks for not explicitly stating the Gauss–Markov assumptions (or forgetting the normality assumption).
- In part (c), there were two common mistakes. First, some candidates used the critical value 1.96 instead of 2.042 for a small sample of $n = 32$. Second, many candidates did not discuss the use of the confidence interval in deciding to reject/not reject the null hypothesis. Some used the default value of zero instead of the estimate of interest at hand (-1).
- In part (d) i., most candidates correctly commented that the interaction term helps indicate the differences between those who went to the restaurant and those who did not. In part ii. candidates struggled. Many used a two-sided t test while a one-sided test was expected; others omitted some components of the testing procedure. Some candidates opted for an F test, which tested whether the unrestricted model with the interaction $D_i \log(price_i)$ was equivalent to the restricted model without the interaction. This method does not directly answer the question at hand.

UNIVERSITY OF LONDON

EC2020 ZA

BSc DEGREES AND GRADUATE DIPLOMAS IN ECONOMICS, MANAGEMENT, FINANCE AND THE SOCIAL SCIENCES, THE DIPLOMA IN ECONOMICS AND SOCIAL SCIENCES AND THE CERTIFICATE IN EDUCATION IN SOCIAL SCIENCES

Summer 2021 Online Assessment Instructions

EC2020 Elements of Econometrics

The assessment will be an **open-book take-home online assessment within a 6-hour window**. The requirements for this assessment remain the same as the closed-book exam, with an expected time/effort of 3 hours and 15 minutes (inclusive of 15 minutes reading time).

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from **Section A** (8 marks each) and **THREE** from **Section B** (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Extracts from statistical tables are given after the final question on this paper.

You should complete your assessment using pen and paper. Please use **BLACK INK only.**

Handwritten work then needs to be scanned, converted to PDF and then uploaded to the VLE as **ONE individual file** including the coversheet. Each scanned sheet should have your **candidate number written clearly at the top**. Please **do not write your name anywhere** on any sheet.

The paper will be available at 15:00 (BST) on Thursday 10 June 2021.

You have **until 21:00 (BST) on Thursday 10 June 2021** to upload your file into the VLE submission portal. However, you are advised not to leave your submission to the last minute.

The assessment has been designed with a duration of 6 hours to provide a more flexible window in which to complete the assessment and to appropriately test the course learning outcomes. As an open-book exam, the expected amount of effort required to complete all questions and upload your answers during this window is no more than 3 hours and 15 minutes. Organise your time well.

© University of London 2021

You may use *any* calculator for any appropriate calculations, but you may not use any computer software to obtain solutions. Credit will only be given if all workings are shown.

You are assured that there will be no benefit in you going beyond the expected 3 hours and 15 minutes of effort. Your assessment has been carefully designed to help you show what you have learned in the hours allocated.

This is an open book assessment and as such you may have access to additional materials including but not limited to subject guides and any recommended reading. But the work you submit is expected to be 100% your own. Therefore, unless instructed otherwise, you must not collaborate or confer with anyone during the assessment. The University of London will carry out checks to ensure the academic integrity of your work. Many students that break the University of London's assessment regulations did not intend to cheat but did not properly understand the University of London's regulations on referencing and plagiarism. The University of London considers all forms of plagiarism, whether deliberate or otherwise, a very serious matter and can apply severe penalties that might impact on your award. The University of London 2020-21 *Procedure for the consideration of Allegations of Assessment Offences* is available online at:

Assessment Offence Procedures - University of London

SECTION A

Answer all questions from this section.

1. Consider the following multivariate regression to explain housing price $price$, using the total area $(totsize)$, the number of bedrooms $(bedrooms)$, and the house location $(urban)$:

$$price = \beta_0 + \beta_1 totsize + \beta_2 bedrooms + \beta_3 urban + u. \tag{1.1}$$

The error term u has a variance $var(u) = \sigma^2 \log(totsize)$, where σ^2 is an unknown constant.

(a) **(3 marks)** Describe a weighted least squares (WLS) estimator to address heteroskedasticity in Regression (1.1). Explain the intuition behind this estimator.

(b) **(2 marks)** Discuss the validity of the hypothesis test of $H_0 : \beta_1 = 0$ using the WLS estimator.

(c) **(3 marks)** Explain why the residuals from an OLS estimation of Model (1.1) sum to zero, whereas the residuals from the WLS estimation in (a) may not.

2. Let us consider the following stationary time series model

$$y_t = \alpha + \rho y_{t-1} + \beta x_t + \varepsilon_t, \qquad |\rho| < 1, \qquad t = 1, ..., T.$$

The zero mean error ε_t exhibits autocorrelation that can be represented well by a stationary AR(2) process. You may assume that ε_t and x_s are independent for all t, s and $y_0 = 0$.

(a) **(3 marks)** Discuss how you can test H_0 : no autocorrelation against H_A : AR(2).

(b) **(3 marks)** Briefly discuss the properties of the OLS estimator when applied to the above model (unbiasedness, consistency). Support your answers with suitable arguments (no derivations expected).

(c) **(2 marks)** What is the purpose of heteroskedasticity and autocorrelation robust (HAC) standard errors? Discuss whether HAC standard errors can resolve the main problem associated with estimating the model using OLS.

3. Consider the model

$$\textit{paid} = \beta_0 + \beta_1 \textit{nwifeinc} + \beta_2 \textit{educ} + \beta_3 \textit{exper} + \beta_4 \textit{exper}^2 + u, \tag{3.1}$$

where *paid* = 1 if the woman reports working for a wage outside the home, and zero otherwise, *nwifeinc* is other sources of income, *educ* is years of education, and *exper* is past years of labor experience. You may assume that $E\left(u|\textit{nwifeinc,educ,exper}\right) = 0$

(a) **(2 marks)** Provide an economic interpretation of the parameter β_2.

Suppose that *npaid* =1 if the woman is not working for a wage outside the home, and zero otherwise. You have been provided with parameter estimates of model

$$\textit{npaid} = \alpha_0 + \alpha_1 \textit{nwifeinc} + \alpha_2 \textit{educ} + \alpha_3 \textit{exper} + \alpha_4 \textit{exper}^2 + e. \tag{3.2}$$

(b) **(3 marks)** Discuss how you can obtain the parameter estimates of the β parameters using this regression.

Perhaps surprisingly, the R^2 of both models will be identical as both the total sum of squares and the residual sum of squares of both models are the same.

(c) **(3 marks)** Prove that the total sum of squares of both models will be identical.

4. Consider the following bivariate regression model

$$y_i = \alpha + \beta x_i + u_i,$$

for $i = 1, ..., n$. The regressor is stochastic and exhibits variability in the sample, i.e., $\sum_{i=1}^{n}(x_i - \bar{x})^2 \neq 0$. The error term u_i is i.i.d. with mean zero and constant variance of σ^2 for any value of x_i. Given a random sample of observations $\{(y_{i,}, x_i)\}_{i=1}^{n}$ from the population of interest, consider the following estimator

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} x_i^2 y_i}{\sum_{i=1}^{n} x_i^3}.$$

(a) **(3 marks)** Under which assumption is $\tilde{\beta}$ an unbiased estimator of β?

(b) **(5 marks)** Let us assume that the assumption in (a) is satisfied so that $\tilde{\beta}$ is unbiased. Will this assumption also ensure that $\tilde{\beta}$ is consistent? In your answer, discuss the difference between unbiasedness and consistency.

5. Let us consider the following cross-sectional model for household consumption

$$C = \beta_0 + \beta_1 Y + \beta_2 Y^2 + u,$$

where Y denotes household income, $E(u) = 0$ and $E\left(u^2\right) = \sigma^2$. We will assume that the error is independent of Y (and therefore Y^2 as well). You are provided with a random sample $\{(C_i, Y_i)\}_{i=1}^n$.

(a) **(2 marks)** Discuss the following statement "The above specification captures diminishing marginal effect of income on consumption".

(b) **(2 marks)** What will happen to the OLS parameter estimates when we measure consumption and income in £1,000 instead of £? Explain your answer.

(c) **(4 marks)** Let the above model represent the true relationship between consumption and income. Discuss the effect of ignoring the quadratic term on the OLS estimator of β_1. Will the bias be upward or downward? Explain your answer.

SECTION B

Answer three questions from this section.

6. A logit regression model and a Linear Probability Model (LPM) are used to explain the mortgage denial rates amongst a random sample of 1234 young families in Britain, 35% of whom belong to a minority ethnicity group. There are 520 successful mortgage applications and 714 denied applications, with an average salary of £33,000 and an average deposit of £50,000. The regression results are as follows:

$$\widehat{denial}_i = \Lambda\left(\underset{(0.011)}{-0.032} - \underset{(0.052)}{0.103}\, salary_i + \underset{(0.054)}{0.240}\, minority_i - \underset{(0.044)}{0.095}\, deposit_i\right) \quad \text{(Logit)}$$

$$\widehat{denial}_i = \underset{(0.022)}{-0.103} - \underset{(0.152)}{0.263}\, salary_i + \underset{(0.024)}{0.062}\, minority_i - \underset{(0.144)}{0.295}\, deposit_i \quad \text{(LPM)}$$

where $\Lambda(z) = \exp(z)/(1+\exp(z))$ is the logistic function. The usual standard errors for the Logit model and the robust standard errors for the LPM are reported in parentheses.

The variable $denial_i$ equals to 1 if the mortgage application was denied, 0 otherwise. The variables $salary_i$ and $deposit_i$ are the total annual salary and the downpayment by family i in £10,000. The variable $minority_i$ equals to 1 if the applicant belongs to a minority group, 0 otherwise.

(a) **(5 marks)** Describe the estimator underlying the logit model and discuss its properties. **Note:** clearly indicate the likelihood function used, but a detailed derivation of the estimator is not expected.

(b) **(5 marks)** Discuss the advantages and drawbacks of using the LPM, rather than the logit model, to model the mortgage denial rates.

(c) **(5 marks)** Discuss how you can test the null hypothesis that the ceteris paribus effect of $salary$ on the probability of denial is equal to -0.25. Clearly state the alternative hypothesis, the test statistic, and rejection rule.

(d) **(5 marks)** Using the logit model, what is the marginal effect of being a minority on mortgage denial rate at the mean values of the explanatory variables. Discuss the limitation(s) of the marginal effect obtained and propose an alternative that overcomes such limitation(s). **Note:** there is no need to calculate the marginal effect using the alternative approach.

7. Let us consider the relationship between the natural logarithm of GDP, GDP_t, and the lagged long-term interest rates $rate_t$, and $rate_{t-1}$.

   (a) Assume that GDP_t and $rate_t$ are $I(1)$.

      i. **(4 marks)** If GDP_t and $rate_t$ are not co-integrated, what are the properties of an OLS estimator in the regression of GDP_t on $rate_t$?

      ii. **(4 marks)** Discuss the implications of the existence of cointegration between GDP_t and $rate_t$ on the short-run and long-run relationship of two variables.

      iii. **(7 marks)** Consider the following regression model, where ε_t has mean zero and is uncorrelated with GDP_{t-1}, GDP_{t-2}, $rate_t$, and $rate_{t-1}$:

$$GDP_t = \delta + \theta_1 GDP_{t-1} + \theta_2 GDP_{t-2} + \phi_1 rate_t + \phi_2 rate_{t-1} + \varepsilon_t.$$

      Derive the Error Correction Model (ECM) representation of the equation, and discuss the long-run equilibrium relation for GDP_t and $rate_t$. Interpret the parameters of the model and discuss the properties of the OLS estimator of the ECM.

   (b) **(5 marks)** Using the UK data from 1988Q4 to 2021Q1, a researcher estimates regressions with the growth rate of GDP, $growth_t = GDP_t - GDP_{t-1}$, as the dependent variable, and the growth of physical capital $capital_t$, the unemployment rate $unemploy_t$, the growth of money supply $msupply_t$, and the growth of productivity, $prod_t$, as the regressors. Assuming these variables are $I(0)$, the researcher reports three estimation results, one with the full sample, and two with sub-periods split at 2006Q1. Robust standard errors are in parentheses, n is the number of observations, SSR is the sum of squared residuals, $\log L$ is the maximised log-likelihood, R^2 is the R-squared.

   Describe how to test whether the regression parameters are the same for both periods. Carefully indicate the alternative hypothesis, the test-statistic, and rejection rule.

	1988Q4-2021Q1	1988Q4-2006Q1	2006Q2-2021Q1
	(1)	(2)	(3)
$capital_t$	0.113 (0.002)	0.033 (0.012)	0.163 (0.052)
$unemploy_t$	−0.005 (0.014)	− 0.08 (0.024)	0.005 (0.013)
$msupply_t$	−0.211 (0.054)	−0.142 (0.084)	−0.335 (0.014)
$prod_t$	0.116 (0.024)	−0.014 (0.004)	0.125 (0.046)
constant	0.003 (0.002)	0.013 (0.005)	0.001 (0.002)
n	130	70	60
SSR	0.288	0.226	0.033
R^2	0.67	0.45	0.47
$\log L$	367.868	179.521	205.485

8. It is postulated that more open economies have lower inflation rates. A proposed model looks at a cross section of n countries using time averages:

$$\textit{infl}_i = \beta_0 + \beta_1 \log(\textit{inc}_i) + \beta_2 \textit{open}_i + u_i, \quad i = 1, 2, ..., n. \tag{8.1}$$

infl$_i$ is average inflation rate, *open*$_i$ is the measure of openness of the economy (average share of imports in GDP), and *inc*$_i$ is average per capita income, and i denotes a particular country.

(a) **(4 marks)** The model (8.1) was estimated by OLS and provided $\hat{\beta}_2 = -.215$ (with $SE(\hat{\beta}_2) = .095$). If openness of the economy itself depends on inflation rate (*infl*$_i$), what sort of problems will this cause for the OLS estimator of (8.1)? Explain your answer.

(b) **(5 marks)** Suppose that the equation

$$\textit{open}_i = \alpha_0 + \alpha_1 \textit{infl}_i + \alpha_2 \log(\textit{inc}_i) + \alpha_3 \log(\textit{land}_i) + v_i, \quad i = 1, 2, ..., n. \tag{8.2}$$

land$_i$ is the land area of the country, holds in addition to (8.1). Could this relation be helpful in providing a suitable method of estimation of β_2 in (8.1)? What is the relevant estimation method and what conditions must be satisfied for it to be valid?

(c) **(6 marks)** Using the cross section of countries, the following regression was obtained by OLS, with robust standard errors reported in parentheses:

$$\widehat{\textit{open}}_i = \underset{(15.85)}{117.08} + \underset{(1.493)}{.546} \log(\textit{inc}_i) - \underset{(.81)}{7.57} \log(\textit{land}_i). \tag{8.3}$$

Discuss whether (and if so how) we can use this equation to verify whether the conditions in (b) are satisfied and how this equation can help us to estimate β_2.

(d) **(3 marks)** Some economists claim that inflation is I(1). Below is an estimated time series model for US inflation, with robust standard errors reported in parentheses:

$$\widehat{\Delta \textit{infl}}_t = \underset{(.21)}{.51} - \underset{(.04)}{.11} \textit{infl}_{t-1} - \underset{(.08)}{.19} \Delta \textit{infl}_{t-1} - \underset{(.07)}{.26} \Delta \textit{infl}_{t-2} + \underset{(.09)}{.20} \Delta \textit{infl}_{t-3} + \underset{(.08)}{.01} \Delta \textit{infl}_{t-4},$$

with $t = 5, ..., T$. Conduct an ADF for the presence of a unit root in inflation using a 5% level of significance. The ADF critical value equals -2.86. Clearly indicate the null and the alternative hypotheses, the test statistic and explain the conclusion.

(e) **(2 marks)** Briefly indicate whether you think your result in (d) has a bearing on the cross-sectional results in (a) - (c). Explain.

9. Let us consider a dataset containing information on government expenditure on education ($EducE$), gross domestic product (GDP), population (Pop), and capital stock ($Capital$) for 153 countries in 2019. Economic theory suggests that government expenditure on education per capita has a positive effect on per capita GDP. That is,

$$perGDP_i = \beta_1 + \beta_2 perEducE_i + \beta_3 perCapital_i + u_i,$$

where $perGDP_i = \frac{GDP_i}{Pop_i}$, $perEducE_i = \frac{EducE_i}{Pop_i}$, and $perCapital_i = \frac{Capital_i}{Pop_i}$, and i denotes a particular country.

(a) **(6 marks)** Explain intuitively why we should be worried about the presence of heteroskedasticity in this model. Briefly indicate **three** distinct methods to potentially deal with heteroskedasticity in this situation.

(b) **(5 marks)** Discuss how you would implement the White test for the presence of heteroskedasticity in this case. Briefly discuss how the White test differs from the Goldfeld-Quandt test.

(c) **(4 marks)** Discuss two reasons why we may expect the OLS estimator for β_2 to be neither unbiased nor consistent. Provide supportive arguments for each.

(d) **(5 marks)** The dataset also has a dummy variable $Democratic_i$ which equals 1 if country i is democratic, 0 otherwise. Discuss how you could test whether the effect of government expenditure on education per capita is higher for democratic countries. Clearly indicate the regression you would run and the test you would use.
Comment: You may assume that the OLS estimator of the regression you would run is consistent, but there may be heteroskedasticity in your model.

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	5·0	1·6449	3·0	1·8808	2·0	2·0537	1·0	2·3263	0·10	3·0902
45	0·1257	4·8	1·6646	2·9	1·8957	1·9	2·0749	0·9	2·3656	0·09	3·1214
40	0·2533	4·6	1·6849	2·8	1·9110	1·8	2·0969	0·8	2·4089	0·08	3·1559
35	0·3853	4·4	1·7060	2·7	1·9268	1·7	2·1201	0·7	2·4573	0·07	3·1947
30	0·5244	4·2	1·7279	2·6	1·9431	1·6	2·1444	0·6	2·5121	0·06	3·2389
25	0·6745	4·0	1·7507	2·5	1·9600	1·5	2·1701	0·5	2·5758	0·05	3·2905
20	0·8416	3·8	1·7744	2·4	1·9774	1·4	2·1973	0·4	2·6521	0·01	3·7190
15	1·0364	3·6	1·7991	2·3	1·9954	1·3	2·2262	0·3	2·7478	0·005	3·8906
10	1·2816	3·4	1·8250	2·2	2·0141	1·2	2·2571	0·2	2·8782	0·001	4·2649
5	1·6449	3·2	1·8522	2·1	2·0335	1·1	2·2904	0·1	3·0902	0·0005	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

UNIVERSITY OF LONDON

EC2020 ZB

BSc DEGREES AND GRADUATE DIPLOMAS IN ECONOMICS, MANAGEMENT, FINANCE AND THE SOCIAL SCIENCES, THE DIPLOMA IN ECONOMICS AND SOCIAL SCIENCES AND THE CERTIFICATE IN EDUCATION IN SOCIAL SCIENCES

Summer 2021 Online Assessment Instructions

EC2020 Elements of Econometrics

The assessment will be an **open-book take-home online assessment within a 6-hour window**. The requirements for this assessment remain the same as the closed-book exam, with an expected time/effort of 3 hours and 15 minutes (inclusive of 15 minutes reading time).

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from **Section A** (8 marks each) and **THREE** from **Section B** (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Extracts from statistical tables are given after the final question on this paper.

You should complete your assessment using pen and paper. Please use **BLACK INK only.**

Handwritten work then needs to be scanned, converted to PDF and then uploaded to the VLE as **ONE individual file** including the coversheet. Each scanned sheet should have your **candidate number written clearly at the top**. Please **do not write your name anywhere** on any sheet.

The paper will be available at 06:00 (BST) on Thursday 10 June 2021.

You have **until 12 midday (BST) on Thursday 10 June 2021** to upload your file into the VLE submission portal. However, you are advised not to leave your submission to the last minute.

The assessment has been designed with a duration of 6 hours to provide a more flexible window in which to complete the assessment and to appropriately test the course learning outcomes. As an open-book exam, the expected amount of effort required to complete all questions and upload your answers during this window is no more than 3 hours and 15 minutes. Organise your time well.

© University of London 2021

You may use *any* calculator for any appropriate calculations, but you may not use any computer software to obtain solutions. Credit will only be given if all workings are shown.

You are assured that there will be no benefit in you going beyond the expected 3 hours and 15 minutes of effort. Your assessment has been carefully designed to help you show what you have learned in the hours allocated.

This is an open book assessment and as such you may have access to additional materials including but not limited to subject guides and any recommended reading. But the work you submit is expected to be 100% your own. Therefore, unless instructed otherwise, you must not collaborate or confer with anyone during the assessment. The University of London will carry out checks to ensure the academic integrity of your work. Many students that break the University of London's assessment regulations did not intend to cheat but did not properly understand the University of London's regulations on referencing and plagiarism. The University of London considers all forms of plagiarism, whether deliberate or otherwise, a very serious matter and can apply severe penalties that might impact on your award. The University of London 2020-21 *Procedure for the consideration of Allegations of Assessment Offences* is available online at:

Assessment Offence Procedures - University of London

SECTION A

Answer all questions from this section.

1. Consider the following regression model

$$Y_i = \beta_0 + \beta_1 \log(X_i) + u_i, \qquad i = 1, ..., n.$$

The error term is i.i.d. with zero mean and variance equal to σ^2 and the error and regressors are independent. We assume $X_i > 0$ for all i.

(a) **(4 marks)** Suppose you are given a random sample of observations $\{(Y_i, X_i)\}_{i=1}^n$. Derive the OLS estimator of β_1 and provide an interpretation of this parameter.

(b) **(4 marks)** Unfortunately you do not observe X_i, instead you observe X_i^* which equals

$$X_i^* = X_i \exp(\varepsilon_i),$$

where ε_i has zero mean and is independent of u_i and X_i (and therefore $\log(X_i)$). You are given a random sample of observations $\{(Y_i, X_i^*)\}_{i=1}^n$. Formulate the estimable regression model and show that the OLS estimator of β_1 is inconsistent, resulting in an underestimate in absolute size of the coefficient on average.

2. Consider the regression model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u, \text{ with } E(u) = 0.$$

You may assume that the regressors are non-stochastic. Suppose the error has a normal distribution and exhibits heteroskedasticity, i.e., $Var(u_i)$ varies with x_{1i} and/or x_{2i}, for all i. You are provided with a random sample $\{(y_i, x_{1i}, x_{2i})\}_{i=1}^n$.

(a) **(4 marks)** Discuss how, with the help of heteroskedasticity robust standard errors, you would test the null hypothesis $H_0^{(a)} : \beta_1 = 1$ against the one-sided alternative $H_A^{(a)} : \beta_1 > 1$ when you use OLS to estimate β_1. In your answer explain the need for robust standard errors.

(b) **(4 marks)** Suppose you are told that $Var(u_i) = c^2 x_{1i}^2$ where c is an unknown constant and $x_{1i} \neq 0$ for all i. Discuss how this knowledge about the form of heteroskedasticity allows you to (i) obtain a more efficient estimator of β_1 and (ii) hence conduct a test of the null hypothesis $H_0 : \beta_1 = 0$ that is more powerful.

3. A researcher tries two regression models to describe the relationship between a variable $X > 0$ and a variable $Y > 0$:

$$\log Y = \alpha_1 + \alpha_2 \log X + u, \tag{3.1}$$

$$\log(XY) = \beta_1 + \beta_2 \log X + v. \tag{3.2}$$

(a) **(2 marks)** Explain whether model (3.2) is a restricted version of (3.1) or not.

(b) **(2 marks)** Using a random sample of size n, the researcher obtains $(\hat{\alpha}_1, \hat{\alpha}_2, \hat{\beta}_1, \hat{\beta}_2)$ as the OLS estimators for $(\alpha_1, \alpha_2, \beta_1, \beta_2)$. Let $(\hat{u}, \hat{v})$ be the residuals of the OLS estimations. Prove that $\hat{u} = \hat{v}$.

(c) **(4 marks)** Assume that the regression equation (3.1) describes the true relationship between X and Y and there exists a variable W uncorrelated with Y. What are the properties of the OLS estimator for γ_2 if we use the following regression specification.

$$\log Y = \gamma_1 + \gamma_2 \log X + \gamma_3 W + w? \tag{3.3}$$

Explain your answer intuitively.

4. (a) **(4 marks)** What are the conditions for a series to be covariance stationary? Consider the following time series model for $\{x_t\}_{t=1}^T$:

$$x_t = x_{t-1} + \lambda^2 \varepsilon_t + \lambda \varepsilon_{t-1} + \varepsilon_{t-2},$$

where ε_t is i.i.d with mean zero and variance σ^2, for $t = 1, ..., T$, and $x_0 = 0$. Discuss whether there are any values of λ so that x_t is covariance stationary.

(b) **(4 marks)** Consider the following time series model for $\{y_t\}_{t=1}^T$:

$$y_t = \theta_0 + \theta_1(\rho - 1)t + \rho y_{t-1} + \varepsilon_t,$$

where ε_t is i.i.d with mean zero and variance σ^2, $t = 1, ..., T$, and $y_0 = 0$. How can we test whether $\rho = 1$? Clearly specify the null and alternative hypothesis, the test statistic, and the rejection rule.

5. Consider the following regression model

$$y_t = \rho y_{t-1} + \gamma_1 z_t + \gamma_2 z_{t-1} + \gamma_3 z_{t-2} + \gamma_4 z_{t-3} + \varepsilon_t \quad \text{with } |\rho| < 1,$$

where ε_t is an MA(1) process, i.e.,

$$\varepsilon_t = u_t + \theta u_{t-1}.$$

u_t is a white noise process generated randomly from a normal distribution with mean zero and constant variance that is independent of $y_{t-1}, y_{t-2,}, ..., z_t, z_{t-1}, ...$. You have a random sample and can use T observations, which we label $t = 1, 2, ..., T$ for convenience, for the estimation. Assume there is no perfect multicollinearity.

(a) **(4 marks)** Briefly discuss the finite-sample and large-sample properties of the OLS estimator for ρ from the regression model where $\theta \neq 0$.

(b) **(4 marks)** Let $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$, so that

$$y_t = \rho y_{t-1} + \varepsilon_t \quad \text{with } |\rho| < 1.$$

Discuss how you can obtain a consistent estimator for ρ that recognizes $\theta \neq 0$. Show that this estimator is consistent. Clearly annotate your proof.

SECTION B

Answer three questions from this section.

6. Consider the following model to determine the effectiveness of vaccination on the containment of measles, a dangerously contagious but vaccine-preventable disease, in a developing country:

$$measles_i = \alpha_0 + \alpha_1 \log(vaccinated)_i + \alpha_2 children_i + \beta_3 avginc_i + \beta_4 city_i + u_i, \qquad (6.1)$$

where the variance of the error term u_i is assumed constant and equal to σ_u^2, and i denotes a particular district; $measles_i$ is the number of reported measles cases per 10,000 residents, $vaccinated_i$ is the reported number of vaccinated residents per 10,000 residents, $children_i$ is the number of children per 10,000 residents, $avginc_i$ is the average level of income, and $city_i$ equals to 1 if the district is in a city, 0 otherwise. You have a random sample of 450 districts.

(a) **(2 marks)** What is the interpretation of the coefficient α_1? What is the expected sign of α_1?

Epidemiological studies suggest that areas with high rates of infection are more heavily targeted with vaccination programmes, hence these areas will have more people vaccinated. The following model describes this relationship:

$$\log(vaccinated)_i = \gamma_0 + \gamma_1 measles_i + \gamma_3 vaccentre_i + \gamma_4 com_i + \gamma_5 city_i + v_i, \qquad (6.2)$$

where the variance of the error term v_i is assumed constant and equal to σ_v^2, $vaccentre_i$ is the number of vaccination centres available per 10,000 residents in district i, com_i is a variable that measures the communication and media environment regarding disease containment issues. Assume that $children$, $avginc$, $city$, $vaccentre$, and com are exogenous.

(b) **(7 marks)** Identify and discuss **two** likely problems with estimating α_1 by OLS. Clearly indicate the direction of bias, if any.

(c) **(7 marks)** In light of your answers in (b), describe an appropriate estimator for α_1. Briefly discuss and justify the necessary assumptions, if any.

(d) **(4 marks)** Another researcher argues that there should be no statistical difference between the estimates for α_1 obtained from OLS and from the estimator in (c). Describe a test that might help the researcher statistically justify the argument. Clearly indicate the null, the alternative, test statistic, test distribution, and rejection rule. Are there any limitations of the test?

7. Let $X_1, X_2, ..., X_n$ be a random sample of size n from the exponential distribution with parameter θ. The probability density function is specified as

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}},\ x \geq 0 \text{ and } \theta > 0.$$

Note that $E(X_i) = \theta$ and $Var(X_i) = \theta^2$.

In this question we will consider various estimators of θ.

(a) Let us first consider a sequence of estimators given by

$$\widehat{\theta}_c = c\overline{X},$$

where c is a positive constant and $\overline{X} = n^{-1} \sum_{i=1}^{n} X_i$.
For example, $\widehat{\theta}_1 = \overline{X}$ and $\widehat{\theta}_{0.5} = 0.5\bar{X}$.

i. **(6 marks)** For any given c, determine the bias, variance, and mean squared error (MSE) of the estimator $\widehat{\theta}_c$.

Note that $MSE(\widehat{\theta}) = E(\widehat{\theta} - \theta)^2 = Bias^2(\widehat{\theta}) + Var(\widehat{\theta})$.

ii. **(3 marks)** Show that $c^* = \dfrac{n}{n+1}$ minimizes the MSE of $\widehat{\theta}_c$.
Hint: You will need to look at the derivative of $MSE(\widehat{\theta})$ with respect to c.

iii. **(5 marks)** Is the estimator $\widehat{\theta}_{c^*} = (n+1)^{-1} \sum_{i=1}^{n} X_i$ unbiased? Consistent? Prove your answers.

(b) **(6 marks)** Derive the MLE estimator for θ, $\hat{\theta}_{MLE}$. Briefly discuss what desirable properties this estimator has.

8. Let us consider a dataset containing information on government expenditure on education ($EducE$), gross domestic product (GDP), population (Pop), and capital stock ($Capital$) for 153 countries in 2019. Economic theory suggests that government expenditure on education per capita has a positive effect on per capita GDP. That is,

$$perGDP_i = \beta_1 + \beta_2 perEducE_i + \beta_3 perCapital_i + u_i,$$

where $perGDP_i = \frac{GDP_i}{Pop_i}$, $perEducE_i = \frac{EducE_i}{Pop_i}$, and $perCapital_i = \frac{Capital_i}{Pop_i}$, and i denotes a particular country.

(a) **(6 marks)** Explain intuitively why we should be worried about the presence of heteroskedasticity in this model. Briefly indicate **three** distinct methods to potentially deal with heteroskedasticity in this situation.

(b) **(5 marks)** Discuss how you would implement the White test for the presence of heteroskedasticity in this case. Briefly discuss how the White test differs from the Goldfeld-Quandt test.

(c) **(4 marks)** Discuss two reasons why we may expect the OLS estimator for β_2 to be neither unbiased nor consistent. Provide supportive arguments for each.

(d) **(5 marks)** The dataset also has a dummy variable $Democratic_i$ which equals 1 if country i is democratic, 0 otherwise. Discuss how you could test whether the effect of government expenditure on education per capita is higher for democratic countries. Clearly indicate the regression you would run and the test you would use.
Comment: You may assume that the OLS estimator of the regression you would run is consistent, but there may be heteroskedasticity in your model.

9. We are interested in explaining the binary variable *vhappy*, a dummy variable that denotes whether an individual considers him/herself "very happy" or not (1 = yes, 0 = no).

   We consider the following socio-demographic variables: *occattend* and *regattend* (which are dummy variables indicating whether the individual occasionally or regularly attends church, where the excluded dummy indicates that the individual never attends church), *income* (family income in 1,000US$), *unemp10* (dummy indicating whether the individual has been unemployed in the last 10 years), and *educ* (years of education completed). Let z be a linear function of the variables that determine the probability of being very happy, where:

$$z = \beta_0 + \beta_1 occattend + \beta_2 regattend + \beta_3 \log(income) + \beta_4 unemp10 + \beta_5 educ.$$

   The β parameters have been estimated based on a Linear Probability Model (LPM) and Logit Model. Two sets of results (labelled specification (A) and (B)) are obtained based on a random sample of 9,963 observations from US. The usual standard errors for the Logit model and the robust standard errors for the LPM are reported in parentheses.

	LPM (A)	LPM (B)	Logit (A)	Logit (B)
constant	.080 (.032)	.088 (.032)	−2.074 (.178)	−2.032 (.176)
occattend	−.005 (.010)	–	−.023 (.051)	–
regattend	.093 (.014)	–	.420 (.065)	–
$\log$(*income)*	.061 (.012)	.064 (.012)	.362 (.070)	.379 (.069)
unemp10	−.097 (.010)	−.103 (.010)	−.487 (.050)	−.514 (.049)
educ	.008 (.002)	.007 (.002)	.036 (.008)	.034 (.008)
R^2	.023	.019		
$\log L$			−5890.6	−5998.9

   (a) **(5 marks)** Discuss clearly how the parameter estimates for β are obtained for the Linear Probability Model and the Logit model. No technical details expected, but intuition rewarded.

(b) For this part use specification (A).

i. **(5 marks)** Obtain for both models the predicted probability of happiness for an individual who regularly attends church, has a family income of £150,000, has not been unemployed in the last 10 years and has 13 years of education. *Hint*: $\log(150) \simeq 5$.

ii. **(5 marks)** Obtain for both models what the effect of an additional year of education is for this individual and state whether the impact of a further additional year is the same.

(c) You would like to test the joint significance of the church attendance variables *occattend* and *regattend*.

i. **(2 marks)** Explain why you **cannot** conduct this test using the LPM based on the reported results.

ii. **(3 marks)** Explain how you would conduct this test using the Logit Model based on the reported results.

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	2·60	0·99534	2·75	0·99702	2·90	0·99813	3·05	0·99886	3·20	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	2·65	0·99598	2·80	0·99744	2·95	0·99841	3·10	0·99903	3·25	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	2·70	0·99653	2·85	0·99781	3·00	0·99865	3·15	0·99918	3·30	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	3·263	0·9994	3·731	0·99990	3·916	0·99995
3·105	0·9991	3·320	0·9995	3·759	0·99991	3·976	0·99996
3·138	0·9992	3·389	0·9996	3·791	0·99992	4·055	0·99997
3·174	0·9993	3·480	0·9997	3·826	0·99993	4·173	0·99998
3·215	0·9994	3·615	0·9998	3·867	0·99994	4·417	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	5·0	1·6449	3·0	1·8808	2·0	2·0537	1·0	2·3263	0·10	3·0902
45	0·1257	4·8	1·6646	2·9	1·8957	1·9	2·0749	0·9	2·3656	0·09	3·1214
40	0·2533	4·6	1·6849	2·8	1·9110	1·8	2·0969	0·8	2·4089	0·08	3·1559
35	0·3853	4·4	1·7060	2·7	1·9268	1·7	2·1201	0·7	2·4573	0·07	3·1947
30	0·5244	4·2	1·7279	2·6	1·9431	1·6	2·1444	0·6	2·5121	0·06	3·2389
25	0·6745	4·0	1·7507	2·5	1·9600	1·5	2·1701	0·5	2·5758	0·05	3·2905
20	0·8416	3·8	1·7744	2·4	1·9774	1·4	2·1973	0·4	2·6521	0·01	3·7190
15	1·0364	3·6	1·7991	2·3	1·9954	1·3	2·2262	0·3	2·7478	0·005	3·8906
10	1·2816	3·4	1·8250	2·2	2·0141	1·2	2·2571	0·2	2·8782	0·001	4·2649
5	1·6449	3·2	1·8522	2·1	2·0335	1·1	2·2904	0·1	3·0902	0·0005	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

Examiners' commentaries 2021

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2020–21. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

General remarks

Learning outcomes

At the end of the course, and having completed the Essential reading and activities, you should be able to:

- describe and apply the classical regression model and its application to cross-section data
- describe and apply the:
  - Gauss–Markov conditions and other assumptions required in the application of the classical regression model
  - reasons for expecting violations of these assumptions in certain circumstances
  - tests for violations
  - potential remedial measures, including, where appropriate, the use of instrumental variables
- recognise and apply the advantages of logit, probit and similar models over regression analysis when fitting binary choice models
- competently use regression, logit and probit analysis to quantify economic relationships using standard regression programmes (Stata and EViews) in simple applications
- describe and explain the principles underlying the use of maximum likelihood estimation
- apply regression analysis to time-series models using stationary time series, with awareness of some of the econometric problems specific to time series applications (for example, autocorrelation) and remedial measures
- recognise the difficulties that arise in the application of regression analysis to nonstationary time series, know how to test for unit roots, and know what is meant by cointegration.

Common mistakes committed by candidates

Due to the impact of the Covid-19 pandemic, the examinations again took place online this year. Different from the 2019/20 academic year, the examinations were no longer 24-hour, instead candidates had a six-hour window to complete and submit their work. The two-zone system also returned this year, in which two different sets of questions were used for Zone A and Zone B, with one common question (Question 9 Zone A and Question 8 Zone B). Realising many challenges due to the pandemic that candidates had to face, we commend the performance of many candidates this year. However, a good number of candidates submitted work that revealed a serious lack of effort despite the open-book nature of the assessment. We discuss below common mistakes made by candidates in both zones, most of them are repeated year after year. We then provide specific comments for each question.

Candidates should realise that $\frac{1}{n}\sum_{i=1}^{n} u_i$ is not the same as $\mathrm{E}(u_i)$. So, while we typically assume $\mathrm{E}(u_i) = 0$, this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} u_i = 0$. Also, while we may be happy to assume $\mathrm{E}(x_i u_i) = 0$ (uncorrelatedness between the errors and regressors), this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} x_i u_i = 0$. Note that:

- both $\frac{1}{n}\sum_{i=1}^{n} u_i$ and $\frac{1}{n}\sum_{i=1}^{n} x_i u_i$ are random variables, which take the value 0 with probability 0 (continuous random variables)!

- $\mathrm{E}(u_i) = 0$ and $\mathrm{E}(x_i u_i) = 0$ are fixed, not stochastic!

The differences between sample and population moments need to come across clearly when looking at unbiasedness and making consistency arguments. In both cases, we first simplify our estimator (plug in the true model) to obtain:

$$\widehat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \beta + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \quad \text{with } x_i = X_i - \bar{X}.$$

- For *unbiasedness*, clearly indicate that you want to show that $\mathrm{E}(\widehat{\beta}) = \beta$. Unbiasedness does not follow from $\sum_{i=1}^{n} x_i u_i = 0$, instead it follows from $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) = 0$.

  If we treat x_i as fixed:

  $$\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) \equiv \mathrm{E}\left(\sum_{i=1}^{n} d_i u_i\right) = \sum_{i=1}^{n} d_i \mathrm{E}(u_i)$$

  and then unbiasedness follows as $\mathrm{E}(u_i) = 0$.

- For *consistency*, clearly indicate that you want to show that $\mathrm{plim}(\widehat{\beta}) = \beta$. Using the plim

properties, we show:

$$\begin{aligned}
\operatorname{plim} \widehat{\beta} = \beta + \operatorname{plim} \left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \right) &= \beta + \frac{\operatorname{plim} \left(\frac{1}{n} \sum_{i=1}^{n} x_i u_i \right)}{\operatorname{plim} \left(\frac{1}{n} \sum_{i=1}^{n} x_i^2 \right)} \\
&\equiv \beta + \frac{\operatorname{plim}(\text{Sample Cov}(x, u))}{\operatorname{plim}(\text{Sample Var}(x))} \\
&= \beta + \frac{\operatorname{Cov}(x, u)}{\operatorname{Var}(x)} \quad \text{using the law of large numbers}
\end{aligned}$$

where $\operatorname{Cov}(x, u) = 0$ and $\operatorname{Var}(x) > 0$, ensuring we get consistency.

- Remember, the law of large numbers ensures that sample averages converge to their population analogues.

A large number of candidates are not able to clearly distinguish between sample variance and covariance, and population variance and covariance.

The use of $\operatorname{Cov}(X, Y)$ and $\operatorname{Var}(X)$ should be restricted to describing the population covariance and variances, respectively, with definitions:

$$\operatorname{Cov}(X, Y) = \operatorname{E}((X - \operatorname{E}(X))(Y - \operatorname{E}(Y))) = \operatorname{E}(XY) - \operatorname{E}(X)\operatorname{E}(Y)$$

and:

$$\operatorname{Var}(X) = \operatorname{E}((X - \operatorname{E}(X))^2) = \operatorname{E}(X^2) - (\operatorname{E}(X))^2$$

(you also may denote $\operatorname{Cov}(X, Y) = \sigma_{XY}$ and $\operatorname{Var}(X) = \sigma_X^2$). They are typically unknown, but fixed, quantities.

The sample covariance and variance are estimators of the population covariance and variance, respectively. They are defined as:

$$\text{Sample Cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})$$

and:

$$\text{Sample Var}(X) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

(you also may use $\widehat{\sigma}_{XY}$ and $\widehat{\sigma}_X^2$). You can compute them given the data.

With a slight abuse of notation, we often divide by n instead, which is irrelevant if we let n be large. The division by $n - 1$ is a finite sample issue only (unbiasedness).

The sample covariance and variance show up in our definition of the OLS estimator of the slope in the simple linear regression model, not the population covariance and variance, as:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = \frac{\text{Sample Cov}(X, Y)}{\text{Sample Var}(X)} \neq \frac{\operatorname{Cov}(X, Y)}{\operatorname{Var}(X)}.$$

Treating them as being the same results in incorrect analyses and candidates lose significant marks.

$$\widehat{\beta}_{IV} = \frac{\sum_{i=1}^{n} (Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (Z_i - \bar{Z})(X_i - \bar{X})} = \frac{\text{Sample Cov}(Z_i, Y_i)}{\text{Sample Cov}(Z_i, X_i)} \neq \frac{\operatorname{Cov}(Z_i, Y_i)}{\operatorname{Cov}(Z_i, X_i)}.$$

It was very common to misunderstand the central limit theorem. While it provides the distribution of the *estimator* $\widehat{\beta}$, no matter what the true distribution of the errors is, it does *not* say that the distribution of the error term is approximately normal in large samples.

Many candidates simply described multicollinearity as independent variables being correlated. Although we accept answers discussing high correlation amongst the independent variables as an alternative to a close linear relation amongst the regressors, simply taking correlation amongst regressors as the definition of multicollinearity is incorrect. Many candidates incorrectly commented that multicollinearity causes biasedness and/or inconsistency.

Autocorrelation and heteroskedasticity were tricky topics for a good number of candidates. Many did not explain well the need to use robust standard errors or an asymptotic version of testing under autocorrelation and/or heteroskedasticity. Most also lost marks when discussing the properties of an estimator that can help solve these problems. Some incorrectly suggested that fixing autocorrelation will also resolve inconsistency and/or biasedness.

When discussing the importance of distinguishing trend stationary and difference stationary processes, two points are expected: discussing the implication for the long-run behaviour of the processes, and discussing statistical implications (the consistency of the estimators). Most candidates discussed the former very well, only a few touched upon the latter.

Related, candidates often lost marks for the interpretation of the ECM model, particularly when discussing the properties of the estimator when the variables are stationary or not. Candidates often struggled with the derivation of an ECM model. A good number of candidates avoided this question, others who attempted it did not do very well.

A good number of candidates seemed to be confused between the Dickey–Fuller test and an Augmented Dickey–Fuller (ADF). The difference between the two tests is in the test equation. In the augmented specification, we add sufficient lagged differences of the residuals to ensure that there is no autocorrelation left in the test equation regression. For this reason, some used the test equation for the ADF when the DF was called for.

Candidates struggled to give competent answers to the interpretation of empirical results. When interpreting an empirical result you should discuss the significance of the coefficients, magnitude, and sign of the coefficients. Related, *ceteris paribus* (or holding other factors constant) is needed to ensure that the estimate can be interpreted as the marginal change in the dependent variable given a change in the independent variable of interest. Many candidates lost marks from these mistakes.

Others struggled with interpreting coefficients from a non-linear regression (for example, log-linear, linear-log, or log-log).

Candidates particularly struggled with the difference between percentage points and % changes when it comes to the interpretation of coefficients in binary choice models. A coefficient in such a model gives the percentage point difference in the outcome variable when the regressor changes by 1 unit. The % change is a *relative* change that depends on the baseline value of the outcome variable (for example, the realisation when a regressor equals zero in the case of a dummy variable regressor). For example, if we have initially 50% of a sample buying ecological apples, a 10% increase will mean now $(1 + 0.1) \times 50 = 55\%$ of the sample buying ecological apples. A 10% point increase instead will result in $50 + 10 = 60\%$ of the sample choosing to buy ecological apples.

When discussing the marginal effect after a Probit estimation, candidates often did not specify the term z in the pdf $\phi(z)$. Some used the coefficients from the LPM regressions to calculate the marginal effects after a Probit/Logit estimation. A popular, related mistake was that candidates did not realise the inherent heteroskedasticity of LPM.

A number of candidates seemed to be confused between the conditions for a valid instrument (exclusion restriction and relevance) and factors related to obtaining a more precise IV estimate (a higher sample size, a higher sample variability of X_i, a lower variance of the error term, and a stronger correlation between X_i and Z_i). More seriously, most candidates did not realise the use of an IV estimator when the regression has a lagged dependent variable as a regressor.

One statistical way to test for an endogeneous variable is the Durbin–Wu–Hausman test. The test statistically compares the OLS and IV estimates, assuming we have valid instruments. Many candidates realised this but failed to discuss the test correctly (mistaking the test as a Chow or a t test). Many did not realise that we can only use the Hausman test if we have more instruments than regressors that need instrumenting.

Exogeneity and endogeneity are often very simply explained as variables 'outside of the model' or 'inside the model', respectively. While this approach is not wrong; it is not a particularly clear or helpful definition in general contexts that are not related to simultaneous equations models. Defining these concepts in terms of the relationship between regressor and the error term is much clearer and naturally leads to a way to show exogeneity or endogeneity by considering the covariance between regressors and error terms.

Candidates often missed that all exogenous regressors and instruments have to be included in the first stage when performing a TSLS estimation. Some candidates use the fitted values of the endogenous variable obtained from the first stage as the dependent variable in the second stage. This mistake is a serious misunderstanding.

When deriving the reduced forms, candidates should check conditions for their existence, in particular, whether the denominator of the expression is different from zero.

Gauss–Markov conditions were used in many places in the examination, for example, when discussing the BLUE property of an estimator or when conducting hypothesis tests. The Gauss–Markov conditions have to be explicitly specified. Only writing that the Gauss–Markov conditions hold is not sufficient. A good practice that many candidates have adopted is to begin the script by explicitly providing the Gauss–Markov conditions and refer back to them thereafter. Many candidates did not mention either the Gauss–Markov conditions or the normality assumption when conducting the t test.

When conducting hypothesis testing, ensure that you clearly indicate the null and alternative hypotheses (in terms of the true parameters, say β_1), the test statistic (in terms of the parameter estimates, here $\widehat{\beta}_1$), its distribution (with degrees of freedom), the rejection rule (one-sided or two-sided) for a given significance level (typically 5%) with suitable critical values, and provide an interpretation of your result. Candidates often lost marks for missing parts of these steps. Choosing the critical value appropriate to the sample size is also important. Several candidates used 1.96 (a critical value for very large degrees of freedom) in a small sample instead of the appropriate one. This mistake also cost marks. Related, many did not know when to use an asymptotic version of a test.

Another common mistake is to conduct a two-sided test when the statement was worded with a clear direction of the relationship. Many candidates fail to provide all the necessary specifics of the testing procedure (testing distribution, the critical value, the assumptions, the rejection rule, and the conclusion) for a t test, an F test, or a Chow test. These often resulted in lost marks.

Just as in previous years, many candidates did not answer all parts of a question. Make sure you read the questions properly and provide all details that are requested. Not answering a question will automatically earn you a zero. Some candidates probably did not have a good examination plan.

Key steps to improvement

Essential reading for this course includes the subject guide and the following:

- Dougherty, C. *Introduction to econometrics*. (Oxford: Oxford University Press, 2016) 5th edition [ISBN 9780199676828]; http://oxfordtextbooks.co.uk/orc/dougherty5e/

Apart from the Essential readings you should do some supplementary reading. One very good book at the same level is:

- Gujarati, D.N. and D.C. Porter *Basic econometrics*. (McGraw–Hill, 2009, International edition) 5th edition [ISBN 9780071276252].

To understand the subject clearly it is important to supplement Dougherty's *Introduction to econometrics* (fifth edition) with the subject guide **EC2020 Elements of econometrics** (2016), especially Chapter 10 which covers maximum likelihood estimation. It is very important to carefully go through the subject guide. The subject guide contains solutions to the questions given in the main textbook and also some additional questions and solutions. Working through these will improve your understanding of the subject.

The chapter in the subject guide on maximum likelihood (Chapter 10) includes some additional theory which has not been covered in the main textbook. It is important to read the additional theory given in the subject guide to have a better understanding of the principles of maximum likelihood and tests based on the likelihood function.

Please check the VLE course page for resources for this subject such as a downloadable copy of the subject guide **EC2020 Elements of econometrics** (2016), PowerPoint slideshows that provide a graphical treatment of the topics covered in the textbook, datasets and statistical tables. Candidates should utilise datasets using standard regression programmes (STATA or EViews). This will help in the understanding of the subject.

Examination revision strategy

Many candidates are disappointed to find that their examination performance is poorer than they expected. This may be due to a number of reasons, but one particular failing is '**question spotting**', that is, confining your examination preparation to a few questions and/or topics which have come up in past papers for the course. This can have serious consequences.

We recognise that candidates might not cover all topics in the syllabus in the same depth, but you need to be aware that examiners are free to set questions on **any aspect** of the syllabus. This means that you need to study enough of the syllabus to enable you to answer the required number of examination questions.

The syllabus can be found in the Course information sheet available on the VLE. You should read the syllabus carefully and ensure that you cover sufficient material in preparation for the examination. Examiners will vary the topics and questions from year to year and may well set questions that have not appeared in past papers. Examination papers may legitimately include questions on any topic in the syllabus. So, although past papers can be helpful during your revision, you cannot assume that topics or specific questions that have come up in past examinations will occur again.

If you rely on a question-spotting strategy, it is likely you will find yourself in difficulties when you sit the examination. We strongly advise you not to adopt this strategy.

Examiners' commentaries 2021

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2020–21. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone A

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from **Section A** (8 marks each) and **THREE** from **Section B** (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Section A

Answer all questions from this section.

Question 1

Consider the following multivariate regression to explain housing price *price*, using the total area (*totsize*), the number of bedrooms (*bedrooms*); and the house location (*urban*):

$$price = \beta_0 + \beta_1 totsize + \beta_2 bedrooms + \beta_3 urban + u. \tag{1.1}$$

The error term u has a variance $\text{var}(u) = \sigma^2 \log(totsize)$, where σ^2 is an unknown constant.

(a) **Describe a weighted least squares (WLS) estimator to address heteroskedasticity in Regression (1.1). Explain the intuition behind this estimator.**

(3 marks)

(b) **Discuss the validity of the hypothesis test of $H_0 : \beta_1 = 0$ using the WLS estimator.**

(2 marks)

(c) Explain why the residuals from an OLS estimation of Model (1.1) sum to zero, whereas the residuals from the WLS estimation in (a) may not.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 1.5, 3.2, 7.1, and 7.3.

Dougherty, C. Subject guide (2016): Chapters 1, 3, 7 and Chapter 15.15.

Approaching the question

(a) With $\text{var}(u) = \sigma^2 \log(totsize) = \sigma^2 h(totsize)$, where $h(totsize)$ is the heteroskedasticity function. We first transform the equation by dividing the original equation by $\sqrt{h(totsize)} = \sqrt{\log(totsize)}$:

$$\frac{price}{\sqrt{\log(totsize)}} = \beta_0 \frac{1}{\sqrt{\log(totsize)}} + \beta_1 \frac{totsize}{\sqrt{\log(totsize)}} + \beta_2 \frac{bedrooms}{\sqrt{\log(totsize)}} + \beta_3 \frac{urban}{\sqrt{\log(totsize)}} + \frac{u}{\sqrt{\log(totsize)}}.$$

The transformed equation now has an error term that is homoskedastic.

- The key intuition behind weighting is that WLS weights observation i by a factor of $1/\sqrt{\log(totsize)}$, and it gives the highest weights to the most informative and reliable observations (those with the lowest values of $\text{var}(u)$.
- The WLS requires us to use OLS to estimate this transformed model (no intercept).

(b) The test statistics from the WLS are exactly valid when the error term is normally distributed and asymptotically valid under non-normality.

(c) In the OLS estimation, there is an intercept, the first-order condition of the minimisation problem of the sum of squared residuals will give us the zero-sum of residuals. Since the WLS estimation does not contain an intercept, the minimisation of the objective function will not guarantee the condition.

Comments:

- The most common mistake was incorrect derivations of the WLS or failing to recognise that WLS is simply OLS, and in this case, without an intercept.
- Candidates also lost marks on (b) for not recognising that the test statistic from the WLS is now valid. Some incorrectly commented that WLS is inefficient and/or biased and raises the likelihood of a Type I error.
- Part (c) is not a challenging question, but many failed to relate the use of first-order conditions when deriving an OLS estimator and to recognise that the key intuition for a zero-sum of residuals comes from including the intercept.

Question 2

Let us consider the following stationary time series model

$$y_t = \alpha + \rho y_{t-1} + \beta x_t + \varepsilon_t, \quad |\rho| < 1, \quad t = 1, \ldots, T.$$

The zero mean error ε_t exhibits autocorrelation that can be represented well by a stationary AR(2) process. You may assume that ε_t and x_s are independent for all t, s and $y_0 = 0$.

(a) Discuss how you can test $\mathbf{H_0}$: no autocorrelation against $\mathbf{H_1}$: AR(2).

(3 marks)

(b) Briefly discuss the properties of the OLS estimator when applied to the above model (unbiasedness, consistency). Support your answers with suitable arguments (no derivations expected).

(3 marks)

(c) What is the purpose of heteroskedasticity and autocorrelation robust (HAC) standard errors? Discuss whether HAC standard errors can resolve the main problem associated with estimating the model using OLS.

(2 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 12.1 and 12.2.

Dougherty, C. Subject guide (2016): Chapter 12.

Approaching the question

(a) To test H_0 : no autocorrelation against H_1 : AR(2) we use the Breusch–Godfrey test. To obtain our test statistic we need to run the following regression:

$$\widehat{\varepsilon}_t = \gamma_1 + \gamma_2\widehat{\varepsilon}_{t-1} + \gamma_3\widehat{\varepsilon}_{t-2} + \gamma_4 y_{t-1} + \gamma_5 x_t + u_t$$

where $\widehat{\varepsilon}_t$ are the OLS residuals of the model. Our test is a joint test of the hypothesis: $\mathrm{H}_0 : \gamma_2 = \gamma_3 = 0$ against $\mathrm{H}_1 : \gamma_2 \neq 0$ and/or $\gamma_3 \neq 0$. We can use the F test or nR^2.

(b) When we apply OLS to the above model in the presence of autocorrelation, AR(2), we will get a biased and inconsistent estimator. The presence of a lagged dependent variable is already enough to tell us that we will get a biased estimator [it violates C.7].

If in addition the errors exhibit autocorrelation we will get inconsistency as well as that will result in $\mathrm{Cov}(y_{t-1}, \varepsilon_t) \neq 0$. Intuivitely, lagging our model shows that y_{t-1} depends on ε_{t-1}. Given the AR(2) structure of our error ($\varepsilon_t = \rho_1\varepsilon_{t-1} + \rho_2\varepsilon_{t-2} + e_t$), ε_t also depends on ε_{t-1}.

(c) HAC robust standard errors are required to ensure our standard errors are valid in the presence of autocorrelation and heteroskedasticity. HAC robust standard errors are used when we apply OLS in the presence of autocorrelation, it does not solve the problem of inconsistency of OLS.

Comments:

- Several candidates did not recognise the need to use the Breusch–Godfrey test but used a Durbin–Watson test, or had an incorrect testing equation (for example, not using lags of the residuals).
- Candidates recognised the bias and inconsistency due to the inclusion of lagged dependent variables, but many forgot the autocorrelation and its consequences.
- A good number of candidates suggested that the estimator will become asymptotically unbiased or consistent with large samples, which is not correct without any correction to the estimator. Some suggested incorrectly that once the autocorrelation is fixed the biasedness would also be fixed.
- Part (c) was in general well done. Marks were often lost for poor explanation.

Question 3

Consider the model

$$\boldsymbol{paid = \beta_0 + \beta_1 nwifeinc + \beta_2 educ + \beta_3 exper + \beta_4 exper^2 + u,} \tag{3.1}$$

where $paid = 1$ if the woman reports working for a wage outside the home, and zero otherwise, $nwifeinc$ is other sources of income, $educ$ is years of education, and $exper$ is past years of labour experience. You may assume that $\mathrm{E}(u \mid nwifeinc, educ, exper) = 0$.

(a) Provide an economic interpretation of the parameter β_2.

(2 marks)

Suppose that $npaid = 1$ if the woman is not working for a wage outside the home, and zero otherwise. You have been provided with parameter estimates of the model

$$\boldsymbol{npaid = \alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper^2 + e.} \tag{3.2}$$

(b) Discuss how you can obtain the parameter estimates of the β parameters using this regression.

(3 marks)

Perhaps surprisingly, the R^2 of both models will be identical as both the total sum of squares and the residual sum of squares of both models are the same.

(c) Prove that the total sum of squares of both models will be identical.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 1.6, and 2.2, and 3.5.

Dougherty, C. Subject guide (2016): Chapter 3

Approaching the question

(a) It shows the effect of one additional year of education on the probability of working for wage earned outside the home, *ceteris paribus* (holding everything else constant). Mathematically, it is $\Delta P(paid = 1 \mid x)/\Delta educ$.

(b) Let us replace $paid = 1 - npaid$ in our first equation and rewrite:

$$1 - npaid = \beta_0 + \beta_1 nwifeinc + \beta_2 educ + \beta_3 exper + \beta_4 exper^2 + u$$

or:

$$npaid = 1 - \beta_0 - \beta_1 nwifeinc - \beta_2 educ - \beta_3 exper - \beta_4 exper^2 - u.$$

Hence we have established:

$$\alpha_0 = 1 - \beta_0, \quad \alpha_1 = -\beta_1, \quad \alpha_2 = -\beta_2, \quad \alpha_3 = -\beta_3, \quad \alpha_4 = -\beta_4, \quad e = -u.$$

We can use this one-to-one relation to estimate our β parameters as:

$$\widehat{\beta}_0 = 1 - \widehat{\alpha}_0, \quad \widehat{\beta}_1 = -\widehat{\alpha}_1, \quad \widehat{\beta}_2 = -\widehat{\alpha}_2, \quad \widehat{\beta}_3 = -\widehat{\alpha}_3, \quad \widehat{\beta}_4 = -\widehat{\alpha}_4$$

and also $\widehat{u} = -\widehat{e}$ so that the residual sum of squares of both regressions will be equal as well $\left(\sum_{i=1}^{n} \widehat{u}_i^2 = \sum_{i=1}^{n} \widehat{e}_i^2\right)$.

(c) Let us start with the total sum of squares of the original model:

$$TSS_{paid} = \sum_{i=1}^{n}(paid_i - \overline{paid})^2.$$

Now we want to replace $paid_i$ with $1 - npaid_i$.

First we observe that:

$$\overline{paid} = \frac{1}{n}\sum_{i=1}^{n} paid_i = \frac{1}{n}\sum_{i=1}^{n}(1 - npaid_i) = 1 - \frac{1}{n}\sum_{i=1}^{n} npaid_i = 1 - \overline{npaid}.$$

From that, we can rewrite:

$$\begin{aligned}\sum_{i=1}^{n}(1 - npaid_i - (1 - \overline{npaid}))^2 &= \sum_{i=1}^{n}(\overline{npaid} - npaid_i)^2 \\ &= \sum_{i=1}^{n}(npaid_i - \overline{npaid})^2 \\ &= TSS_{npaid}.\end{aligned}$$

Comments:

- Part (a) most marks were lost for forgetting the *ceteris paribus* condition.
- Candidates used many ways to show the equivalence of the two regressions, and were rewarded accordingly for their effort. Some lost marks for incorrect annotations, but not significantly.

Question 4

Consider the following bivariate regression model

$$\boldsymbol{y_i = \alpha + \beta x_i + u_i,}$$

for $i = 1, \ldots, n$. The regressor is stochastic and exhibits variability in the sample, i.e. $\sum_{i=1}^{n}(x_i - \bar{x})^2 \neq 0$. The error term u_i is i.i.d. with mean zero and constant variance of σ^2 for any value of x_i. Given a random sample of observations $\{(y_i, x_i)\}_{i=1}^{n}$ from the population of interest, consider the following estimator

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} x_i^2 y_i}{\sum_{i=1}^{n} x_i^3}.$$

(a) Under which assumption is $\tilde{\beta}$ an unbiased estimator of β?

(3 marks)

(b) Let us assume that the assumption in (a) is satisfied so that $\tilde{\beta}$ is unbiased. Will this assumption also ensure that $\tilde{\beta}$ is consistent? In your answer, discuss the difference between unbiasedness and consistency.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters R14, 3.3, 8.2, and 8.3.

Dougherty, C. Subject guide (2016): Chapters 2 and 8.

Approaching the question

(a) For unbiasedness, we need to show that $\mathrm{E}(\tilde{\beta}) = \beta$. Let us first consider the setting where the regressors are not stochastic. The expectation of $\tilde{\beta}$ is then given by:

$$\begin{aligned}
\mathrm{E}(\tilde{\beta}) = \mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i^2 y_i}{\sum_{i=1}^{n} x_i^3}\right) &= \mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i^2(\alpha + \beta x_i + u_i)}{\sum_{i=1}^{n} x_i^3}\right) && \text{(replacing } y_i \text{ from the model)} \\
&= \mathrm{E}\left(\frac{\alpha\left(\sum_{i=1}^{n} x_i^2\right) + \beta\left(\sum_{i=1}^{n} x_i^3\right) + \sum_{i=1}^{n} x_i^2 u_i}{\sum_{i=1}^{n} x_i^3}\right) \\
&= \beta + \alpha\left(\frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^3}\right) + \frac{\sum_{i=1}^{n} x_i^2 \mathrm{E}(u_i)}{\sum_{i=1}^{n} x_i^3} \\
&= \beta + \alpha\left(\frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^3}\right). && \text{(because } \mathrm{E}(u_i) = 0\text{)}
\end{aligned}$$

Thus, in general $\tilde{\beta}$ is a biased estimator unless the population intercept $\alpha = 0$. We rule out the case of $\sum_{i=1}^{n} x_i^2 = 0$ due to the sample variability in x_i.

To deal with the stochastic nature of the regressors, candidates can use the fact that under independence of errors and regressors we can rewrite:

$$\mathrm{E}(\tilde{\beta}) = \beta + \alpha\,\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^3}\right) + \sum_{i=1}^{n} \mathrm{E}\left(\frac{x_i^2}{\sum_{i=1}^{n} x_i^3}\right)\mathrm{E}(u_i) = \beta + \alpha\left(\frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^3}\right) \neq \beta$$

unless $\alpha = 0$.

(b) Consistency requires that as the sample size increases, the estimator gets arbitrarily close to the true parameter (or it converges in probability to the true parameter). It is a large sample property and we need to show that:

$$\mathrm{plim}(\tilde{\beta}) = \beta.$$

Unbiasedness is a finite sample property, which requires that the estimator on average is the true parameter when we repeatedly draw an infinite number of samples of size n.

When the estimator is unbiased, we must have $\alpha = 0$. From (a), we can rewrite the estimator as:

$$\tilde{\beta} = \beta + \alpha\left(\frac{\sum_{i=1}^{n} x_i^2}{\sum_{i=1}^{n} x_i^3}\right) + \frac{\sum_{i=1}^{n} x_i^2 u_i}{\sum_{i=1}^{n} x_i^3} = \beta + \frac{\sum_{i=1}^{n} x_i^2 u_i}{\sum_{i=1}^{n} x_i^3}.$$

The probability limit of $\tilde{\beta}$ is:

$$\begin{aligned}
\operatorname{plim}(\tilde{\beta}) &= \beta + \operatorname{plim}\left(\frac{\sum_{i=1}^{n} x_i^2 u_i}{\sum_{i=1}^{n} x_i^3}\right) \\
&= \beta + \frac{\operatorname{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2 u_i\right)}{\operatorname{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^3\right)} && \text{(properties of plim)} \\
&= \beta + \frac{\mathrm{E}(x_i^2 u_i)}{\mathrm{E}(x_i^3)} && \text{(from the Law of Large Numbers)} \\
&= \beta + \frac{x_i^2 \mathrm{E}(u_i)}{\mathrm{E}(x_i^3)} \\
&= \beta. && \text{(as } \mathrm{E}(u_i) = 0\text{)}
\end{aligned}$$

Therefore, when $\tilde{\beta}$ is an unbiased estimator, it is also consistent.

Comments:

- As discussed in the general comments, some candidates mistook the meanings of unbiasedness and inconsistency or did not utilise the assumptions to justify their derivations.
- Some lost marks due to incorrect derivations.
- The discussion on the differences between unbiasedness and inconsistency also cost some candidates marks for a lack of clarity or misunderstanding. Some candidates misused the LLN and the CLT and instead called them the properties of plim.

Question 5

Let us consider the following cross-sectional model for household consumption

$$C = \beta_0 + \beta_1 Y + \beta_2 Y^2 + u,$$

where Y denotes household income, $\mathrm{E}(u) = 0$ and $\mathrm{E}(u^2) = \sigma^2$. We will assume that the error is independent of Y (and therefore Y^2 as well). You are provided with a random sample $\{(C_i, Y_i)\}_{i=1}^{n}$.

(a) **Discuss the following statement 'The above specification captures diminishing marginal effect of income on consumption'.**

(2 marks)

(b) **What will happen to the OLS parameter estimates when we measure consumption and income in £1,000 instead of £? Explain your answer.**

(2 marks)

(c) **Let the above model represent the true relationship between consumption and income. Discuss the effect of ignoring the quadratic term on the OLS estimator of β_1. Will the bias be upward or downward? Explain your answer.**

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 2.2, 3.2, 4.3, 6.2.

Dougherty, C. Subject guide (2016): Chapters 2, 4 and 6.

Approaching the question

(a) This allows us to assume that there will be a diminishing effect of income on consumption which would require both $\beta_1 > 0$ and $\beta_2 < 0$.

(b) The parameter β_1 remains unchanged, but β_0 and β_2 need to be rescaled. Let $\tilde{C} = C/1{,}000$ and $\tilde{Y} = Y/1{,}000$ denote our new variables. Let's substitute them:

$$1{,}000\tilde{C} = \beta_0 + \beta_1 1{,}000\tilde{Y} + \beta_2(1{,}000\tilde{Y})^2 + u$$

$$\tilde{C} = \frac{\beta_0}{1{,}000} + \beta_1\tilde{Y} + 1{,}000\beta_2\tilde{Y}^2 + u.$$

The intercept will be divided by 1,000 whereas β_2 will be multiplied by 1,000.

(c) If we estimate:

$$C = \beta_0 + \beta_1 Y + e \qquad \text{(short equation)}$$

we will get the classical OVB problem. With:

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(C_i - \bar{C})}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} = \beta_1 + \beta_2\frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(Y_i^2 - \bar{Y}^2)}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2} + \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(u_i - \bar{u})}{\sum_{i=1}^{n}(Y_i - \bar{Y})^2}$$

and (assuming Y_i not stochastic and using $\mathrm{E}(u) = 0$):

$$\mathrm{E}(\widehat{\beta}_1) = \beta_1 + \underbrace{\beta_2\frac{\text{Sample Cov}(Y_i, Y_i^2)}{\text{Sample Var}(Y_i)}}_{\text{OVB}}.$$

As we expect that Sample $\mathrm{Cov}(Y_i, Y_i^2) > 0$ and $\beta_2 < 0$; the bias will be negative.

Comments:

- Some candidates lost marks for not offering a good discussion on the intuition behind the marginal effect of income being diminishing.
- Most candidates lost marks due to algebraic mistakes or misinterpreting the parameters of the reparameterised equation.
- Part (c) proved to be a challenge as it required intuitive explanations for the direction of the bias, which some missed. Most did not realise the use of the OLS estimator to work out the bias. Some misused the notations of sample covariance and population covariance. Once these steps were done well, candidates found the bias direction correctly.

Section B

Answer three questions from this section

Question 6

A logit regression model and a Linear Probability Model (LPM) are used to explain the mortgage denial rates amongst a random sample of 1,234 young families in Britain, 35% of whom belong to a minority ethnicity group. There are 520 successful mortgage applications and 714 denied applications, with an average salary of £33,000 and an average deposit of £50,000. The regression results are as follows:

$$\widehat{denial}_i = \Lambda\left(\underset{(0.011)}{-0.032} - \underset{(0.052)}{0.103}\,salary_i + \underset{(0.054)}{0.240}\,minority_i - \underset{(0.044)}{0.095}\,deposit_i\right) \qquad \text{(Logit)}$$

$$\widehat{denial}_i = \underset{(0.022)}{-0.103} - \underset{(0.152)}{0.263}\,salary_i + \underset{(0.024)}{0.062}\,minority_i - \underset{(0.144)}{0.295}\,deposit_i \qquad \text{(LPM)}$$

where $\Lambda(z) = \exp(z)/(1 + \exp(z))$ is the logistic function. The usual standard errors for the Logit model and the robust standard errors for the LPM are reported in parentheses.

The variable $denial_i$ equals to 1 if the mortgage application was denied, 0 otherwise. The variables $salary_i$ and $deposit_i$ are the total annual salary and the downpayment by family i in £10,000. The variable $minority_i$ equals to 1 if the applicant belongs to a minority group, 0 otherwise.

(a) **Describe the estimator underlying the logit model and discuss its properties. Note: clearly indicate the likelihood function used, but a detailed derivation of the estimator is not expected.**

(5 marks)

(b) **Discuss the advantages and drawbacks of using the LPM, rather than the logit model, to model the mortgage denial rates.**

(5 marks)

(c) **Discuss how you can test the null hypothesis that the *ceteris paribus* effect of *salary* on the probability of denial is equal to −0.25. Clearly state the alternative hypothesis, the test statistic, and rejection rule.**

(5 marks)

(d) **Using the logit model, what is the marginal effect of being a minority on mortgage denial rate at the mean values of the explanatory variables. Discuss the limitation(s) of the marginal effect obtained and propose an alternative that overcomes such limitation(s). Note: there is no need to calculate the marginal effect using the alternative approach.**

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 10.1, 10.2 and 10.6.

Dougherty, C. Subject guide (2016): Chapter 10.

Approaching the question

(a) The probability of denial for each individual is:

$$p_i(denial_i = 1) = \widehat{denial}_i = \Lambda(\alpha_0 + \alpha_1 salary_i + \alpha_2 minority_i + \alpha_3 deposit_i)$$

and the probability of a successful application is:

$$p_i(denial_i = 0) = 1 - \Lambda(\alpha_0 + \alpha_1 salary_i + \alpha_2 minority_i + \alpha_3 deposit_i)$$

where $\Lambda(z) = \exp(z)/(1 + \exp(z))$.

The likelihood function for this sample is:

$$\begin{aligned} L(\alpha_1, \alpha_2, \alpha_3 \,|\, \text{data}) &= \prod_{denial} p_i(denial = 1) \prod_{success} p_i(denial = 0) \\ &= \prod_{denial} \Lambda(\alpha_0 + \alpha_1 salary_i + \alpha_2 minority_i + \alpha_3 deposit_i) \\ &\quad \times \prod_{success} (1 - \Lambda(\alpha_0 + \alpha_1 salary_i + \alpha_2 minority_i + \alpha_3 deposit_i)). \end{aligned}$$

To estimate the parameters, we can use the Maximum Likelihood Estimator that chooses $\{\alpha_j\}$ to maximise the joint probability of the outcomes.

There are no mathematical formulae for the estimates – the popular method is to adopt an iterative trial-and-error process.

In large samples, the MLEs are consistent, asymptotically normally distributed, and asymptotically efficient.

(b) The key problem with the LPM is that the predicted probabilities of denial obtained from the LPM can be larger than 1 or smaller than 0, which is non-sensical. This problem is very much likely to happen, given the estimates from the LPM model.

The logit model addresses this concern as the probability is defined using the cdf ($\Lambda(z)$); which ensures the predicted probabilities are contained between 0 and 1.

The LPM implicitly assumes that the marginal effect of minority (or any variables) is constant over its entire range. Even though it is easier and straightforward to interpret, this assumption might not be intuitively appealing. Since the sample is large, we do not have to worry about the error term of the LPM not having a normal distribution.

Candidates can also discuss the easy implementation of the LPM compared to Logit (which is no longer a concern due to computability nowadays, or LPM will inherently be heteroskedastic).

(c) We need to do a hypothesis test on whether the LPM coefficient on $salary_i$ is equal to -0.25. The Logit parameter estimate does not have this required interpretation for the test, so should not have been used.

The null hypothesis is $\alpha_1 = -0.25$, the alternative hypothesis is $\alpha_1 \neq -0.25$. The appropriate test for the LPM estimate is the z test (asymptotic t test) and we need to use the robust standard error (heteroskedasticity). The z test is:

$$\frac{\widehat{\alpha} + 0.25}{\text{se}(\widehat{\alpha}_1)} = -0.09.$$

Under the null hypothesis, the z test has a $N(0, 1)$ distribution. At the 5% significance level and two-tailed test, the critical values are ± 1.96. If the test statistic is larger than the (absolute) critical value, we reject the null hypothesis, in favour of statistical evidence for racial discrimination in the mortgage approval process. We cannot reject the hypothesis that the effect of salary on the probability of denial is -0.25.

(d) The marginal effect of interest will tell us how the probability of denial, $p_i(denial_i = 1)$, for minority applicants differs from that of non-minority applicants. The marginal effect at the mean values of the explanatory variables is:

$$\frac{\partial p_i(denial_i = 1)}{\partial minority} = 0.240 \times f(z)$$

where $f(z) = \mathrm{e}^{-z}/(1+\mathrm{e}^{-z})^2$ is the pdf of the logistic function, and:

$$\begin{aligned} z &= -0.032 - 0.103\overline{salary} + 0.240\overline{minority} - 0.095\overline{deposit} \\ &= -0.032 - 0.103 \times 3.3 + 0.24 \times 0.35 - 0.095 \times 5 \\ &= -0.763. \end{aligned}$$

Plugging $z = -0.763$ into the pdf of the logistic function, $f(z) = \mathrm{e}^{-z}/(1+\mathrm{e}^{-z})^2 = 0.217$. Thus the marginal effect of interest is $0.240 \times 0.217 = 0.052$. Candidates are not expected to calculate the numbers. [Alternative approach simply compares the probability at averages with $minority = 1$ versus 0.]

The problem with this calculation is that its interpretation is non-sensical as there can be no person that is of 35% minority. The interpretation for an individual with average characteristics might not be of any practical use. The alternative is to look at the average marginal effect over all individuals (here the sample size is $n = 1{,}234$), which can be calculated as follows:

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\partial p_i(denial_i = 1)}{\partial minority} = \frac{1}{n}\sum_{i=1}^{n} f(z_i) \times 0.240$$

where:

$$z_i = \alpha_0 + \alpha_1 salary_i + \alpha_2 minority_i + \alpha_3 deposit_i$$

and $f(z)$ is the pdf.

Comments:

- Most candidates lost marks on (a) for poor explanations on the intuition behind the MLE or missing one or more properties of the MLE. Some failed to define the probability of denial/success, or correctly define the likelihood function.
- Most candidates did well in (b) identifying the advantages and drawbacks of the LPM, particularly the non-sensical interpretation. Some went beyond and discussed the increased computability and discussed the similarity in terms of the results between an LPM and Logit. A good number of candidates missed the fact that a Logit/Probit model's cdf allows for non-constant marginal effects.
- For (c), the key part is to realise that we needed to use an asymptotic t test (z test) for a testing procedure with the LPM estimates. Once this point is correctly identified, the test specifics should be discussed, which was missed by some candidates. Some candidates used the logit estimates instead, but still got marks for a correct procedure. Some commented that due to heteroskedasticity, the LPM estimates cannot be used to do testing, which is only correct when inference is done with usual standard errors. As soon as robust standard errors are used (and they are reported), the inference is valid asymptotically.
- For (d), the most common mistake, which is unfortunate, was to use incorrect estimates to calculate the marginal effects. Candidates used both the methods mentioned in the solution, and both were rewarded. Some failed to suggest the use of the average marginal effect over all individuals as an alternative to looking at the average individual.

Question 7

Let us consider the relationship between the natural logarithm of GDP, GDP_t, and the lagged long-term interest rates $rate_t$, and $rate_{t-1}$.

(a) Assume that GDP_t and $rate_t$ are $I(1)$.

i. If GDP_t and $rate_t$ are not co-integrated, what are the properties of an OLS estimator in the regression of GDP_t on $rate_t$?

(4 marks)

ii. Discuss the implications of the existence of cointegration between GDP_t and $rate_t$ on the short-run and long-run relationship of two variables.

(4 marks)

iii. Consider the following regression model, where ε_t has mean zero and is uncorrelated with GDP_{t-1}, GDP_{t-2}, $rate_t$, and $rate_{t-1}$:

$$GDP_t = \delta + \theta_1 GDP_{t-1} + \theta_2 GDP_{t-2} + \phi_1 rate_t + \phi_2 rate_{t-1} + \varepsilon_t.$$

Derive the Error Correction Model (ECM) representation of the equation, and discuss the long-run equilibrium relation for GDP_t and $rate_t$. Interpret the parameters of the model and discuss the properties of the OLS estimator of the ECM.

(7 marks)

(b) Using the UK data from 1988Q4 to 2021Q1, a researcher estimates regressions with the growth rate of GDP, $growth_t = GDP_t - GDP_{t-1}$, as the dependent variable, and the growth of physical capital $capital_t$, the unemployment rate $unemploy_t$, the growth of money supply $msupply_t$, and the growth of productivity, $prod_t$, as the regressors. Assuming these variables are $I(0)$, the researcher reports three estimation results, one with the full sample, and two with sub-periods split at 2006Q1. Robust standard errors are in parentheses, n is the number of observations, SSR is the sum of squared residuals, $\log L$ is the maximised log-likelihood, R^2 is the R-squared. Describe how to test whether the regression parameters are the same for both periods. Carefully indicate the alternative hypothesis, the test statistic, and rejection rule.

	1988Q4–2021Q1 (1)	1988Q4–2006Q1 (2)	2006Q2–2021Q1 (3)
$capital_t$	0.113 (0.002)	0.033 (0.012)	0.163 (0.052)
$unemploy_t$	−0.005 (0.014)	− 0.08 (0.024)	0.005 (0.013)
$msupply_t$	−0.211 (0.054)	−0.142 (0.084)	−0.335 (0.014)
$prod_t$	0.116 (0.024)	−0.014 (0.004)	0.125 (0.046)
constant	0.003 (0.002)	0.013 (0.005)	0.001 (0.002)
n	130	70	60
SSR	0.288	0.226	0.033
R^2	0.67	0.45	0.47
$\log L$	367.868	179.521	205.485

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 5.4, 11.2, 13.1, 13.2 and 13.6.

Dougherty, C. Subject guide (2016): Chapters 5 and 13.

Approaching the question

(a) i. Candidates are expected to discuss the following points:

* If the two variables are not co-integrated, the difference between GDP_t and $rate_t$ could become very large, with no tendency for them to come back together. Technically, there exists no linear combination of the variables that becomes stationary $I(0)$.

  The regression result, if any, is spurious, and it tells us nothing meaningful about the relationship between the variables.

* The OLS estimator will not be consistent, the R^2 is superficially high. The error term is likely $I(1)$ and the statistical inference becomes invalid. Any significant parameter is meaningless.

ii. Two $I(1)$ processes are co-integrated when there is a long-term relationship between these two.

Co-integration tells us nothing about the short-run dynamics between the two variables, but it indicates that there must be some short-term forces that maintain the long-term relationship.

This implies that it should be possible to construct a more comprehensive model that combines the short-run and long-run dynamics of the relationship. For example, we can conduct an Error Correction Model that describes the relationship.

iii. We derive the ECM by adding and subtracting the lagged term of GDP_t:

$$GDP_t = \delta + \theta_1 GDP_{t-1} + \theta_2 GDP_{t-2} + \phi_1 rate_t + \phi_2 rate_{t-1} + \varepsilon_t$$

$$GDP_t - GDP_{t-1} = \delta + (\theta_1 - 1)GDP_{t-1} + \theta_2 GDP_{t-2} + \phi_1 rate_t + \phi_2 rate_{t-1} + \varepsilon_t$$

$$\begin{aligned} GDP_t - GDP_{t-1} = {} & \delta + (\theta_1 + \theta_2 - 1)GDP_{t-1} + \theta_2(GDP_{t-2} - GDP_{t-1}) \\ & + \phi_1(rate_t - rate_{t-1}) + (\phi_2 + \phi_1)rate_{t-1} + \varepsilon_t. \end{aligned}$$

Simplify the expression, we have the following ECM model:

$$\begin{aligned} \Delta GDP_t = {} & -\theta_2 \Delta GDP_{t-1} + \phi_1 \Delta rate_t \\ & + (\theta_1 + \theta_2 - 1)\left(GDP_{t-1} - \frac{\phi_2 + \phi_1}{1 - \theta_1 - \theta_2} rate_{t-1} - \frac{\delta}{1 - \theta_1 - \theta_2}\right) + \varepsilon_t. \end{aligned}$$

The error correction term, which measures the disequilibrium at time $t-1$, when $\Delta GDP_t = \Delta rate_t = 0$, is:

$$GDP_t - \frac{\phi_2 + \phi_1}{1 - \theta_1 - \theta_2} rate_t - \frac{\delta}{1 - \theta_1 - \theta_2}.$$

The parameter $(\theta_1 + \theta_2 - 1)$ indicates how quickly the variable GDP_t will error correct in the long run: if GDP_{t-1} is larger than the equilibrium value, then $\Delta GDP_t < 0$ to bring GDP_t back to equilibrium.

The parameters $-\theta_2$ and ϕ_1 are the direct effects of a change in the past GDP level and the current interest rate has on the the current GDP level.

The error term ε_t is now a white noise (because of the co-integration), an OLS estimator will be valid and super-consistent.

(b) We will need to use a Chow test for this exercise.

- The null hypothesis is that the parameters are the same for the two regression models, the alternative is otherwise.
- The test statistic is an F statistic, that uses the SSR from the combined data, and from each sub-groups; the number of parameters (including the intercept, which is 5), and the number of observations (130):

$$F = \frac{(0.288 - (0.226 + 0.033))/5}{(0.226 + 0.033)/(130 - 10)} = 2.687.$$

The test statistic follows the F distribution with 5 (number of parameters) and $n - 2k = 120$ degrees of freedom, under the assumption that the error term is i.i.d. with unknown variance. The critical value is 1.896 from the F table.

- Since the F statistic is larger than the critical value, we reject the null hypothesis, in favour of the alternative.
- We conclude that there is evidence that the parameters from the two models are different.

Comments:

- For (a) i., most candidates realised the properties of an OLS estimator with two variables that are not co-integrated. Fewer commented on the fact that in the long run the variables would not tend to each other. Some discussed the properties of the R^2 and got rewarded.
- For (a) ii., candidates did less well in this question but most recognised the existence of the long-term relationship between the two variables and the existence of an ECM.
- For (a) iii., only a few candidates got the correct derivation of the ECM. Many did not interpret the parameters correctly nor intuitively. Only a few candidates realised that OLS becomes valid again due to the presence of white noise error terms.
- Candidates realised the use of the Chow test for (b) and did well there. Most, however, lost marks for the lack of explanations or incorrect test specifics.

Question 8

It is postulated that more open economies have lower inflation rates. A proposed model looks at a cross-section of n countries using time averages:

$$infl_i = \beta_0 + \beta_1 \log(inc_i) + \beta_2 open_i + u_i, \quad i = 1, 2, \ldots, n. \tag{8.1}$$

$infl_i$ is the average inflation rate, $open_i$ is the measure of openness of the economy (average share of imports in GDP), and inc_i is average per capita income, and i denotes a particular country.

(a) The model (8.1) was estimated by OLS and provided $\widehat{\beta}_2 = -0.215$ (with $\text{SE}(\widehat{\beta}_2) = 0.095$). If openness of the economy itself depends on inflation rate ($infl_i$), what sort of problems will this cause for the OLS estimator of (8.1)? Explain your answer.

(4 marks)

(b) Suppose that the equation

$$open_i = \alpha_0 + \alpha_1 infl_i + \alpha_2 \log(inc_i) + \alpha_3 \log(land_i) + v_i, \quad i = 1, 2, \ldots, n \tag{8.2}$$

$land_i$ is the land area of the country, holds in addition to (8.1). Could this relation be helpful in providing a suitable method of estimation of β_2 in (8.1)? What is the relevant estimation method and what conditions must be satisfied for it to be valid?

(5 marks)

(c) Using the cross-section of countries, the following regression was obtained by OLS, with robust standard errors reported in parentheses:

$$\widehat{open}_i = \underset{(15.85)}{117.08} + \underset{(1.493)}{0.546} \log(inc_i) - \underset{(0.81)}{7.57} \log(land_i). \tag{8.3}$$

Discuss whether (and if so how) we can use this equation to verify whether the conditions in (b) are satisfied and how this equation can help us to estimate β_2.

(6 marks)

(d) Some economists claim that inflation is $I(1)$. Below is an estimated time series model for US inflation, with robust standard errors reported in parentheses:

$$\widehat{\Delta infl}_t = \underset{(0.21)}{0.51} - \underset{(0.04)}{0.11}\, infl_{t-1} - \underset{(0.08)}{0.19}\, \Delta infl_{t-1} - \underset{(0.07)}{0.26}\, \Delta infl_{t-2} + \underset{(0.09)}{0.20}\, \Delta infl_{t-3} + \underset{(0.08)}{0.01}\, \Delta infl_{t-4},$$

with $t = 5, \ldots, T$. Conduct an ADF for the presence of a unit root in inflation using a 5% level of significance. The ADF critical value equals -2.86. Clearly indicate the null and the alternative hypotheses, the test statistic and explain the conclusion.

(3 marks)

(e) Briefly indicate whether you think your result in (d) has a bearing on the cross-sectional results in (a)–(c). Explain.

(2 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 9.1, 9.2 and 13.2.

Dougherty, C. Subject guide (2016): Chapter 9 and 13.

Approaching the question

(a) The problem referred to here is 'simultaneity' or reverse causality. It causes a correlation between *open* and u which will make our OLS estimator biased and inconsistent.

(b) This equation shows that the original equation is exact identified (system of simultaneous equations), whereby we can use $\log(land_i)$ as an instrument for $open_i$. The estimation method we will be using to deal with the problem is Instrumental Variables. The conditions required are validity ($\text{Cov}(\log(land), u) = 0$, and $\log(land_i)$ is uncorrelated with the error in the equation) and relevance ($\text{Cov}(\log(land), open) \neq 0$, the instrument $\log(land)$ is correlated with the endogenous regressor *open*. We also have to assume that $\log(land)$ does not directly affect *infl*, which is the exclusion restriction.

(c) The regression shows that the parameter on $\log(land)$ is statistically significant (using a t test) so that the relevance requirement is satisfied. We cannot test the validity of the instrument (here exact identified) as the overidentification test is not applicable. The fitted value of this regression can be used as our instrument for *open* or equivalently be used in our 2SLS formulation of our IV estimator. We can perform OLS on:

$$infl_i = \beta_0 + \beta_1 \log(inc_i) + \beta_2 \widehat{open}_i + error_i.$$

(d) We would test $\text{H}_0 : infl_{-1} = 0$ (unit root, nonstationary, persistent) against $\text{H}_1 : infl_{-1} < 0$ (no unit root, stationary, weak dependent).

We use the statistic:

$$\frac{\widehat{\beta}_{infl_{-1}}}{\text{se}(\widehat{\beta}_{infl_{-1}})}.$$

Its realisation here equals $-0.11/0.04 = -2.75$. Under the null (non-stationarity) its distribution is not standard and has been labelled the Dickey–Fuller distribution. At the 5% level of significance we should reject H_0 if the realisation is less than -2.86. We will therefore not be able to reject the null of the presence of a unit root, finding support of *infl* being $I(1)$.

(e) When we use time series data having $I(1)$ regressors and $I(1)$ dependent variables it may give rise to spurious regression (meaningless). However, here we do not consider a time series regression; instead we run a regression using time averages across countries.

Whereas the dependence that the process $\{infl_t\}$ exhibits over time may be strong for each country, this does not tell us that the averages of $\{infl_t\}$ across countries exhibit strong correlation.

Comments:

- For (a), most candidates realised the simultaneity and discussed the problems (biased and inconsistent) associated with it well.

- For (b), even though most realised the need for an IV to fix the problems identified in (a), several failed to use and explain why log($land$) was an instrument. The discussion on the three IV conditions was not done well in general, mostly due to candidates simply repeating the textbook discussions without any applications to the problem at hand. Some went beyond and derived the reduced forms to discuss the points, which was rewarded.
- Only some candidates realised that the test was for the relevance requirement for (c). But more candidates discussed the test well. Only a few realised that due to exact identification, we cannot test for the IV validity. Some candidates discussed the 2SLS procedure incorrectly.
- Most candidates did well for (d), some lost marks for missing one of the test specifics (for example, the test distribution is the Dickey–Fuller distribution).
- Only a few candidates attempted this question, but those who attempted did very well. Some did not realise the spurious regression issue.

Question 9

Let us consider a dataset containing information on government expenditure on education ($EducE$), gross domestic product (GDP), population (Pop), and capital stock ($Capital$) for 153 countries in 2019. Economic theory suggests that government expenditure on education per capita has a positive effect on per capita GDP. That is,

$$perGDP_i = \beta_1 + \beta_2 perEducE_i + \beta_3 perCapital_i + u_i,$$

where $perGDP_i = \dfrac{GDP_i}{Pop_i}$, $perEducE_i = \dfrac{EducE_i}{Pop_i}$, and $perCapital_i = \dfrac{Capital_i}{Pop_i}$, and i denotes a particular country.

(a) **Explain intuitively why we should be worried about the presence of heteroskedasticity in this model. Briefly indicate three distinct methods to potentially deal with heteroskedasticity in this situation.**

(6 marks)

(b) **Discuss how you would implement the White test for the presence of heteroskedasticity in this case. Briefly discuss how the White test differs from the Goldfeld–Quandt test.**

(5 marks)

(c) **Discuss two reasons why we may expect the OLS estimator for β_2 to be neither unbiased nor consistent. Provide supportive arguments for each.**

(4 marks)

(d) **The dataset also has a dummy variable $Democratic_i$ which equals 1 if country i is democratic, 0 otherwise. Discuss how you could test whether the effect of government expenditure on education per capita is higher for democratic countries. Clearly indicate the regression you would run and the test you would use.**

Comment: **You may assume that the OLS estimator of the regression you would run is consistent, but there may be heteroskedasticity in your model.**

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 4.3, 6.1, 7.2, 7.3 , 8.2 and 8.3.

Dougherty, C. Subject guide (2016): Chapters 4, 7, and 8.

Approaching the question

(a) Candidates were expected to discuss the intuition behind heteroskedasticity in the model, not the mathematical derivation.

We suspect heteroskedasticity because richer countries with a higher absolute GDP per capita, will have a larger budget to distribute and just have greater flexibility of how much they can spend on education. This flexible constraint may result in large variations in the education budget, or the error terms for richer countries and larger education expenses might have a larger variance.

Countries with a lower GDP per capita may have smaller budget options and the education budget is likely to vary less, resulting in smaller variances of their error terms. The relationship, however, is not necessarily linear or proportional.

Candidates can discuss the inefficiency and invalid test statistics due to heteroskedasticity.

There are three popular ways to deal with this situation:

- Use the White heteroskedasticity robust standard errors.
- Use a Generalised Least Squares estimator (GLS or FGLS) to reweight the observations. Since the heteroskedasticity form is unknown, there is a risk of misspecification.
- From economic theory, a logarithm rescaling can be used to reduce the large variances between large and small economies. That is, we can run the following estimation instead:

$$\log(perGDP_i) = \alpha_1 + \alpha_2 \log(perEducE_i) + \alpha_3 \log(perCapital_i) + u_i.$$

(b) Candidates are expected to do the following steps:

- Estimate the original model by OLS, as usual. Obtain the OLS residuals $\widehat{u}_i$. Compute the squared residuals $\widehat{u}_i^2$.
- Run the following regression and keep its R^2, $R^2_{\widehat{u}^2}$:

$$\begin{aligned}\widehat{u}_i^2 = \delta_0 + \delta_1 perEducE_i + \delta_2 perCapital_i + \delta_3 perEducE_i^2 + \delta_4 perCapital_i^2 \\ + \delta_5 perCapital_i \times perEduc_i + error.\end{aligned}$$

- The null hypothesis is $\mathrm{H}_0 : \delta_1 = \cdots = \delta_5 = 0$ and H_1 : otherwise. It's equivalent to a test of $k = 5$ restrictions.
- The test statistic is either the F statistic:

$$F = \frac{R^2_{\widehat{u}^2}/k}{1 - R^2_{\widehat{u}^2}/(n-k-1)}$$

  or the LM statistic:

$$LM = nR^2_{\widehat{u}^2}$$

  with $n = 153$.
- Under the null, the F statistic follows the $F_{k,\,n-k-1}$ distribution or the χ^2_5 distribution. If the statistic exceeds the critical value from the respective distribution, we reject the null hypothesis of homoskedasticity.
- The Goldfeld–Quandt test assumes a monotonic form of heteroskedasticity (the standard deviation of the error distribution is proportional to one of the explanatory variables) and will not work well for non-monotonic forms of heteroskedasticity, which we should expect from the discussion about heteroskedasticity in question (a). The general White test addresses this concern. One potential concern about the White test is that it has low power, especially when there are many explanatory variables as the degrees of freedom will get smaller.

(c) Two potential endogeneity issues:

- Omitted variable bias: there might be an important variable that can explain both $perGDP$ and $perEducE$, for example, some unobserved culture of the country, or the political institution.
- Simultaneity bias: there might also be another structural equation in which $perGDP$ can explain $perEducE$ – richer countries may spend relatively more on education. Candidates can also discuss classical measurement errors in the regressor as another source of endogeneity.

(d) We expect to run the following regression, using the White robust standard error:

$$perGDP_i = \theta_1 + \theta_2 perEducE_i + \theta_3 perCapital_i + \theta_4 Democratic_i$$
$$+ \theta_5 Democratic_i \times perEducE_i + u_i.$$

The differential effect of $perEduc$ on $perGDP$ for democratic countries is captured by the coefficient θ_5.

We can simply do a t test, using the robust standard error, of the null hypothesis that $\mathrm{H}_0 : \theta_5 = 0$ and $\mathrm{H}_1 : \theta_5 > 0$ because we expect $\theta_2 > 0$.

Comments:

- Part (a) was attempted well in general, with candidates correctly explaining why heteroskedasticity would be a problem to our estimator. Many further explained why we would be worried about the presence of heteroskedasticity in our set-up. The remedies of heteroskedasticity were less well done. Most often suggested the use of WLS and GLS, referring to whether the form of heteroskedasticity is known. Even though this answer was correct, the two methods cannot be counted as distinct. Others suggested collecting a different set of data free of heteroskedasticity or with little variations, which would not be a sensible thing to do. The least popular method suggested was to use the logarithms of the variables. Some suggested different ways to detect it, not to solve it, which was not the point of the question. Finally, some only listed the methods without actually discussing them, hence losing some marks.
- Part (b) was also attempted well. Most lost marks due to incorrect test specifics or incomplete steps. The most common mistake was to omit the cross-products of the variables. Others did not distinguish well the difference between the White test and the Goldfeld–Quandt test, often referring to their different test equations as the key distinction; while we should focus on the assumptions behind the form of heteroskedasticity or the power of the tests.
- Part (c) posed some challenges. Several candidates suggested multicollinearity or heteroskedasticity as a cause for biasedness or inconsistency, which was not correct. Others, despite correctly identifying the OVB, reverse causality, or measurement errors, did not provide any explanations or arguments. Some went beyond and discussed the three mentioned problems and discussed them very well.
- Part (d) required candidates to think about the use of the interaction and its interpretation. Many realised this and did well in describing the new regression with an interaction. Some lost marks for failing to use the robust standard errors for the test or lacking test specifics of the z test. Some did not realise that we should conduct a one-tailed test and θ_2 should be positive. Others proposed the use of F tests or Chow tests, without supportive arguments. Some proposed to split the data into two samples of democratic and non-democratic and then use the Chow test. This approach will help test the equivalence of the parameters, but not necessarily whether democratic countries would benefit more from education expenses.

Examiners' commentaries 2021

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2020–21. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone B

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from **Section A** (8 marks each) and **THREE** from **Section B** (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Section A

Answer all questions from this section.

Question 1

Consider the following regression model

$$Y_i = \beta_0 + \beta_1 \log(X_i) + u_i, \quad i = 1, \ldots, n.$$

The error term is i.i.d. with zero mean and variance equal to σ^2 and the error and regressors are independent. We assume $X_i > 0$ for all i.

(a) **Suppose you are given a random sample of observations $\{(Y_i, X_i)\}_{i=1}^n$. Derive the OLS estimator of β_1 and provide an interpretation of this parameter.**

(4 marks)

(b) **Unfortunately you do not observe X_i, instead you observe X_i^* which equals**

$$X_i^* = X_i \exp(\varepsilon_i),$$

where ε_i has zero mean and is independent of u_i and X_i (and therefore $\log(X_i)$). You are given a random sample of observations $\{(Y_i, X_i^*)\}_{i=1}^n$.

Formulate the estimable regression model and show that the OLS estimator of β_1 is inconsistent, resulting in an underestimate in absolute size of the coefficient on average.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 3.3, 4.2, 8.2 and 8.4.

Dougherty, C. Subject guide (2016): Chapters 4 and 8.

Approaching the question

(a) To facilitate our discussion, let $Z_i = \log(X_i)$. To obtain the OLS estimators we should solve:

$$\min_{b_0,b_1} \sum_{i=1}^{n} (Y_i - b_0 - b_1 Z_i)^2.$$

This requires us to solve the first-order conditions:

$$-2\sum_{i=1}^{n}(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 Z_i) = 0$$

$$-2\sum_{i=1}^{n}(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 Z_i)Z_i = 0.$$

Solving these equation yields:

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1 \bar{Z} = \bar{Y} - \widehat{\beta}_1 \overline{\log(X)}$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n}(Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(Z_i - \bar{Z})^2} = \frac{\sum_{i=1}^{n}(\log(X_i) - \overline{\log(X)})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(\log(X_i) - \overline{\log(X)})^2}.$$

Observe:

$$\beta_1 = \frac{\Delta Y}{\Delta \log X} = \frac{\Delta Y}{\%\Delta X}$$

for a small change in X. Therefore, if you change X by 1%, we expect Y to change by $\beta_1/100$ units.

(b) First candidates should recognise that:

$$\log X_i^* = \log X_i + \varepsilon_i.$$

Plugging in $\log X_i = \log X_i^* - \varepsilon_i$, they should then specify:

$$Y_i = \beta_0 + \beta_1(\log X_i^* - \varepsilon_i) + u_i$$

or:

$$Y_i = \beta_0 + \beta_1 \log X_i^* + v_i$$

where $v_i = u_i - \beta_1 \varepsilon_i$. For the consistency of the estimator, notice that:

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^{n}(Z_i^* - \bar{Z}^*)(Y_i - \bar{Y})}{\sum_{i=1}^{n}(Z_i^* - \bar{Z}^*)^2} = \beta_1 + \frac{\sum_{i=1}^{n}(Z_i^* - \bar{Z}^*)(v_i - \bar{v})}{\sum_{i=1}^{n}(Z_i^* - \bar{Z}^*)^2}.$$

Using properties of the plim operator and law of large numbers we obtain:

$$\text{plim}\ \widehat{\beta}_1 = \beta_1 + \frac{\text{plim}\ \frac{1}{n}\sum_{i=1}^{n}(Z_i^* - \bar{Z}^*)(v_i - \bar{v})}{\text{plim}\ \frac{1}{n}\sum_{i=1}^{n}(Z_i^* - \bar{Z}^*)^2} = \beta_1 + \frac{\text{Cov}(Z_i^*, v_i)}{\text{Var}(Z_i^*)}.$$

Showing $\text{Var}(Z_i^*) = \text{Var}(Z_i) + \text{Var}(\varepsilon_i)$ (by independence of $Z_i = \log(X_i)$ and ε_i) and:

$$\text{Cov}(Z_i^*, v_i) = \text{Cov}(Z_i + \varepsilon_i, u_i - \beta_1\varepsilon_i) = -\beta_1\text{Var}(\varepsilon_i)$$

(by independence of ε_i with $\log(X_i)$ and u_i) yields the expected result:

$$\text{plim}\ \widehat{\beta}_1 = \beta_1 - \beta_1\frac{\text{Var}(\varepsilon_i)}{\text{Var}(Z_i) + \text{Var}(\varepsilon_i)} = \beta_1\frac{\text{Var}(Z_i)}{\text{Var}(Z_i) + \text{Var}(\varepsilon_i)}.$$

Since:

$$0 < \frac{\text{Var}(Z_i)}{\text{Var}(Z_i) + \text{Var}(\varepsilon_i)} < 1$$

the estimator exhibits attenuation bias where the estimated effect is smaller in magnitude than it should be.

Comments:

- For (a), candidates often lost marks for skipping the interpretation part of the question. Some mistook the average of $\log(X)$ with the log of $\bar{X}$. Some candidates interpreted the coefficient as the semi-elasticity of Y with respect to X, which was an alternative. Others lost marks for incorrect derivations of the estimator.
- For (b), most candidates lost marks during the derivation of inconsistency: some used incorrect annotations, some missed the LLN or mistook it as a plim property. Some excellent scripts discussed the attenuation bias and the direction.

Question 2

Consider the regression model

$$\boldsymbol{y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u, \quad \text{with } E(u) = 0.}$$

You may assume that the regressors are non-stochastic. Suppose the error has a normal distribution and exhibits heteroskedasticity, i.e. $\text{Var}(u_i)$ varies with x_{1i} and/or x_{2i}, for all i. You are provided with a random sample $\{(y_i, x_{1i}, x_{2i})\}_{i=1}^n$.

(a) **Discuss how, with the help of heteroskedasticity robust standard errors, you would test the null hypothesis $H_0^{(a)} : \beta_1 = 1$ against the one-sided alternative $H_1^{(a)} : \beta_1 > 1$ when you use OLS to estimate β_1. In your answer explain the need for robust standard errors.**

(4 marks)

(b) **Suppose you are told that $\text{Var}(u_i) = c^2 x_{1i}^2$ where c is an unknown constant and $x_{1i} \neq 0$ for all i. Discuss how this knowledge about the form of heteroskedasticity allows you to (i) obtain a more efficient estimator of β_1 and (ii) hence conduct a test of the null hypothesis $H_0 : \beta_1 = 0$ that is more powerful.**

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters R.11, 7.1 and 7.3.

Dougherty, C. Subject guide (2016): Chapter 7.

Approaching the question

(a) We would use the test statistic:

$$\frac{\widehat{\beta}_1 - 1}{\text{se}(\widehat{\beta}_1)}.$$

As there is a presence of heteroskedaticity, we cannot use the usual standard errors as they would invalidate the test. Instead we need to apply the White heteroskedasticity robust standard errors. We will want to reject H_0 when the realisation of our test statistic exceeds the critical value at a given level of significance (one-sided). This is an **asymptotic test** (we do not need to assume normality of the errors) and we can use the critical values of the $N(0,1)$ distribution. (Candidates may call it an asymptotic t test, but it is not a finite sample test!)

(b) Now we will apply WLS (as it will give us the BLUE for our parameters). Specifically we use:

$$\frac{y}{x_1} = \beta_0 \frac{1}{x_1} + \beta_1 + \beta_2 \frac{x_2}{x_1} + \frac{u}{x_1}$$

or:

$$y^* = \beta_0 x_1^* + \beta_1 + \beta_2 x_2^* + v$$

with $y^* = y/x_1$, $x_1^* = 1/x_1$ and $v = u/x_1$.

Clearly now our error satisfies the Gauss–Markov assumption (in particular $\text{Var}(v) = \text{Var}(u)/x_1^2 = c^2$ (a constant, hence homoskedastic)).

We can estimate this model by OLS and its standard errors are valid. The test statistic:

$$\frac{\widehat{\beta}_1 - 0}{\text{se}(\widehat{\beta}_1)}$$

now has the t_{n-3} degrees of freedom, and we should reject if the realisation of:

$$\frac{\widehat{\beta}_1 - 0}{\text{se}(\widehat{\beta}_1)} > t_{n-3,\,\alpha}.$$

Candidates could formulate a two-tailed test in which a critical value of $t_{n-3,\,\alpha/2}$ is used.

A test that uses a more efficient estimator permits us to reject H_0 more easily, which will ensure:

$$\frac{\widehat{\beta}_1 - 0}{\text{se}(\widehat{\beta}_1)}$$

is larger, as this also ensures that we reject H_0 earlier when the null is false so this ensures we have a more powerful test.

Comments:

- For (a), candidates often lost marks for insufficient discussions of the need to use a robust standard error, or for forgetting the use of asymptotic tests under heteroskedasticity. Again, many lost marks for missing some test specifics or incorrect rejection rule or test distribution.
- For (b), most realised the need to use WLS and the WLS equation. Candidates often lost marks for not commenting on the efficiency loss or the fact that WLS is BLUE. Some again lost marks for the incorrect testing procedure.

Question 3

A researcher tries two regression models to describe the relationship between a variable $X > 0$ and a variable $Y > 0$:

$$\log Y = \alpha_1 + \alpha_2 \log X + u, \tag{3.1}$$

$$\log(XY) = \beta_1 + \beta_2 \log X + v. \tag{3.2}$$

(a) Explain whether model (3.2) is a restricted version of (3.1) or not.

(2 marks)

(b) Using a random sample of size n, the researcher obtains $(\widehat{\alpha}_1, \widehat{\alpha}_2, \widehat{\beta}_1, \widehat{\beta}_2)$ as the OLS estimators for $(\alpha_1, \alpha_2, \beta_1, \beta_2)$. Let $(\widehat{u}, \widehat{v})$ be the residuals of the OLS estimations. Prove that $\widehat{u} = \widehat{v}$.

(2 marks)

(c) Assume that the regression equation (3.1) describes the true relationship between X and Y and there exists a variable W uncorrelated with Y. What are the properties of the OLS estimator for γ_2 if we use the following regression specification?

$$\log Y = \gamma_1 + \gamma_2 \log X + \gamma_3 W + w. \tag{3.3}$$

Explain your answer intuitively.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 1.4, 1.5, 4.2, 6.3.

Dougherty, C. Subject guide (2016): Chapters 1, 4 and 6.

Approaching the question

(a) Model (3.2) can be rewritten as:

$$\log Y = \beta_1 + (\beta_2 - 1) \log X + v.$$

This is simply a reparameterised version of model (3.1) with $\beta_1 = \alpha_1$ and $\alpha_2 = \beta_2 - 1$. It is not a restricted version because it does not impose any restriction on the values of α_1 or α_2. Candidates may interpret the restrictions imposed in the sense that $\alpha_2 = \beta_2 - 1$ and $\beta_1 = \alpha_1$.

(b) We first need to show that $\widehat{\beta}_1 = \widehat{\alpha}_1$ and $\widehat{\alpha}_2 = \widehat{\beta}_2 - 1$. Let Z $= XY$, and let $z = \log(Z)$, $x = \log(X)$ and $y = \log(Y)$. We have $z = y + x$ and $\bar{z} = \bar{y} + \bar{x}$.

The OLS estimator for $\widehat{\beta}_2$ is:

$$\begin{aligned}
\widehat{\beta}_2 = \frac{\sum(x_i - \bar{x})(z_i - \bar{z})}{\sum(x_i - \bar{x})^2} &= \frac{\sum(x_i - \bar{x})[y_i + x_i - (\bar{y} + \bar{x})]}{\sum(x_i - \bar{x})^2} \\
&= \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sum(x_i - \bar{x})^2} + \frac{\sum(x_i - \bar{x})^2}{\sum(x_i - \bar{x})^2} \\
&= \widehat{\alpha}_2 + 1.
\end{aligned}$$

To express $\widehat{\alpha}_1$ in terms of $\widehat{\beta}_1$, notice that:

$$\begin{aligned}
\widehat{\beta}_1 = \bar{z} - \widehat{\beta}_2 \bar{x} &= (\bar{y} + \bar{x}) - (\widehat{\alpha}_2 + 1)\bar{x} \\
&= \bar{y} + \bar{x} - \widehat{\alpha}_2 \bar{x} - \bar{x} \\
&= \bar{y} - \widehat{\alpha}_2 \bar{x} \\
&= \widehat{\alpha}_1.
\end{aligned}$$

Now, we can rewrite $\widehat{v}$ with the new expressions as:

$$\begin{aligned}\widehat{v} = \widehat{z} - z &= \widehat{\beta}_1 + \widehat{\beta}_2 x - (y - x) \\ &= \widehat{\beta}_1 + (\widehat{\beta}_2 + 1)x - y \\ &= \widehat{\alpha}_1 + \widehat{\alpha}_2 x - y \\ &= \widehat{y} - y \\ &= \widehat{u}.\end{aligned}$$

Alternatively, candidates can explain that since (3.1) is just a reparameterised version of (3.2) (and vice versa), the residuals are equal.

(c) Because W does not appear in the true regression model, the estimator for its coefficient is unbiased and has an expectation of zero, when added to the regression. The estimator for γ_2 remains unbiased and the standard errors remain valid.

If W is correlated with X, there will be a loss of efficiency as the variance of $\widehat{\gamma}_2$ is larger as the model now includes a noisy part unrelated to the outcome variable (mathematically, the variance of $\widehat{\gamma}_2$ will be inflated by a factor of $1/(1 - \rho^2_{W,\log X})$, where $\rho^2_{W,\log X}$ is the correlation between W and $\log X$.

If X and W are uncorrelated, there will be no loss of efficiency.

Comments:

- For (a), candidates can simply recognise that the model is a reparameterised version of (3.2). Marks were awarded for answers discussing the reparameterisation and the restrictions of the two models' parameters.
- Part (b) followed directly from (a), in which candidates explained the equivalence of the parameters after the reparameterisation. Some lost marks due to a lack of explanation. Some arrived at $\sum \widehat{u} = \sum \widehat{v}$, and claimed to have proven the question. Unfortunately, this is not the case.
- Part (c) required more intuition to think about the situation when we included uncorrelated covariates. Most realised that the estimator remained unbiased and the statistical inference remained valid. Fewer realised that there would be a loss of efficiency. Only a very few realised that the final point is that when X and W are uncorrelated, there will be no loss of efficiency. Candidates were not penalised for not recognising the latter.

Question 4

(a) What are the conditions for a series to be covariance stationary? Consider the following time series model for $\{x_t\}_{t=1}^T$:

$$x_t = x_{t-1} + \lambda^2 \varepsilon_t + \lambda \varepsilon_{t-1} + \varepsilon_{t-2},$$

where ε_t is i.i.d. with mean zero and variance σ^2, for $t = 1, \ldots, T$, and $x_0 = 0$. Discuss whether there are any values of λ so that x_t is covariance stationary.

(4 marks)

(b) Consider the following time series model for $\{y_t\}_{t=1}^T$:

$$y_t = \theta_0 + \theta_1(\rho - 1)t + \rho y_{t-1} + \varepsilon_t,$$

where ε_t is i.i.d. with mean zero and variance σ^2, $t = 1, \ldots, T$, and $y_0 = 0$. How can we test whether $\rho = 1$? Clearly specify the null and alternative hypothesis, the test statistic, and the rejection rule.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 11.4, 13.1, and 13.5.

Dougherty, C. Subject guide (2016): Chapters 11 and 13.

Approaching the question

(a) For a process x_t to be stationary, its expected value, variance, and covariance of x_t and x_{t+s} are independent of t. We need to find λ so that all the terms are independent of time.

Lagging one period and then substituting back to the model, until time $t = 2$, we have:

$$x_t = x_1 + \lambda^2\varepsilon_t + (\lambda + \lambda^2)\varepsilon_{t-1} + (1 + \lambda + \lambda^2)(\varepsilon_{t-2} + \cdots + \varepsilon_2) + (\lambda + 1)\varepsilon_1 + \varepsilon_0.$$

The variance of x_t is:

$$\mathrm{Var}(x_t) = \sigma^2[\lambda^4 + (\lambda + 1)^2\lambda^2 + (\lambda^2 + \lambda + 1)^2(t - 3) + 1].$$

For the variance to be independent of t, we must have $\lambda^2 + \lambda + 1 = 0$. There is no such λ because $\lambda^2 + \lambda + 1 = (\lambda + 1/2)^2 + 3/4 > 0$ for all λ.

Candidates can mention that the expectation of x_t is simply x_0 and independent of time regardless of λ (because $\mathrm{E}(\varepsilon_t) = 0$).

(b) We must first rewrite the regression model for a unit root test with drift and deterministic time trend:

$$\Delta y_t = \theta_1 + \theta(\rho - 1)t + (\rho - 1)y_{t-1} + \varepsilon_t.$$

We run the following equivalent regression:

$$\Delta y_t = \theta_1 + \beta_1 t + \beta_2 y_{t-1} + \varepsilon_t.$$

The null hypothesis that $\rho = 1$ is equivalent to a joint test of the coefficients on the time trend and y_{t-1} being zero. $\mathrm{H}_0 : \rho = 1$ (or $\beta_2 = 0$). The alternative hypothesis is $\mathrm{H}_1 : \rho < 1$ (or $\beta_2 < 0$). The test statistic:

$$\frac{\widehat{\beta}_2}{\mathrm{se}(\widehat{\beta}_2)}.$$

Under the usual assumptions, the test statistic follows a Dickey–Fuller distribution reported in the Dickey–Fuller table. We will reject the null if:

$$\frac{\widehat{\beta}_2}{\mathrm{se}(\widehat{\beta}_2)} < \tau$$

where τ is the Dickey–Fuller critical value (with trend and intercept). Note that we cannot use a t distribution because under the null the distribution is not Student's t distribution.

Comments:

- For (a), most candidates could discuss the conditions for stationarity. The key challenge here was to see if there would be any value of λ that satisfies the condition. Several candidates offered a good attempt at deriving the variance of x_t, some made some algebraic mistakes but most who attempted arrived at the same conclusion, which was rewarded accordingly.
- Part (b) was a more standard question, which tested candidates' knowledge of the Dickey–Fuller test. Most did well in this part. Some lost marks for not recognising the use of a Dickey–Fuller test, others lost marks for not taking the difference of y_t as part of the test. A few used the t distribution instead of the Dickey–Fuller distribution, which was incorrect.

Question 5

Consider the following regression model

$$y_t = \rho y_{t-1} + \gamma_1 z_t + \gamma_2 z_{t-1} + \gamma_3 z_{t-2} + \gamma_4 z_{t-3} + \varepsilon_t \quad \text{with } |\rho| < 1,$$

where ε_t is an MA(1) process, i.e.:

$$\varepsilon_t = u_t + \theta u_{t-1}.$$

u_t is a white noise process generated randomly from a normal distribution with mean zero and constant variance that is independent of $y_{t-1}, y_{t-2}, \ldots, z_t, z_{t-1}, \ldots$. You have a random sample and can use T observations, which we label $t = 1, 2, \ldots, T$ for convenience, for the estimation. Assume there is no perfect multicollinearity.

(a) Briefly discuss the finite-sample and large-sample properties of the OLS estimator for ρ from the regression model where $\theta \neq 0$.

(4 marks)

(b) Let $\gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = 0$, so that

$$y_t = \rho y_{t-1} + \varepsilon_t \quad \text{with } |\rho| < 1.$$

Discuss how you can obtain a consistent estimator for ρ that recognises $\rho \neq 0$. Show that this estimator is consistent. Clearly annotate your proof.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 11.4, 11.5, 11.6.

Dougherty, C. Subject guide (2016): Chapter 11.

Approaching the question

(a) Because y_{t-1} and ε_t are correlated (candidates needed to show this), we have the problem of endogeneity.

The OLS estimator will be both biased and inconsistent.

We have the problem of serial correlation. Since the error term ε_t is an MA(1) process, the autocorrelation function (ACF) for lag 1 is non-zero ($\rho_1 = \theta/(1+\theta^2)$), or observations of the error terms that are one period apart will be correlated (details of this discussion are not needed).

The OLS estimator is inefficient and the OLS standard errors are wrong.

(b) To obtain a consistent estimator for ρ, we need to use an instrument for y_{t-1}. The instrument should be relevant (correlated with y_{t-1}), valid (uncorrelated with the error terms), and should not appear itself in the regression equation.

One possible and popular choice is the lagged y_{t-2}, which in general satisfies the conditions. If we choose y_{t-2} as the instrument, the IV estimator has the following form:

$$\widehat{\rho}_{IV} = \frac{\sum_{t=1}^{T} y_{t-2} y_t}{\sum_{t=1}^{T} y_{t-2} y_{t-1}} = \frac{\text{Sample Cov}(y_{t-2}, y_t)}{\text{Sample Cov}(y_{t-2}, y_{t-1})}.$$

Under the assumptions, we want to show plim $\rho_{IV} = \rho$:

$$\begin{aligned}
\text{plim}(\widehat{\rho}_{IV}) &= \text{plim}\left(\frac{\sum_{t=1}^{T}(\rho y_{t-1} + \varepsilon_t)y_{t-2}}{\sum_{t=1}^{T} y_{t-1}y_{t-2}}\right) \\
&= \rho + \frac{\text{plim}\frac{1}{T}\sum_{t=1}^{T}\varepsilon_t y_{t-2}}{\text{plim}\frac{1}{T}\sum_{t=1}^{T} y_{t-1}y_{t-2}} \\
&= \rho + \frac{\text{Cov}(\varepsilon_t, y_{t-2})}{\text{Cov}(y_{t-1}, y_{t-2})} \\
&= \rho.
\end{aligned}$$

The first equality plugs in the true model, the second equality follows the properties of plim and Slutsky theorem. The third equality uses law of large numbers (LLN) and the final equality uses the assumptions indicated above to make ε_t and the instrument uncorrelated.

Note that, we should not have the averages in the IV estimator because there is no intercept in the regression model.

Comments:

- Most candidates skipped part (b) or did not find the correct approach to the question. It was a surprise to the examiners because part (a) set candidates up for the task in part (b) – how to solve the endogeneity problem in a model with only the lagged dependent variable. Some realised the need to modify the OLS estimator to consistently estimate ρ, however, most ended up proving incorrectly that an OLS estimator will just be consistent, which contradicted the conclusion in (a). A very few candidates realised the use of an IV estimator and discussed it well. Some lagged the dependent variables even further to resolve the autocorrelation problem but did not discuss why this new estimator will be consistent.
- For part (a), most candidates did well recognising the biasedness and inconsistency. Some forgot the autocorrelation problem and did not mention that the OLS standard errors will be invalid, which is of secondary concern to the inconsistency problem.

Section B

Answer three questions from this section

Question 6

Consider the following model to determine the effectiveness of vaccination on the containment of measles, a dangerously contagious but vaccine-preventable disease, in a developing country:

$$\boldsymbol{measles_i = \alpha_0 + \alpha_1 \log(vaccinated)_i + \alpha_2 children_i + \beta_3 avginc_i + \beta_4 city_i + u_i}, \quad \textbf{(6.1)}$$

where the variance of the error term u_i is assumed constant and equal to σ_u^2, and i denotes a particular district; $measles_i$ is the number of reported measles cases per 10,000 residents, $vaccinated_i$ is the reported number of vaccinated residents per 10,000 residents, $children_i$ is the number of children per 10,000 residents, $avginc_i$ is the average level of income, and $city_i$ equals to 1 if the district is in a city, 0 otherwise. You have a random sample of 450 districts.

(a) What is the interpretation of the coefficient α_1? What is the expected sign of α_1?

(2 marks)

Epidemiological studies suggest that areas with high rates of infection are more heavily targeted with vaccination programmes, hence these areas will have more people vaccinated. The following model describes this relationship:

$$\log(vaccinated)_i = \gamma_0 + \gamma_1 measles_i + \gamma_3 vaccentre_i + \gamma_4 com_i + \gamma_5 city_i + v_i, \qquad (6.2)$$

where the variance of the error term v_i is assumed constant and equal to σ_v^2, $vaccentre_i$ is the number of vaccination centres available per 10,000 residents in district i, com_i is a variable that measures the communication and media environment regarding disease containment issues. Assume that *children*, *avginc*, *city*, *vaccentre*; and *com* are exogenous.

(b) Identify and discuss <u>two</u> likely problems with estimating α_1 by OLS. Clearly indicate the direction of bias, if any.

(7 marks)

(c) In light of your answers in (b), describe an appropriate estimator for α_1. Briefly discuss and justify the necessary assumptions, if any.

(7 marks)

(d) Another researcher argues that there should be no statistical difference between the estimates for α_1 obtained from OLS and from the estimator in (c). Describe a test that might help the researcher statistically justify the argument. Clearly indicate the null, the alternative, test statistic, test distribution, and rejection rule. Are there any limitations of the test?

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 8.1, 8.3, 8.4, and 8.5.

Dougherty, C. Subject guide (2016): Chapter 8.

Approaching the question

(a) Everything else equal (*ceteris paribus*), a 1% increase in the number of vaccinated people per 10,000 residents will on average lead to $\alpha_1/100$ change in the number of measles per 10,000 residents in the district. We should expect that more vaccinated people should reduce the rate of measles transmissions, so we expect $\alpha_1 < 0$.

(b) We are concerned about at least two endogeneity issues.

The most likely concern is simultaneity bias, in which the number of measles cases and vaccinated residents are jointly determined: while more vaccinated people will help slow down the increase of measles incidents, rational residents will take more vaccines when they see a higher number of measles cases in the area.

- If we substitute Model 1 into Model 2, we will see that $\log(vaccinated_i)$ depends on $\gamma_1 u$. Because we expect $\gamma_1 > 0$, if the structural error (v) in Model 2 is uncorrelated with u, then $\text{Cov}(\log(vaccinated), u) = \gamma_1 \text{Var}(u) > 0$ or $\log(vaccinated_i)$ is positively correlated with the error term u, giving rise to simultaneity bias in Model 1 if estimated by OLS.
- Thus the OLS estimator for α_1 will be upwards biased, towards zero since $\alpha_1 < 0$. That is, we are likely to underestimate the effectiveness of vaccination in reducing measles cases. Candidates could also write down the reduced form to derive the formula of the bias.

Of secondary importance, we might be worried about measurement error issues. The number of vaccinated residents could be mismeasured randomly, causing attenuation bias to the OLS estimator of α_1. Classical measurement error (or random noise) introduced to the number of reported measles cases does not cause bias, but will inflate the standard error of interest. Candidates can also discuss a selection bias or OVB here.

(c) In light of the endogeneity issues and the specification of Model 2, we can use instruments to consistently estimate the parameter α_1 in Model 1.

There are two viable candidates for the instruments from Model 2, $vaccentre_i$ and com_i. Note that, $city_i$ is not a valid instrument.

For these to be valid instruments, we must assume that the two variables do not appear, in addition to $\log(vaccinated_i)$ in Model 1. It is a reasonable assumption as the number of vaccination centers and communication effectiveness should not directly affect the number of measles cases, except through the vaccination programmes.

But more importantly, we must assume that these variables are not correlated with any other factors that can affect $measles_i$ unobserved in u.

Finally, we must assume that the two instrumental variables have some partial effect on $\log(vaccinated_i)$, which is likely to hold.

With the availability of these two instruments, Model 1 is over-identified and we must use a 2SLS procedure to estimate α_1. The steps are as follows:

First, we estimate the first stage by regressing $\log(vaccinated_i)$ on all exogenous variables, and then get the fitted value $\log(\widehat{vaccinated_i})$:

$$\log(\widehat{vaccinated_i}) = \widehat{\pi}_0 + \widehat{\pi}_1 vaccentre_i + \widehat{\pi}_2 com_i + \widehat{\pi}_3 city_i + \widehat{\pi}_4 children_i + \widehat{\pi}_5 avginc_i.$$

Second, we use $\log(\widehat{vaccinated_i})$ instead of $\log(vaccinated_i)$ in Model 1 and run OLS to consistently estimate the α_1 parameter, i.e.:

$$measles_i = \alpha_0 + \alpha_{12SLS}\log(\widehat{vaccinated_i}) + \alpha_2 children_i + \beta_3 avginc_i + \beta_4 city_i + w_i.$$

(d) The appropriate test is the Durbin–Wu–Hausman test or Hausman test.

The null hypothesis is that the estimated coefficients are distributed independently of the error term u and the difference in the OLS and 2SLS coefficients are random, not systematic. The alternative is the difference between the two sets of coefficients is systematic. Under the null hypothesis, the 2SLS is consistent; under the alternative, the OLS estimate, but efficient and unbiased under the null.

Under the null, the test statistic is distributed as a χ^2 statistic with degrees of freedom equal to the number of coefficients being compared (2 in this case). If the test statistic exceeds its critical value, given the significance level, we can reject the null hypothesis and support the alternative that we should use the IV estimator rather than the OLS.

The main limitation of the test is that if the instrument has a weak predictive power, the power of the Hausman test will suffer.

Comments:

- For part (a), most candidates got the correct prediction for the sign, but some interpreted the coefficient incorrectly, either missing the unit or a wrong approach to a linear-log regression.
- For part (b), surprisingly, a majority of the candidates did not get two endogeneity issues. Most recognised the reverse causality (simultaneity) but did not provide a good discussion of why we should be concerned about simultaneity in this model. Some discussed the OVB and the measurement error concern, which was also rewarded. However, neither of these concerns was of primary importance, our focus should be on simultaneity. Worryingly, a

good number of candidates suggested incorrectly multicollinearity or heteroskedasticity as the sources for endogeneity. Some candidates derived the reduced forms to discuss the problems and were well-rewarded.

- For part (c), most who correctly identified the endogeneity in (b) realised the need to use an IV for this question. Most marks were lost due to having an incorrect 2SLS procedure, or unclear conditions for a valid IV. Some did not specify the instruments to be used and lost marks. Some suggested the use of a 'proxy variable' or use some 'omitted variables' to resolve the endogeneity problem. Others, following their incorrect answers in (b), suggested using robust standard errors or using a new dataset. These received few marks.
- A good number of candidates did not recognise the need to use the Wu–Hausman test and discussed some other irrelevant testing procedures (Chow test, or t test). Some identified the Hausman test but employed a t test as the procedure. Only a few discussed the limitations of the test.

Question 7

Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the exponential distribution with parameter θ. The probability density function is specified as

$$f(x) = \frac{1}{\theta}\,\mathrm{e}^{-x/\theta}, \quad x \geq 0 \text{ and } \theta > 0.$$

Note that $\mathrm{E}(X_i) = \theta$ and $\mathrm{Var}(X_i) = \theta^2$.

In this question we will consider various estimators of θ.

(a) Let us first consider a sequence of estimators given by

$$\widehat{\theta}_c = c\bar{X},$$

where c is a positive constant and $\bar{X} = n^{-1}\sum_{i=1}^{n} X_i$.

For example, $\widehat{\theta}_1 = \bar{X}$ and $\widehat{\theta}_{0.5} = 0.5\bar{X}$.

i. For any given c, determine the bias, variance, and mean squared error (MSE) of the estimator $\widehat{\theta}_c$.
Note that $\mathrm{MSE}(\widehat{\theta}) = \mathrm{E}((\widehat{\theta} - \theta)^2) = \mathrm{Bias}^2(\widehat{\theta}) + \mathrm{Var}(\widehat{\theta})$.

(6 marks)

ii. Show that $c^* = \dfrac{n}{n+1}$ minimises the MSE of $\widehat{\theta}_c$.
***Hint:* You will need to look at the derivative of $\mathrm{MSE}(\widehat{\theta})$ with respect to c.**

(3 marks)

iii. Is the estimator $\widehat{\theta}_{c^*} = (n+1)^{-1}\sum_{i=1}^{n} X_i$ unbiased? Consistent? Prove your answers.

(5 marks)

(b) Derive the MLE estimator for θ, $\widehat{\theta}_{MLE}$. Briefly discuss what desirable properties this estimator has.

(6 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters R14, R15, and 10.6.

Dougherty, C. Subject guide (2016): Chapter 10.

Approaching the question

(a) Candidates are expected to derive the following terms:

i. **Mean**: Using the fact that c is a constant:

$$\mathrm{E}(\widehat{\theta}_c) = \mathrm{E}(c\bar{X}) = c\,\mathrm{E}(\bar{X}) = c\theta$$

as:

$$\mathrm{E}(\bar{X}) = n^{-1}\sum_{i=1}^{n}\mathrm{E}(X_i) = \theta.$$

So the **bias** is:

$$c\theta - \theta = (c-1)\theta.$$

Variance: Using the fact that c is a constant:

$$\mathrm{Var}(\widehat{\theta}_c) = \mathrm{Var}(c\bar{X}) = c^2\mathrm{Var}(\bar{X}) = \frac{c^2\theta^2}{n}$$

as:

$$\mathrm{Var}(\bar{X}) = n^{-2}\mathrm{Var}\left(\sum_{i=1}^{n}X_i\right) = n^{-2}\sum_{i=1}^{n}\mathrm{Var}(X_i)$$

by independence and $\mathrm{Var}(X_i) = \theta^2$. Finally, we can find the **MSE** as follows:

$$\begin{aligned}\mathrm{MSE}(\bar{X}) &= \mathrm{Var}(\bar{X}) + (\mathrm{E}(\widehat{\theta}_c) - \theta)^2\\ &= \frac{c^2\theta^2}{n} + ((c-1)\theta)^2\\ &= \frac{c^2\theta^2}{n} + (c-1)^2\theta^2.\end{aligned}$$

ii. We should look at the value of c that will set the first derivative of $\mathrm{MSE}(\bar{X})$ w.r.t. c equal to zero:

$$\frac{\mathrm{d}(c^2\theta^2/n + (c-1)^2\theta^2)}{\mathrm{d}c} = \frac{2c\theta^2}{n} + 2(c-1)\theta^2$$

$$\frac{2c^*\theta^2}{n} + 2(c^*-1)\theta^2 = 0$$

$$2\theta^2\left(\frac{c^*}{n} + c^* - 1\right) = 0$$

$$c^*\left(\frac{1}{n} + 1\right) = 1 \quad\Rightarrow\quad c^*\left(\frac{n+1}{n}\right) = 1$$

hence $c^* = n/(n+1)$.

The second-order condition is obviously positive, hence c^* minimises the MSE.

iii. We want to look at the property of:

$$\widehat{\theta}_{c^*} = \frac{n}{n+1}\bar{X} = (n+1)^{-1}\sum_{i=1}^{n}X_i.$$

The estimator is biased:

$$\mathrm{E}(\widehat{\theta}_{c^*}) = (n+1)^{-1}\sum_{i=1}^{n}\mathrm{E}(X_i) = \frac{n}{n+1}\theta \neq \theta.$$

Candidates can use sufficient conditions to prove consistency. That is:

$$\lim_{n\to\infty}\mathrm{E}(\widehat{\theta}_{c^*}) = \theta$$

and:

$$\lim_{n\to\infty} \operatorname{Var}(\widehat{\theta}_{c^*}) = \lim_{n\to\infty} \left(\left(\frac{n}{n+1} \right)^2 \frac{\theta^2}{n} \right) = \lim_{n\to\infty} \left(\frac{n}{(n+1)^2} \theta^2 \right) = 0$$

or alternatively show:

$$\operatorname{plim} \widehat{\theta}_{c^*} = \operatorname{plim} \frac{n}{n+1} \times \operatorname{plim} \bar{X} = 1 \times \theta = \theta$$

where LLN ensures that:

$$\operatorname{plim} \bar{X} = \mathrm{E}(X_i) = \theta.$$

(b) To obtain the MLE, we first define the joint density (defines the likelihood function). Given independence we have:

$$\begin{aligned} f(X_1, \ldots, X_n) = \prod_{i=1}^{n} f(X_i, \theta) &= \frac{1}{\theta} \mathrm{e}^{-X_1/\theta} \frac{1}{\theta} \mathrm{e}^{-X_2/\theta} \cdots \frac{1}{\theta} \mathrm{e}^{-X_n/\theta} \\ &= \theta^{-n} \mathrm{e}^{-\frac{1}{\theta} \sum X_i} \\ &= L(\theta; X_1, \ldots, X_n). \end{aligned}$$

We then specify the log likelihood function:

$$\log L(\theta; X_1, \ldots, X_n) = -n \log(\theta) - \frac{1}{\theta} \sum X_i.$$

The MLE requires us to solve the first-order condition:

$$\frac{\mathrm{d} \log L(\theta)}{\mathrm{d}\theta} = -\frac{n}{\theta} + \frac{1}{\theta^2} \sum X_i$$

for $\widehat{\theta}_{MLE}$, we have:

$$-\frac{n}{\widehat{\theta}_{MLE}} + \frac{1}{\widehat{\theta}^2_{MLE}} \sum X_i = 0.$$

Multiplying this equation by $\widehat{\theta}^2_{MLE}$ yields:

$$-n\widehat{\theta}_{MLE} + \sum X_i = 0 \quad \Rightarrow \quad \widehat{\theta}_{MLE} = \frac{\sum X_i}{n} = \bar{X}.$$

The MLE is consistent, asymptotically normal and asymptotically efficient.

Comments:

- Most candidates did well for part (a), some marks were lost due to poor annotations or a lack of clarification. Some made algebraic mistakes but were only penalised once.
- Most recognised the need to take a first-order condition and got the correct answers in part (b).
- Most did well in the biasedness proof for part (c). The consistency proof was relatively less well done, most marks were lost for not explaining the steps or mistaking the LLN with the CLT or a plim property.
- Most derived the correct log-likelihood function and solved the first-order condition correctly. Marks were lost mostly for poor annotations or insufficient discussions of the MLE properties.

Question 8

Let us consider a dataset containing information on government expenditure on education ($EducE$), gross domestic product (GDP), population (Pop), and capital stock ($Capital$) for 153 countries in 2019. Economic theory suggests that government expenditure on education per capita has a positive effect on per capita GDP. That is,

$$perGDP_i = \beta_1 + \beta_2 perEducE_i + \beta_3 perCapital_i + u_i,$$

where $perGDP_i = \dfrac{GDP_i}{Pop_i}$, $perEducE_i = \dfrac{EducE_i}{Pop_i}$, and $perCapital_i = \dfrac{Capital_i}{Pop_i}$, and i denotes a particular country.

(a) **Explain intuitively why we should be worried about the presence of heteroskedasticity in this model. Briefly indicate three distinct methods to potentially deal with heteroskedasticity in this situation.**

(6 marks)

(b) **Discuss how you would implement the White test for the presence of heteroskedasticity in this case. Briefly discuss how the White test differs from the Goldfeld–Quandt test.**

(5 marks)

(c) **Discuss two reasons why we may expect the OLS estimator for β_2 to be neither unbiased nor consistent. Provide supportive arguments for each.**

(4 marks)

(d) **The dataset also has a dummy variable $Democratic_i$ which equals 1 if country i is democratic, 0 otherwise. Discuss how you could test whether the effect of government expenditure on education per capita is higher for democratic countries. Clearly indicate the regression you would run and the test you would use.**

Comment: **You may assume that the OLS estimator of the regression you would run is consistent, but there may be heteroskedasticity in your model.**

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 4.3, 6.1, 7.2, 7.3 , 8.2 and 8.3.

Dougherty, C. Subject guide (2016): Chapters 4, 7, and 8.

Approaching the question

(a) Candidates were expected to discuss the intuition behind heteroskedasticity in the model, not the mathematical derivation.

We suspect heteroskedasticity because richer countries with a higher absolute GDP per capita, will have a larger budget to distribute and just have greater flexibility of how much they can spend on education. This flexible constraint may result in large variations in the education budget, or the error terms for richer countries and larger education expenses might have a larger variance.

Countries with a lower GDP per capita may have smaller budget options and the education budget is likely to vary less, resulting in smaller variances of their error terms. The relationship, however, is not necessarily linear or proportional.

Candidates can discuss the inefficiency and invalid test statistics due to heteroskedasticity.

There are three popular ways to deal with this situation:

- Use the White heteroskedasticity robust standard errors.
- Use a Generalised Least Squares estimator (GLS or FGLS) to reweight the observations. Since the heteroskedasticity form is unknown, there is a risk of misspecification.
- From economic theory, a logarithm rescaling can be used to reduce the large variances between large and small economies. That is, we can run the following estimation instead:

$$\log(perGDP_i) = \alpha_1 + \alpha_2 \log(perEducE_i) + \alpha_3 \log(perCapital_i) + u_i.$$

(b) Candidates are expected to do the following steps:

- Estimate the original model by OLS, as usual. Obtain the OLS residuals $\widehat{u}_i$. Compute the squared residuals $\widehat{u}_i^2$.
- Run the following regression and keep its R^2, $R^2_{\widehat{u}^2}$:

$$\widehat{u}_i^2 = \delta_0 + \delta_1 perEducE_i + \delta_2 perCapital_i + \delta_3 perEducE_i^2 + \delta_4 perCapital_i^2$$
$$+ \delta_5 perCapital_i \times perEduc_i + error.$$

- The null hypothesis is $\mathrm{H}_0 : \delta_1 = \cdots = \delta_5 = 0$ and H_1 : otherwise. It's equivalent to a test of $k = 5$ restrictions.
- The test statistic is either the F statistic:

$$F = \frac{R^2_{\widehat{u}^2}/k}{1 - R^2_{\widehat{u}^2}/(n-k-1)}$$

  or the LM statistic:

$$LM = nR^2_{\widehat{u}^2}$$

  with $n = 153$.
- Under the null, the F statistic follows the $F_{k,\,n-k-1}$ distribution or the χ^2_5 distribution. If the statistic exceeds the critical value from the respective distribution, we reject the null hypothesis of homoskedasticity.
- The Goldfeld–Quandt test assumes a monotonic form of heteroskedasticity (the standard deviation of the error distribution is proportional to one of the explanatory variables) and will not work well for non-monotonic forms of heteroskedasticity, which we should expect from the discussion about heteroskedasticity in question (a). The general White test addresses this concern. One potential concern about the White test is that it has low power, especially when there are many explanatory variables as the degrees of freedom will get smaller.

(c) Two potential endogeneity issues:

- Omitted variable bias: there might be an important variable that can explain both $perGDP$ and $perEducE$, for example, some unobserved culture of the country, or the political institution.
- Simultaneity bias: there might also be another structural equation in which $perGDP$ can explain $perEducE$ – richer countries may spend relatively more on education. Candidates can also discuss classical measurement errors in the regressor as another source of endogeneity.

(d) We expect to run the following regression, using the White robust standard error:

$$perGDP_i = \theta_1 + \theta_2 perEducE_i + \theta_3 perCapital_i + \theta_4 Democratic_i$$
$$+ \theta_5 Democratic_i \times perEducE_i + u_i.$$

The differential effect of $perEduc$ on $perGDP$ for democratic countries is captured by the coefficient θ_5.

We can simply do a t test, using the robust standard error, of the null hypothesis that $\mathrm{H}_0 : \theta_5 = 0$ and $\mathrm{H}_1 : \theta_5 > 0$ because we expect $\theta_2 > 0$.

Comments:

- Part (a) was attempted well in general, with candidates correctly explaining why heteroskedasticity would be a problem to our estimator. Many further explained why we would be worried about the presence of heteroskedasticity in our set-up. The remedies of heteroskedasticity were less well done. Most often suggested the use of WLS and GLS, referring to whether the form of heteroskedasticity is known. Even though this answer was correct, the two methods cannot be counted as distinct. Others suggested collecting a different set of data free of heteroskedasticity or with little variations, which would not be a sensible thing to do. The least popular method suggested was to use the logarithms of the variables. Some suggested different ways to detect it, not to solve it, which was not the point of the question. Finally, some only listed the methods without actually discussing them, hence losing some marks.

- Part (b) was also attempted well. Most lost marks due to incorrect test specifics or incomplete steps. The most common mistake was to omit the cross-products of the variables. Others did not distinguish well the difference between the White test and the Goldfeld–Quandt test, often referring to their different test equations as the key distinction; while we should focus on the assumptions behind the form of heteroskedasticity or the power of the tests.

- Part (c) posed some challenges. Several candidates suggested multicollinearity or heteroskedasticity as a cause for biasedness or inconsistency, which was not correct. Others, despite correctly identifying the OVB, reverse causality, or measurement errors, did not provide any explanations or arguments. Some went beyond and discussed the three mentioned problems and discussed them very well.

- Part (d) required candidates to think about the use of the interaction and its interpretation. Many realised this and did well in describing the new regression with an interaction. Some lost marks for failing to use the robust standard errors for the test or lacking test specifics of the z test. Some did not realise that we should conduct a one-tailed test and θ_2 should be positive. Others proposed the use of F tests or Chow tests, without supportive arguments. Some proposed to split the data into two samples of democratic and non-democratic and then use the Chow test. This approach will help test the equivalence of the parameters, but not necessarily whether democratic countries would benefit more from education expenses.

Question 9

We are interested in explaining the binary variable *vhappy*, a dummy variable that denotes whether an individual considers him/herself 'very happy' or not (1 = yes, 0 = no).

We consider the following socio-demographic variables: *occattend* and *regattend* (which are dummy variables indicating whether the individual occasionally or regularly attends church, where the excluded dummy indicates that the individual never attends church), *income* (family income in 1,000US$), *unemp*10 (dummy indicating whether the individual has been unemployed in the last 10 years), and *educ* (years of education completed). Let z be a linear function of the variables that determine the probability of being very happy, where:

$$z = \beta_0 + \beta_1 occattend + \beta_2 regattend + \beta_3 \log(income) + \beta_4 unemp10 + \beta_5 educ.$$

The β parameters have been estimated based on a Linear Probability Model (LPM) and Logit Model. Two sets of results (labelled specification (A) and (B)) are obtained based on a random sample of 9,963 observations from US. The usual standard errors for the Logit model and the robust standard errors for the LPM are reported in parentheses.

	LPM (A)	LPM (B)	Logit (A)	Logit (B)
constant	0.080 (0.032)	0.088 (0.032)	−2.074 (0.178)	−2.032 (0.176)
occattend	−0.005 (0.010)	—	−0.023 (0.051)	–
regattend	0.093 (0.014)	—	0.420 (0.065)	—
log(*income*)	0.061 (0.012)	0.064 (0.012)	0.362 (0.070)	0.379 (0.069)
*unemp*10	−0.097 (0.010)	−0.103 (0.010)	−0.487 (0.050)	−0.514 (0.049)
educ	0.008 (0.002)	0.007 (0.002)	0.036 (0.008)	0.034 (0.008)
R^2	0.023	0.019		
$\log L$			−5,890.6	−5,998.9

(a) **Discuss clearly how the parameter estimates for β are obtained for the Linear Probability Model and the Logit model. No technical details expected, but intuition rewarded.**

(5 marks)

(b) **For this part use specification (A).**

i. **Obtain for both models the predicted probability of happiness for an individual who regularly attends church, has a family income of £150,000, has not been unemployed in the last 10 years and has 13 years of education. *Hint:* $\log(150) \simeq 5$.**

(5 marks)

ii. **Obtain for both models what the effect of an additional year of education is for this individual and state whether the impact of a further additional year is the same.**

(5 marks)

(c) **You would like to test the joint significance of the church attendance variables *occattend* and *regattend*.**

i. **Explain why you cannot conduct this test using the LPM based on the reported results.**

(2 marks)

ii. **Explain how you would conduct this test using the Logit Model based on the reported results.**

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 10.1, 10.2, and 10.6.

Dougherty, C. Subject guide (2016): Chapter 10.

Approaching the question

(a) To obtain the LPM parameter estimates we simply apply OLS where the dependent variable is the dummy variable *vhappy*. As the LPM assumes a linear relation between $\mathrm{E}(y)$ and the explanatory variables, minimising the residual sum of squares seems a reasonable approach for its estimation (it fits the linear regression framework). (Care should be taken for the computation of the standard errors as there will be heteroskedasticity, candidates are not asked to discuss this.)

To obtain the Logit parameter estimates we use the MLE. We need to maximise the log-likelihood function which here is given by:

$$\begin{aligned}\log L &= \sum_{i=1}^{n}\{vhappy_i \log(P(vhappy_i = 1 \,|\, x_i) + (1 - vhappy_i)\log(P(vhappy_i = 0 \,|\, x_i)\} \\ &= \sum_{i=1}^{n}\{vhappy_i \log(\Lambda(z_i)) + (1 - vhappy_i)\log(1 - \Lambda(z_i))\}\end{aligned}$$

where:

$$z_i = \beta_0 + \beta_1 occattend_i + \beta_2 regattend_i + \beta_3 \log(income_i) + \beta_4 unemp10_i + \beta_5 educ_i$$

and Λ is the cdf of the logistic distribution where $\Lambda(z) = \exp(z)/(1 + \exp(z))$. While formally we solve the first-order condition, intuitively we try to choose our β in such a way that the predicted probability of being very happy is high for individuals who are happy and the predicted probability of not being very happy is high for individuals who are not happy.

(b) Candidates are expected to discuss the following points:

i. For LPM:

$$P(\widehat{vhappy} = 1 \,|\, x) = 0.080 + 0.093 + 0.061 \times 5 + 0.008 \times 13 = 0.582.$$

For Logit, the predicted probability is:

$$\begin{aligned}P(\widehat{vhappy} = 1 \,|\, x) &= \Lambda(-2.074 + 0.420 + 0.362 \times 5 + 0.036 \times 13) \\ &= \Lambda(0.624) \\ &= \frac{\exp(0.624)}{1 + \exp(0.624)} \\ &= 0.651.\end{aligned}$$

ii. For the LPM, the effect of an additional year equals 0.008 (or 0.8% points); for a further year, this is the same (LPM marginal effect is constant). For the Logit model the effect can be obtained by comparing the probabilities:

$$\begin{aligned}\Lambda(0.66) - \Lambda(0.624) &= \Lambda(-2.074 + 0.420 + 0.362 \times 5 + 0.036 \times 14) \\ &\quad - \Lambda(-2.074 + 0.420 + 0.362 \times 5 + 0.036 \times 13) \\ &= 0.659 - 0.651 \\ &= 0.008\end{aligned}$$

but can use the result given by:

$$0.036 \times \frac{\exp(z)}{(1 + \exp(z))^2} = 0.036 \times \frac{\exp(0.624)}{(1 + \exp(0.624))^2} = 0.008.$$

For a further change in the regressor, the effect from the LPM will be smaller as the Logit marginal effect is diminishing.

(c) You would like to test the joint significance of the church attendance variables *occattend* and *regattend*.

i. In the presence of heteroskedasticity we cannot use the F test as the standard errors are invalid. Even if ignoring the problem of heteroskedasticity, neither the SSR nor the covariance between the parameter estimates is provided.

ii. We can use the LR test here. $\mathrm{H}_0 : \beta_{occattend} = \beta_{regattend} = 0$ against $\mathrm{H}_1 : \beta_{occattend} \neq 0$ and/or $\beta_{regattend} \neq 0$.
The test statistic equals:

$$LR = 2(\log L^{UR} - \log L^{R}) = 2 \times (-5{,}890.6 - -5{,}998.9) = 216.6.$$

At the 5% level of significance we should reject if $LR > 5.991$. Under the null the test is distributed asymptotically as χ^2_2.

Comments:

- Candidates generally did well in part (a). Most discussed well the intuitions behind LPM and logit, with good explanations and even derivations of the likelihood. Some lost marks for incorrect statements or insufficient discussion.
- Part (b) in general was straightforward and well-attempted by candidates. Common mistakes included using the wrong estimates for the calculations (for example, using logit estimates for the LPM calculation, and vice versa). Some used incorrect marginal effect formulae and lost marks as a result.
- Candidates who attempted part (c) did very well, particularly for realising the heteroskedasticity inherent in the LPM. Some lost marks in (c) ii. for not recording the need to use the LR test instead of an F test for tests with logit estimates. Others lost marks for not calculating the test statistic and only writing out formulae instead.

Examiners' commentaries 2022 (October)

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2021–22. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

General remarks

Learning outcomes

At the end of the course, and having completed the Essential reading and activities, you should be able to:

- describe and apply the classical regression model and its application to cross-section data
- describe and apply the:
  - Gauss–Markov conditions and other assumptions required in the application of the classical regression model
  - reasons for expecting violations of these assumptions in certain circumstances
  - tests for violations
  - potential remedial measures, including, where appropriate, the use of instrumental variables
- recognise and apply the advantages of logit, probit and similar models over regression analysis when fitting binary choice models
- competently use regression, logit and probit analysis to quantify economic relationships using standard regression programmes (Stata and EViews) in simple applications
- describe and explain the principles underlying the use of maximum likelihood estimation
- apply regression analysis to time-series models using stationary time series, with awareness of some of the econometric problems specific to time series applications (for example, autocorrelation) and remedial measures
- recognise the difficulties that arise in the application of regression analysis to nonstationary time series, know how to test for unit roots, and know what is meant by cointegration.

Common mistakes committed by candidates

The resit examination, first of its kind, took place online in October 2022. There was only one examination paper for both Zone A and Zone B. Similar to last year, we put serious effort to ensure the examination structure and the expected level of effort was consistent and comparable to those in normal years (the examination should take candidates three hours to complete). More focus was put on questions that require candidates to evaluate or discuss the properties of estimators/test procedures, as well as to ask candidates to explain their answers in a more intuitive way and provide real-life examples. We discuss below common mistakes made by candidates in both zones, most of them are repeated year after year. We then provide specific comments for each question.

Candidates should realise that $\frac{1}{n}\sum_{i=1}^{n} u_i$ is not the same as $\mathrm{E}(u_i)$. So, while we typically assume $\mathrm{E}(u_i) = 0$, this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} u_i = 0$. Also, while we may be happy to assume $\mathrm{E}(x_i u_i) = 0$ (uncorrelatedness between the errors and regressors), this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} x_i u_i = 0$. Misusing these concepts will be penalised.

- Both $\frac{1}{n}\sum_{i=1}^{n} u_i$ and $\frac{1}{n}\sum_{i=1}^{n} x_i u_i$ are random variables, which take the value 0 with probability 0 (continuous random variables)!

- $\mathrm{E}(u_i) = 0$ and $\mathrm{E}(x_i u_i) = 0$ are fixed quantities, not stochastic!

The differences between sample and population moments need to come across clearly when looking at unbiasedness and making consistency arguments. In both cases, we first simplify our estimator (plug in the true model) to obtain:

$$\widehat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \beta + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \quad \text{with } x_i = X_i - \bar{X}.$$

- For *unbiasedness*, clearly indicate that you want to show that $\mathrm{E}(\widehat{\beta}) = \beta$. Unbiasedness does not follow from $\sum_{i=1}^{n} x_i u_i = 0$, instead it follows from $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) = 0$.

  If we treat x_i as fixed:

$$\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) \equiv \mathrm{E}\left(\sum_{i=1}^{n} d_i u_i\right) = \sum_{i=1}^{n} d_i \mathrm{E}(u_i)$$

  and then unbiasedness follows as $\mathrm{E}(u_i) = 0$.

- For *consistency*, clearly indicate that you want to show that $\mathrm{plim}(\widehat{\beta}) = \beta$. Using the plim

properties, we show:

$$\begin{aligned}
\operatorname{plim} \widehat{\beta} = \beta + \operatorname{plim} \left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \right) &= \beta + \frac{\operatorname{plim} \left(\frac{1}{n} \sum_{i=1}^{n} x_i u_i \right)}{\operatorname{plim} \left(\frac{1}{n} \sum_{i=1}^{n} x_i^2 \right)} \\
&\equiv \beta + \frac{\operatorname{plim}(\text{Sample Cov}(x, u))}{\operatorname{plim}(\text{Sample Var}(x))} \\
&= \beta + \frac{\operatorname{Cov}(x, u)}{\operatorname{Var}(x)} \quad \text{using the law of large numbers}
\end{aligned}$$

where $\operatorname{Cov}(x, u) = 0$ and $\operatorname{Var}(x) > 0$, ensuring we get consistency.

- Remember, the law of large numbers ensures that sample averages converge to their population analogues.

A large number of candidates are not able to clearly distinguish between sample variance and covariance, and population variance and covariance.

The use of $\operatorname{Cov}(X, Y)$ and $\operatorname{Var}(X)$ should be restricted to describing the population covariance and variances, respectively, with definitions:

$$\operatorname{Cov}(X, Y) = \operatorname{E}((X - \operatorname{E}(X))(Y - \operatorname{E}(Y))) = \operatorname{E}(XY) - \operatorname{E}(X)\operatorname{E}(Y)$$

and:

$$\operatorname{Var}(X) = \operatorname{E}((X - \operatorname{E}(X))^2) = \operatorname{E}(X^2) - (\operatorname{E}(X))^2$$

(you also may denote $\operatorname{Cov}(X, Y) = \sigma_{XY}$ and $\operatorname{Var}(X) = \sigma_X^2$). They are typically unknown, but fixed, quantities.

The sample covariance and variance are estimators of the population covariance and variance, respectively. They are defined as:

$$\text{Sample Cov}(X, Y) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})$$

and:

$$\text{Sample Var}(X) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

(you also may use $\widehat{\sigma}_{XY}$ and $\widehat{\sigma}_X^2$). You can compute them given the data.

With a slight abuse of notation, we often divide by n instead, which is irrelevant if we let n be large. The division by $n - 1$ is a finite sample issue only (unbiasedness).

The sample covariance and variance show up in our definition of the OLS estimator of the slope in the simple linear regression model, not the population covariance and variance, as:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = \frac{\text{Sample Cov}(X, Y)}{\text{Sample Var}(X)} \neq \frac{\operatorname{Cov}(X, Y)}{\operatorname{Var}(X)}.$$

It was very common to misunderstand the central limit theorem. While it provides the distribution of the *estimator* $\widehat{\beta}$, no matter what the true distribution of the errors is, it does *not* say that the distribution of the error term is approximately normal in large samples.

Many candidates simply described multicollinearity as independent variables being correlated. Although we accept answers discussing high correlation among the independent variables as an

alternative to a close linear relation amongst the regressors, simply taking correlation among regressors as the definition of multicollinearity is incorrect. Many candidates incorrectly commented that multicollinearity causes biasedness and/or inconsistency.

Many candidates made the classic error of not recognising that $\mathrm{E}(X/Y) \neq \mathrm{E}(X)/\mathrm{E}(Y)$, in general.

Only a few candidates realised the differences in the proof for unbiasedness of an estimator under the assumption of non-stochastic and stochastic (random variable) regressors in Question 1. When we assume a non-stochastic X, we can treat a function, say, a_i of $X_1, X_2, \ldots, X_n$ as a constant so that we can have $\mathrm{E}(a_i u_i) = a_i \mathrm{E}(u_i) = 0$. When X is assumed random (or stochastic), a_i is no longer a fixed constant but a random variable. We must first condition on $X = (X_1, X_2, \ldots, X_n)$ to take a_i out of the conditional expectation: $\mathrm{E}(a_i u_i \mid X) = a_i \mathrm{E}(u_i \mid X) = 0$. The Law of Iterated Expectation, $\mathrm{E}(\mathrm{E}(y \mid X)) = \mathrm{E}(y)$ where y is a random variable, should then be used to obtain the unconditional expectation of the estimator. Among those who realised the need on conditioning on a_i, many did not use the LIE to arrive at the final condition for unbiasedness, $\mathrm{E}(\widehat{\beta}) = \beta$.

Autocorrelation and heteroskedasticity were tricky topics for a good number of candidates. Many did not explain well the need to use robust standard errors or an asymptotic version of testing under autocorrelation and/or heteroskedasticity. Most also lost marks when discussing the properties of an estimator that can help solve these problems. Some incorrectly suggested that fixing autocorrelation will also resolve inconsistency and/or biasedness.

When discussing the importance of distinguishing trend stationary and difference stationary processes, two points are expected: discussing the implication for the long-run behaviour of the processes, and discussing statistical implications (the consistency of the estimators). Most candidates discussed the former very well, only a few touched upon the latter.

Related, candidates often lost marks for the interpretation of the ECM model, particularly when discussing the properties of the estimator when the variables are stationary or not. Candidates often struggled with the derivation of an ECM model. A good number of candidates avoided this question, others who attempted did not do very well.

A good number of candidates seemed to be confused between the Dickey–Fuller test and an Augmented Dickey–Fuller (ADF). The difference between the two tests is in the test equation. In the augmented specification, we add sufficient lagged differences of the residuals to ensure that there is no autocorrelation left in the test equation regression. For this reason, some used the test equation for the ADF when the DF was called for. Related, others forgot that even though we can use a very similar test statistic to the t statistic in the Dickey–Fuller stationarity test (we call it τ statistic), it does not follow a t distribution but the Dickey–Fuller distribution instead.

Candidates struggled to give competent answers to the interpretation of empirical results. When interpreting an empirical result you should discuss the significance of the coefficients, magnitude, and sign of the coefficients. Related, *ceteris paribus* (or holding other factors constant) is needed to ensure that the estimate can be interpreted as the marginal change in the dependent variable given a change in the independent variable of interest.

Related, many candidates failed to discuss problems surrounding the omitting of important, relevant explanatory variables and irrelevant explanatory variable. While omitting the former will cause (omitted variable) bias, omitting the latter will only lead to the loss of efficiency.

Others struggled with interpreting coefficients from a non-linear regression (for example, log-linear, linear-log, or log-log).

Candidates particularly struggled with the difference between percentage points and % changes when it comes to the interpretation of coefficients in binary choice models. A coefficient in such a model gives the percentage point difference in the outcome variable when the regressor changes by 1 unit. The % change is a *relative* change that depends on the baseline value of the outcome variable (for example, the realisation when a regressor equals zero in the case of a dummy variable regressor). For example, if we have initially 50% of a sample buying ecological apples, a 10% increase will mean now $(1 + 0.1) \times 50 = 55\%$ of the sample buying ecological apples. A 10% point increase instead will

result in $50 + 10 = 60\%$ of the sample choosing to buy ecological apples.

In the case of dummy variables as regressors, candidates often missed that the coefficients have to be interpreted relative to the base case which is the left out dummy variable (see appropriate questions in the examination for details).

When discussing the marginal effect after a Probit estimation, candidates often did not specify the term z in the pdf $\phi(z)$. Some used the coefficients from the LPM regressions to calculate the marginal effects after a Probit/Logit estimation. A popular, related mistake was that candidates did not realise the inherent heteroskedasticity of LPM.

A number of candidates seemed to be confused between the conditions for a valid instrument (exclusion restriction and relevance) and factors related to obtaining a more precise IV estimate (a higher sample size, a higher sample variability of X_i, a lower variance of the error term, and a stronger correlation between X_i and Z_i). More seriously, most candidates did not realise the need to use an IV estimator when the regression has a lagged dependent variable as a regressor in the presence of autocorrelation.

Exogeneity and endogeneity are often very simply explained as variables 'outside of the model' or 'inside the model', respectively. While this approach is not wrong; it is not a particularly clear or helpful definition in general contexts that are not related to simultaneous equations models. Defining these concepts in terms of the relationship between regressor and the error term is much clearer and naturally leads to a way to show exogeneity or endogeneity by considering the covariance between regressors and error terms.

Candidates often missed that all exogenous regressors and instruments have to be included in the first stage when performing a TSLS estimation. Some candidates use the fitted values of the endogenous variable obtained from the first stage as the dependent variable in the second stage. This mistake is a serious misunderstanding. It is fine to indicate that this fitted value can be used as an instrument of the endogenous variable, but is not really the way the second stage of TSLS is typically understood.

When asked about identification of structural form equations, candidates often compared the number of excluded regressors to the number of endogenous variables in general, without pointing out which variables are the excluded regressors and could therefore be used as an instrument. Candidates were also often unclear of the definitions of over-identification, under-identification, and exact identification. The best approach to this type of questions is to discuss the endogenous and exogenous variables, and then think about the available exogenous variables that can be used as instruments. Some candidates discussed the rank condition, which is a necessary and sufficient condition for identification, but requires knowledge of matrix algebra.

When deriving the reduced forms, candidates should check conditions for their existence, in particular, whether the denominator of the expression is different from zero.

Gauss–Markov conditions were used in many places in the examination, for example, when discussing the BLUE property of an estimator or when conducting hypothesis tests. The Gauss–Markov conditions have to be explicitly specified. Only writing that the Gauss–Markov conditions hold is not sufficient. A good practice that many candidates have adopted is to begin the script by explicitly providing the Gauss–Markov conditions and refer back to them thereafter. Many candidates did not mention either the Gauss–Markov conditions or the normality assumption when conducting the t test.

When conducting hypothesis testing, ensure that you clearly indicate the null and alternative hypotheses (in terms of the true parameters, say β_1), the test statistic (in terms of the parameter estimates, here $\widehat{\beta}_1$), its distribution (with degrees of freedom), the rejection rule (one-sided or two-sided) for a given significance level (typically 5%) with suitable critical values, and provide an interpretation of your result. Candidates often lost marks for missing parts of these steps. Choosing the critical value appropriate to the sample size is also important. Several candidates used 1.96 (a critical value for very large degrees of freedom) in a small sample instead of the appropriate one. This mistake also cost marks. Related, many did not know when to use an asymptotic version of a

test – for example, when normality of the errors is not assumed, or when we use robust standard errors.

Another common mistake is to conduct a two-sided test when the statement was worded with a clear direction of the relationship. Many candidates fail to provide all the necessary specifics of the testing procedure (testing distribution, the critical value, the assumptions, the rejection rule, and the conclusion) for a t test, an F test, or a Chow test.

Just as in previous years, many candidates did not answer all parts of a question. Make sure you read the questions properly and provide all details that are requested. Not answering a question will automatically earn you a zero. Some candidates probably did not have a good examination plan.

Key steps to improvement

Essential reading for this course includes the subject guide and the following:

- Dougherty, C. *Introduction to econometrics*. (Oxford: Oxford University Press, 2016) 5th edition [ISBN 9780199676828]; http://oxfordtextbooks.co.uk/orc/dougherty5e/

Apart from the Essential readings you should do some supplementary reading. One very good book at the same level is:

- Gujarati, D.N. and D.C. Porter *Basic econometrics*. (McGraw–Hill, 2009, International edition) 5th edition [ISBN 9780071276252].

To understand the subject clearly it is important to supplement Dougherty's *Introduction to econometrics* (fifth edition) with the subject guide **EC2020 Elements of econometrics** (2016), especially Chapter 10 which covers maximum likelihood estimation. It is very important to carefully go through the subject guide. The subject guide contains solutions to the questions given in the main textbook and also some additional questions and solutions. Working through these will improve your understanding of the subject.

The chapter in the subject guide on maximum likelihood (Chapter 10) includes some additional theory which has not been covered in the main textbook. It is important to read the additional theory given in the subject guide to have a better understanding of the principles of maximum likelihood and tests based on the likelihood function.

Please check the VLE course page for resources for this subject such as a downloadable copy of the subject guide **EC2020 Elements of econometrics** (2016), PowerPoint slideshows that provide a graphical treatment of the topics covered in the textbook, datasets and statistical tables. Candidates should utilise datasets using standard regression programmes (STATA or EViews). This will help in the understanding of the subject.

Please note that the subject guide for EC2020 Elements of econometrics has been updated for the 2022–23 academic year and will be based on the textbook 'Introductory Econometris: A Modern Approach' by Wooldridge.

Examination revision strategy

Many candidates are disappointed to find that their examination performance is poorer than they expected. This may be due to a number of reasons, but one particular failing is '**question spotting**', that is, confining your examination preparation to a few questions and/or topics which have come up in past papers for the course. This can have serious consequences.

We recognise that candidates might not cover all topics in the syllabus in the same depth, but you need to be aware that examiners are free to set questions on **any aspect** of the syllabus. This means that you need to study enough of the syllabus to enable you to answer the required number of examination questions.

The syllabus can be found in the Course information sheet available on the VLE. You should read the syllabus carefully and ensure that you cover sufficient material in preparation for the examination. Examiners will vary the topics and questions from year to year and may well set questions that have not appeared in past papers. Examination papers may legitimately include questions on any topic in the syllabus. So, although past papers can be helpful during your revision, you cannot assume that topics or specific questions that have come up in past examinations will occur again.

If you rely on a question-spotting strategy, it is likely you will find yourself in difficulties when you sit the examination. We strongly advise you not to adopt this strategy.

Examiners' commentaries 2022 (October)

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2021–22. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from **Section A** (8 marks each) and **THREE** from **Section B** (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Section A

Answer all questions from this section.

Question 1

AirBnB is an online marketplace for lodging, mainly homestays for vacation rentals, and tourism activities. A researcher tries the following regression models to determine the factors affecting AirBnB renting price in the UK, where $AirBnB$ is the average monthly rent paid on an AirBnB rental units in a UK city. Let $size$ denote the room size in m^2, $avginc$ the average city income in thousands GBP, and $tourist$ the tourist population as a percentage of the total population:

$$\log(AirBnB) = \beta_0 + \beta_1 \log(size) + \beta_2 \log(avginc) + \beta_3 tourist + u.$$

The equation estimated using data from 66 cities is as follows, with the standard errors reported in parenthesis:

$$\log(\widehat{AirBnB}) = \underset{(.234)}{.013} + \underset{(.009)}{.011} \log(size) + \underset{(.081)}{.507} \log(avginc) + \underset{(.0007)}{.0016} tourist$$

$$n = 66, \ R^2 = .658.$$

(a) You are interested in testing the following hypothesis $\mathbf{H_0 : 50\beta_1 = \beta_2}$. Discuss how to implement the test for $\mathbf{H_0}$, clearly describing any additional information you would require to perform the test.

(4 marks)

(b) A colleague claims that since the standard error for log(*size*) is unreasonably large, the model evidently suffers from heteroskedasticity. Discuss whether the colleague is correct or not. Indicate how you should proceed with your test of $\mathbf{H_0}$ in light of the presence of heteroskedasticity (of unknown form) if the colleague is correct.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 2.1, 2.2, 7.1, 7.2.

Dougherty, C. Subject guide (2016): Chapters 2 and 7.

Approaching the question

(a) The t statistic to test the null hypothesis H_0 is written as:

$$t = \frac{50\widehat{\beta}_1 - \widehat{\beta}_2}{se(50\widehat{\beta}_1 - \widehat{\beta}_2)}.$$

From the above output, we know $50\widehat{\beta}_1 - \widehat{\beta_2}$. On the other hand, the standard error of $50\widehat{\beta}_1 - \widehat{\beta}_2$ is given by:

$$se(50\widehat{\beta}_1 - \widehat{\beta}_2) = \sqrt{2{,}500 \times (se(\widehat{\beta}_1))^2 + (se(\widehat{\beta}_2))^2 - 100 \times s_{12}}$$

where s_{12} is the estimate of $\mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_2)$. To implement the above t test, we need s_{12}.

Under the Classical Linear Regression Assumptions (Gauss–Markov assumptions and normality) (candidates must state these somewhere in the script), we know that $t \sim t_{62}$ under H_0. Once we get s_{12}, we can implement the t test.

Let $\mathrm{H}_1 : 50\beta_1 \neq \beta_2$ be the two-sided alternative and set the significance level at 5%. The rejection rule is $|t| > c$, where c is the 97.5th percentile of the t_{62} distribution.

(b) An unreasonably larger standard error might give some informal sign of heteroskedasticity. We cannot rely on this as there are many other reasons for a large standard error (mis-specification, multicollinearity, or low variations in the regressors. Instead, one must use classical tests such as the White test, or the Breusch–Pagan, or Goldfeld–Quandt tests.

If there is indeed heteroskedasticity, we should use White's heteroskedasticity robust standard errors. Using the robust standard errors we then apply an asymptotic t test (does not require the assumption of normality of the errors), and the critical values must be obtained from the $N(0, 1)$ distribution instead of the t distribution.

Comments:

The question is a standard hypothesis test (in this case a t test) and discussing the problems of heteroskedasticity. Marks were lost due to missing test specifics or unclear explanation. In particular, some were not clear on the classical tests for heteroskedasticity or the asymptotic version of the t test in (b).

Question 2

Consider the following regression equation:

$$y_t = \beta_0 + \beta_1 z_t + \beta_2 z_{t-1} + \varepsilon_t, \tag{2.1}$$

where ε_t can be expressed as follows:

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t, \quad \text{with } |\rho| < 1 \tag{2.2}$$

and v_t is a white noise process generated randomly from a normal distribution with mean zero and constant variance. We assume that $\mathrm{E}(\varepsilon_t \mid z_t, z_{t-1}, \ldots) = 0$ and $\mathrm{Var}(\varepsilon_t \mid z_t, z_{t-1}, \ldots) = \sigma^2$. You have a random sample and can use T observations for the estimation, which we label $t = 1, 2, \ldots, T$ for convenience. Assume there is no perfect multicollinearity.

(a) You are concerned that the error terms ε_t exhibit autocorrelation. Explain the meaning of the term 'autocorrelation'. What consequences does the presence of autocorrelation have for the OLS estimation of the regression equation (2.1)?

(2 marks)

(b) Discuss the following statement: 'Multicollinearity causes a more severe problem to the OLS estimator than autocorrelation does.'

(2 marks)

(c) What is the name we have for the process underlying the error terms? Discuss how we can test for the presence of autocorrelation in this model.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 11.4, 12.1, 12.3, 13.1, 2, 3 and 5.

Dougherty, C. Subject guide (2016): Chapters 11, 12 and 13.

Approaching the question

(a) Autocorrelation refers to the assumption that the error terms over time are correlated, that is $\mathrm{E}(\varepsilon_t\varepsilon_s) \neq 0$ for some $t \neq s$. The presence of autocorrelation itself does not affect the unbiasedness or consistency of our OLS estimator. However, autocorrelation invalidates the standard errors, meaning that tests using the conventional standard errors are invalid. The OLS estimator is also no longer BLUE (loss of efficiency).

(b) Multicollinearity (often known as near multicollinearity) indicates a close (high) linear relation between the regressors. The consequence of having multicollinearity is simply that the standard errors become higher, but *remain valid*. Autocorrelation leads to *invalid standard errors*, so the statement is false. Also, whereas individual parameter estimates (for example, short-run effect) cannot be estimated very precisely, the long-run effect can.

(c) The error terms follows a stationary AR(1) process. To test for autocorrelation, the null hypothesis is $\mathrm{H}_0 : \rho = 0$ against $\mathrm{H}_1 : \rho \neq 0$. The testing equation is obtained from using the OLS residuals: $\widehat{\varepsilon}_t = y_t - (\widehat{\beta}_0 + \widehat{\beta}_1 z_t + \widehat{\beta}_2 z_{t-1})$. The testing equation is:

$$\widehat{\varepsilon}_t = \gamma_0 + \gamma_1 z_t + \gamma_2 z_{t-1} + \rho\widehat{\varepsilon}_{t-1} + u_t.$$

- We must use the asymptotic t statistics $\widehat{\rho}/se(\widehat{\rho})$ for a large sample.
- At the 5% level of significance, we should reject if $|\widehat{\rho}/se(\widehat{\rho})| > 1.96$, which follows a $N(0, 1)$ distribution).
- Candidates can instead use the LM test statistics $(T-1)R^2 \sim \chi^2_1$.

Comments:

Part (a) is a standard question about the definition and properties of autocorrelation. Most candidates did well with the definition but gave some unclear comments on the properties of an OLS under autocorrelation.

Part (b) is another standard question about multicollinearity, most did very well in the definition but unclear with the explanation on the invalidity of the standard errors.

Part (c) is about identifying an AR(1) model and perform the appropriate test. Marks were lost due to missing steps or an incorrect test distribution. Some candidates discussed the LM test, which were marked and rewarded accordingly.

Question 3

Consider the following regression model:

$$y_i = \alpha + \beta x_i^* + u_i,$$

where x_i^*, is a stochastic regressor, which has a non-zero mean and constant variance $\sigma^2_{x^*}$. The error term u has mean zero and constant variance σ^2_u and is distributed independently of x_i^*.

The stochastic regressor cannot be measured accurately, and you are only able to obtain a random sample $\{(y_i, x_i)\}_{i=1}^n$, where:

$$x_i = x_i^* + v_i.$$

The measurement error v_i, is understood to have a non-zero mean, m, and a constant variance σ^2_v. While it is reasonable to assume that the measurement error is distributed independently of u_i, we expect the measurement error exhibits correlation with x_i^*. That is, the measurement error is not completely random. (It does not satisfy the classical error in variables (CEV) assumptions.)

(a) Derive the probability limit of the OLS estimator for β obtained using the observable random sample. Discuss how your result contrasts with the setting where the measurement error is purely random.

(5 marks)

(b) Discuss how you can deal with the problem in question (a)? What type of information would you need to help you implement this approach.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 8.3, 8.4.

Dougherty, C. Subject guide (2016): Chapter 8.

Approaching the question

(a) Substituting to have:

$$\begin{aligned} y_i &= \alpha + \beta(x_i - v_i) + u_i \\ &= \alpha + \beta x_i + u_i - \beta v_i \\ &= \alpha + \beta x_i + \epsilon_i \end{aligned}$$

where $\epsilon_i = u_i - \beta v_i$. The OLS estimator for β from this equation is:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}.$$

The probability limit of $\widehat{\beta}$ is:

$$\begin{aligned}
\text{plim } \widehat{\beta} &= \frac{\text{plim } n^{-1}\sum_{i=1}^{n}(y_i - \bar{y})(x_i - \bar{x})}{\text{plim } n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{\text{plim } n^{-1}\sum_{i=1}^{n}(\beta(x_i - \bar{x}) + (\epsilon_i - \bar{\epsilon}))(x_i - \bar{x})}{\text{plim } n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \frac{\text{plim } \beta \text{ plim } n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 + \text{plim } n^{-1}\sum_{i=1}^{n}(\epsilon_i - \bar{\epsilon})(x_i - \bar{x})}{\text{plim } n^{-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \\
&= \beta + \frac{\text{plim SampCov}(\epsilon_i, x_i)}{\text{plim SampVar}(x_i)} \\
&= \beta + \frac{\text{Cov}(\epsilon_i, x_i)}{\text{Var}(x_i)} \\
&= \beta + \frac{\text{Cov}(u_i - \beta v_i, x_i^* + v_i)}{\text{Var}(x_i^* + v_i)} \\
&= \beta + \frac{\text{Cov}(u_i, x_i^*) - \beta\text{Cov}(v_i, x_i^*) + \text{Cov}(u_i, v_i) - \beta\text{Cov}(v_i, v_i)}{\text{Var}(x_i^*) + \text{Var}(v_i) + 2\text{Cov}(x_i^*, v_i)} \\
&= \beta\left(1 - \frac{\text{Cov}(v_i, x_i^*) + \sigma_v^2}{\sigma_{x^*}^2 + \sigma_v^2 + 2\text{Cov}(x_i^*, v_i)}\right) \\
&= \beta\left(\frac{\sigma_{x^*}^2 - \text{Cov}(v_i, x_i^*)}{\sigma_{x^*}^2 + \sigma_v^2 + 2\text{Cov}(x_i^*, v_i)}\right)
\end{aligned}$$

where the last equalities follow the assumption that u_i is independent of x_i^*, v_i is distributed independently of u_i, $\text{Cov}(v, x^*)$ is the covariance between x^* and v. The results rely on law of large numbers, that ensure:

$$\text{plim } n^{-1}\sum_{i=1}^{n} z_i = \text{E}(z_i).$$

The OLS estimator is not consistent. The non-zero mean of the measurement error does not have any impact on $\widehat{\beta}$ and this inconsistency result.

When $\text{Cov}(v_i, x_i^*) = 0$, we get the well-known attenuation bias result:

$$\text{plim } \widehat{\beta} = \beta\left(\frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_v^2}\right) \quad \text{where} \quad \frac{\sigma_{x^*}^2}{\sigma_{x^*}^2 + \sigma_v^2} < 1.$$

(b) There are two methods: either collect a better set of data with a more accurate measure of the variables, or find an instrumental variable to consistently estimate the parameter. Both are difficult in reality: the former asks for a new data collection, the latter requires some understanding of the nature of the measurement error and a potential source of instrumental variable.

Comments:

The question is about measurement errors and deriving the properties of the OLS.

For part (a), candidates lost marks for not using the plim in their proof or missing steps. Some made unfortunate algebraic mistakes and arrived at an incorrect bias formula. Some lacked explanations of steps. Some did not discuss what happens to OLS if we have random noise (attenuation bias).

For part (b), most realised the use of an IV to solve this measurement error, a few discussed the collection of a new dataset.

Question 4

Let us consider the following system of simultaneous equations that describe the equilibrium condition of wheat consumption in a country:

$$\boldsymbol{wheat = \alpha_1 price + \alpha_2 GDP + u_1,} \tag{4.1}$$

$$\boldsymbol{wheat = \beta_1 price + \beta_2 sunshine + \beta_3 sunshine^2 + \beta_4 flood + u_2,} \tag{4.2}$$

where (u_1, u_2) are i.i.d. errors with zero mean and constant variances, let $\mathrm{Cov}(u_1, u_2) \neq 0$. Let *wheat* denote per capita consumption of wheat (in kilograms) at the country aggregate level, *price* denote the domestic price per kilogram of wheat, *GDP* denote per capita income of the country, and *sunshine* denote the level of sunshine hours (or solar radiation) and *flood* denote the number of severe flooding events during the last wheat-growing season in the country.

(a) Which of the two equations is the supply equation, and which is the demand equation? Explain your answer.

(2 marks)

(b) Explain the concepts of endogenous versus exogenous variables. Briefly show estimating (4.1) by OLS is problematic. Hint: you are expected to derive the reduced form equation for *price*.

(4 marks)

(c) What is it meant by 'under-identification' and 'over-identification'? Is Equation (4.1) under-identified or over-identified? Briefly explain your answer.

(2 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 9.1, 9.3.

Dougherty, C. Subject guide (2016): Chapter 9.

Approaching the question

(a) Equation (4.1) is the demand equation because *GDP* is a demand factor that can explain the aggregate demand of wheat consumption. Equation (4.2) is a supply equation because *sunshine* and *flood* are supply factors that can explain the aggregate supply of wheat in the country.

(b) Exogenous variables are those whose values are determined outside of the model. *GDP*, *sunshine* and *flood* are such variables.

Endogenous variables are those whose values are determined interactively within the model (the system of equations). In our case, *wheat* and *price* are endogenous because their values

are determined through the models. To demonstrate the endogeneity of *price* we can express *price* in terms of the exogenous variable by equating (4.1) and (4.2):

$$\alpha_1 price + \alpha_2 GDP + u_1 = \beta_1 price + \beta_2 sunshine + \beta_3 sunshine^2 + \beta_4 flood + u_2$$

$$\Rightarrow \quad price = \frac{\beta_2 sunshine + \beta_3 sunshine^2 + \beta_4 flood - \alpha_2 GDP + u_2 - u_1}{\alpha_1 - \beta_1}.$$

This reduced form exists because it is reasonable to assume $\alpha_1 \neq \beta_1$. We can clearly see that now $\text{Cov}(price, u_1) \neq 0$ because in the reduced form of price there is a component of u_1. Thus Assumption B.7 that the disturbance term is independently distributed of the explanatory variable is violated, OLS estimator would be inconsistent.

(c) If there are more exogenous variables available to work as instrumental variables than the number of endogenous variables, the equation is called 'over-identified'. If there are fewer of such exogenous variables, the equation is called 'under-identified'. There are 3 potential exogenous variables (*sunshine*, $sunshine^2$ and *flood*) to be used as instruments for price in (4.1), (4.1) is over-identified.

Comments:

Most recognised the problem of simultaneity and associated issues (joint determination and inconsistency). Many gave great real-life examples. Marks were lost mostly on the explanations. The identification of the supply and demand equation was also done well.

Some lost marks in (b) for not clearly showing the relationship between *price* and u, and others for not commenting that this would violate the exogeneity assumption.

Most did well in (c). The definition of the over-identification problem was done well and candidates were able to correctly recognise the problem.

Question 5

Consider the following time series model for $\{y_t\}_{t=1}^T$:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t + \lambda \varepsilon_{t-1},$$

where t is a deterministic trend, ε_t is i.i.d. with mean zero and variance σ^2, for $t = 1, \ldots, T$, and $\beta_1 \neq 0$.

(a) Show that y_t is non-stationary.

(2 marks)

(b) Show that the autocovariance $\text{Cov}(y_{t+h}, y_t) = 0$ for $h > 1$, which ensures that y_t is weakly dependent.

(2 marks)

(c) The above process is also called a trend stationary process. Explain this terminology and compare/contrast it with a difference stationary process. Why is it important to distinguish them from one another?

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 11.2, 11.3, 12.1, 12.2, 13.1 and 13.5.

Dougherty, C. Subject guide (2016): Chapters 11, 12, and 13.

Approaching the question

(a) For a process y_t to be stationary, its expected value, variance, and covariance of y_t and y_{t+s} are independent of t. We only need to check each of the conditions.

The expected value of y_t is dependent of time, hence y_t is non-stationary:

$$\mathrm{E}(y_t) = \mathrm{E}(\beta_0 + \beta_1 t + \varepsilon_t + \lambda\varepsilon_{t-1}) = \beta_0 + \beta_1 t + 0 + 0 = \beta_0 + \beta_1 t.$$

(b) We check the covariance $\mathrm{Cov}(y_{t+h}, y_t)$, where $h > 1$:

$$\begin{aligned}\mathrm{Cov}(y_{t+h}, y_t) &= \mathrm{Cov}(\beta_0 + \beta_1(t+h) + \varepsilon_{t+h} + \lambda\varepsilon_{t+h-1}, \beta_0 + \beta_1 t + \varepsilon_t + \lambda\varepsilon_{t-1}) \\ &= \mathrm{Cov}(\varepsilon_{t+h} + \lambda\varepsilon_{t+h-1}, \varepsilon_t + \lambda\varepsilon_{t-1}).\end{aligned}$$

By definition of the covariance the additive non-stochastic terms vanish. Since $\mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0$, $i \neq j$, then we should note that the covariance must be equal to zero for $h > 1$. In particular, with $h = 1$, we get:

$$\mathrm{Cov}(y_{t+1}, y_t) = \mathrm{Cov}(\varepsilon_{t+1} + \lambda\varepsilon_t, \varepsilon_t + \lambda\varepsilon_{t-1}) = \lambda\mathrm{Var}(\varepsilon_t) \neq 0$$

but for $h = 2$, we get:

$$\mathrm{Cov}(y_{t+2}, y_t) = \mathrm{Cov}(\varepsilon_{t+2} + \lambda\varepsilon_{t+1}, \varepsilon_t + \lambda\varepsilon_{t-1}) = 0$$

for $h > 2$ we do not get any corresponding error terms either.

A time series is weakly dependent as the correlation between y_t and future periods y_{t+h} tends to zero asymptotically; with the periods increasing to infinity, the dependency between the two values tends to zero.

(c) A trend stationary process if by detrending, we can make it stationary. A difference stationary process if by differencing, we can make it stationary. It is important to distinguish the two types as they have implications for the long-run behaviours of a process after a shock to the process. After a shock, trend stationary processes will always revert to the trend in the long run (they would exhibit weak dependence). Difference stationary processes, however, never recover from shocks to the system. Shocks to the system will have a permanent effect on the process and the process exhibits strong dependence.

Comments:

Part (a) is a standard derivation question to show stationarity. Most candidates could give out the conditions for stationarity but struggled with correct steps of the derivation (despite the clear hint). The key is to show that $\mathrm{E}(y)$ is dependent on the time t.

Part (b) is about checking the covariance between two values of y_t, the key is to show that when the distance is larger than 1, the covariance is 0, while for succeeding terms the covariance should be non-zero. Many candidates got this right and explained the steps well.

Part (c) is a standard question on the differences between trend stationary and difference stationary series. Some marks were lost due to lack of clarity in explanation of the long-term behaviour of the two series.

Section B

Answer three questions from this section

Question 6

Let us consider the following model that describes the relationship between the stock price $price_t$ and the log of quarterly dividends $\log div_t$ of a company at quarter t:

$$price_t = \beta_0 + \beta_1 price_{t-1} + \beta_2 \log div_t + \beta_3 \log div_{t-1} + u_t. \tag{6.1}$$

We assume that $price_t$ and $\log div_t$ are both $I(1)$ variables, $|\beta_1| < 1$, and the error term u_t is independent of the regressors and has a zero mean.

(a) **Explain the concepts of spurious relationship and co-integration. Clearly discuss why it is important to test whether the relationship in (6.1) is spurious or cointegrating. *Hint:* You are expected to discuss the statistical consequences of spurious relationship and co-integrating on an OLS estimation of (6.1).**

(6 marks)

(b) **You are told that (6.1) can also be fitted by an Error Correction Model (ECM) that has the following structure:**

$$\Delta price_t = \rho(price_{t-1} - \gamma_1 - \gamma_2 \log div_{t-1}) + \gamma_3 \Delta \log div_t + u_t. \tag{6.2}$$

Explain what an Error Correction Model is and the intuition behind the ECM. How does the existence of the ECM relate to your answer in (a). Express ρ, γ_1, γ_2, γ_3 in terms of $(\beta_0, \beta_1, \beta_2, \beta_3)$ and carefully interpret the various components of this ECM.

(6 marks)

(c) **What are the advantages and challenges of estimating (6.2) using OLS? Discuss a procedure using OLS that can fit the ECM in (6.2).**

(6 marks)

(d) **Following the procedure in (c), we obtain an estimate for ρ of 0.5. Using this estimate, discuss the adjustment process of the stock price and dividend of the company towards the long-run equilibrium.**

(2 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 8.2, 8.3, 9.3, 11.4, 11.6, 13.4, 13.5.

Dougherty, C. Subject guide (2016): Chapter 8, 9, 11, 13.

Approaching the question

(a) Candidates are expected to discuss the following points:

Spurious regression refers to the scenario where two unrelated series give the impression that they are related, purely because of the fact that both are trending – that is, stock price and dividends are entirely unrelated but they follow a trend or there is another omitted variable that drives both variables and it appears that there is a meaningful relationship between the two series. Running a regression of Y_t on X_t (spurious) has a large R^2 and a seemingly significant slope parameter. The statistical problem is that here the error term is $I(1)$ as well and statistical inference is invalid. The OLS regression is meaningless: the OLS parameter estimates will not be consistent (and already biased), the R^2 will be high and the DW test statistic is small (as evidence of strong dependence in the errors).

On the other hand, cointegration describes the situation when we do have a long-run relationship between our $I(1)$ variables X_t and Y_t. Statistically, a linear combination of these $I(1)$ variables will become stationary $I(0)$. For example, if X_t and Y_t are $I(1)$ processes, but $Y_t - \beta X_t$ is stationary, then we can say X and Y are co-integrated. Cointegration means there is a long-term relationship between the two non-stationary variables and we can use OLS to fit a co-integrating relationship. The key statistical consequence is that OLS estimators will be super-consistent: even if there is any system of simultaneous relationships, any bias will tend to zero asymptotically.

(b) ECM is a model that describes the short-run and long-run dynamics between two cointegrating variables. If there is an ECM model for X and Y, then X and Y must be co-integrated. On its own, a co-integrating relationship sheds no light on short-run dynamics, but its very existence indicates that there must be some short-term forces that are responsible for keeping the relationship intact. This implies that it should be possible to construct a more comprehensive model that combines short-run and long-run dynamics. We can derive the ECM by rewriting (6.1) as follows:

$$\Delta price_t = (\beta_1 - 1)\left(price_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta_1}\log div_{t-1}\right) + \beta_3 \Delta \log div_t + u_t.$$

Candidates are expected to interpret the following components:

- The term:
$$price_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta_1}\log div_{t-1}$$
is the error correction term, measuring the disequilibrium at time $t-1$.
- $\rho = (\beta_1 - 1)$ indicates how fast the two processes are coming back to the steady state. It is sometimes called the adjustment coefficient. When the variables are out of the long-run equilibrium (for example, the error correction term is negative or positive), economic forces captured by this adjustment coefficient will push the model back to its long-run equilibrium.
- $\gamma_3 = \beta_2$ is the direct effect a change in $\log div$ has on *price*.

(c) Advantages: Since the variables are $I(1)$, all of the terms now in (6.2) are stationary, it means (6.2) can be estimated by an OLS. The challenge of doing so is that the β parameters from (6.1) to construct the error correction term are unknown hence this term is unobservable. One procedure that can overcome this challenge is the Engle–Granger two-step procedure, that uses the values of the parameters estimated in the co-integrating regression to compute the co-integrating term. The process is as follows:

- Firstly, we estimate the cointegating relationship between $price_t$ and $\log div_t$ by OLS to obtain the residuals, called $diseq_{t-1}$, where:
$$\widehat{diseq}_{t-1} = price_{t-1} - \widehat{\gamma}_1 - \widehat{\gamma}\log div_{t-1}.$$
- Using $\widehat{diseq}_{t-1}$ we run OLS on the error correction equation:
$$\Delta price_t = \rho\widehat{diseq}_{t-1} + \gamma_3 \Delta X_t + u_t.$$

(d) The speed of adjustment towards the long-run equilibrium is determined by the magnitude of ρ: an estimate of 0.5 implies that roughly one half of the disequilibrium error is corrected or eliminated in one period (or one quarter) – that is any movement of the error correction term above or below will be corrected after roughly two quarters.

Comments:

Part (a) asks about the definition and explanation of spurious relationship, leading to a discussion of co-integration. Candidates lost some marks for the lack of clarity of their definition.

Marks were also lost for the reasons behind the need to test for co-integration and key properties of the OLS.

Part (b) is a standard question about ECM and its derivation. Some candidates lost marks for the incorrect formula of ECM (missing some terms, such as the time lag difference in the div_t variable etc.). The interpretation of the terms was unclear in some cases.

Part (c) is about the estimation using OLS of an ECM. Only a few recognised the challenge of using OLS in this case. Most mentioned correctly the use of the Engle–Granger method and discussed it quite well.

Part (d) relies on the identification and interpretation of the ECM model. The key is to recognise that the speed is captured in the ρ term. Once candidates recognised it, the following interpretation was done well.

Question 7

Consider the following OLS estimation of a model to explain the labour demand of a large FTSE100 company using 120 observations from 2012m1 to 2021m12, all variables are calculated at the end of month t:

$$\widehat{hires}_t = \underset{(1.02)}{0.96} + \underset{(0.27)}{0.74}\log(profit_t) - \underset{(0.32)}{0.79}\log(research_t) + \underset{(0.12)}{0.34}\,news_t, \qquad (7.1)$$

$$n = 120,\ R^2 = 0.39,\ SSR = 1.59,\ F_{3,\,116} = 3.99$$

where $hires$ is the number of newly recruited employees in thousands (treated as a continuous variable), $profit_t$ is the profit before tax in millions of GBP, $research_t$ is the expenditure on research and development in millions of GBP, and $news_t$ is the number of newsletters issued by the company. The standard errors are reported in the parentheses, SSR is the Sum of Squared Residuals, and the F statistic for the significance of the regression is provided.

(a) **What is the interpretation of the coefficient on $\log(profit_t)$? Is the sign of the coefficient as you would expect? Explain your answer. Recall: hires is measured in thousands of employees.**

(2 marks)

(b) **Looking at the estimates, a colleague claims that the effect of profit before tax ($profit$) on the labour demand is exactly offset by the effect of research and development expenses ($research$) on the stock price. Describe a suitable test to examine this claim. Clearly specify the null and the alternative hypothesis and the assumptions you need to make. Indicate additional information, if necessary, to conduct such a test.**

(5 marks)

(c) **Another colleague gets hold of a variable $estimatedprofit_t$ which captures the average predicted profit of the company predicted by stock analysts. The colleague wanted to include $estimatedprofit_t$ in the model. He argued that in the past $estimatedprofit_t$ can predict the profit before tax result extremely well. Discuss the reasons behind the colleague's suggestion. What are the likely effects of the inclusion of $estimatedprofit_t$ on the properties of the OLS estimator of the parameter coefficient on $\log(profit)$? Explain your answers intuitively.**

(5 marks)

(d) **The research manager is concerned that the regression model is likely to suffer from heteroskedasticty. Clearly explain the term heteroskedasticity. What are going to be the consequences of heteroskedasticty on the OLS estimation of (7.1)?**

(3 marks)

(e) Discuss a suitable test for the presence of heteroskedasticity when the form of heteroskedasticity is unknown.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 4.2, 4.3, 5.2, 6.1, 6.2, 6.3, 6.5.

Dougherty, C. Subject guide (2016): Chapter 4, 5 and 6.

Approaching the question

(a) 1% increase in the profit is on average associated with an increase of $0.74 \times 1{,}000/100 = 7.4$ newly recruited employees. This is expected as higher profits can allow the company to further expand.

(b) The colleague claims that $\beta_1 + \beta_2 = 0$, where β_1, β_2 are the population parameters for $\log(profit)$ and $\log(research)$. The null hypothesis is $\mathrm{H}_0 : \beta_2 = -\beta_1$. The restriction is $\beta_2 = -\beta_1$, and the alternative is otherwise. At this point, there are three approaches:

- Candidates can do a t test, in which the covariance between $\mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_2)$ is needed, using:

$$t = \frac{\widehat{\beta}_1 + \widehat{\beta}_2}{se(\widehat{\beta}_1 + \widehat{\beta}_2)}.$$

  The t test follows a t distribution and details of the test are expected.
- Candidates also can create a new variable $X = \log(profit) - \log(research)$ and run the following model:

$$hire_i = \alpha + \theta \log(profit) + \beta_2 X + \beta_3 news_t + error$$

  then do a t test on $\theta = 0$.
- Candidates can also describe an F test with the restriction of $\beta_2 = -\beta_1$.
- Candidates need to clearly indicate the use of CLM assumptions, Gauss–Markov + normality.

(c) The colleague is concerned about the omission of a relevant explanatory variable, for example, the profit after tax (call it X). If the true model that explains the company's stock price has other variables not included in the estimations, the regression coefficients are going to be biased and the standard errors are invalid. The intuitive reason behind the bias is that in addition to the direct effect, the included variables have an apparent indirect effect as a consequence of acting as a proxy for the missing, omitted variable X. The strength of this proxy effect depends on two factors, the strength of the effect of X on the stock price, and the ability of the included variables to mimic X. If X and the included variables are correlated, the apparent indirect effect will be non-zero, causing the direct effect of the included variables to be affected, causing the OLS estimator to be biased.

However, including this predicted profits variable is unlikely to help because it is going to be highly correlated with the profit before tax. So including the predicted profits variable will not address the omission of a relevant variable problem, but will cause the OLS estimation to be less precise or the estimates to be more volatile (changing with small changes in the data).

(d) Heteroskedasticity refers to the situation when the variance of the error terms is dependent on the regressors. Heteroskedasticity in itself does not affect the unbiasedness nor consistency of an OLS estimator. The standard errors, however, now become invalid and the conventional testing inference will become invalid.

(e) Candidates are expected to do the following steps:

- Estimate the original model by OLS, as usual. Obtain the OLS residuals $\widehat{u}_i$. Compute the squared residuals $\widehat{u}_i^2$.
- Run the following regression and keep its R-squared, $R^2_{\widehat{u}^2}$:

$$\begin{aligned}\widehat{u}_i^2 = \delta_0 &+ \delta_1 \log(\textit{profit}) + \delta_2 \log(\textit{research}) + \delta_3 \textit{news} + \beta_4 \log(\textit{profit})^2 \\ &+ \delta_5 \log(\textit{research})^2 + \delta_6 \textit{news}^2 + \delta_7 \log(\textit{profit}) \times \log(\textit{research}) \\ &+ \delta_8 \log(\textit{profit}) \times \textit{news} + \delta_9 \log(\textit{research}) \times \textit{news} + \textit{error}.\end{aligned}$$

- The null hypothesis is $\mathrm{H}_0 : \delta_1 = \cdots = \delta_9 = 0$ and H_1 : otherwise. It's equivalent to a test of $k = 9$ restrictions.
- The test statistic is either the F statistic:

$$F = \frac{R^2_{\widehat{u}^2}/k}{(1 - R^2_{\widehat{u}^2})/(n-k-1)}$$

  or the LM statistic $LM = nR^2_{\widehat{u}^2}$ with $n = 130$.
- Under the null, the F statistic follows the $F_{k,\,n-k-1}$ distribution, and the LM statistic the χ^2_9 distribution. If the statistic exceeds the critical value from the respective distribution, we reject the null hypothesis of homoskedasticity.

Comments:

For part (a), most candidates got the correct prediction for the sign, but some interpreted the coefficient incorrectly, either missing the unit or a wrong approach to a linear-log regression.

For part (b), candidates were expected to perform a usual t test and explain how to obtain the test specifics. Most did well but as usual, some lost marks on missing some specifications. Some candidates did not notice the setting of a two-sided test or failed to discuss how to calculate the standard error. Some used an F test, which was also correct. Only a few mentioned and listed the Gauss–Markov assumptions clearly.

For part (c), most correctly identified the problem of omitting an important variable and an irrelevant variable and discussed intelligently the omitted variable bias problem. A good number of candidates realise $\textit{estimatedprofit}_t$ is a highly correlated covariate, hence creates some multicollinearity problem. Only a few carried on and discussed what would happen in this case (losing efficiency and precision).

For part (d), a discussion of heteroskedasticity is expected and many did well.

Part (e) asks a typical and classical test for heteroskedasticity. A clear description of steps was expected, some missed a few test specifics or forgot to explain how to obtain the residual for the testing equation.

Question 8

Consider the following regression model that describes the demand for Money in the USA:

$$\log(M_t^*) = \beta_0 + \beta_1 \log(R_t) + \beta_2 \log(Y_t) + u_t, \tag{8.1}$$

where, M_t^* is the desired (unobserved) or long run demand for Money (the real cash balances), R_t is the long-term interest (%, measured by the 90-day corporate rate of interest), and Y_t is the aggregate real national income (measured by the GDP), and u_t is an i.i.d. error term, and independent of $M_t, M_{t-1}, \ldots, R_t, R_{t-1}, \ldots, Y_t, Y_{t-1}$. The unobserved desired demand variable $\log(M_t^*)$ is determined by an adaptive process as follows:

$$\log(M_t) - \log(M_{t-1}) = \delta[\log(M_t^*) - \log(M_{t-1})], \tag{8.2}$$

where M_t is the actual demand for money at time t.

(a) What is the interpretation of δ in (8.2)? Demonstrate that the following short-run model of the demand function for money can be derived from models (8.1) and (8.2).

$$\log(M_t) = \gamma_0 + \gamma_1 \log(R_t) + \gamma_2 \log(Y_t) + \gamma_3 \log(M_{t-1}) + v_t, \tag{8.3}$$

where v_t is a disturbance term.

(5 marks)

(b) Carefully discuss the properties of the OLS estimators for (8.3) in terms of unbiasedness, consistency. Are your OLS standard errors valid? Provide suitable arguments for your claims, but no formal derivations are expected.

(5 marks)

(c) Discuss how we can test for the presence of autocorrelation in (8.3). We can assume the alternative vt can be approximated by a stationary AR(1).

(5 marks)

(d) Instead of fitting model (8.3), a researcher estimates the following model without $\log(M_{t-1})$.

$$\log(M_t) = \alpha_0 + \alpha_1 \log(R_t) + \alpha_2 \log(Y_t) + e_t, \tag{8.4}$$

Discuss the properties of the OLS estimators for model (8.4).

Hint: **Prove that the error term in model (8.4) exhibits autocorrelation.**

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 2.1, 2.2, 2.3, and 7.1, 7.2, 7.3.

Dougherty, C. Subject guide (2016): Chapters 2 and 7.

Approaching the question

(a) Rearrange (8.2), we have:

$$\begin{aligned}\log(M_t) &= \delta \log(M_t^*) + (1-\delta)\log(M_{t-1}) \\ &= \delta(\beta_0 + \beta_1 \log(R_t) + \beta_2 \log(Y_t) + u_t) + (1 - \delta \log(M_{t-1})) \\ &= \delta\beta_0 + \delta\beta_1 \log(R_t) + \delta\beta_2 \log(Y_t) + (1-\delta)\log(M_{t-1}) + \delta u_t.\end{aligned}$$

δ captures the percentage of the differences between the actual and desired real cash balances (demand for money) and is eliminated within a single period (here a year). It tells us how quickly the actual and desired demand for money converge.

(b) Because of the presence of $\log(M_{t-1})$ in the above model, we can not establish unbiasedness.

The assumptions for consistency are satisfied because for this we only need to ensure that $v_t = \delta u_t$ is uncorrelated with the regressors $\log(M_{t-1})$, $\log(R_t)$, $\log(Y_t)$ and X_{t-1}. We have the assumption of independence of the errors so they are uncorrelated. As long as there is no autocorrelation, we will also not have any correlation between $\log(M_{t-1})$ and u_t.

We don't need to use robust standard errors as there is no heteroskedasticity or autocorrelation.

(c) As the alternative v_t follows an AR(1) process, it can be written as $v_t = \rho v_{t-1} + \varepsilon_t$, where ε_t is a white noise and $|\rho| < 1$. The null hypothesis is $\mathrm{H}_0 : \rho = 0$ against $\mathrm{H}_1 : \rho \neq 0$. We will proceed as follows:

1. Run OLS on the original equation to obtain the residuals $\widehat{v}_t$, such that:

$$\widehat{v}_t = Y_t - \widehat{\gamma}_0 - \widehat{\gamma}_1 \log(R_t) - \widehat{\gamma}_2 \log(Y_t) - \widehat{\beta}_3 \log(M_{t-1}).$$

2. By the assumption that the alternative is a stationary AR(1) process, the test equation is:

$$\widehat{v}_t = \alpha_0^* + \alpha_1^* \log(R_t) + \alpha_2^* \log(Y_t) + \alpha_3^* \log(M_{t-1}) + \rho \widehat{v}_{t-1} + \varepsilon_t.$$

3. Finally, using the sample size, n, and its R^2 (goodness of fit measure), we obtain our test statistic $nR^2 \overset{a}{\sim} \chi_1^2$. We should reject the null of no autocorrelation if nR^2 exceeds the critical value of the χ_1^2 distribution at a given level of significance. Or, we use the asymptotic t test, given by $\widehat{\rho}/se(\widehat{\rho}) \overset{a}{\sim} N(0,1)$. We should reject the null if the absolute value of the t test exceeds the critical value of $N(0,1)$ at a given level of significance.

(d) The errors in model (8.4) that excluded the relevant variable $\log(M_{t-1})$ is:

$$e_t = \gamma_3 \log(M_{t-1}) + v_t$$

where $v_t = \delta u_t$ is assumed as a white noise. The omission will lead to autocorrelation. Indeed, we can show that, for example, $\text{Cov}(e_{t+1}, e_t) \neq 0$. Using our definition of e_t and e_{t+1}, we have:

$$\begin{aligned}
\text{Cov}(e_{t+1}, e_t) &= \text{Cov}(\gamma_3 \log(M_t) + v_{t+1}, \gamma_3 \log(M_{t-1}) + v_t) \\
&= \gamma_3^2 \text{Cov}(\log(M_t), \log(M_{t-1})) + \gamma_3 \text{Cov}(\log(M_t), v_t) \\
&\quad + \gamma_3 \text{Cov}(v_{t+1}, \log(M_{t-1})) + \text{Cov}(v_{t+1}, v_t) \\
&= \gamma_3 \text{Cov}(\log(M_t), v_t).
\end{aligned}$$

The last equality holds because v_t is white noise (i.i.d. errors unrelated to anything in the past) hence $\text{Cov}(v_{t+1}, v_t) = 0$ and $\text{Cov}(u_{t+1}, \log(M_{t-1})) = 0$. $\text{Cov}(\log(M_t), v_t) \neq 0$ because of the relationship between the two variables from (8.3). Therefore, we have $\text{Cov}(e_{t+1}, e_t) \neq 0$, or the error term exhibits autocorrelation. With the presence of autocorrelation, the weak exogeneity assumption will not hold in this model with a lagged dependent variable, the OLS estimators are no longer consistent. The standard errors are now invalid too (but really irrelevant).

Comments:

Part (a) requires the candidate to rearrange the equation to get the interpretation right. Some misinterpreted by not using elasticity, which is a result of a log-log equation. Some lost marks on not commenting on the speed of the convergence of the two matters. Perhaps the mixture of time series (superfically) and cross-sectional settings put off many candidates from attempting it.

Part (b) requires a discussion of OLS when having a lagged outcome (unbiasedness and consistency). Some candidates lost marks for missing either point. Some incorrectly mentioned the need of robust standard error. Note, we assumed away autocorrelation and heteroskedasticity in this case.

Part (c) is a standard testing question, most got the steps right, but some lost marks on the test distribution (chi-squared).

Part (d) is about autocorrelation, similar to one of the questions in Section A. Some did not make clear the definition and the proof. Some did not comment on the fact that OLS is inconsistent as a result.

Question 9

Let us consider a model where we want to explain whether a married woman works for wages or not. We may define the dependent variable paid to equal 1 when the woman works for wages, and zero otherwise, and consider the model:

$$\boldsymbol{Pr(paid = 1 \mid x) = G(\alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper^2)}, \qquad (9.1)$$

where G is the logistic cumulative density function where $G(z) = \exp(z)/(1 + \exp(z))$. The explanatory variables are *nwifeinc* denoting other sources of income in the household (in \$000), *educ* years of education, and *exper* years of past labour experience. Experience enters both linearly and quadratically.

(a) **Discuss the advantages and drawbacks of using the Logit model instead of a Linear Probability Model (LPM) model to explain whether a married woman works for wages or not. In your answer you are expected to introduce the LPM model.**

(5 marks)

(b) **Provide the log-likelihood for estimating the parameters in (9.1). Briefly discuss what happens to the parameters if instead of *paid*, the qualitative coding of the dependent variable is now *paid** as follows:**

$$paid^* = \begin{cases} 2 & \text{if married woman works for wages} \\ 1 & \text{if married women does not work for wages.} \end{cases}$$

(5 marks)

(c) **Discuss how you can obtain the partial effect of experience on the probability of an average woman working for wages. Briefly comment on how you can obtain the average partial effect (APE) for experience using your MLE parameter estimates.**

Hint: **The logistic probability density function is given $g(z) = G'(z) = \exp(z)/(1 + \exp(z))^2$.**

(5 marks)

(d) **A colleague claims that the marginal effect of education, *educ*, on whether a married woman is paid is significantly different for women living in the urban and rural areas. Assume you have a binary variable *urban* equal to 1 when the woman lives in a city, and 0 otherwise. Discuss how you can help the colleague examine the claim. You are expected to discuss a formal test in your answer.**

Hint: You are expected to modify (9.1) to permit the marginal effect of education to differ for women in urban and rural areas.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 4.1–4.3, 10.1, 10.3, 10.6.

Dougherty, C. Subject guide (2016): Chapters 4 and 10.

Approaching the question

(a) LPM is simply the OLS estimation of the following model:

$$paid = \alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper^2 + u.$$

When using OLS when the dependent variable is binary $(0, 1)$ (also called a LPM). Since $\mathrm{E}(y \mid x) = Pr(y = 1 \mid x)$, the probability is assumed to be linear in the parameters and this is a strong assumption to make.

The Logit model has various advantages over the LPM: Predicted probabilities are restricted to lie in $[0, 1]$, the MLE underlying Logit is (asymptotically) efficient whereas OLS (LPM) will be inefficient. The MLE automatically deals with heteroskedasticity, which is a problem under the LPM. Finally, the marginal effect is allowed to depend on individual characteristics, unlike being constant in the LPM.

The main drawback(s) of the Logit model relative to the LPM is that the coefficients cannot be directly interpreted as the marginal effects of the regressor(s) of interest, we need to compute the predicted probabilities using the Logit specification. Finally, it is also computationally more demanding (not nowadays).

(b) The probability of a married woman does not have paid work is:

$$Pr(paid = 0 \,|\, x) = 1 - G(\alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper^2).$$

The likelihood function for this sample is:

$$\begin{aligned} L(\alpha \,|\, \text{data}) &= \prod_{paid} Pr(paid = 1 \,|\, x) \prod_{unpaid} Pr(paid = 0 \,|\, x) \\ &= \prod_{paid} G(\alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper^2) \\ &\quad \times \prod_{unpaid} [1 - G(\alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper^2)]. \end{aligned}$$

We can then write down the log-likelihood function, $\log L(\alpha \,|\, x)$.

The likelihood function is identical as $Pr(paid = 1 \,|\, x) = Pr(paid^* = 2 \,|\, x)$:

$$\begin{aligned} L(\alpha^* \,|\, \text{data}) &= \prod_{paid} Pr(paid^* = 2 \,|\, x) \prod_{unpaid} Pr(paid^* = 1 \,|\, x) \\ &= \prod_{paid} G(\alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper^2) \\ &\quad \times \prod_{unpaid} [1 - G(\alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper^2)]. \end{aligned}$$

(c) The marginal effect of interest will tell us how the probability of working for wages $Pr(paid_i = 1)$ differs for women with different amounts of experience. The marginal effect is:

$$\frac{\partial Pr(paid = 1 \,|\, x)}{\partial exper} = (\alpha_3 + 2\alpha_4) \times g(z)$$

where $g(z) = \mathrm{e}^{-z}/(1 + \mathrm{e}^{-z})^2$ is the PDF of the logistic function, and:

$$z = \alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper^2.$$

We can compute the marginal effect for each individual. For the average woman, we can take the averages of the regressors (characteristics of the average woman), and then substitute these averages into the marginal effect formula.

The APE reports the average marginal effect over all individuals (here the sample size is n), which can be calculated as follows:

$$\begin{aligned} \frac{1}{n} \sum_{i=1}^{n} \frac{\partial Pr(paid_i = 1 \,|\, x_i)}{\partial exper} &= \frac{1}{n} \sum_{i=1}^{n} G(z_i) \times [1 - G(z_i)] \times (\alpha_3 + 2\alpha_4) \\ &= \frac{1}{n} \sum_{i=1}^{n} g(z_i) \times (\alpha_3 + 2\alpha_4). \end{aligned}$$

(d) Candidates are expected to discuss the inclusion of the dummy variable *urban* in the regression (including an interaction term is fine, but not needed). Since under probit, the marginal effects depend on individual characteristics so for the same woman the effect of education will now depend on whether she is in the rural or urban area when the dummy is included. The Probit regression is:

$$Pr(paid = 1 \mid x) = G(\theta_0 + \theta_1 nwifeinc + \theta_2 educ + \alpha_3 exper + \theta_4 exper^2 + \theta urban).$$

The null hypothesis is $\theta = 0$, against otherwise. We obtain the estimate θ and use the Wald test with the Wald statistic of $W = \widehat{\theta}/se(\widehat{\theta})$. This statistic follows a $N(0,1)$ distribution and we can simply compare the statistic with the critical value at a chosen significance level. If we reject the null, we have evidence that the effect of education is different for women in the urban area.

Comments:

Part (a) was well attempted in general. Marks were lost due to some small mistakes or a lack of explanation. The key is to explain the LPM as an OLS and the advantages and drawbacks associated with using OLS in this setting.

Part (b) was less well done, with the key being to write and explain the log-likelihood clearly.

Part (c) posed some challenges. Those who attempted did quite well, providing different forms of marginal effects used in interpreting Probit/Logit. The APE formula written by some candidates was missing notations or unclear with the final result $(\alpha_3 + 2\alpha_4)$.

Part (d) was an interesting question on the non-linear interpretation of a Probit/Logit. Most included an interaction term, which should not be needed and strictly not advised. Some realised that the Probit setup already allows for the non-linearity. The test specifics were discussed well but some missed one or two specifics. It is highly recommended that candidates learn well the testing procedures, perhaps mechanically, and apply them in the future.

UNIVERSITY OF LONDON

EC2020

BSc DEGREES AND GRADUATE DIPLOMAS IN ECONOMICS, MANAGEMENT, FINANCE AND THE SOCIAL SCIENCES, THE DIPLOMA IN ECONOMICS AND SOCIAL SCIENCES AND THE CERTIFICATE IN EDUCATION IN SOCIAL SCIENCES

Summer 2022 Online Assessment Instructions

EC2020 Elements of econometrics

Monday 30 May 2022: 09:00 - 13:15 (BST)

The assessment will be a **closed-book take-home online assessment within a 4-hour and 15-minute window**. The expected time/effort to answer all questions is **3 hours and 15 minutes**.

Candidates should answer EIGHT of the following NINE questions: all FIVE from Section A (8 marks each) and THREE from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Extracts from statistical tables are given after the final question on this paper.

You should complete this paper using pen and paper. Please use BLACK INK only.

Handwritten work then needs to be scanned, converted to PDF and then uploaded to the exam platform as ONE individual file. Please ensure that your candidate number is written

You may use any calculator for any appropriate calculations, but you may not use any computer software to obtain solutions. Credit will only be given if all workings are shown.

You have until 13:15 (BST) on Monday 30 May 2022 to submit your answers. However, you are advised not to leave your submission to the last minute in order to allow sufficient time to submit your work.

If you think there is any information missing or any error in any question, then you should indicate this but proceed to answer the question stating any assumptions you have made.

© University of London 2022

The assessment has been designed with a duration of 4 hours and 15 minutes to provide a more flexible window in which to complete the assessment. As a closed-book exam, the expected amount of effort required to complete all questions is no more than 3 hours and 15 minutes. Organise your time well. You are assured that in terms of answering all questions, there will be no benefit in you going beyond the expected 3 hours and 15 minutes of effort. Your assessment has been carefully designed to help you show what you have learned in the hours allocated.

This is a closed-book assessment but we recognise that some students have the ability to memorise sentences or entire paragraphs of text. As such, under closed-book conditions, examiners will be within their rights to consider plagiarism may have taken place if they encounter any strings of sentences taken from ANY source which are closely paraphrased or directly quoted and which are not referenced.

You are not expected to compile a full bibliography for any answer, but if you are recalling a direct quote from ANY source (including books, journals, internet sources, the subject guide, or lecture notes), then you must use quotation marks and add a reference to identify the author, and preferably also the year. If you are paraphrasing closely you must also attempt to reference where possible, and at least provide a basic reference to indicate the author. Please ensure that ahead of taking your examination, you have read the additional guidance on referencing under exam conditions which is available at:

Recognised Teaching Centre students - Avoiding plagiarism: Referencing in EMFSS examinations (london.ac.uk)

2U Online students – course wall on VLE (Digital Campus)

By accessing this question paper, you agree not to commit any assessment offence. Assessment offences include (but are not limited to) **committing plagiarism** and the **use or access of any paid-for or any other services offering live assistance** during an examination. You must not confer with anyone else during a live examination; and we take conferring to include any exchange of information or discussion about the assessment with others in any way that could potentially give you or another student an advantage in the examination. As such, any exchanging with others of exam questions; or any accessing of websites, blogs, forums **or any other form of oral or written communication with others which involves any discussion of live examination questions or potential answers/solutions to exam questions** will be considered an assessment offence.

The University of London will conduct checks to ensure the academic integrity of your work. Many students that break the University of London's assessment regulations did not intend to cheat but did not properly understand the University of London's regulations on referencing and plagiarism. **The University of London considers all forms of plagiarism, whether deliberate or otherwise, a very serious matter and can apply severe penalties that might impact on your award.**

© University of London 2022

The University of London's *Procedure for the Consideration of Allegations of Assessment Offences* is available online at:

Assessment Offence Procedures - University of London

© University of London 2022

UNIVERSITY OF LONDON

EC2020 Elements of econometrics

Candidates should answer EIGHT of the following NINE questions: all FIVE from Section A (8 marks each) and THREE from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Extracts from statistical tables are given after the final question on this paper.

Please find questions on the following page.

© University of London 2022

SECTION A

Answer all questions from this section.

1. Consider the following bivariate regression model:

$$Y_i = \beta\left(\frac{1}{X_i}\right) + u_i,$$

for a given random sample of observations $\{(Y_{i,}, X_i)\}_{i=1}^n$. The regressor is stochastic, whose sample variance is not 0, and $X_i \neq 0$ for all i. We may assume $E(u_i|\mathbf{X}) = 0$, where $\mathbf{X} = (X_1, ..., X_n)$.

(a) **(5 marks)** Is the following estimator

$$\tilde{\beta} = \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n Y_i^2}.$$

an unbiased estimator for β?
Hint: in your answer you need to treat x_i as a random variable, carefully derive $E[\tilde{\beta}|\mathbf{X}]$ first!

(b) **(3 marks)** You are advised that an unbiased estimator for β is given by

$$\hat{\beta} = \frac{\sum_{i=1}^n \left(\frac{Y_i}{X_i}\right)}{\sum_{i=1}^n \left(\frac{1}{X_i^2}\right)}.$$

Discuss how you can obtain this estimator. Is this estimator BLUE? Provide suitable arguments to support your answers.

2. Consider the following regression model:

$$y_i^* = \alpha + \beta x_i^* + u_i.$$

where the error term u has mean zero and variance σ_u^2, and u is independently distributed of x. You are told that both y_i^* and x_i^* are subject to the same measurement error w_i.

Instead of observing $\{(y_i^*, x_i^*)\}$, you are given a random sample $\{(y_i, x_i)\}_{i=1}^n$,where:

$$y_i = y_i^* + w_i,$$

and

$$x_i = x_i^* + w_i.$$

The measurement error w_i has zero mean, and is assumed to be distributed independently of u_i, x_i^* , and y_i^*.

(a) **(5 marks)** Let $\hat{\beta}$ be the OLS estimator of the slope of the linear regression of y_i on x_i with an intercept. Demonstrate that $\hat{\beta}$ is an asymptotically biased estimator of β. What is the sign of the bias?

(b) **(3 marks)** Discuss the following statement: Measurement error in regressor x^* poses a more serious problem than measurement error in the dependent variable y^*. Support your answer with suitable argument. No technical derivations expected.

3. A researcher considers the following regression model to determine the factors affecting rent rates in the UK, where $rent$ be the average monthly rent paid on rental units in a UK city. Let pop denote the total city population, $avginc$ the average city income, and $pctstu$ the student population as a percentage of the total population:

$$\log(rent) = \beta_0 + \beta_1 \log(pop) + \beta_2 \log(avginc) + \beta_3 pctstu + u.$$

The equation estimated using data from 64 cities is as follows, with standard errors reported in the parenthesis.

$$\begin{aligned} \widehat{\log(rent)} &= \underset{(.844)}{.043} + \underset{(.039)}{.066} \log(pop) + \underset{(.081)}{.507} \log(avginc) + \underset{(.0017)}{.0056} pctstu. \\ & \quad n = 64, \qquad R^2 = .458. \end{aligned}$$

(a) **(4 marks)** You are interested in testing the following joint hypothesis $H_0 : \beta_1 = \beta_2 = 0$. Explain how to implement the test for H_0, clearly describing any additional information you would require to perform the test.

(b) **(4 marks)** You are told that the regression model suffers from heteroskedasticity. Clearly explain what heteroskedasticity means and briefly indicate what implication this has for your test in (a). Discuss how you may address this problem. You are not expected to provide a test here.

4. (a) **(5 marks)** Consider the following time series model for $\{y_t\}_{t=1}^T$:

$$y_t = y_{t-1} + \varepsilon_t + \lambda \varepsilon_{t-1},$$

where ε_t is i.i.d with mean zero and variance σ^2, for $t = 1, ..., T$. Let $y_0 = 0$. Demonstrate that y_t is non-stationary unless $\lambda = -1$. In your answer, clearly provide the conditions for a covariance stationary process.

Hint: Apply recursive substitution to express y_t in terms of current and lagged errors.

(b) **(3 marks)** Briefly discuss the problem of applying the Dickey Fuller test when testing for a unit root when the model of a time series x_t is given by:

$$x_t = \rho x_{t-1} + u_t,$$

where the error term u_t exhibits autocorrelation. Clearly state what the null, alternative hypothesis, and the test statistics are for your test.

5. Let us consider the following stationary time series model:

$$y_t = \alpha + \rho y_{t-1} + \beta_1 x_t + \beta_2 x_{t-1} + \varepsilon_t, \qquad |\rho| < 1, \qquad t = 2, ..., T.$$

(a) **(2 marks)** Clearly indicate what assumptions will ensure that the OLS estimators $(\hat{\alpha}, \hat{\rho}, \hat{\beta}_1, \hat{\beta}_2)$ are consistent. You may assume stationarity and weak dependence needed to rely on usual law of large numbers are satisfied. A proof of consistency is not expected.

(b) **(4 marks)** Let us assume the conditions for consistency in (a) are satisfied. Define the short and long run effect that x has on y. Discuss how you can obtain a consistent estimator of these effects using $(\hat{\alpha}, \hat{\rho}, \hat{\beta}_1, \hat{\beta}_2)$, and prove your claim.

(c) **(2 marks)** Discuss the following statement: "If the x process is slowly changing over time it is difficult to get a precise estimator of the short run effect."
(*Hint:* You may for simplicity consider the model where $\rho = 0$).

SECTION B

Answer three questions from this section.

6. Consider the following OLS estimation of a model to explain the stock prices of a FTSE100 company using 120 observations from 2012m1 to 2021m12, all variables are calculated at the end of month t:

$$\widehat{\log(\textit{stock}_t)} = \underset{(1.12)}{0.86} + \underset{(0.24)}{0.54} \log(\textit{profit}_t) - \underset{(0.30)}{0.65} \log(\textit{research}_t) - \underset{(0.12)}{1.34} \log(\textit{marketing}_t) \quad (6.1)$$
$$n = 120,\ R^2 = 0.34, SSR = 1.29, F^3_{116} = 3.89.$$

where $stock_t$ is the stock price in GBP (British Sterling). $profit_t$ (the profit before tax in millions of GBP), $research_t$ (expenditure on research and development), and $marketing_t$ (expenditure on marketing) are measured in millions GBP. Standard errors are reported in parentheses, SSR is the Sum of Squared Residuals, and the F statistic for the significance of the regression is provided.

(a) **(5 marks)** What is the interpretation of the coefficient on $\log(\textit{profit}_t)$? Is the sign of the coefficient as you would expect? Explain your answer. Would your interpretation change if the profit is now measured in thousands of GBP? Explain your answer.
*Hint: a million is thousand*thousand*.

(b) **(5 marks)** Looking at the estimates, a colleague claims that the effect of marketing expenses is *more* than twice as large as the effect of research and development expenses on the stock price. Describe a suitable test to examine this claim. Clearly specify the null and the alternative hypothesis and assumptions underlying your test. Indicate additional information, if necessary, to conduct such a test.

(c) **(5 marks)** Another colleague gets hold of a variable $small_t$ which captures the monthly stock price of a small, random firm in Bulgaria from 2012m1 to 2021m12. Excited with the discovery, the colleague insists on including $small_t$ in the model. Discuss the statistical reasoning behind including additional variables in the model. What are the likely effects of the inclusion of $small_t$ on the properties of the OLS estimators of the parameters of the model? Explain your answers intuitively. *Hints:* The Bulgarian company is completely unrelated with the FTSE company in the model.

(d) **(5 marks)** Your research manager believes that the stock price action of the FTSE100 company behaves differently after 2016m6. He claims that the coefficients of the regressors in (6.1) for the months after 2016m6 are different from the coefficients in (6.1) for the months in and before 2016m6. Discuss how you can test whether the manager is correct. Clearly specify the null and the alternative hypothesis. Indicate additional information, if necessary, to conduct such a test.

7. A researcher models the relationship between the expenditure of a company, S_t in period t and the expected profit, π^*_{t+1}, in period $t+1$ as follows:

$$S_t = \beta_0 + \beta_1 \pi^*_{t+1} + \beta_2 r_t + u_t, \tag{7.1}$$

where r_t is the borrowing interest rate set by the central bank (measured in percentage) and u_t is an i.i.d. error term with

$$E(u_t | S_{t-1}, S_{t-2}, ..., \pi_t, \pi_{t-1}, ..., r_t, r_{t-1}, ...) = 0.$$

The expected profit is determined by the following adaptive expectation process:

$$\pi^*_{t+1} - \pi^*_t = \theta(\pi_t - \pi^*_t), \tag{7.2}$$

where π_t is the actual profit realised at time t. Using quarterly data from a US company, the researcher obtains the following estimates from using OLS:

$$\widehat{S}_t = \underset{(0.142)}{0.36} + \underset{(0.54)}{0.94}\pi_t - \underset{(2.85)}{34.65} r_t + \underset{(0.85)}{0.65} S_{t-1}, \tag{7.3}$$
$$n = 240,\ R^2 = 0.56.$$

(a) **(6 marks)** What is the interpretation of θ in (7.2)?. Using the regression results in (7.3) obtain an estimate for θ.

Hint: Use (7.1) and (7.2) to express S_t as follows:

$$S_t = \alpha_0 + \alpha_1 \pi_t + \alpha_2 r_t + \alpha_3 S_{t-1} + v_t, \tag{7.4}$$

where $v_t = u_t - (1-\theta)u_{t-1}$.

(b) **(5 marks)** You are concerned that the estimate for θ obtained in (a) is not suitable. Demonstrate formally that the OLS estimator of (7.4) will be inconsistent. *Hints:* You are not expected to look at the consistency proof for the α parameters explicitly.

(c) **(4 marks)** Discuss how you can use an IV estimator to obtain a consistent estimator for the α parameters and hence obtain a consistent estimator for θ.

(d) **(5 marks)** Suppose a suitable univariate model for S_t is given by:

$$S_t = \lambda_1 + \lambda_2 S_{t-1} + \gamma t + e_t, \tag{7.5}$$

where t is a deterministic trend and e_t is white noise, an i.i.d error term with zero mean and constant variance that is independent of S_{t-1}. Discuss how to test whether the expenditure process S_t has a unit root. Clearly indicate the null and the alternative hypothesis.

8. Let us consider a model where we want to explain whether a married woman works for wages or not. We may define the dependent variable *paid* to equal 1 when the woman works for wages, and zero otherwise, and consider the model

$$\textit{paid} = \alpha_0 + \alpha_1 \textit{nwifeinc} + \alpha_2 \textit{educ} + \alpha_3 \textit{exper} + \alpha_4 \textit{exper}^2 + u, \tag{8.1}$$

where *nwifeinc* is other sources of income (in \$000), *educ* is years of education, and *exper* is past years of labour experience. You may assume that $E\left(u|\textit{nwifeinc},\ \textit{educ},\ \textit{exper}\right) = 0$.

(a) **(2 marks)** Provide an economic interpretation of the parameter α_2. What should be the sign of α_2?

(b) **(3 marks)** What would happen to the parameter estimates if the qualitative outcome was coded differently, in particular:

$$\begin{aligned} \textit{paid}^* &= 2 \text{ if married woman works for wages} \\ &= 1 \text{ if married women does not work for wages,} \end{aligned}$$

and the following regression was estimated by OLS

$$\textit{paid}^* = \alpha_0 + \alpha_1^* \textit{nwifeinc} + \alpha_2 \textit{educ} + \alpha_3^* \textit{exper} + \alpha_4^* \textit{exper}^2 + u^*. \tag{8.2}$$

(c) **(6 marks)** Discuss why the following statements are False. Suggest suitable changes that will make the statements correct. Details supporting your claims will be rewarded.

(i) The R-squared obtained from the OLS estimation of (8.2) is larger than the R-squared obtained from the OLS estimation of (8.1).

(ii) To test whether the non-education variables $(nwifeinc, exper, exper^2)$ are jointly significant in (8.1), we can use the usual F-test that compares the restricted and unrestricted residual sum of squares.

(d) **(7 marks)** Instead of using OLS, a researcher proposes we should use the Probit model. Introduce the model and discuss the benefits and drawback of using this approach. In your answer you are expected to provide the main equation underlying the Probit model and the log-likelihood function.

(e) **(2 marks)** Discuss the following statement: "A probit model that uses $nwifeinc$, $educ$, $exper$, and $exper^2$ to explain whether a married woman works for wages or not, will exhibit diminishing effects of additional years of education on the probability of working for wages." In your answer clearly indicate how to obtain the partial effect under consideration.

9. Let us consider the following system of simultaneous equations

$$y_1 = \alpha_1 y_2 + \alpha_2 x_1 + \alpha_3 x_2 + u_1 \tag{9.1}$$

$$y_2 = \beta_1 y_1 + \beta_2 x_1 + \beta_3 x_3 + \beta_4 x_4 + u_2, \tag{9.2}$$

where (u_1, u_2) are i.i.d. errors with zero mean, $Var\,(u_1) = \sigma_1^2$, $Var(u_2) = \sigma_2^2$, and let $Cov(u_1, u_2) = 0$. The endogenous variables are (y_1, y_2) and (x_1, x_2, x_3, x_4) are exogenous variables.

(a) **(4 marks)** Explain the simultaneity issue associated with the above simultaneous equation model (SEM) intuitively (no derivations expected). Give an example that fits such a structure.

(b) **(7 marks)** You are told that in the above SEM there exists a perfect linear relation between x_2 and x_3, in particular,

$$x_3 = 2x_2.$$

Obtain the reduced form equations for y_1 and y_2, recognising this exact linear relation between x_2 and x_3.

(c) **(4 marks)** Use the result obtained in (b) to discuss the identification of the two structural equations. Clearly state whether the equations are over identified, exact (just) identified, or under identified. *Hint:* Your answer is expected to discuss what conditions need to be satisfied to ensure that we can use the observable data to estimate the parameters consistently.

(d) **(5 marks)** Discuss how you can obtain a consistent estimator of the α parameters using 2SLS. Provide details of the estimator.

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	**2·55**	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	**2·60**	0·99534	**2·75**	0·99702	**2·90**	0·99813	**3·05**	0·99886	**3·20**	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	**2·65**	0·99598	**2·80**	0·99744	**2·95**	0·99841	**3·10**	0·99903	**3·25**	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918	**3·30**	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	**3·263**	0·9994	**3·731**	0·99990	**3·916**	0·99995
3·105	0·9991	**3·320**	0·9995	**3·759**	0·99991	**3·976**	0·99996
3·138	0·9992	**3·389**	0·9996	**3·791**	0·99992	**4·055**	0·99997
3·174	0·9993	**3·480**	0·9997	**3·826**	0·99993	**4·173**	0·99998
3·215	0·9994	**3·615**	0·9998	**3·867**	0·99994	**4·417**	0·99999
			0·9999		0·99995		1·00000

When $x > 3·3$ the formula $1-\Phi(x) \doteqdot \frac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\frac{1}{x^2}+\frac{3}{x^4}-\frac{15}{x^6}+\frac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	**5·0**	1·6449	**3·0**	1·8808	**2·0**	2·0537	**1·0**	2·3263	**0·10**	3·0902
45	0·1257	**4·8**	1·6646	**2·9**	1·8957	**1·9**	2·0749	**0·9**	2·3656	**0·09**	3·1214
40	0·2533	**4·6**	1·6849	**2·8**	1·9110	**1·8**	2·0969	**0·8**	2·4089	**0·08**	3·1559
35	0·3853	**4·4**	1·7060	**2·7**	1·9268	**1·7**	2·1201	**0·7**	2·4573	**0·07**	3·1947
30	0·5244	**4·2**	1·7279	**2·6**	1·9431	**1·6**	2·1444	**0·6**	2·5121	**0·06**	3·2389
25	0·6745	**4·0**	1·7507	**2·5**	1·9600	**1·5**	2·1701	**0·5**	2·5758	**0·05**	3·2905
20	0·8416	**3·8**	1·7744	**2·4**	1·9774	**1·4**	2·1973	**0·4**	2·6521	**0·01**	3·7190
15	1·0364	**3·6**	1·7991	**2·3**	1·9954	**1·3**	2·2262	**0·3**	2·7478	**0·005**	3·8906
10	1·2816	**3·4**	1·8250	**2·2**	2·0141	**1·2**	2·2571	**0·2**	2·8782	**0·001**	4·2649
5	1·6449	**3·2**	1·8522	**2·1**	2·0335	**1·1**	2·2904	**0·1**	3·0902	**0·0005**	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

Examiners' commentaries 2022

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2021–22. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

General remarks

Learning outcomes

At the end of the course, and having completed the Essential reading and activities, you should be able to:

- describe and apply the classical regression model and its application to cross-section data
- describe and apply the:
  - Gauss–Markov conditions and other assumptions required in the application of the classical regression model
  - reasons for expecting violations of these assumptions in certain circumstances
  - tests for violations
  - potential remedial measures, including, where appropriate, the use of instrumental variables
- recognise and apply the advantages of logit, probit and similar models over regression analysis when fitting binary choice models
- competently use regression, logit and probit analysis to quantify economic relationships using standard regression programmes (Stata and EViews) in simple applications
- describe and explain the principles underlying the use of maximum likelihood estimation
- apply regression analysis to time-series models using stationary time series, with awareness of some of the econometric problems specific to time series applications (for example, autocorrelation) and remedial measures
- recognise the difficulties that arise in the application of regression analysis to nonstationary time series, know how to test for unit roots, and know what is meant by cointegration.

Common mistakes committed by candidates

Due to the Covid-19 pandemic, the examination took place online again this year. There was only one examination paper for both Zone A and Zone B. Similar to last year, we put serious effort to ensure the examination structure and the expected level of effort were comparable to those in normal years (the examination should take candidates three hours to complete). Realising many challenges due to the pandemic that candidates had to face, we commend the performance of many candidates this year. However, a good number of candidates submitted an attempt that revealed a serious lack of effort and preparation. We discuss below common mistakes made by candidates, most of them are repeated year after year. We then provide specific comments for each question.

Candidates should realise that $\frac{1}{n}\sum_{i=1}^{n} u_i$ is not the same as $\mathrm{E}(u_i)$. So, while we typically assume $\mathrm{E}(u_i) = 0$, this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} u_i = 0$. Also, while we may be happy to assume $\mathrm{E}(x_i u_i) = 0$ (uncorrelatedness between the errors and regressors), this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} x_i u_i = 0$. Note that:

- both $\frac{1}{n}\sum_{i=1}^{n} u_i$ and $\frac{1}{n}\sum_{i=1}^{n} x_i u_i$ are random variables, which take the value 0 with probability 0 (continuous random variables)!

- $\mathrm{E}(u_i) = 0$ and $\mathrm{E}(x_i u_i) = 0$ are fixed, not stochastic!

The differences between sample and population moments need to come across clearly when looking at unbiasedness and making consistency arguments. In both cases, we first simplify our estimator (plug in the true model) to obtain:

$$\widehat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \beta + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \quad \text{with } x_i = X_i - \bar{X}.$$

- For *unbiasedness*, clearly indicate that you want to show that $\mathrm{E}(\widehat{\beta}) = \beta$. Unbiasedness does not follow from $\sum_{i=1}^{n} x_i u_i = 0$, instead it follows from $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) = 0$.

  If we treat x_i as fixed:

  $$\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) \equiv \mathrm{E}\left(\sum_{i=1}^{n} d_i u_i\right) = \sum_{i=1}^{n} d_i \mathrm{E}(u_i)$$

  and then unbiasedness follows as $\mathrm{E}(u_i) = 0$.

- For *consistency*, clearly indicate that you want to show that $\mathrm{plim}(\widehat{\beta}) = \beta$. Using the plim

properties, we show:

$$\begin{aligned}
\operatorname{plim} \widehat{\beta} = \beta + \operatorname{plim} \left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \right) &= \beta + \frac{\operatorname{plim}\left(\frac{1}{n} \sum_{i=1}^{n} x_i u_i \right)}{\operatorname{plim}\left(\frac{1}{n} \sum_{i=1}^{n} x_i^2 \right)} \\
&\equiv \beta + \frac{\operatorname{plim}(\text{Sample Cov}(x,u))}{\operatorname{plim}(\text{Sample Var}(x))} \\
&= \beta + \frac{\operatorname{Cov}(x,u)}{\operatorname{Var}(x)} \quad \text{using the law of large numbers}
\end{aligned}$$

where $\operatorname{Cov}(x,u) = 0$ and $\operatorname{Var}(x) > 0$, ensuring we get consistency.

- Remember, the law of large numbers ensures that sample averages converge to their population analogues.

A large number of candidates are not able to clearly distinguish between sample variance and covariance, and population variance and covariance.

The use of $\operatorname{Cov}(X,Y)$ and $\operatorname{Var}(X)$ should be restricted to describing the population covariance and variances, respectively, with definitions:

$$\operatorname{Cov}(X,Y) = \operatorname{E}((X - \operatorname{E}(X))(Y - \operatorname{E}(Y))) = \operatorname{E}(XY) - \operatorname{E}(X)\operatorname{E}(Y)$$

and:

$$\operatorname{Var}(X) = \operatorname{E}((X - \operatorname{E}(X))^2) = \operatorname{E}(X^2) - (\operatorname{E}(X))^2$$

(you also may denote $\operatorname{Cov}(X,Y) = \sigma_{XY}$ and $\operatorname{Var}(X) = \sigma_X^2$). They are typically unknown, but fixed, quantities.

The sample covariance and variance are estimators of the population covariance and variance, respectively. They are defined as:

$$\text{Sample Cov}(X,Y) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})$$

and:

$$\text{Sample Var}(X) = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

(you also may use $\widehat{\sigma}_{XY}$ and $\widehat{\sigma}_X^2$). You can compute them given the data.

With a slight abuse of notation, we often divide by n instead, which is irrelevant if we let n be large. The division by $n-1$ is a finite sample issue only (unbiasedness).

The sample covariance and variance show up in our definition of the OLS estimator of the slope in the simple linear regression model, not the population covariance and variance, as:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} = \frac{\text{Sample Cov}(X,Y)}{\text{Sample Var}(X)} \neq \frac{\operatorname{Cov}(X,Y)}{\operatorname{Var}(X)}.$$

It was very common to misunderstand the central limit theorem. While it provides the distribution of the *estimator* $\widehat{\beta}$, no matter what the true distribution of the errors is, it does *not* say that the distribution of the error term is approximately normal in large samples.

Many candidates made the classic error of not recognising that $\operatorname{E}(X/Y) \neq \operatorname{E}(X)/\operatorname{E}(Y)$, in general.

Only a few candidates realised the differences in the proof for unbiasedness of an estimator under the assumption of non-stochastic and stochastic (random variable) regressors in Question 1. When

we assume a non-stochastic X, we can treat a function, say, a_i of $X_1, X_2, \ldots, X_n$ as a constant so that we can have $\mathrm{E}(a_i u_i) = a_i \mathrm{E}(u_i) = 0$. When X is assumed random (or stochastic), a_i is no longer a fixed constant but a random variable. We must first condition on $X = (X_1, X_2, \ldots, X_n)$ to take a_i out of the conditional expectation: $\mathrm{E}(a_i u_i \mid X) = a_i \mathrm{E}(u_i \mid X) = 0$. The Law of Iterated Expectation, $\mathrm{E}(\mathrm{E}(y \mid X)) = \mathrm{E}(y)$ where y is a random variable, should then be used to obtain the unconditional expectation of the estimator. Among those who realised the need on conditioning on a_i, many did not use the LIE to arrive at the final condition for unbiasedness, $\mathrm{E}(\widehat{\beta}) = \beta$.

Autocorrelation and heteroskedasticity were tricky topics for a good number of candidates. Many did not explain well the need to use robust standard errors or an asymptotic version of testing under autocorrelation and/or heteroskedasticity. Most also lost marks when discussing the properties of an estimator that can help solve these problems. Some incorrectly suggested that fixing autocorrelation will also resolve inconsistency and/or biasedness.

When discussing the importance of distinguishing trend stationary and difference stationary processes, two points are expected: discussing the implication for the long-run behaviour of the processes, and discussing statistical implications (the consistency of the estimators). Most candidates discussed the former very well, only a few touched upon the latter.

A good number of candidates seemed to be confused between the Dickey–Fuller test and an Augmented Dickey–Fuller (ADF). The difference between the two tests is in the test equation. In the augmented specification, we add sufficient lagged differences of the residuals to ensure that there is no autocorrelation left in the test equation regression. For this reason, some used the test equation for the ADF when the DF was called for. Related, others forgot that even though we can use a very similar test statistic to the t statistic in the Dickey–Fuller stationarity test (we call it τ statistic), it does not follow a t distribution but the Dickey–Fuller distribution instead.

Candidates struggled to give competent answers to the interpretation of empirical results. When interpreting an empirical result you should discuss the significance of the coefficients, magnitude, and sign of the coefficients. Related, *ceteris paribus* (or holding other factors constant) is needed to ensure that the estimate can be interpreted as the marginal change in the dependent variable given a change in the independent variable of interest. Many candidates lost marks from these mistakes.

Related, many candidates failed to discuss problems surrounding the omitting of important, relevant explanatory variables and irrelevant explanatory variable. While omitting the former will cause (omitted variable) bias, omitting the latter will only lead to the loss of efficiency.

Others struggled with interpreting coefficients from a non-linear regression (for example, log-linear, linear-log, or log-log).

Candidates particularly struggled with the difference between percentage points and % changes when it comes to the interpretation of coefficients in binary choice models. A coefficient in such a model gives the percentage point difference in the outcome variable when the regressor changes by 1 unit. The % change is a *relative* change that depends on the baseline value of the outcome variable (for example, the realisation when a regressor equals zero in the case of a dummy variable regressor). For example, if we have initially 50% of a sample buying ecological apples, a 10% increase will mean now $(1 + 0.1) \times 50 = 55\%$ of the sample buying ecological apples. A 10% point increase instead will result in $50 + 10 = 60\%$ of the sample choosing to buy ecological apples.

When discussing the marginal effect after a Probit estimation, candidates often did not specify the term z in the pdf $\phi(z)$. Some used the coefficients from the LPM regressions to calculate the marginal effects after a Probit/Logit estimation. A popular, related mistake was that candidates did not realise the inherent heteroskedasticity of LPM.

A number of candidates seemed to be confused between the conditions for a valid instrument (exclusion restriction and relevance) and factors related to obtaining a more precise IV estimate (a higher sample size, a higher sample variability of X_i, a lower variance of the error term, and a stronger correlation between X_i and Z_i). More seriously, most candidates did not realise the need to use an IV estimator when the regression has a lagged dependent variable as a regressor in the presence of autocorrelation.

Exogeneity and endogeneity are often very simply explained as variables 'outside of the model' or 'inside the model', respectively. While this approach is not wrong; it is not a particularly clear or helpful definition in general contexts that are not related to simultaneous equations models. Defining these concepts in terms of the relationship between regressor and the error term is much clearer and naturally leads to a way to show exogeneity or endogeneity by considering the covariance between regressors and error terms.

Candidates often missed that all exogenous regressors and instruments have to be included in the first stage when performing a TSLS estimation. Some candidates use the fitted values of the endogenous variable obtained from the first stage as the dependent variable in the second stage. This mistake is a serious misunderstanding. It is fine to indicate that this fitted value can be used as an instrument of the endogenous variable, but is not really the way the second stage of TSLS is typically understood.

When asked about identification of structural form equations, candidates often compared the number of excluded regressors to the number of endogenous variables in general, without pointing out which variables are the excluded regressors and could therefore be used as an instrument. Candidates were also often unclear of the definitions of over-identification, under-identification, and exact identification. The best approach to this type of questions is to discuss the endogenous and exogenous variables, and then think about the available exogenous variables that can be used as instruments. Some candidates discussed the rank condition, which is a necessary and sufficient condition for identification, but requires knowledge of matrix algebra.

When deriving the reduced forms, candidates should check conditions for their existence, in particular, whether the denominator of the expression is different from zero.

Gauss–Markov conditions were used in many places in the examination, for example, when discussing the BLUE property of an estimator or when conducting hypothesis tests. The Gauss–Markov conditions have to be explicitly specified. Only writing that the Gauss–Markov conditions hold is not sufficient. A good practice that many candidates have adopted is to begin the script by explicitly providing the Gauss–Markov conditions and refer back to them thereafter. Many candidates did not mention either the Gauss–Markov conditions or the normality assumption when conducting the t test.

When conducting hypothesis testing, ensure that you clearly indicate the null and alternative hypotheses (in terms of the true parameters, say β_1), the test statistic (in terms of the parameter estimates, here $\widehat{\beta}_1$), its distribution (with degrees of freedom), the rejection rule (one-sided or two-sided) for a given significance level (typically 5%) with suitable critical values, and provide an interpretation of your result. Candidates often lost marks for missing parts of these steps. Choosing the critical value appropriate to the sample size is also important. Several candidates used 1.96 (a critical value for very large degrees of freedom) in a small sample instead of the appropriate one. This mistake also cost marks. Related, many did not know when to use an asymptotic version of a test.

Another common mistake is to conduct a two-sided test when the statement was worded with a clear direction of the relationship. Many candidates fail to provide all the necessary specifics of the testing procedure (testing distribution, the critical value, the assumptions, the rejection rule, and the conclusion) for a t test, an F test, or a Chow test. These often resulted in lost marks.

Just as in previous years, many candidates did not answer all parts of a question. Make sure you read the questions properly and provide all details that are requested. Not answering a question will automatically earn you a zero. Some candidates probably did not have a good examination plan.

Key steps to improvement

Essential reading for this course includes the subject guide and the following:

- Dougherty, C. *Introduction to econometrics*. (Oxford: Oxford University Press, 2016) 5th

edition [ISBN 9780199676828]; http://oxfordtextbooks.co.uk/orc/dougherty5e/

Apart from the Essential readings you should do some supplementary reading. One very good book at the same level is:

- Gujarati, D.N. and D.C. Porter *Basic econometrics*. (McGraw–Hill, 2009, International edition) 5th edition [ISBN 9780071276252].

To understand the subject clearly it is important to supplement Dougherty's *Introduction to econometrics* (fifth edition) with the subject guide **EC2020 Elements of econometrics** (2016), especially Chapter 10 which covers maximum likelihood estimation. It is very important to carefully go through the subject guide. The subject guide contains solutions to the questions given in the main textbook and also some additional questions and solutions. Working through these will improve your understanding of the subject.

The chapter in the subject guide on maximum likelihood (Chapter 10) includes some additional theory which has not been covered in the main textbook. It is important to read the additional theory given in the subject guide to have a better understanding of the principles of maximum likelihood and tests based on the likelihood function.

Please check the VLE course page for resources for this subject such as a downloadable copy of the subject guide **EC2020 Elements of econometrics** (2016), PowerPoint slideshows that provide a graphical treatment of the topics covered in the textbook, datasets and statistical tables. Candidates should utilise datasets using standard regression programmes (STATA or EViews). This will help in the understanding of the subject.

Please note that the subject guide for EC2020 Elements of econometrics has been updated for the 2022–23 academic year and will be based on the textbook 'Introductory Econometris: A Modern Approach' by Wooldridge.

Examination revision strategy

Many candidates are disappointed to find that their examination performance is poorer than they expected. This may be due to a number of reasons, but one particular failing is '**question spotting**', that is, confining your examination preparation to a few questions and/or topics which have come up in past papers for the course. This can have serious consequences.

We recognise that candidates might not cover all topics in the syllabus in the same depth, but you need to be aware that examiners are free to set questions on **any aspect** of the syllabus. This means that you need to study enough of the syllabus to enable you to answer the required number of examination questions.

The syllabus can be found in the Course information sheet available on the VLE. You should read the syllabus carefully and ensure that you cover sufficient material in preparation for the examination. Examiners will vary the topics and questions from year to year and may well set questions that have not appeared in past papers. Examination papers may legitimately include questions on any topic in the syllabus. So, although past papers can be helpful during your revision, you cannot assume that topics or specific questions that have come up in past examinations will occur again.

If you rely on a question-spotting strategy, it is likely you will find yourself in difficulties when you sit the examination. We strongly advise you not to adopt this strategy.

Examiners' commentaries 2022

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2021–22. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE).

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2016). You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from **Section A** (8 marks each) and **THREE** from **Section B** (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Section A

Answer all questions from this section.

Question 1

Consider the following bivariate regression model:

$$Y_i = \beta\left(\frac{1}{X_i}\right) + u_i,$$

for a given random sample of observations $\{(Y_i, X_i)\}_{i=1}^n$. The regressor is stochastic, whose sample variance is not 0, and $X_i \neq 0$ for all i. We may assume $\mathrm{E}(u_i \mid X) = 0$, where $\mathrm{X} = (X_1, \ldots, X_n)$.

(a) **Is the following estimator**

$$\widetilde{\beta} = \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n Y_i^2}$$

an unbiased estimator of β?

Hint: **in your answer you need to treat x_i as a random variable, carefully derive $\mathrm{E}[\widetilde{\beta} \mid \mathrm{X}]$ first!**

(5 marks)

(b) You are advised that an unbiased estimator of β is given by

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}\left(\frac{Y_i}{X_i}\right)}{\sum_{i=1}^{n}\left(\frac{1}{X_i^2}\right)}.$$

Discuss how you can obtain this estimator. Is this estimator BLUE? Provide suitable arguments to support your answers.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 2.1, 2.2, 4.1, 4.2, 8.1.

Dougherty, C. Subject guide (2016): Chapters 3 and 8.

Approaching the question

(a) Candidates need to condition on $\mathbf{X}$, because $\mathbf{X}$ is now assumed stochastic (random). Not taking the condition will be penalised heavily. With that in mind, we rewrite the estimator as:

$$\widetilde{\beta} = \frac{\sum_{i=1}^{n} X_i Y_i}{\sum_{i=1}^{n} Y_i^2} = \frac{n\beta}{\sum_{i=1}^{n} Y_i^2} + \frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} Y_i^2} \qquad (*)$$

where the first equality follows from the definition of $\widetilde{\beta}$, and the second equality follows from:

$$Y_i = \beta\left(\frac{1}{X_i}\right) + u_i.$$

Let $\mathbf{X} = (X_1, \ldots, X_n)$. Taking the conditional expectation, we have:

$$\mathrm{E}[\widetilde{\beta}\,|\,\mathbf{X}] = \mathrm{E}\left[\frac{n\beta}{\sum_{i=1}^{n} Y_i^2} + \frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} Y_i^2}\,\Big|\,\mathbf{X}\right] = n\beta\mathrm{E}\left[\frac{1}{\sum_{i=1}^{n} Y_i^2}\,\Big|\,\mathbf{X}\right] + \mathrm{E}\left[\frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} Y_i^2}\,\Big|\,\mathbf{X}\right]$$

where the first equality follows from (*), and the second equality follows from the property of conditional expectation.

Now, notice that $\mathrm{E}(X/Y) \neq \mathrm{E}(X)/\mathrm{E}(Y)$, in general, unless $\mathrm{Cov}(X/Y, Y) = 0$, we cannot go further from here. Therefore:

$$\mathrm{E}[\widetilde{\beta}\,|\,\mathbf{X}] \neq \beta$$

and $\widetilde{\beta}$ is biased, unless:

$$n\mathrm{E}\left[\frac{1}{\sum_{i=1}^{n} Y_i^2}\,\Big|\,\mathbf{X}\right] = 1 \quad \text{and} \quad \mathrm{E}\left[\frac{\sum_{i=1}^{n} X_i u_i}{\sum_{i=1}^{n} Y_i^2}\,\Big|\,\mathbf{X}\right] = 0$$

as unbiasedness would require (law of iterated expectations):

$$\mathrm{E}(\widetilde{\beta}) \equiv \mathrm{E}\left(\mathrm{E}[\widetilde{\beta}\,|\,\mathbf{X}]\right) = \beta.$$

The conclusion is that $\widetilde{\beta}$ is a biased estimator, in general.

(b) Candidates are expected to recognise that this estimator is obtained by regressing Y_i on $1/X_i$ (without intercept). A derivation of this OLS estimator should be shown, for example. The OLS without intercept will minimise:

$$RSS = \sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2 = \sum_{i=1}^{n}\left(Y_i - \widehat{\beta}\left(\frac{1}{X_i}\right)\right)^2.$$

Then candidates are expected to show or discuss the first-order condition to derive the estimator $\widehat{\beta}$.

The Gauss–Markov theorem states that OLS is BLUE when all Gauss–Markov assumptions are satisfied. In general, we do not expect the estimator to be BLUE because we have not made any assumptions on homoskedasticity.

Comments:

Part (a) was quite unusual in two senses: we asked for a proof with stochastic regressors and we use a non-standard estimator. Most candidates lost marks for not realising that $\mathrm{E}(X)/\mathrm{E}(Y) \neq \mathrm{E}(X/Y)$, in general. Most also failed to take into account the hint provided to treat the regressors as stochastic. Many overcomplicated the process and led to unnecessary complications and an incorrect conclusion. Noticeably, some mistook the estimator and treated the Y_i term in the denominator as X_i, which would have been a standard discussion indeed. Those candidates were awarded a few marks if they solved it well. For completeness, we provide below the solution, had the estimator been $\widehat{\beta} = \dfrac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2}$ instead.

We rewrite the estimator as:

$$\widetilde{\beta} = \frac{\sum_{i=1}^n X_i Y_i}{\sum_{i=1}^n X_i^2} = \frac{n\beta}{\sum_{i=1}^n X_i^2} + \frac{\sum_{i=1}^n X_i u_i}{\sum_{i=1}^n X_i^2} \qquad (**)$$

where the first equality follows from the definition of $\widetilde{\beta}$, and the second equality follows from:

$$Y_i = \beta\left(\frac{1}{X_i}\right) + u_i.$$

Let $\mathbf{X} = (X_1, \ldots, X_n)$. Taking the conditional expectation, we have:

$$\begin{aligned}
\mathrm{E}[\widetilde{\beta} \,|\, \mathbf{X}] = \mathrm{E}\left[\frac{n\beta}{\sum_{i=1}^n X_i^2} + \frac{\sum_{i=1}^n X_i u_i}{\sum_{i=1}^n X_i^2} \,\middle|\, \mathbf{X}\right] &= \frac{n\beta}{\sum_{i=1}^n X_i^2} + \frac{\mathrm{E}[\sum_{i=1}^n X_i u_i \,|\, \mathbf{X}]}{\sum_{i=1}^n X_i^2} \\
&= \frac{n\beta}{\sum_{i=1}^n X_i^2} + \frac{\sum_{i=1}^n X_i \mathrm{E}(u_i \,|\, \mathbf{X})}{\sum_{i=1}^n X_i^2} \\
&= \frac{n\beta}{\sum_{i=1}^n X_i^2}
\end{aligned}$$

where the first equality follows from (**), the second and third equalities follow from the property of conditional expectation, and the fourth equality follows from the population model. Therefore:

$$\mathrm{E}[\widetilde{\beta} \,|\, \mathbf{X}] = \frac{n\beta}{\sum_{i=1}^n X_i^2} \neq \beta$$

and $\widetilde{\beta}$ is biased, unless:

$$\mathrm{E}\left[\frac{n}{\sum_{i=1}^n X_i^2}\right] = 1$$

as unbiasedness would require (law of iterated expectations):

$$\mathrm{E}(\widetilde{\beta}) \equiv \mathrm{E}\left(\mathrm{E}[\widetilde{\beta} \,|\, \mathbf{X}]\right) = \beta.$$

For part (b), most candidates lost marks for simply stating that this estimator is an OLS without showing how to get there. Many discussed the Gauss–Markov theorem. Marks were deducted if candidates made no distinct discussion on heteroskedasticity and simply assumed the Gauss–Markov theorem.

Question 2

Consider the following regression model:

$$\boldsymbol{y_i^* = \alpha + \beta x_i^* + u_i}$$

where the error term u has mean zero and variance σ_u^2, and u is independently distributed of x. You are told that both y_i^* and x_i^* are subject to the same measurement error w_i.

Instead of observing $\{(y_i^*, x_i^*)\}$, you are given a random sample $\{(y_i, x_i)\}_{i=1}^n$, where:

$$y_i = y_i^* + w_i,$$

and

$$x_i = x_i^* + w_i.$$

The measurement error w_i has zero mean, and is assumed to be distributed independently of u_i, x_i^*, and y_i^*.

(a) Let $\widehat{\beta}$ be the OLS estimator of the slope of the linear regression of y_i on x_i with an intercept. Demonstrate that $\widehat{\beta}$ is an asymptotically biased estimator of β. What is the sign of the bias?

(5 marks)

(b) Discuss the following statement: Measurement error in regressor x^* poses a more serious problem than measurement error in the dependent variable y^*. Support your answer with suitable argument. No technical derivations expected.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 8.3, 8.4.

Dougherty, C. Subject guide (2016): Chapter 8.

Approaching the question

(a) Substituting for the population model we have:

$$y_i - w_i = \alpha + \beta(x_i - w_i) + u_i.$$

Thus:

$$\begin{aligned} y_i &= \alpha + \beta x_i + u_i + (1-\beta)w_i \\ &= \alpha + \beta x_i + v_i \end{aligned}$$

where $v_i = u_i + (1-\beta)w_i$. The OLS estimator $\widehat{\beta}$ is:

$$\widehat{\beta} = \beta + \frac{\sum_{i=1}^n (x_i - \bar{x})(v_i - \bar{v})}{\sum_{i=1}^n (x_i - \bar{x})^2}.$$

Take the probability limit of $\widehat{\beta}$, we have:

$$\begin{aligned} \text{plim } \widehat{\beta} &= \beta + \frac{\text{plim } n^{-1}\sum_{i=1}^n (x_i - \bar{x})(v_i - \bar{v})}{\text{plim } n^{-1}\sum_{i=1}^n (x_i - \bar{x})^2} \\ &= \beta + \frac{\text{Cov}(x, v)}{\text{Var}(x)} \\ &= \widehat{\beta} + \frac{\text{Cov}(x^* + w, u + (1-\beta)w)}{\text{Var}(x^* + w)} \\ &= \beta + \frac{\text{Cov}(x^*, u) + \text{Cov}(w, u) + (1-\beta)[\text{Cov}(x^*, w) + \text{Cov}(w, w)]}{\text{Var}(x^*) + 2\text{Cov}(x^*, w) + \text{Var}(w)} \\ &= \beta + (1-\beta)\frac{\text{Var}(w)}{\text{Var}(x^*) + \text{Var}(w)}. \end{aligned}$$

The last equality comes from the fact that w is independent of x^* and u, while $\text{Cov}(x^*, u) = 0$. Unless $\beta = 1$, the estimator $\widehat{\beta}$ is asymptotically biased. The sign of the bias depends on the sign of $(1-\beta)$.

(b) It depends on the nature of the measurement errors. Typically measurement error in the dependent variable is seen as less of a problem, contributing only to the error term, leaving the estimator consistent but less precise. However, if the measurement error in the dependent variable is correlated with the regressor, then also measurement error in the dependent variable will yield inconsistency. Measurement error in the explanatory variables will lead to inconsistency in general (attenuation bias).

Comments:

For part (a), candidates lost marks for not using the plim in their proof or missing steps. Some made unfortunate algebraic mistakes and arrived at an incorrect bias formula. For part (b), most realised the differences in the consequences of measurement errors in the dependent variable and in the explanatory variable. Almost no candidate discussed the non-random measurement error in the dependent variable.

Question 3

A researcher considers the following regression model to determine the factors affecting rent rates in the UK, where *rent* is the average monthly rent paid on rental units in a UK city. Let *pop* denote the total city population, *avginc* the average city income, and *pctstu* the student population as a percentage of the total population:

$$\log(rent) = \beta_0 + \beta_1 \log(pop) + \beta_2 \log(avginc) + \beta_3 pctstu + u.$$

The equation estimated using data from 64 cities is as follows, with standard errors reported in the parentheses.

$$\log(rent) = \underset{(.844)}{.043} + \underset{(.039)}{.066} \log(pop) + \underset{(.081)}{.507} \log(avginc) + \underset{(.0017)}{.0056} pctstu.$$

$$n = 64, \ R^2 = .458.$$

(a) You are interested in testing the following joint hypothesis $H_0 : \beta_1 = \beta_2 = 0$. Explain how to implement the test for H_0, clearly describing any additional information you would require to perform the test.

(4 marks)

(b) You are told that the regression model suffers from heteroskedasticity. Clearly explain what heteroskedasticity means and briefly indicate what implication this has for your test in (a). Discuss how you may address this problem. You are not expected to provide a test here.

(4 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 2.1, 2.2, 7.1, 7.2.

Dougherty, C. Subject guide (2016): Chapters 2 and 7.

Approaching the question

(a) The R^2 reported above can be used as the one for the unrestricted model. So the additional information we need is the R^2 for the restricted regression model:

$$\log(rent) = \beta_0 + \beta_3 pctstu + u.$$

Let R_r^2 be the R^2 obtained from this regression, and let R_{ur}^2 be the R^2 obtained from the original regression (i.e. $R_{ur}^2 = .458$). Then the null hypothesis H_0 can be tested by the following F statistic:

$$F = \frac{(R_{ur}^2 - R_r^2)/2}{(1 - R_{ur}^2)/60}.$$

Under the Classical Linear Regression Assumptions (Gauss–Markov assumptions and normality) (candidates must state these somewhere in the script), the F statistic follows the $F_{2,60}$ distribution under H_0. So, at the significance level 0.05 (say), the rejection rule is given by $F > c$ where c is the 95th percentile of the $F_{2,60}$ distribution.

(b) Heteroskedasticity refers to the situation when the variance of the error terms (or the outcome variable) is dependent ons the regressors. Under the presence of heteroskedasticity, the usual standard errors from the OLS regressions become invalid and the F test results are no longer valid. To resolve this problem, we can use the White robust standard errors, which will give the valid results for the standard error calculations. For the F test, it is only asymptotically valid, and we need to use the chi-squared distribution for the test statistic. Candidates can simply mention the robust F statistic to get the mark.

Comments:

The question is a standard hypothesis testing and discussing the problems of heteroskedasticity. Marks were lost due to missing test specifics or unclear explanation.

Question 4

(a) Consider the following time series model for $\{y_t\}_{t=1}^T$

$$y_t = y_{t-1} + \varepsilon_t + \lambda\varepsilon_{t-1},$$

where ε_t is i.i.d. with mean zero and variance σ^2, for $t = 1, \ldots, T$. Let $y_0 = 0$. Demonstrate that y_t is non-stationary unless $\lambda = -1$. In your answer, clearly provide the conditions for a covariance stationary process.

Hint: **Apply recursive substitution to express y_t in terms of current and lagged errors.**

(5 marks)

(b) Briefly discuss the problem of applying the Dickey–Fuller test when testing for a unit root when the model of a time series x_t is given by:

$$x_t = \rho x_{t-1} + u_t,$$

where the error term u_t exhibits autocorrelation. Clearly state what the null, alternative hypothesis, and the test statistics are for your test.

(3 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 11.4, 12.1, 12.3, 13.1 and 13.5.

Dougherty, C. Subject guide (2016): Chapters 11, 12, and 13.

Approaching the question

(a) For a process y_t to be stationary, its expected value, variance, and covariance of y_t and y_{t+s} are independent of t. We need to find λ so that all the terms are independent of time.

- Lagging one period and then substituting back to the model, until time $t = 0$, we have:

$$\begin{aligned} y_t &= \varepsilon_t + \lambda\varepsilon_{t-1} + y_{t-1} \\ &= \varepsilon_t + (\lambda+1)\varepsilon_{t-1} + \lambda\varepsilon_{t-2} + y_{t-2} \\ &= \varepsilon_t + (\lambda+1)\varepsilon_{t-1} + (\lambda+1)\varepsilon_{t-2} + \lambda\varepsilon_{t-3} + y_{t-3} \\ &\vdots \\ &= \varepsilon_t + \sum_{s=1}^{t-1}(\lambda+1)\varepsilon_{t-s} + \lambda\varepsilon_0 + y_0 \\ &= \varepsilon_t + \sum_{s=1}^{t-1}(\lambda+1)\varepsilon_{t-s} + \lambda\varepsilon_0. \end{aligned}$$

- The expectation of y_t is zero, which does not depend on time.
- The variance of y_t is:

$$\mathrm{Var}(y_t) = \sigma^2[1 + (t-1)(1+\lambda)^2 + \lambda^2]$$

  using the rules of sum of variance and independence of errors. This variance does not depend on t when $\lambda = -1$.
- Finally, we consider $\mathrm{Cov}(y_{t+h}, y_t)$. Here:

$$\begin{aligned} \mathrm{Cov}(y_{t+h}, y_t) &= \mathrm{Cov}\left(\varepsilon_{t+h} + \sum_{s=1}^{t+h-1}(\lambda+1)\varepsilon_{t+h-s} + \lambda\varepsilon_0, \varepsilon_t + \sum_{s=1}^{t-1}(\lambda+1)\varepsilon_{t-s} + \lambda\varepsilon_0\right) \\ &= (\lambda+1)\mathrm{Var}(\varepsilon_t) + (\lambda+1)^2\sum_{s=1}^{t-1}\mathrm{Var}(\varepsilon_{t-s}) + \lambda^2\mathrm{Var}(\varepsilon_0) \\ &= \sigma^2[\lambda + 1 + (t-1)(\lambda+1)^2 + \lambda^2]. \end{aligned}$$

  This covariance also does not depend on t when $\lambda = -1$.

(b) We must first rewrite the regression model for a unit root test:

$$\Delta x_t = \alpha + (\rho - 1)x_{t-1} + u_t.$$

We run the following equivalent regression:

$$\Delta x_t = \alpha + \gamma x_{t-1} + u_t.$$

The null hypothesis that $\rho = 1$ is equivalent to a test of $\mathrm{H}_0 : \rho = 1$ (or $\gamma = 0$). The alternative hypothesis: $\rho < 1$ (or $\gamma < 0$). The main problem with the test is autocorrelation. The test statistic is:

$$\frac{\widehat{\gamma}}{s.e.(\widehat{\gamma})}.$$

Under the usual assumptions, the test statistic follows a Dickey–Fuller distribution reported in the Dickey–Fuller table. We will reject the null if $\widehat{\gamma}/s.e(\widehat{\gamma}) < \tau$, where τ is the Dickey–Fuller critical value at the 5% significance level (-2.86). Note that we cannot use a t distribution because under the null the distribution is not the Student's t distribution.

The main problem with a DF test with autocorrelation is that the standard errors are invalid and we must get rid of autocorrelation before performing the test. A common approach is to perform an ADF test and include lags into the regression. Instead of presenting the above DF test, candidates can present an ADF test assuming the autocorrelation has an AR(1) form.

Comments:

Part (a) is a standard derivation question to show stationarity. Most candidates were able to provide the conditions for stationarity but struggled with correct steps of the derivation (despite

the clear hint). Part (b) is about the Dickey–Fuller test and problems when there is autocorrelation. Only a few candidates realised we need to get rid of autocorrelation before the test, or to use the ADF version, assuming a form of the autocorrelation.

Question 5

Let us consider the following stationary time series model:

$$\boldsymbol{y_t = \alpha + \rho y_{t-1} + \beta_1 x_t + \beta_2 x_{t-1} + \varepsilon_t, \quad |\rho| < 1,\ t = 2, \ldots, T.}$$

(a) Clearly indicate what assumptions will ensure that the OLS estimators $(\widehat{\alpha}, \widehat{\rho}, \widehat{\beta}_1, \widehat{\beta}_2)$ are consistent. You may assume stationarity and weak dependence needed to rely on usual law of large numbers are satisfied. A proof of consistency is not expected.

(2 marks)

(b) Let us assume the conditions for consistency in (a) are satisfied. Define the short and long run effect that x has on y. Discuss how you can obtain a consistent estimator of these effects using $(\widehat{\alpha}, \widehat{\rho}, \widehat{\beta}_1, \widehat{\beta}_2)$, and prove your claim.

(4 marks)

(c) Discuss the following statement: 'If the x process is slowly changing over time it is difficult to get a precise estimator of the short run effect.'

(Hint: You may for simplicity consider the model where $\rho = 0$).

(2 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 11.3, 11.4, 11.5, 11.7.

Dougherty, C. Subject guide (2016): Chapter 11.

Approaching the question

(a) We would require the following assumptions, all of which rely on large sample properties.

The regressors are weakly persistent (C2), no multicollinearity (C3). For consistency, ε_t: the regressors are weak dependent and (covariance) stationary, and we need to assume weak exogeneity: because of the lagged outcome variable in the regression, the error term must not exhibit autocorrelation (candidates can explain the weak exogeneity assumption, or discuss the non-autocorrelation assumption). (For unbiasedness, the error terms must be independent of all regressors, current, past and future (C7). Because of the presence of y_{t-1}, this assumption cannot hold.)

(b) The short-term effect of x_t on y_t is simply β_1. The cumulative (long-term) effect of x on y could be represented as:

$$y = \alpha + \rho y + (\beta_1 + \beta_2)x.$$

The long-run effect of x on y is $(\beta_1 + \beta_2)/(1 - \rho)$, this is the effect of a one-unit change in equilibrium x on equilibrium y. Since the estimators of the coefficients are consistent, we can simply obtain these effects. Candidates are expected to show the derivation of these effects.

(c) When $\rho = 0$, the long-run effect simply becomes $\beta_1 + \beta_2$. Because x_t and x_{t-1} are likely to be highly correlated, the OLS estimation in the short run might suffer from multicollinearity, which will reduce the precision of the estimates. In the long run, it is more likely to get a more precise estimate. Intuitively, an effect is more likely to be more detectable over the long run.

Comments:

For part (a), most candidates did well recognising the assumptions needed. For part (b), marks were only lost due to lack of explanation. For part (c), candidates mostly did not recognise the new long-run effect and then discussed the multicollinearity between x_t and x_{t-1} when x_t changes very slowly. Admittedly, this interpretive question was tricky.

Section B

Answer three questions from this section

Question 6

Consider the following OLS estimation of a model to explain the stock prices of a FTSE100 company using 120 observations from 2012m1 to 2021m12, all variables are calculated at the end of month t:

$$\log(stock_t) = \underset{(1.12)}{0.86} + \underset{(0.24)}{0.54}\log(profit_t) - \underset{(0.30)}{0.65}\log(research_t) - \underset{(0.12)}{1.34}\log(marketing_t) \tag{6.1}$$

$$n = 120,\ R^2 = 0.34,\ SSR = 1.29,\ F_{3,116} = 3.89$$

where $stock_t$ is the stock price in GBP (British Sterling). $profit_t$ (the profit before tax in millions of GBP), $research_t$ (expenditure on research and development), and $marketing_t$ (expenditure on marketing) are measured in millions GBP. Standard errors are reported in parentheses, SSR is the Sum of Squared Residuals, and the F statistic for the significance of the regression is provided.

(a) **What is the interpretation of the coefficient on $\log(profit_t)$? Is the sign of the coefficient as you would expect? Explain your answer. Would your interpretation change if the profit is now measured in thousands of GBP? Explain your answer.**

Hint: a million is thousand $*$ thousand.

(5 marks)

(b) **Looking at the estimates, a colleague claims that the effect of marketing expenses is *more* than twice as large as the effect of research and development expenses on the stock price. Describe a suitable test to examine this claim. Clearly specify the null and the alternative hypothesis and assumptions underlying your test. Indicate additional information, if necessary, to conduct such a test.**

(5 marks)

(c) **Another colleague gets hold of a variable $small_t$ which captures the monthly stock price of a small, random firm in Bulgaria from 2012m1 to 2021m12. Excited with the discovery, the colleague insists on including $small_t$ in the model. Discuss the statistical reasoning behind including additional variables in the model. What are the likely effects of the inclusion of $small_t$ on the properties of the OLS estimators of the parameters of the model? Explain your answers intuitively. *Hint:* The Bulgarian company is completely unrelated with the FTSE company in the model.**

(5 marks)

(d) **Your research manager believes that the stock price action of the FTSE100 company behaves differently after 2016m6. He claims that the coefficients of the regressors in (6.1) for the months after 2016m6 are different from the coefficients in (6.1) for the months in and before 2016m6. Discuss how you can**

test whether the manager is correct. Clearly specify the null and the alternative hypothesis. Indicate additional information, if necessary, to conduct such a test.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 4.3, 5.4, 6.1, 6.2, 6.3, 6.5.

Dougherty, C. Subject guide (2016): Chapter 4 and 6 (with some reading on the Chow test in Chapter 5).

Approaching the question

(a) Everything else equal (*ceteris paribus*), a 1% increase in the monthly profits is associated with on average an increase of 0.54% in the monthly stock price. The change in the unit of measurement would not affect the interpretation of the coefficient, because we can think of the interpretation as the elasticity of the stock price. Another way to explain is to mathematically derive the new estimates (not necessary).

(b) Let β_2 and β_3 be the population parameters for the coefficients of $\log(research_t)$ and $\log(marketing_t)$, respectively. We are interested in testing $\mathrm{H}_0 : \beta_3 = 2\beta_2$ against the alternative hypothesis $\mathrm{H}_1 : \beta_3 > 2\beta_2$.

The t statistic to test the null hypothesis is written as:

$$t = \frac{\widehat{\beta}_3 - 2\widehat{\beta}_2}{s.e.(\widehat{\beta_3} - 2\widehat{\beta}_2)}.$$

From the above output, we know $\widehat{\beta}_3 - 2\widehat{\beta}_2$. On the other hand, the standard error of $\widehat{\beta}_3 - 2\widehat{\beta}_2$ is given by:

$$se(\widehat{\beta}_3 - 2\widehat{\beta}_2) = \sqrt{se(\widehat{\beta}_3)^2 + 4se(\widehat{\beta}_2)^2 - 4s_{23}}$$

where s_{23} is the estimate of $\mathrm{Cov}(\widehat{\beta}_2, \widehat{\beta}_3)$. To implement the above t test, we need to know s_{23}. Under the Classical Linear Regression Assumptions (Gauss–Markov + normality), we know that $t \sim t_{120-4}$. Once we get s_{23}, we can implement the one-sided t test. With the significance level at 5%, the rejection rule is $t > c$, where c is the 95th percentile of the t_{116} distribution.

(c) The colleague is concerned about the omission of a relevant explanatory variable, for example, the revenue (call it X). If the true model that explains the company's stock price has other variables not included in the estimated model, the regression coefficients are going to be biased and the standard errors will be invalid. The intuitive reason behind the bias is that in addition to the direct effect, the included variables have an apparent indirect effect as a consequence of acting as a proxy for the missing, omitted variable X. The strength of this proxy effect depends on two factors, the strength of the effect of X on the stock price, and the ability of the included variables to mimic X. If X and the included variables are correlated, the apparent indirect effect will be non-zero, causing the direct effect of the included variables to be affected, causing the OLS estimator to be biased. If the variable $small_t$ is indeed a relevant variable, it will help reduce the bias. However, it is more likely that the variable is an irrelevant variable as this small company in Bulgaria (a random company), does not have any explanatory power and should not be included in the regression model. If insisted, the biasedness of the coefficients in general remains unaffected. However, the estimator is less efficient: the standard errors remain valid, but needlessly large. The intuitive reason is that the model is taking the information in $small_t$ that should not be in the true model, hence adding more noise. If $small_t$ and the included regressors are not correlated, which is likely to be the case, there will be no loss of efficiency.

(d) The appropriate test is the Chow test. We have to split the data into two groups, from 2012m1 to 2016m6, and 2016m7 to 2021m12, then run the following regressions:

$$\log(stock_t) = \alpha_1 + \alpha_2 \log(profit_t) + \alpha_3 \log(research_t) + \alpha_4 \log(marketing_t) + u$$

where, $t = 2012\text{m}1, \ldots, 2016\text{m}6$, and:

$$\log(stock_t) = \gamma_1 + \gamma_2 \log(profit_t) + \gamma_3 \log(research_t) + \gamma_4 \log(marketing_t) + v$$

where $t = 2016\text{m}7, \ldots, 2021\text{m}12$. The null hypothesis of the Chow test is $\text{H}_0 : \alpha_i = \gamma_i$, with $i = 1, 2, 3, 4$, the alternative hypothesis is otherwise.

The test statistic is an F statistic, that uses the sum of squared residuals from the combined data SSR, from the first data group, SSR_1, and from the second data group, SSR_2. The numbers of observations are $n = 120$, $n_1 = 66$ and $n_2 = 54$, the total number of parameters, including the intercept, is 4. The Chow test statistic is:

$$F \text{ statistic} = \frac{SSR - (SSR_1 + SSR_2))/4}{SSR_1 + SSR_2)/(120 - 2 \times 4)}.$$

Under CLM assumptions (Gauss–Markov + normality). The critical value is from $F_{4,\,120-2\times4}$.

If the F statistic is larger than the critical F-value, we reject the null hypothesis, suggesting that the manager is correct (the model coefficients are different before and after 2016m6). If we cannot reject the null hypothesis, the manager's suggestion cannot be concluded.

Comments:

For part (a), most candidates got the correct prediction for the sign, but some interpreted the coefficient incorrectly, either missing the units or a wrong approach to a log-log regression. The discussion, either intuitive or mathematical, on the new interpretation of the estimate when changing the units was sometimes lacklustre. For part (b), candidates were expected to perform a usual t test and explain how to obtain the test specifics. Most did well, but as usual some lost marks on missing some specifications. Some candidates did not notice the setting of a one-sided test. For part (c), most who correctly identified the problem of omitting an important variable discussed intelligently the omitted variable bias problems. A good number of candidates realised $small_t$ is not such a variable, and an irrelevant one instead. Only a few carried on and discussed what would happen in this case (losing efficiency). Part (d) is a classic set-up for a Chow test, however many skipped this part of the question. Once correctly identified, candidates did very well. Some explained an incorrect version of the F test of coefficient equalities, or missed test specifications.

Question 7

A researcher models the relationship between the expenditure of a company, S_t in period t and the expected profit, π^*_{t+1}, in period $t+1$ as follows:

$$S_t = \beta_0 + \beta_1 \pi^*_{t+1} + \beta_2 r_t + u_t, \tag{7.1}$$

where r_t is the borrowing interest rate set by the central bank (measured in percentage) and u_t is an i.i.d. error term with:

$$\mathrm{E}(u_t \mid S_{t-1}, S_{t-2}, \ldots, \pi_t, \pi_{t-1}, \ldots, r_t, r_{t-1}, \ldots) = 0.$$

The expected profit is determined by the following adaptive expectation process:

$$\pi^*_{t+1} - \pi^*_t = \theta(\pi_t - \pi^*_t), \tag{7.2}$$

where π_t is the actual profit realised at time t. Using quarterly data from a US company, the researcher obtains the following estimates from using OLS:

$$\widehat{S}_t = \underset{(0.142)}{0.36} + \underset{(0.54)}{0.94}\,\pi_t - \underset{(2.85)}{34.65} r_t + \underset{(0.85)}{0.65}\,S_{t-1}, \tag{7.3}$$

$$n = 240, \;\; R^2 = 0.56.$$

(a) **What is the interpretation of θ in (7.2)? Using the regression results in (7.3) obtain an estimate of θ.**

Hint: **Use (7.1) and (7.2) to express S_t as follows:**

$$S_t = \alpha_0 + \alpha_1 \pi_t + \alpha_2 r_t + \alpha_3 S_{t-1} + v_t, \tag{7.4}$$

where $v_t = u_t - (1-\theta)u_{t-1}$.

(6 marks)

(b) **You are concerned that the estimate of θ obtained in (a) is not suitable. Demonstrate formally that the OLS estimator of (7.4) will be inconsistent.**
Hint: **You are not expected to look at the consistency proof for the α parameters explicitly.**

(5 marks)

(c) **Discuss how you can use an IV estimator to obtain a consistent estimator of the α parameters and hence obtain a consistent estimator of θ.**

(4 marks)

(d) **Suppose a suitable univariate model for St is given by:**

$$S_t = \lambda_1 + \lambda_2 S_{t-1} + \gamma t + e_t, \tag{7.5}$$

where t is a deterministic trend and e_t is white noise, an i.i.d. error term with zero mean and constant variance that is independent of S_{t-1}. Discuss how to test whether the expenditure process S_t has a unit root. Clearly indicate the null and the alternative hypothesis.

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 8.5, 9.3, 11.4, 11.6, 13.4, 13.5.

Dougherty, C. Subject guide (2016): Chapter 8, 9, 11, 13.

Approaching the question

(a) There is an omission of the term r_{t-1} in the expressions (7.2) and (7.3). No candidates were penalised for not discussing this term in their derivation.

With that remark, the adaptive expectations process can be rewritten as:

$$\pi^*_{t+1} = \theta\pi_t + (1-\theta)\pi^*_t.$$

From (7.1), lagging by one time period, we have:

$$S_{t-1} = \beta_0 + \beta_1 \pi^*_t + \beta_2 r_{t-1} + u_{t-1} \tag{*}$$

$$\rightarrow \quad \beta_1 \pi^*_t = S_{t-1} - \beta_0 - \beta_2 r_{t-1} - u_{t-1}.$$

Substituting these expressions into (7.1), we have:

$$\begin{aligned} S_t &= \beta_0 + \beta_1(\theta\pi_t + (1-\theta)\pi^*_t) + \beta_2 r_t + u_t \\ &= \beta_0 + \beta_1\theta\pi_t + (1-\theta)(S_{t-1} - \beta_0 - \beta_2 r_{t-1} - u_{t-1}) + \beta_2 r_t + u_t \\ &= \beta_0\theta + \beta_1\theta\pi_t + \beta_2 r_t + (1-\theta)\beta_2 r_{t-1} + (1-\theta)S_{t-1} + u_t - (1-\theta)u_{t-1}. \end{aligned} \tag{**}$$

So, (7.5) is equivalent to (**) when we have $\alpha_0 = \theta\beta_0$, $\alpha_1 = \theta\beta_1$, $\alpha_2 = \beta_2$, $\alpha_3 = (1-\theta)$, and $v_t = u_t - (1-\theta)u_{t-1}$. There is also a coefficient of $(1-\theta)\beta_2$ for r_{t-1} that is omitted from (7.2) and (7.3).

Another (more common) approach is to go from $\theta\pi_t = \pi^*_{t+1} - (1-\theta)\pi^*_t$ (which is observable) and then obtain $(1-\theta)S_{t-1}$. Both methods get the same marks.

θ measures how fast the speed of adjustment of expected profits is between two periods. From (*) we must have $(1-\theta) = 0.65$ or $\theta = 0.35$, which lies between $(0, 1)$. The number is qualitatively large.

(b) In (7.4), the regressor S_{t-1} is partly dependent on u_{t-1} from (*). If the error terms exhibit autocorrelation, v_t will contain u_{t-1}, v_t is also correlated with S_{t-1}. Candidates are expected to show this claim algebraically. The assumption of contemporaneous exogeneity (part 2 of C7) is violated hence the estimators are not consistent. Part 1 of C7 is also violated so the OLS estimator is always biased.

(c) The endogenous variable in the equation is S_{t-1}. An IV estimator would be appropriate here. The typical choice of instrument would be lagged dependent variables (S_{t-2}). There are two conditions: (1) The instrument has to be strongly correlated with S_{t-1}, which suggests that past expenditure is strongly correlated with current expenditure, and (2) the instrument should be uncorrelated with the error terms. Candidates can show, by lagging one time period on (7.4), that this condition is likely to be satisfied. With the conditions satisfied, the IV estimator can be obtained through a 2SLS procedure.

Stage 1: Estimate S_{t-1} on S_{t-2} and r_t and π_t. Obtain the predicted values of $\widehat{S}_{t-1}$.

Stage 2: Estimate (7.4), replace S_{t-1} by $\widehat{S}_{t-1}$.

(d) We can simply use the Dickey–Fuller test with a deterministic trend with a drift (if $\lambda_1 \neq 0$). Taking differences of (7.5):

$$S_t = \lambda_1 + \lambda_2 S_{t-1} + \gamma t + e_t$$

we get:

$$\Delta S_t = \lambda_1 + (\lambda_2 - 1)S_{t-1} + \gamma t + e_t. \qquad (***)$$

The null hypothesis is $\mathrm{H}_0 : \phi_1 = \lambda_2 - 1 = 0$ (unit root, non-stationary, persistent), against $\mathrm{H}_1 : \phi_1 = \lambda_2 - 1 < 0$ (no unit root, stationary, weak dependent). Now, we can use the usual t test.

After running the testing equation (***). The test-statistic is:

$$\left| \frac{\widehat{\phi}_1}{se(\widehat{\phi}_1)} \right|.$$

Under the null hypothesis, the distribution is no longer a t distribution, but one called the Dickey–Fuller distribution. We reject the null if the t statistic is smaller than the critical value at the 5% level. Note the alternative is one sided ($\phi_1 < 0$). If the process S_t is confirmed non-stationary, the relationship we observe in (7.4) can be spurious and hence meaningless.

Comments:

There was an omission of the term r_{t-1} in (7.2) and (7.3), which should not affect the derivation and progress of the question. Most candidates who attempted the question would realise the omission and (correctly) moved on and assumed this omission away. Similar to previous years, most candidates avoided this type of long questions that involve long derivations with non-standard set-ups. Those who attempted would actually do very well.

Most candidates did well for part (a), some marks were lost due to poor annotations or a lack of clarifications. Some made algebraic mistakes but were only penalised once. Most recognised in part (b) the problem of having the lagged outcome variable and then discussed how S_{t-1} is correlated with v_t. Some lost marks for unclear explanation or forgetting the inconsistency problem. The key word here is 'formally', which means we would expect some mathematical proof of the correlation. Most did well in part (c). Marks were lost for not realising the use of other lagged explanatory variables as instruments, or lack of discussions on the IV assumptions. Most attempted to discuss the Dickey–Fuller test in part (d). Marks were lost for not adding the trend, or missing test specifications, or not using a one-sided test.

Question 8

Let us consider a model where we want to explain whether a married woman works for wages or not. We may define the dependent variable *paid* to equal 1 when the woman works for wages, and zero otherwise, and consider the model:

$$paid = \alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper^2 + u, \tag{8.1}$$

where *nwifeinc* is other sources of income (in \$000), *educ* is years of education, and *exper* is past years of labour experience. You may assume that $\mathrm{E}(u \mid nwifeinc, educ, exper) = 0$.

(a) **Provide an economic interpretation of the parameter α_2. What should be the sign of α_2?**

(2 marks)

(b) **What would happen to the parameter estimates if the qualitative outcome was coded differently, in particular:**

$$paid^* = \begin{cases} 2 & \text{if married woman works for wages} \\ 1 & \text{if married women does not work for wages} \end{cases}$$

and the following regression was estimated by OLS:

$$paid^* = \alpha_0 + \alpha_1^* nwifeinc + \alpha_2 educ + \alpha_3^* exper + \alpha_4^* exper^2 + u^*? \tag{8.2}$$

(3 marks)

(c) **Discuss why the following statements are False. Suggest suitable changes that will make the statements correct. Details supporting your claims will be rewarded.**

i. **The R-squared obtained from the OLS estimation of (8.2) is larger than the R-squared obtained from the OLS estimation of (8.1).**

ii. **To test whether the non-education variables (*nwifeinc*, *exper*, *exper*2) are jointly significant in (8.1), we can use the usual F test that compares the restricted and unrestricted residual sum of squares.**

(6 marks)

(d) **Instead of using OLS, a researcher proposes we should use the Probit model. Introduce the model and discuss the benefits and drawback of using this approach. In your answer you are expected to provide the main equation underlying the Probit model and the log-likelihood function.**

(7 marks)

(e) **Discuss the following statement: 'A probit model that uses *nwifeinc*, *educ*, *exper*, and *exper*2 to explain whether a married woman works for wages or not, will exhibit diminishing effects of additional years of education on the probability of working for wages.' In your answer clearly indicate how to obtain the partial effect under consideration.**

(2 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 10.1, 10.3, 10.6.

Dougherty, C. Subject guide (2016): Chapters 10.

Approaching the question

(a) A 1-year increase in education is associated with on average α_2 change in the probability of a married woman having a paid job, holding everything else constant. We should expect α_2 to be positive (more educated means more likely to get a paid job). Candidates can also suggest an alternative explanation for a negative coefficient but it has to be reasonable.

(b) We can see that $paid^* = paid + 1$, so we simply added a constant 1 to the outcome variable in (8.1). The slope coefficients (non-intercept) estimates are unchanged. The intercept coefficient will go up by 1. Details (derivations) are expected.

(c) i. The statement is wrong because the two R-squareds are the same. The only thing we did was to reparameterise the dependent variable by adding a constant (1) into it. The variations in the dependent variable will not change (because after substracting the mean, the deviations will be the same), and the model does not fundamentally change. The R-squared remains.

ii. Because of heteroskedasticity underlying the LPM, the standard errors are invalid and we cannot use the standard errors to conduct the hypothesis tests. There are two solutions, either we report the robust standard errors or to use the LR test. For the use of LPM estimates, we need the F test with a chi-squared distribution (asymptotic tests). For the LR test, we need $LR = 2(\log L^{UR} - \log L^R)$ and a chi-squared distribution with degrees of freedom of 3.

(d) The researcher should formulate the following form: the probability of a married woman working for a wage is:

$$Pr(paid = 1 \,|\, x) = \Phi(\alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper2)$$

where $\Phi(\cdot)$ is the standard normal cumulative distribution function, x is the set of the regressors. The probability of a married woman does not have paid work is:

$$Pr(paid = 0 \,|\, x) = 1 - \Phi(\alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper2).$$

The likelihood function for this sample is:

$$\begin{aligned} L(\alpha \,|\, data) &= \prod_{paid} Pr(paid = 1 \,|\, x) \prod_{unpaid} Pr(paid = 0 \,|\, x) \\ &= \prod_{paid} \Phi(\alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper2) \\ &\quad \times \prod_{unpaid} [1 - \Phi(\alpha_0 + \alpha_1 nwifeinc + \alpha_2 educ + \alpha_3 exper + \alpha_4 exper2)]. \end{aligned}$$

We can then write down the log-likelihood function $\log L(\alpha \,|\, x)$. To estimate the parameters, we can use the maximum likelihood estimator that chooses $\{\alpha_j\}$ to maximise the joint probability of the outcomes. The intuition is that we try to choose the $\widehat{\alpha}$ such that the $\widehat{Pr}(paid = 1)$ is high for paid individuals, and low for the unpaid. There are no mathematical formulae for the estimates – the popular method is to adopt an iterative trial-and-error process.

Benefits of Probit (MLE): non-constant marginal effect (more realistic), predicted probabilities within $(0, 1)$, and automatically resolve the heteroskedasticity issue. Furthermore, in large samples, the MLEs are consistent, asymptotically normally distributed, and asymptotically efficient. (Candidates are only expected to provide the first three benefits).

Drawbacks: more difficult to interpret and used to be more computationally challenging.

(e) Because the probability of observing a paid or unpaid event given the regressor is the cumulative density evaluated at each point of the regressors, the probit functional form allows for the effects of the regressors to have diminishing returns as the predicted probability approaches to 0 or 1. So the statement is correct.

Comments:

Part (a) was well-attempted in general. Marks were lost due to some small mistakes or a lack of explanation. Part (b) relatively well-attempted. The answer is straightforward but some candidates lost marks on the explanation. Part (c) posed some challenges. Most did not explain well the inherent problem of heteroskedasticity in an LPM, nor proposed a good alternative. Part (d) required the candidates to discuss the MLE procedure, with some good intuition. Some candidates did not provide a discussion on its advantages and drawbacks. Part (e) was the least attempted. Those who attempted did quite well, providing different forms of marginal effects used in interpreting Probit/Logit.

Question 9

Let us consider the following system of simultaneous equations:

$$y_1 = \alpha_1 y_2 + \alpha_2 x_1 + \alpha_3 x_2 + u_1 \tag{9.1}$$

$$y_2 = \beta_1 y_1 + \beta_2 x_1 + \beta_3 x_3 + \beta_4 x_4 + u_2, \tag{9.2}$$

where (u_1, u_2) are i.i.d. errors with zero mean, $\text{Var}(u_1) = \sigma_1^2$, $\text{Var}(u_2) = \sigma_2^2$, and let $\text{Cov}(u_1, u_2) = 0$. The endogenous variables are (y_1, y_2) and (x_1, x_2, x_3, x_4) are exogenous variables.

(a) **Explain the simultaneity issue associated with the above simultaneous equation model (SEM) intuitively (no derivations expected). Give an example that fits such a structure.**

(4 marks)

(b) **You are told that in the above SEM there exists a perfect linear relation between x_2 and x_3, in particular,**

$$x_3 = 2x_2.$$

Obtain the reduced form equations for y_1 and y_2, recognising this exact linear relation between x_2 and x_3.

(7 marks)

(c) **Use the result obtained in (b) to discuss the identification of the two structural equations. Clearly state whether the equations are over identified, exact (just) identified, or under identified. *Hint:* Your answer is expected to discuss what conditions need to be satisfied to ensure that we can use the observable data to estimate the parameters consistently.**

(4 marks)

(d) **Discuss how you can obtain a consistent estimator of the α parameters using 2SLS. Provide details of the estimator.**

(5 marks)

Reading for this question

Dougherty, C. *Introduction to econometrics* (fifth edition): Chapters 9.1, 9.3.

Dougherty, C. Subject guide (2016): Chapter 9.

Approaching the question

(a) The problem here is reverse causality: both equations have their own *ceteris paribus* interpretation, the simultaneity gives rise to a correlation between the two endogenous variables that are jointly determined in this system which makes it that when we estimate these structural equations in isolation this will give rise to inconsistency. One example is police and crime.

(b) Substitute (9.2) to (9.1) we have:

$$y_1 = \alpha_1(\beta_1 y_1 + \beta_2 x_1 + \beta_3 x_3 + \beta_4 x_4 + u_2) + \alpha_2 x_1 + \alpha_3 x_2 + u_1$$

$$(1 - \alpha_1\beta_1)y_1 = (\alpha_1\beta_2 + \alpha_2)x_1 + \alpha_3 x_2 + \alpha_1\beta_3 x_3 + \alpha_1\beta_4 x_4 + u_1 + \alpha_1 u_2$$

$$y_1 = \frac{(\alpha_1\beta_2 + \alpha_2)x_1 + \alpha_3 x_2 + \alpha_1\beta_3 x_3 + \alpha_1\beta_4 x_4 + u_1 + \alpha_1 u_2}{(1 - \alpha_1\beta_1)}$$

and for y_2, from (9.1) to (9.2):

$$\begin{aligned} y_2 &= \beta_1 y_1 + \beta_2 x_1 + \beta_3 x_3 + \beta_4 x_4 + u_2 \\ &= \beta_1(\alpha_1 y_2 + \alpha_2 x_1 + \alpha_3 x_2 + u_1) + \beta_2 x_1 + \beta_3 x_3 + \beta_4 x_4 + u_2 \\ &= \frac{(\beta_1\alpha_2 + \alpha_2)x_1 + \beta_1\alpha_3 x_2 + \beta_3 x_3 + \beta_4 x_4 + u_2 + \beta_1 u_1}{(1 - \alpha_1\beta_1)}. \end{aligned}$$

Now using the assumption that $x_3 = 2x_2$, we have:

$$\begin{aligned} y_1 &= \frac{(\alpha_1\beta_2 + \alpha_2)x_1 + \alpha_3 x_2 + \alpha_1\beta_3 x_3 + \alpha_1\beta_4 x_4 + u_1 + \alpha_1 u_2}{(1 - \alpha_1\beta_1)} \\ &= \frac{(\alpha_1\beta_2 + \alpha_2)x_1 + (\alpha_3 + 2\alpha_1\beta_3)x_2 + \alpha_1\beta_4 x_4 + u_1 + \alpha_1 u_2}{(1 - \alpha_1\beta_1)} \end{aligned}$$

and for y_2:

$$y_2 = \frac{(\beta_1\alpha_2 + \alpha_2)x_1 + (\beta_1\alpha_3 + 2\beta_3)x_2 + \beta_4 x_4 + u_2 + \beta_1 u_1}{(1 - \alpha_1\beta_1)}.$$

The reduced forms are given in the following forms:

$$y_1 = \pi_{11}x_1 + \pi_{12}x_2 + \pi_{13}x_4 + v_1$$

$$y_2 = \pi_{21}x_1 + \pi_{22}x_2 + \pi_{23}x_4 + v_2.$$

(c) Equation(9.1) is just identified as we have x_4 as the exogenous variable to work as an instrument for the only endogenous variable (y_2). x_2 and x_3 cannot be used as instruments as they appear in both the equations. The conditions are: β_4 is statistically significant (strongly predicts y_2), and the exogeneity of x_4 holds ($\text{Cov}(x_4, u_1) = 0$). Equation (9.2) is underidentified because we cannot use any of the regressors as an instrument to identify β_1 following the multicollinearity between x_2 and x_3.

(d) We can only obtain a consistent estimator of α_1, by using the Two Stage Least Squares.

(1) Estimate the reduce form of y_2 on x_1, x_2, x_4, and get the following fitted value, $\widehat{y}_2$:

$$\widehat{y}_2 = \widehat{\pi}_0 + \widehat{\pi}_1 x_1 + \widehat{\pi}_2 x_2 + \widehat{\pi}_4 x_4.$$

(2) Plug $\widehat{y}_2$ instead of y_2 into (9.1), and run an OLS estimation to obtain the consistent estimators:

$$y_1 = \textit{constant} + \alpha_1\widehat{y}_2 + \gamma_1 x_1 + \gamma_2 x_2 + \textit{error}.$$

Candidates can also explain an IV estimator using $\widehat{y}_2$ as the instrument.

Comments:

In part (a), most recognised the problem of reverse causality and associated issues (joint determination and inconsistency). Many gave great real-life examples. Marks were lost mostly on the explanations. In part (b), most got full marks for the derivations. Some lost a few marks on some small errors or explanations of the steps. Candidates were expected to recognise the fact that we can rewrite the equations using only either x_2 or x_3 and neither can be used as an instrument. x_4 is the only viable choice. Marks were lost for incomplete definitions of identification (and under- vs over-). Most did well in part (d). Marks were lost for not including all exogenous variables in both stages, or discussing some convoluted IV estimator without the setting laid out.

EC2020 Elements of Econometrics

Candidates should answer **EIGHT** of the following **NINE** questions: All **FIVE** from Section A (8 marks each) and **THREE** from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than **EIGHT** questions are answered, only the first **EIGHT** questions attempted will be counted.

Please find the questions on the following pages.

© University of London 2023

SECTION A

Answer all questions from this section. Each question carries 8 marks.

1. Consider the following multivariate regression model:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + u_i, \tag{1.1}$$

for a given random sample of n observations $\{(y_{i,}, x_{1i} > 0, x_{2i} > 0)\}_{i=1}^{n}$ under the standard MLR.1-MLR.4 assumptions (linearity in parameters, random sampling, no perfect multicollinearity, and zero conditional mean). Let:

$$z_i = x_{1i}^{20} + \sqrt[23]{x_{1i}}, \quad \bar{z} = \sum_{i=1}^{n} z_i/n,$$

and define the following slope estimator:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n}(z_i - \bar{z})y_i}{\sum_{i=1}^{n}(z_i - \bar{z})x_{1i}}.$$

(a) **(4 marks)** Carefully derive the asymptotic bias in the estimator $\hat{\beta}_1$. What are the conditions for $\hat{\beta}_1$ to be a consistent estimator of β_1? In your answer, clearly explain the concept of consistency.

(b) **(4 marks)** Suppose that the conditional variance of u_i takes the following form:

$$Var(u_i|x_{1i}, x_{2i}) = \delta^2(x_{1i}^{20} + \sqrt[22]{x_{2i}})$$

Describe the problem with the OLS estimator $\hat{\beta}_1^{OLS}$ of β_1. Explain why it is not the BLUE and propose a method to obtain the BLUE of β_1?

2. A researcher is interested in the relationship between (the natural logarithm of) family income $(\log faminc_i)$ and (the natural logarithm of) annual household savings $(\log saving_i)$. Using a random sample of 2023 families in the UK, she estimates the following regression:

$$\log saving_i = \beta_0 + \beta_1 \log faminc_i + \beta_2 children_i + \beta_3 qualification_i + \varepsilon_i, \tag{2.1}$$

where $qualification_i$ is a binary variable of whether the household head has a professional qualification, $children_i$ is the number of children in the household. Assume that the model given by (1.2) satisfies Assumptions MLR.1-MLR.4 (linearity in parameters, random sampling, no perfect multicollinearity, and zero conditional mean).

(a) **(2 marks)** The researcher plans to include the logarithm of the total consumption $\log(cons)$, where $cons = faminc - saving$ into (2.1). Should we be concerned about the multicollinearity problem? In your answer, explain clearly what multicollinearity is and its implication(s).

(b) **(4 marks)** A colleague claims that the benefit of having a professional qualification on the saving level will be offset by having two additional children. Discuss how you can test for this claim, and clearly specify your null hypothesis and other test specifics.

(c) **(2 marks)** Another colleague observes that in the OLS estimation of (2.1), the coefficient estimates for $children_i$ and $qualification_i$ are both not statistically significant. They claim that it is an indicator of heteroscedasticity. Do you agree with this claim? In your answer, state two factors that can affect the statistical significance of a coefficient estimate.

3. A researcher is interested in the relationship between the political institution of a country, measured by a political stability index, $stability_i$, and the number of reported Covid-19 cases per one million residents, $cases_i^*$. She captures the relationship in a simple regression, using data from 145 countries in 2022:

$$cases_i^* = \alpha + \beta stability_i + u_i, \tag{3.1}$$

where u is the error term and is assumed to satisfy SLR.1 to SLR.5.

(a) **(4 marks)** A researcher is worried that the number of cases in a country is likely to be misreported, particularly, a politically less stable country may under-report the number of cases to make the country look better. That is, instead of the true number of cases, $cases^*$, the researcher only obtain a misreported number of cases, $cases_i$:

$$cases_i^* = cases_i + e_i$$

where $e_i + u_i$ is the error negatively correlated with $stability_i$ for each country. Derive the bias associated with this systematic error. What is the sign of the bias?

(b) **(2 marks)** Instead of the error in the number of cases, another researcher is concerned about the objectivity of the political stability index. She argues that the index tends to exaggerate the true level of political stability by 10%. Discuss how this claim can affect the OLS estimator of β. What is the direction of the bias?

(c) **(2 marks)** How would your answer to (b) change if the researcher assumes that the subjectivity of the stability index is purely random? *Hint:* more politically stable countries are thought to report fewer cases as they may contain the virus more effectively.

4. Consider the following time series model:

$$\begin{aligned} y_t &= \alpha + \beta t + \epsilon_t, t = 2..., T, \\ \epsilon_t &= \theta_1 u_t + \theta_2 u_{t-1} + u_{t-2}, \end{aligned} \tag{4.1}$$

where u_t is a white noise error term with a zero mean and a constant variance of δ^2; $\theta_1, \theta_2 \neq 0$.

(a) **(4 marks)** Provide the condition(s) that ensures the stationarity of ϵ_t. In your answer, discuss what it is meant by stationarity and how it is different from weak dependence.

(b) **(2 marks)** Show that under the condition(s) provided in (a), y_t is trend stationary.

(c) **(2 marks)** Given the answer to (a), discuss the problem with an OLS estimation of (4.1). What are the consequences of this problem? Propose a method to address this problem.

5. A researcher is interested in the relationship between the annual salary growth rate $salary_t$ and the number of university graduates, $graduate_i$ in France since 2000. She estimates the following regression model using observations from T quarters:

$$salary_t = \beta_0 + \beta_1 graduate_t + \beta_2 graduate_{t-1} + \theta salary_{t-1} + u_t, \tag{5.1}$$

where $|\theta| < 1, t = 2, ..., T$, and the error term u_t has a mean zero and a constant variance δ^2. The researcher observes that salary growth rates are weakly dependent.

(a) **(2 marks)** Discuss the statistical properties of the OLS estimator for θ. Be clear about whether any of the Gauss-Markov assumptions *are* violated.

(b) **(2 marks)** What is the short-term and long-term effect of the number of graduates on the salary growth rate?

(c) **(4 marks)** Suppose that u_t can be expressed as $u_t = \rho u_{t-1} + e_t$, where e_t is a white noise series. Discuss how this assumption will affect your answer in (a). Propose a method to address the problem.

SECTION B

Answer three questions from this section. Each question carries 20 marks.

6. A team examines the determinants of the number of Covid-19 cases, $cases_i$ per million of residents of 142 countries around the world in 2022. They also have the data on the number of hospitals per million of residents, $hospital_i$, the total government expenses on healthcare, $healthexp_i$ in millions of dollars, and a binary to indicate whether tourism accounts for more than 25% of the country's GDP, $tourism_i$. They obtain the following estimation, using the logarithms of $healthexp_i$:

$$\widehat{cases_i} = \underset{(3.414)}{-5.001} - \underset{(20.231)}{100.11}\log(healthexp_i) - \underset{(22.51)}{15.507}hospital_i + \underset{(4.213)}{10.501}tourism_i, \quad (6.1)$$
$$n = 142, \qquad R^2 = .728.$$

The robust standard errors are reported in the parentheses.

(a) **(5 marks)** What is the interpretation of the coefficient estimates on $\log(healthexp_i)$ and $tourism$? Are the sign and size of the coefficient as you would expect? Explain your answer. Can you interpret any of the estimates as causal effects? Give at least two distinctive reasons for your answer.

(b) **(5 marks)** An intern generates a different variable $nontourism_i = 1 - tourism_i$ to indicate if a country is not dependent on tourism. Explain what will happen to the estimates reported in (6.1) if we:

(i) Replace $tourism_i$ by $nontourism_i$ in (6.1).

(ii) Add $nontourism$ to the regression in (6.1) and drop the intercept. In case (ii), explain whether the interpretation of the coefficient estimate for $tourism_i$ would be different.

(c) **(5 marks)** The team manager would like to test whether being more reliant on tourism would reduce the effect of the government's health expenses on the number of cases. Describe a suitable test. Be clear about any additional information or model specification you would need.

(d) **(5 marks)** Another colleague believes that the number of cases is going to be varying more or less for countries with higher health expenses than for those with lower health expenses. Discuss a statistical test for this claim. Discuss what will happen to the estimation in (6.1) if we fail to reject the null hypothesis of this test.

7. An economist is interested in the relationship between social stability and economic output. She examines the following model for the (logarithm of) output-per-capita GDP_i in a given country i:

$$GDP_i = \alpha + \beta_1 protest_i + \beta_2 public_i + \beta_3 regions_i + u_i, \tag{7.1}$$

where $protest_i$ is the number of public protests happening in the country during a fiscal year, $public_i$ is the per capita government spending (set by law) on social benefit in a single year, and $regions$ is a variable indicating the regional geography of the country (such as Asia, Europe, North America, ...). The economist obtains the data for $n = 150$ countries around the world in 2022.

(a) **(5 marks)** Give two distinct reasons why an OLS regression applied to (7.1) is inappropriate for estimating the causal effect of social instability (measured by $protest_i$) on economic development (GDP_i). Be clear on the direction of bias if any.

(b) **(5 marks)** The data contains the variable, $adverseweather_i$, which indicates the number of adverse weather events happening in the country during the year that prevented large public gatherings. The OLS estimation of $protest_i$ on $adverseweather_i$ with robust standard errors reported in the parentheses gives:

$$\begin{aligned} \widehat{protest}_i &= \underset{(2.214)}{5.13} - \underset{(2.31)}{10.61} adverseweather_i - \underset{(11.51)}{15.57} public_i + \underset{(4.13)}{1.01} regions_i, \quad (7.2) \\ & n = 150, \qquad R^2 = .718, Fstat = 24.33. \end{aligned}$$

The economist aims to use $adverseweather_i$ as the instrumental variable for $protest_i$ in Equation (1). What is the purpose of the estimation in (7.2)? Discuss whether you believe $adverseweather$ is a valid instrument for $protest_i$. In your answer, evaluate the required assumptions for a valid instrument.

(c) **(7 marks)** The economist learns about the following relationship between $protest_i$ and GDP_i from a senior colleague:

$$protest_i = \gamma_0 + \gamma_1 GDP_i + \gamma_2 regions_i + \gamma_3 trust_i + \gamma_4 trust_i \times GDP_i + v_i, \tag{7.3}$$

where $trust_i$ indicates whether the residents of the country overwhelmingly support the government, and $trust_i \times GDP_i$ is the interaction term between two variables. Examine the identification of each structural equation in (7.1) and (7.3). Be clear on the problem of using OLS to estimate the models and whether you would be able to estimate the β_1 or γ_1 parameters in (7.1) consistently. Describe an estimator to estimate the parameters consistently.

(d) **(3 marks)** Demonstrate that $public_i$ and $trust \times public$ are valid instruments to identify the parameters in (7.3). Which estimator would you use to estimate the parameters in (7.3)? *Hint:* You are requested to show how the variables satisfy the conditions for valid instruments.

8. A researcher examines the probability of a random sample of economics professors in Europe that use Twitter to share their professional ideas ($Twitter_i = 1$ if yes, and $= 0$ if otherwise), where i indicates an individual economics professor. She uses OLS with *conventional* standard errors reported in the parentheses to obtain the following OLS regression results:

$$\widehat{Twitter}_i = \underset{(0.114)}{0.138} - \underset{(0.023)}{0.011} age_i - \underset{(0.415)}{0.135} male_i + \underset{(0.312)}{0.754} head_i + \underset{(0.005)}{0.064} pub_i, \quad (8.1)$$
$$n = 2123,\ R^2 = 0.19,$$

where age_i, $male_i$, pub_i, $head_i$ are respectively the professor's age, whether the professor is a male, the number of publications, and whether the professor is currently the head of the department.

(a) **(2 marks)** What is the interpretation of the intercept estimate? Is the sign what you would expect?

(b) **(5 marks)** Looking at the estimates, a colleague claims that being a departmental head has a strong and *statistically significant* effect on the probability of an economics professor using Twitter. Explain why you would be sceptical about this claim. In your answer, propose two adjustments to the OLS estimation that can fix the problem that you identify.

The researcher decides to re-estimate the model using Logit, and obtain the following regression results:

$$\Pr\left(\widehat{Twitter}_i = 1\right) = \Lambda\left(\underset{(0.114)}{0.238} - \underset{(0.013)}{0.021} age_i - \underset{(0.215)}{0.125} male_i + \underset{(0.012)}{0.854} head_i + \underset{(0.022)}{0.054} pub_i\right), \quad (8.2)$$
$$n = 2123,\ PseudoR^2 = 0.034, \log L = -316.60,$$

where is $\Lambda(.)$ the logistic cumulative density function and $\Lambda(z) = \exp(z)/(1 + \exp(z))$, and the conventional standard errors are reported in the parentheses.

(c) **(5 marks)** Explain the advantages and potential drawbacks of using Logit estimation to examine the probability of using Twitter. In your answer, *briefly* describe the estimator behind the Logit estimation and state its statistical properties.

(d) **(5 marks)** Describe how we can test whether age and the number of publications both have no effect on the probability of using Twitter by an economics professor. Be specific about the testing procedure and any additional information required.

(e) **(3 marks)** What is the estimated effect of being a head of the department on using Twitter for a 40-year-old female professor with 10 publications? *Hint:* There is no need to give an exact number, clarity of the computation suffices.

9. Let us consider the following model that describes the relationship between the log of monthly revenue of a retail chain $\log revenue_t$ and the log of monthly expenses on advertisement $\log ad_t$ of a company at quarter t:

$$\log revenue_t = \beta_0 + \beta_1 \log revenue_{t-1} + \beta_2 \log ad_t + \beta_3 \log ad_{t-1} + u_t. \tag{9.1}$$

We assume that $revenue_t$ and $\log ad_t$ are both I(1) variables, $|\beta_1| < 1$, and the error term u_t is independent of the regressors and has a zero mean and a constant variance.

(a) **(5 marks)** Explain the concepts of spurious relationship and co-integration. Clearly discuss why it is important to test whether the relationship in (9.1) is spurious or co-integrating. *Hint:* You are expected to discuss the statistical consequences of spurious relationship and co-integrating on an OLS estimation of (9.1).

(b) **(5 marks)** Show that the model (9.1) can be fitted by an Error Correction Model (ECM) that has the following structure:

$$\Delta \log revenue_t = \rho \left(\log revenue_{t-1} - \gamma_1 - \gamma_2 \log ad_{t-1}\right) + \gamma_3 \Delta \log ad_t + u_t. \tag{9.2}$$

Explain what an Error Correction Model is and how does the existence of the ECM relate to your answer in (a). Express $\rho, \gamma_1, \gamma_2, \gamma_3$ in terms of $(\beta_0, \beta_1, \beta_2, \beta_3)$ and carefully interpret the various components of this ECM.

(c) **(5 marks)** What are the advantages and challenges of estimating (9.2) using OLS? Discuss a procedure using OLS that can fit the ECM in (9.2).

(d) **(5 marks)** A colleague suspects that the error term u_t is following an AR(1) structure. Describe how you can test for this claim, knowing that firms are likely to spend more on advertisement during the summer to prepare for the festive shopping seasons. Based on your answers from (a) to (c), would you expect to reject the null hypothesis of this test? Explain your testing procedure and the rationales behind your answer.

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	**2·55**	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	**2·60**	0·99534	**2·75**	0·99702	**2·90**	0·99813	**3·05**	0·99886	**3·20**	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	**2·65**	0·99598	**2·80**	0·99744	**2·95**	0·99841	**3·10**	0·99903	**3·25**	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918	**3·30**	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	**3·263**	0·9994	**3·731**	0·99990	**3·916**	0·99995
3·105	0·9991	**3·320**	0·9995	**3·759**	0·99991	**3·976**	0·99996
3·138	0·9992	**3·389**	0·9996	**3·791**	0·99992	**4·055**	0·99997
3·174	0·9993	**3·480**	0·9997	**3·826**	0·99993	**4·173**	0·99998
3·215	0·9994	**3·615**	0·9998	**3·867**	0·99994	**4·417**	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	**5·0**	1·6449	**3·0**	1·8808	**2·0**	2·0537	**1·0**	2·3263	**0·10**	3·0902
45	0·1257	**4·8**	1·6646	**2·9**	1·8957	**1·9**	2·0749	**0·9**	2·3656	**0·09**	3·1214
40	0·2533	**4·6**	1·6849	**2·8**	1·9110	**1·8**	2·0969	**0·8**	2·4089	**0·08**	3·1559
35	0·3853	**4·4**	1·7060	**2·7**	1·9268	**1·7**	2·1201	**0·7**	2·4573	**0·07**	3·1947
30	0·5244	**4·2**	1·7279	**2·6**	1·9431	**1·6**	2·1444	**0·6**	2·5121	**0·06**	3·2389
25	0·6745	**4·0**	1·7507	**2·5**	1·9600	**1·5**	2·1701	**0·5**	2·5758	**0·05**	3·2905
20	0·8416	**3·8**	1·7744	**2·4**	1·9774	**1·4**	2·1973	**0·4**	2·6521	**0·01**	3·7190
15	1·0364	**3·6**	1·7991	**2·3**	1·9954	**1·3**	2·2262	**0·3**	2·7478	**0·005**	3·8906
10	1·2816	**3·4**	1·8250	**2·2**	2·0141	**1·2**	2·2571	**0·2**	2·8782	**0·001**	4·2649
5	1·6449	**3·2**	1·8522	**2·1**	2·0335	**1·1**	2·2904	**0·1**	3·0902	**0·0005**	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

© University of London 2023

EC2020 Elements of Econometrics

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from Section A (8 marks each) and **THREE** from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Please find the questions on the following page.

© University of London 2023

SECTION A

Answer all questions from this section. Each question carries 8 marks.

1. Consider the following regression model

$$y_t = \rho y_{t-1} + \gamma_1 z_t + \gamma_2 z_{t-1} + \varepsilon_t \quad \text{with } |\rho| < 1,$$

where ε_t is an MA(1) process, i.e.,

$$\varepsilon_t = u_t + \theta u_{t-1}.$$

u_t is a white noise process generated randomly from a normal distribution with mean zero and constant variance that is independent of $y_{t-1}, y_{t-2,}, ..., z_t, z_{t-1}, ...$ Suppose you obtain a random sample of T observations, which we label $t = 1, 2, ..., T$ for convenience. Assume there is no perfect multicollinearity.

(a) **(4 marks)** Discuss how the knowledge of the MA(1) structure of the error terms would affect the statistical properties of the OLS estimator of ρ? Be clear on unbiasedness, consistency, and efficiency.

(b) **(4 marks)** Assume now y_t and z_t are two I(1) processes. Derive the long-run relationship between y_t and z_t. Discuss any potential problems with using OLS to estimate this relationship. In your answer, clearly distinguish the concepts of co-integration and spurious relationship.

2. Consider the regression model

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u, \text{ with } E(u) = 0.$$

Assume that the regressors are non-stochastic and there is no perfect multicollinearity. Suppose the error has a normal distribution and is independent of the regressors. You are given a random sample of observations $\{(y_i, x_{1i}, x_{2i})\}_{i=1}^n$.

(a) **(2 marks)** Explain the concept of multicollinearity. Why is it important to assume no perfect multicollinearity?

(b) **(2 marks)** Discuss two factors that can affect the standard error of the OLS estimator for β_1.

(c) **(4 marks)** You discover that $Var\,(u_i|x_1, x_2) = \delta^2[x_{1i}^2 + \sqrt{x_{2i}}]$ where δ is an unknown constant and $x_{1i}, x_{2i} > 0$ for all i. Discuss why the OLS estimator for β_1 is problemactic. Propose the BLUE estimator for β_1. In your answer, be clear about the problem(s) with an OLS and potential remedies.

3. Consider the following simple regression model to explain the number of trade partners of a country (Y_i). A researcher uses the amount of Foreign Direct Investment (FDI), denoted $\log(X_i)$, of a country i as the only explanatory variable:

$$Y_i = \beta_0 + \beta_1 \log(X_i) + u_i, \qquad i = 1, ..., n. \tag{3.1}$$

The error term is assumed to be i.i.d. with zero mean and variance equal to σ^2 and independent of the regressor. We assume $X_i > 0$ for all countries i.

(a) **(4 marks)** Suppose you are given a random sample of observations $\{(Y_i, X_i)\}_{i=1}^n$. Derive the OLS estimator of β_1. Provide an interpretation of this parameter.

(b) **(4 marks)** The Finance Minister mentioned that the amount of FDI is often miscalculated, so unfortunately we do not observe the true X_i, instead you observe X_i^* which equals

$$X_i^* = X_i \exp(\varepsilon_i), \tag{3.2}$$

where ε_i has zero mean and is independent of u_i and X_i (and therefore $\log(X_i)$). You are given a random sample of observations $\{(Y_i, X_i^*)\}_{i=1}^n$. Describe the estimable regression model and show that the OLS estimator of β_1 is inconsistent. Comment on the asymptotic bias of this OLS estimator.

4. (a) **(4 marks)** Consider the following time series model for $\{x_t\}_{t=1}^T$:

$$x_t = x_{t-1} + \lambda\varepsilon_t + \varepsilon_{t-1},$$

where ε_t is i.i.d with mean zero and variance σ^2, for $t = 1, ..., T$, and $x_0 = \varepsilon_0 = 0$. Discuss whether there are any values of λ so that x_t is covariance stationary. In your answer, explain the concept of stationarity.

(b) **(4 marks)** Consider the following time series model for $\{y_t\}_{t=1}^T$:

$$y_t = \theta_0 + \theta t + \rho y_{t-1} + \varepsilon_t, |\rho \leq 1,$$

where ε_t is i.i.d with mean zero and variance σ^2, $t = 1, ..., T$, and $y_0 = 0$. How should we test whether $\rho = 1$? Clearly specify the null and alternative hypothesis, the test statistic, and the rejection rule. What does it tell us if we fail to reject the null hypothesis of this test?

5. Consider the following time series model for $\{y_t\}_{t=1}^{T}$:

$$\begin{aligned} y_t &= \alpha + \beta t + u_t, \qquad t = 1, ..., T \\ &\quad \text{with } u_t = \rho u_{t-1} + \varepsilon_t \quad \text{and } |\rho| \leq 1, \end{aligned}$$

where ε_t is an i.i.d. $\left(0, \sigma_\varepsilon^2\right)$ process that is independent of anything in the past (white noise).

(a) **(3 marks)** Show that y_t is trend stationary when $|\rho| < 1$.

(b) **(3 marks)** Show that y_t is difference stationary when $\rho = 1$.

(c) **(2 marks)** Discuss the importance of distinguishing between trend stationary and difference stationary processes.

SECTION B

Answer three questions from this section. Each question carries 20 marks.

6. The Ministry of Finance is interested in how to incentivise workers to save up for pension. A worker i decides to sign up for the pension scheme provided by their employer, $pension_i$ (= 1 if yes, and = 0 otherwise). The ministry randomly provides a set of bonuses to a number of companies around the UK, $bonus_i$ (= 1 if worker i works in a company eligible for the bonus, = 0 otherwise). We also have information on age_i of the worker, $gender_i$ (1 = male, 0 = female), and $graduate_i$ (= 1 if graduated from a university, = 0 otherwise). A team of researchers run the following Probit regression model:

$$\text{Pr}(pension_i = 1) = \Phi\left(\beta_0 + \beta_1 bonus_i + \beta_2 age_i + \beta_3 gender_i + \beta_4 graduate_i\right), \tag{6.1}$$

where $\Phi(.)$ is the cumulative distribution function of the standard normal distribution.

(a) **(5 marks)** Describe how the Probit estimates are calculated. In your answer, specify the log likelihood function involved, the intuition of the estimator, and its statistical properties.

(b) **(5 marks)** A colleague argues that using the Probit regression model is superior to using an OLS to estimate the determinants of the decision to sign up for work pension. Discuss whether you would agree with him, highlighting both the drawbacks and advantages of Probit compared to the OLS.

(c) **(5 marks)** Describe how you can obtain the effect of the bonus on the probability of signing up for work pension for a 45-year-old male, university graduated worker. *Hint:* You are not expected to give a number: clarity of the computations would suffices.

(d) **(5 marks)** Discuss how you can test the joint significance of the workers' characteristics (age, gender, education) using the Probit estimation. Be clear about the test statistic, its distribution, the rejection rule, and any additional information you would require.

7. A team of economists consider the effect of the mask-wearing mandate in 1234 workplaces around Britain in 2022 on the mask-wearing behaviour of employees throughout the year. Let $wear_i$ be a dummy variable indicating whether a worker always wears masks (1 = yes, 0 = no) throughout the year, and $mandate_i$, a binary variable (1 = yes, 0 = no) if there is a mask-wearing policy in the worker i's workplace.

   (a) **(5 marks)** The team obtain the following OLS regression results for the simple linear regression model, with the robust standard errors in parentheses:

$$\begin{aligned} \widehat{\textit{wear}} &= \underset{(0.007)}{0.290} + \underset{(0.009)}{0.078}\textit{mandate} \\ n &= 1234, R^2 = 0.820 \end{aligned} \tag{7.1}$$

   What are the interpretations of the intercept and the coefficient of $mandate$. Provide the (approximate) 95% confidence interval for the coefficient on $mandate$. How can we use this confidence interval to test the hypothesis that $\beta_{\textit{mandate}} = 0$?

The team collect more data on age_i of the worker, $gender_i$ (1 = male, 0 = female), $ethnicity_i$ (1 = white, 0 = non-white), $graduate_i$ (= 1 if graduated from a university, = 0 otherwise), and $parttime$ (= 1 if i works part-time, = 0 if full time). They incorporate the variables and obtain the following estimates, with robust standard errors in the parentheses:

$$\begin{aligned} \widehat{\textit{wear}} &= \underset{(0.019)}{0.201} + \underset{(0.009)}{0.045}\textit{mandate} - \underset{(.009)}{0.033}\textit{gender} + \underset{(.0003)}{0.001}\textit{age} - \underset{(.016)}{0.027}\textit{ethnicity} \\ &\quad \underset{(.014)}{-0.104}\textit{parttime} + \underset{(.019)}{0.310}\textit{graduate} \\ n &= 1234, R^2 = 0.829 \end{aligned} \tag{7.2}$$

   (b) **(5 marks)** Compare the estimates of the coefficient on $mandate$ from (7.1) and that from (7.2). Why are they different? Explain if the OLS estimator for the coefficient of $mandate$ in either of the models is biased.

   (c) **(5 marks)** Discuss a statistical test to examine whether we can drop all the additional variables introduced in (7.2) altogether. Perform the test if you can. If not, specify the required additional information.

   (d) **(5 marks)** A colleague claims that without the robust standard errors, the OLS estimation of both models (7.1) and (7.2) will suffer from heteroskedasticity. Explain why you would agree with them. Clearly show any derivation you would take. Other than using robust standard errors, propose two statistical methods to solve this problem. Explain your answer.

8. A team of economists are interested in explaining the short-term interest rate in the US before the financial crisis of 2007 -2009. They look at the monthly interest rate for 10 years from 1997 to 2007, r_t, given by:

$$r_t - r_{t-1} = \beta_0 + \beta_1 r_{t-1} + \varepsilon_t. \tag{8.1}$$

Assume that for now the error term, which has zero mean and constant variance, ε_t is not serially correlated.

(a) **(3 marks)** Explain why it is important to assume ε_t to be free of serial correlation. In your answer, explain the term and discuss the impact on the OLS estimator for β_1 in (8.1) if this assumption fails.

(b) **(5 marks)** Provide sufficient conditions (on β_1) for r_t to be stationary and weakly dependent. Will the OLS estimator for a linear regression of r_t on r_{t-1} in this case be unbiased and consistent? Explain your answer.

(c) **(4 marks)** A colleague hypothesizes that $\beta_1 = 0$. Explain how you would test this hypothesis. What name do we give the process r_t when $\beta_1 = 0$. In your answer, be clear about the test specifics.

The team then examines the relationship between long-term interest rates R_t and short-term interest rates r_t: by looking at the following model:

$$R_t - r_t = \alpha_0 + \alpha_1 r_t + u_t, \tag{8.2}$$

where u_t is an error term that has zero mean.

(d) **(4 marks)** Let us assume $r_t = \beta_0 + \varepsilon_t$ (that is $\beta_1 = -1$) and ε_t remains serially uncorrelated. How would you test whether u_t is serially correlated? Explain your answer.

(e) **(4 marks)** Let us assume $\beta_1 = 0$. Discuss the statistical properties of the OLS estimators for α_0 and α_1 applied to Equation (8.2) when:

i. if u_t is stationary and weakly dependent?

ii. if u_t is non-stationary and strongly dependent?

9. An economist is interested in the following simultaneous system of equations to describe the demand of supply of hybrid electric in Britain:

$$Q_t = \alpha_0 + \alpha_1 P_t^{elec} + \alpha_2 P_t^{hybrid} + \alpha_3 Y_t + \alpha_4 ad_t + u_t \quad \text{(demand)}$$
$$Q_t = \beta_0 + \beta_1 P_t^{elec} + \beta_3 Subsidy_t + v_t \quad \text{(supply)}$$

where $Q_t =$ per capita quantity of electric cars in year t, $P_t^{elec} =$ the average price of electric cars inflation adjusted, $P_t^{hybrid} =$ the average price of hybrid cars inflation adjusted, $Y_i =$ real per capita disposable income, $ad_t =$ real per capital advertising expenditure by electric car companies, and $Subsidy_t =$ the average green subsidy per car (measured in thousands of GBP).

The endogenous variables in this model are Q_t and P_t^{elec}, and the exogenous variables are P_t^{hybrid}, Y_t, ad_t and $Subsidy_t$. The variance of u_t and v_t are, respectively σ_u^2 and σ_v^2 and $Cov(u_t, v_t) = \sigma_{uv} \neq 0$. We assume that the errors do not exhibit any correlation over time.

(a) **(5 marks)** Derive the reduced form for P_t^{elec}. Explain your steps.

(b) **(5 marks)** The OLS estimation of the demand function, based on annual data from 2007-2022 ($T = 16$), gave the following results (**all variables are in logs** and figures in parentheses are t-ratios).

$$\hat{Q}_t = \underset{(-3.04)}{-13.351} + \underset{(3.21)}{2.148} P_t^{elec} \underset{(-0.25)}{-0.155} P_t^{hybrid} \underset{(7.5)}{+3.224} Y_t \underset{(1.23)}{+0.663} ad_t, \tag{9.1}$$

Are the signs and statistical significance of the estimates what you would expect? Using your result in (a), explain why it is not appropriate to use OLS to estimate the parameters αs and βs.

(c) **(5 marks)** Discuss the difference that distinguishes overidentification from exact identification and under-identification. Explain the terms clearly. Is the supply equation over or under-identified? Provide one set of assumptions that would render the supply equation exactly identified.

(d) **(5 marks)** Using the available information, propose an estimator to consistently estimate the parameters in the supply equation.

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	**2·55**	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	**2·60**	0·99534	**2·75**	0·99702	**2·90**	0·99813	**3·05**	0·99886	**3·20**	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	**2·65**	0·99598	**2·80**	0·99744	**2·95**	0·99841	**3·10**	0·99903	**3·25**	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918	**3·30**	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	**3·263**	0·9994	**3·731**	0·99990	**3·916**	0·99995
3·105	0·9991	**3·320**	0·9995	**3·759**	0·99991	**3·976**	0·99996
3·138	0·9992	**3·389**	0·9996	**3·791**	0·99992	**4·055**	0·99997
3·174	0·9993	**3·480**	0·9997	**3·826**	0·99993	**4·173**	0·99998
3·215	0·9994	**3·615**	0·9998	**3·867**	0·99994	**4·417**	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	**5·0**	1·6449	**3·0**	1·8808	**2·0**	2·0537	**1·0**	2·3263	**0·10**	3·0902
45	0·1257	**4·8**	1·6646	**2·9**	1·8957	**1·9**	2·0749	**0·9**	2·3656	**0·09**	3·1214
40	0·2533	**4·6**	1·6849	**2·8**	1·9110	**1·8**	2·0969	**0·8**	2·4089	**0·08**	3·1559
35	0·3853	**4·4**	1·7060	**2·7**	1·9268	**1·7**	2·1201	**0·7**	2·4573	**0·07**	3·1947
30	0·5244	**4·2**	1·7279	**2·6**	1·9431	**1·6**	2·1444	**0·6**	2·5121	**0·06**	3·2389
25	0·6745	**4·0**	1·7507	**2·5**	1·9600	**1·5**	2·1701	**0·5**	2·5758	**0·05**	3·2905
20	0·8416	**3·8**	1·7744	**2·4**	1·9774	**1·4**	2·1973	**0·4**	2·6521	**0·01**	3·7190
15	1·0364	**3·6**	1·7991	**2·3**	1·9954	**1·3**	2·2262	**0·3**	2·7478	**0·005**	3·8906
10	1·2816	**3·4**	1·8250	**2·2**	2·0141	**1·2**	2·2571	**0·2**	2·8782	**0·001**	4·2649
5	1·6449	**3·2**	1·8522	**2·1**	2·0335	**1·1**	2·2904	**0·1**	3·0902	**0·0005**	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

Examiners' commentaries 2023 (May)

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2022–23. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE). The examinable materials are based on the new subject guide (Schafgans and Komarova, 2022, henceforth referred to as subject guide (2022)) and the new textbook:

- Wooldridge, J. *Introductory Econometrics: A Modern Approach.* (Cengage, 2020), seventh edition [ISBN 9781337558860].

Detailed reading references in this subject guide refer to the edition of the set textbook listed above. A new edition of this textbook may have been published by the time you study this course. You can, however, use a more recent edition; use the detailed chapter and section headings to identify relevant readings. Also check the virtual learning environment (VLE) regularly for updated guidance on readings.

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2022). The subject guide is intended for you to use in conjunction with the main textbook. It will act to highlight important aspects of the content in the chapters, provide further clarifications where deemed useful and also to extend it in certain places. For the final examination, you will need to be familiar with the material in both the textbook and the subject guide. You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

General remarks

Learning outcomes

At the end of the course, and having completed the Essential reading and activities, you should be able to:

- describe and apply the classical regression model and its application to cross-section data
- describe and apply the:
  - Gauss–Markov conditions and other assumptions required in the application of the classical regression model
  - reasons for expecting violations of these assumptions in certain circumstances
  - tests for violations

  - potential remedial measures, including, where appropriate, the use of instrumental variables

- recognise and apply the advantages of logit, probit and similar models over regression analysis when fitting binary choice models
- competently use regression, logit and probit analysis to quantify economic relationships using R
- describe and explain the principles underlying the use of maximum likelihood estimation
- apply regression analysis to time-series models using stationary time series, with awareness of some of the econometric problems specific to time series applications (for example, autocorrelation) and remedial measures
- recognise the difficulties that arise in the application of regression analysis to nonstationary time series, know how to test for unit roots, and know what is meant by cointegration.

Common mistakes committed by candidates

In accordance with the format adopted over the past three years, the May examination was once again conducted online. Two examination papers were administered, one for Zone A and another for Zone B. In the following discussion, we will address the common errors made by candidates in both zones, many of which reoccur on an annual basis despite changes to the subject guide. Subsequently, we will offer specific comments on each question in the examination papers for both zones.

Similar to the previous year, considerable effort was dedicated to ensuring that the examination structure and expected level of effort were consistent with those of previous years. Questions based on derivations and proofs with textbook settings that could easily be found in reference materials were limited. Greater emphasis was placed on questions that required candidates to adopt learned materials to a new setting, or evaluate and discuss the properties of estimators and test procedures. Furthermore, candidates were asked to explain their answers in a more intuitive manner and provide real-life examples.

Outstanding responses would demonstrate a discernible understanding of econometrics principles and accurate implementation in the given context. Additionally, we anticipate critical analysis of the methods or estimators employed, along with clear real-world examples. Attention to details and a thorough adherence to testing or estimation procedures are also expected. We commend the excellent performance of many candidates this year. However, a number of candidates displayed a lack of effort and preparation in their submitted attempt.

Candidates should realise that $\frac{1}{n}\sum_{i=1}^{n} u_i$ is not the same as $\mathrm{E}(u_i)$. So, while we typically assume $\mathrm{E}(u_i) = 0$, this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} u_i = 0$. Also, while we may be happy to assume $\mathrm{E}(x_i u_i) = 0$ (uncorrelatedness between the errors and regressors), this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} x_i u_i = 0$. Note that:

- both $\frac{1}{n}\sum_{i=1}^{n} u_i$ and $\frac{1}{n}\sum_{i=1}^{n} x_i u_i$ are random variables, which take the value 0 with probability 0 (continuous random variables)!
- $\mathrm{E}(u_i) = 0$ and $\mathrm{E}(x_i u_i) = 0$ are fixed, not stochastic!
- If we treat x_i as a random variable, we must use the Law of Iterated Expectations to obtain $\mathrm{E}(\widehat{\beta}) = \mathrm{E}(\mathrm{E}(\widehat{\beta} \mid X))$. Candidates often skipped this part in their proof.

The differences between sample and population moments need to come across clearly when looking at unbiasedness and making consistency arguments. In both cases, we first simplify our estimator (plug in the true model) to obtain:

$$\widehat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \beta + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \quad \text{with } x_i = X_i - \bar{X}.$$

- Note that this method applies for a wide range of estimator set-ups, not only for OLS. So in this case x_i can take any form of a function of a regressor.
- For *unbiasedness*, clearly indicate that you want to show that $\mathrm{E}(\widehat{\beta}) = \beta$. Unbiasedness does not follow from $\sum_{i=1}^{n} x_i u_i = 0$, instead it follows from $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) = 0$.

  If we treat x_i as fixed:

  $$\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) \equiv \mathrm{E}\left(\sum_{i=1}^{n} d_i u_i\right) = \sum_{i=1}^{n} d_i \mathrm{E}(u_i)$$

  and then unbiasedness follows as $\mathrm{E}(u_i) = 0$.
- For *consistency*, clearly indicate that you want to show that $\mathrm{plim}(\widehat{\beta}) = \beta$. Using the plim properties, we show:

  $$\begin{aligned}\mathrm{plim}\ \widehat{\beta} = \beta + \mathrm{plim}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) &= \beta + \frac{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i u_i\right)}{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)} \\ &\equiv \beta + \frac{\mathrm{plim}(\text{Sample Cov}(x,u))}{\mathrm{plim}(\text{Sample Var}(x))} \\ &= \beta + \frac{\mathrm{Cov}(x,u)}{\mathrm{Var}(x)} \quad \text{using the law of large numbers}\end{aligned}$$

  where $\mathrm{Cov}(x,u) = 0$ and $\mathrm{Var}(x) > 0$, ensuring we get consistency.
- Remember, the Law of Large Numbers guarantees that sample averages approach their corresponding population values. In order to maintain consistency in proofs, it is crucial to utilise the LLN to finalise the necessary steps. Likewise, the Law of Iterated Expectations (LIE) plays a significant role in establishing the connection between the conditional expectations and the unconditional expectations of an estimator.

A large number of candidates are not able to clearly distinguish between sample variance and covariance, and population variance and covariance.

The use of $\mathrm{Cov}(X,Y)$ and $\mathrm{Var}(X)$ should be restricted to describing the population covariance and variances, respectively, with definitions:

$$\mathrm{Cov}(X,Y) = \mathrm{E}((X - \mathrm{E}(X))(Y - \mathrm{E}(Y))) = \mathrm{E}(XY) - \mathrm{E}(X)\,\mathrm{E}(Y)$$

and:

$$\mathrm{Var}(X) = \mathrm{E}((X - \mathrm{E}(X))^2) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$$

(you also may denote $\mathrm{Cov}(X,Y) = \sigma_{XY}$ and $\mathrm{Var}(X) = \sigma_X^2$). They are typically unknown, but fixed, quantities.

The sample covariance and variance are estimators of the population covariance and variance, respectively. They are defined as:

$$\text{Sample Cov}(X,Y) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

and:

$$\text{Sample Var}(X) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

(you also may use $\widehat{\sigma}_{XY}$ and $\widehat{\sigma}_X^2$). You can compute them given the data.

With a slight abuse of notation, we often divide by n instead, which is irrelevant if we let n be large. The division by $n-1$ is a finite sample issue only (unbiasedness).

The sample covariance and variance show up in our definition of the OLS estimator of the slope in the simple linear regression model, not the population covariance and variance, as:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{\text{Sample Cov}(X,Y)}{\text{Sample Var}(X)} \neq \frac{\text{Cov}(X,Y)}{\text{Var}(X)}.$$

Treating them as being the same results in incorrect analyses and candidates losing significant marks.

It was very common to misunderstand the central limit theorem. While it provides the distribution of the *estimator* $\widehat{\beta}$, no matter what the true distribution of the errors is, it does *not* say that the distribution of the error term is approximately normal in large samples.

The examination made frequent use of the Gauss–Markov conditions, particularly in relation to discussing the BLUE property of an estimator or in conducting hypotheses tests. It is necessary to explicitly state the Gauss–Markov conditions rather than simply asserting their validity. A recommended approach, often followed by candidates, is to begin by explicitly stating the Gauss–Markov conditions and then refer back to them throughout the analysis. However, many candidates neglected to mention either the Gauss–Markov conditions or the normality assumption when conducting the t test.

Some candidates did not provide satisfactory answers when it came to interpreting empirical results. When engaging in the interpretation of an empirical result, it is essential to analyse the significance, magnitude, and sign of the coefficients. Moreover, it is crucial to establish the principle of *ceteris paribus* or holding all other factors constant. This principle ensures that the estimation can be properly understood as the marginal change in the dependent variable in response to a change in the independent variable of interest. Furthermore, a significant number of candidates failed to address the issue of omitting important and relevant explanatory variables, as well as irrelevant explanatory variables. By neglecting the former, there exists a risk of introducing bias, known as 'omitted variable bias'. On the other hand, omitting the latter would merely result in a loss of efficiency. The concepts of omitted variable bias and the associated OVB formula are significant subjects in the curriculum, particularly in the latest version of the course. It would be advantageous for candidates to engage in more extensive revision on these topics.

It is also important to understand that unbiasedness is a property of an estimator (not an estimate). It may be tempting to say an 'unbiased estimate'. Technically this statement is incorrect. It is the estimator (or the rule or the recipe) used to obtain the estimate that is biased, not the estimate itself.

Numerous candidates appear to encounter difficulties when deciphering coefficients in the context of a non-linear regression, such as log-linear/linear-log or log-log relationships. One helpful strategy for such interpretations is to bear in mind that coefficients pertaining to variables measured in logarithmic form should be interpreted in a relative sense.

In the case of dummy variables as regressors, candidates often missed that the coefficients have to be interpreted relative to the base case which is the left out dummy variable (see appropriate questions in the examinations for details).

Candidates particularly struggled with the difference between percentage points and % changes when it comes to the interpretation of coefficients in binary choice models. A coefficient in such a model gives the percentage point difference in the outcome variable when the regressor changes by 1 unit. The % change is a *relative* change that depends on the baseline value of the outcome variable (for example, the realisation when a regressor equals zero in the case of a dummy variable regressor). For example, if we have initially 50% of a sample buying ecological apples, a 10% increase will mean now $(1 + 0.1) \times 50 = 55\%$ of the sample buying ecological apples. A 10% point increase instead will result in $50 + 10 = 60\%$ of the sample choosing to buy ecological apples.

When conducting hypothesis testing, it is important to ensure that you clearly indicate the null and alternative hypotheses in terms of the true parameters, denoted such as β_1). Additionally, provide the test statistic in terms of the parameter estimator (t-stat, τ-stat, F-stat, LR). Specify the distribution of the test statistic under the null hypothesis and appropriate assumptions with the corresponding degrees of freedom. Clearly state the rejection rule, whether it is one-sided or two-sided, for a given significance level, typically at 5%, using appropriate critical values. Finally, provide a comprehensive interpretation of the result. It should be noted that missing any of these steps may result in a loss of marks. Another important consideration is selecting the critical value that is appropriate for the sample size. Using a critical value of 1.96, which is suitable for a large degree of freedom, in a small sample is an error that should be avoided. Furthermore, it is crucial to understand when to use an asymptotic version of a test, such as when normality of the errors is not assumed or when robust standard errors are employed.

When using confidence intervals in a hypothesis test, candidates should decide if the hypothesised value for the population parameter of interest β lies within the confidence interval. Many candidates used the default value of zero when discussing their confidence interval regardless of the population parameter in question. Another popular mistake is to conduct a two-sided test when the statement was worded with a clear direction of the relationship.

Several candidates inaccurately defined multicollinearity as a correlation between independent variables. While we acknowledge that a high correlation among the independent variables may serve as an indicator of multicollinearity, it is incorrect to solely rely on the correlation between the regressors to define multicollinearity. Additionally, numerous candidates mistakenly claimed that multicollinearity leads to biasedness and/or inconsistency.

Many candidates made the classic error of not recognising that $\mathrm{E}(X/Y) \neq \mathrm{E}(X)/\mathrm{E}(Y)$ in general. We do, however, have $\mathrm{plim}(X/Y) = \mathrm{plim}(X)/\mathrm{plim}(Y)$.

The concepts of autocorrelation and heteroskedasticity proved to be challenging for many candidates. A considerable number struggled to adequately explain the rationale behind employing robust standard errors or asymptotic methods for testing in the presence of autocorrelation and/or heteroskedasticity. Furthermore, many candidates received deductions in their grades when discussing the properties of an estimator that can effectively address these issues. It should be noted that some candidates erroneously proposed that addressing autocorrelation would automatically resolve issues related to inconsistency and biasedness.

Testing and recognising the problem of heteroskedasticity posed challenges, particularly when the scenarios were presented in a more real-life manner. One popular wording was that the variability of outcome variables differed across different groups of one or some of the regressors. A significant number of candidates failed to identify this as the classic problem of heteroskedasticity. Among those who did recognise it, many employed the Goldfeld–Quandt version of the heteroskedasticity test. However, in the new subject guide, the Goldfeld–Quandt test is no longer introduced due to its limitations compared to alternative options such as the Breusch–Pagan and the White tests. These alternatives offer advantages such as relaxing the assumption that the error term variance to be a monotonic function of a regressor. For future iterations of the course, it is recommended that candidates thoroughly study the latest version of the subject guide, as it presents the most up-to-date approach to the discussed materials.

When discussing the significance of differentiating between trend stationary and difference stationary processes, there are two essential aspects to consider: the implications for the long-term behaviour of the processes and the statistical implications, specifically the consistency of the

estimators. While the majority of candidates discussed the former aspect effectively, only a few briefly addressed the latter.

Candidates frequently received lower marks for their interpretation of the Error Correction Model (ECM), especially when discussing the characteristics of the estimator in the context of stationary or non-stationary variables. Many candidates encountered difficulties in deriving an ECM model. A significant portion of the candidates either chose to skip this question entirely or performed poorly when attempting it.

Some candidates appeared to be uncertain about the distinction between the Dickey–Fuller test and the Augmented Dickey–Fuller test (ADF). The discrepancy lies in the formulation of the test equation. In the augmented variant, additional lagged differences of the residuals are incorporated to ensure the absence of remaining autocorrelation in the regression of the test equation. Consequently, some individuals mistakenly utilised the ADF test equation when the DF test equation was intended. Relatedly, others forgot that even though we can use a very similar test statistic to the t statistic in the DF stationarity test (we call it the τ statistic), it does not follow a t distribution but the Dickey–Fuller distribution instead.

When discussing the marginal effect after a Probit estimation, candidates often did not specify the term z in the pdf $\phi(z)$. Some used the coefficients from the LPM regressions to calculate the marginal effects after a Probit/Logit estimation. A popular, related mistake was that candidates did not realise the inherent heteroskedasticity of the LPM.

A number of candidates seemed to be confused between the conditions for a valid instrument (exclusion restriction and relevance) and factors related to obtaining a more precise IV estimate (a higher sample size, a higher sample variability of X_i, a lower variance of the error term, and a stronger correlation between X_i and Z_i). More seriously, most candidates did not realise the need to use an IV estimator when the regression has a lagged dependent variable as a regressor in the presence of autocorrelation.

Note that the IV framework works for both a cross-sectional setting and a time series analysis.

One statistical way to test for an endogeneous variable is the Durbin–Wu–Hausman test. The test statistically compares the OLS and IV estimates, assuming we have valid instruments. Many candidates realised this but failed to discuss the test correctly (mistaking the test as a Chow or a t test). Many did not realise that we can only use the Durbin–Wu–Hausman test if we have more instruments than regressors that need instrumenting. In general, the exogeneity assumption is not testable because the error term is never observable. We must use economic intuitions to justify the validity of exogeneity.

Exogeneity and endogeneity are often very simply explained as variables 'outside of the model' or 'inside the model', respectively. While this approach is not wrong; it is not a particularly clear or helpful definition in general contexts that are not related to simultaneous equations models. Defining these concepts in terms of the relationship between regressor and the error term is much clearer and naturally leads to a way to show exogeneity or endogeneity by considering the covariance between regressors and error terms.

Candidates often missed that all exogenous regressors and instruments have to be included in the first stage when performing a TSLS estimation. Some candidates used the fitted values of the endogenous variable obtained from the first stage as the dependent variable in the second stage. This mistake is a serious misunderstanding. It is fine to indicate that this fitted value can be used as an instrument of the endogenous variable, but this is not really the way the second stage of TSLS is typically understood.

When asked about identification of structural form equations, candidates often compared the number of excluded regressors to the number of endogenous variables in general, without pointing out which variables are the excluded regressors and could therefore be used as an instrument. Candidates were also often unclear of the definitions of over-identification, under-identification, and exact identification. The best approach to this type of question is to discuss the endogenous and exogenous variables, and then think about the available exogenous variables that can be used as instruments.

When deriving the reduced forms, candidates should check conditions for their existence, in particular, whether the denominator of the expression is different from zero.

Just as in previous years, many candidates did not answer all parts of a question. Make sure you read the questions properly and provide all details that are requested. Not answering a question will automatically earn you a zero. Some candidates probably did not have a good examination plan.

Key steps to improvement

Essential reading for this course includes the new subject guide (2022) and the following:

- Wooldridge, J. *Introductory Econometrics: A Modern Approach.* (Cengage, 2020) seventh edition [ISBN 9781337558860].

Please note that as long as you read the Essential reading you are then free to read around the subject area in any text, paper or online resource. You will need to support your learning by reading as widely as possible and by thinking about how these principles apply in the real world. To help you read extensively, you have free access to the virtual learning environment (VLE) and University of London Online Library (see below).

- Dougherty, C. *Introduction to econometrics.* (Oxford: Oxford University Press, 2016) 5th edition [ISBN 9780199676828].
- Gujarati, D.N. and D.C. Porter *Basic econometrics.* (McGraw–Hill, 2009, International edition) 5th edition [ISBN 9780071276252].
- Maddala, G.S. and K. Lahiri *Introduction to econometrics.* (Wiley, 2009) 4th edition [ISBN 9780470015124].
- Stock, J.H. and M.W. Watson *Introduction to econometrics.* (Pearson, 2012) 3rd edition [ISBN 9781408264331].

To understand the subject it is important to supplement the reading of the textbook Wooldridge with the subject guide (Schafgans and Komarova, 2022). The subject guide is intended for you to use in conjunction with the main textbook. It will act to highlight important aspects of the content in the chapters, provide further clarifications where deemed useful and also to extend it in certain places. For the final examination, you will need to be familiar with the material in both the textbook and the subject guide.

The subject guide provides a range of activities that will enable you to test your understanding of the basic ideas and concepts. We want to encourage you to try the exercises you encounter throughout the material before working through the solutions provided on the virtual learning environment (VLE). With econometrics the motto has to be 'practise, practise, practise...'. It is the best way to learn the material and prepare for examinations. The course is rigorous and demanding, but the skills you will be developing will be rewarding and well recognised by future employers.

At the end of each chapter you will find an overview of the material covered, a list of key terms and concepts, and a reminder of the learning outcomes. It concludes with additional examination based exercises that will enable you to test your knowledge and understanding with worked-out solutions. Worked-out solutions for the exercises can be found on the VLE (downloadable pdf files). Candidates should utilise datasets using standard regression programmes (R in our course, but Stata or Python are other popular choices). This will help the understanding of the subject.

Examination revision strategy

Many candidates are disappointed to find that their examination performance is poorer than they expected. This may be due to a number of reasons, but one particular failing is '**question spotting**', that is, confining your examination preparation to a few questions and/or topics which have come up in past papers for the course. This can have serious consequences.

We recognise that candidates might not cover all topics in the syllabus in the same depth, but you need to be aware that examiners are free to set questions on **any aspect** of the syllabus. This means that you need to study enough of the syllabus to enable you to answer the required number of examination questions.

The syllabus can be found in the Course information sheet available on the VLE. You should read the syllabus carefully and ensure that you cover sufficient material in preparation for the examination. Examiners will vary the topics and questions from year to year and may well set questions that have not appeared in past papers. Examination papers may legitimately include questions on any topic in the syllabus. So, although past papers can be helpful during your revision, you cannot assume that topics or specific questions that have come up in past examinations will occur again.

If you rely on a question-spotting strategy, it is likely you will find yourself in difficulties when you sit the examination. We strongly advise you not to adopt this strategy.

Examiners' commentaries 2023 (May)

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2022–23. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE). The examinable materials are based on the new subject guide (Schafgans and Komarova, 2022, henceforth referred to as subject guide (2022)) and the new textbook:

- Wooldridge, J. *Introductory Econometrics: A Modern Approach.* (Cengage, 2020), seventh edition [ISBN 9781337558860].

Detailed reading references in this subject guide refer to the edition of the set textbook listed above. A new edition of this textbook may have been published by the time you study this course. You can, however, use a more recent edition; use the detailed chapter and section headings to identify relevant readings. Also check the virtual learning environment (VLE) regularly for updated guidance on readings.

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2022). The subject guide is intended for you to use in conjunction with the main textbook. It will act to highlight important aspects of the content in the chapters, provide further clarifications where deemed useful and also to extend it in certain places. For the final examination, you will need to be familiar with the material in both the textbook and the subject guide. You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone A

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from **Section A** (8 marks each) and **THREE** from **Section B** (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Section A

Answer all questions from this section. Each question carries 8 marks.

Question 1

A researcher is interested in the relationship between (the natural logarithm of) family income ($\log faminc_i$) and (the natural logarithm of) birthweight of the

family's first child ($\log bweight_i$). Using a random sample of 2023 families in the UK, she estimates the following regression:

$$\log bweight_i = \beta_0 + \beta_1 \log faminc_i + u_i. \tag{1.1}$$

Suppose that the true population model is in fact:

$$\log bweight_i = \beta_0 + \beta_1 \log faminc_i + \beta_2 fatheduc_i + \varepsilon_i, \tag{1.2}$$

where $fatheduc_i$ is the total years of schooling of the father. Assume that the model given by (1.2) satisfies Assumptions MLR.1–MLR.4 (linearity in parameters, random sampling, no perfect multicollinearity, and zero conditional mean).

(a) A colleague obtains the OLS estimates of Equation (1.1) and comments that because the (unreported) R-squared from the OLS estimation of (1.1) is too low, we cannot interpret the estimate in a causal manner. Discuss whether you disagree or agree with this comment. In your answer, clearly derive the bias, if any, of the OLS estimator of β_1.

(4 marks)

(b) Another colleague wants to test whether 20% extra income has the same effect on the child's birth weight as an additional year of paternal education. Discuss how you can test for this claim. Clearly specify your null hypothesis and other test specifics.

(4 marks)

Reading for this question

Subject guide (2022), Chapters 4.2.3, 5.2.2, 7.2.2.

Wooldrige, Sections 3.3, 3.2g, 4.2, 5.2, Appendix B.5e.

Approaching the question

Part (a) is a question about causal interpretation with regards to the independence of error terms (OVB). Candidates were expected to clearly disregard R-squared, which is a goodness-of-fit measure, as having anything to do with causality. The best answers linked the omission of *fatheduc* and explain the bias by either the OLS formula or the OVB formula. Only a few candidates got full marks for this question.

Part (b) is a routine hypothesis testing question, noting that since MLR5 is not given, robust standard errors must be used. As a result, an asymptotic version of the t test is appropriate. Only a handful of candidates noticed this and applied the correct test. As with other testing questions, most marks were lost due to missing details or steps in the testing procedure. Overall, Question 1 was relatively well attempted by candidates.

(a) The causality interpretation or unbiasedness of an estimator is about whether the regressors are independent of the error term or the conditional mean. The R-squared is a goodness-of-fit measure and has nothing to do with this aspect.

However, it is correct that the estimator of β_1 is probably biased because of the omitted

variable $fatheduc_i$. The OLS estimator of β_1 is:

$$\begin{aligned}
\widehat{\beta}_1 &= \frac{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i}) \log bweight_i}{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i})^2} \\
&= \frac{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i})(\beta_0 + \beta_1 \log faminc_i + \beta_2 fatheduc_i + \varepsilon_i)}{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i})^2} \\
&= \beta_1 + \beta_2 \frac{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i}) fatheduc_i + \varepsilon_i}{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i})^2} \\
&\quad + \frac{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i}) \varepsilon_i}{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i})^2}.
\end{aligned}$$

Taking the conditional expectation of the estimator on $X = [\log faminc_i, fatheduc_i]$, for $i = 1, \ldots, n$:

$$\begin{aligned}
\mathrm{E}\left(\widehat{\beta}_1 \mid X\right) &= \beta_1 + \beta_2 \frac{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i}) fatheduc_i}{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i})^2} \\
&\quad + \mathrm{E}\left(\frac{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i}) \varepsilon_i}{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i})^2} \mid X \right) \\
&= \beta_1 + \beta_2 \frac{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i}) fatheduc_i}{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i})^2}.
\end{aligned}$$

The equality follows MLR.2 and MLR.4. The bias of the estimator following (1.1) is:

$$\beta_2 \frac{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i}) fatheduc_i}{\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i})^2}.$$

This term can be 0 (or no bias) if:

$$\sum_{i=1}^n (\log faminc_i - \overline{\log faminc_i}) fatheduc_i = \mathrm{SampleCov}(\log faminc, fatherduc) = 0.$$

Otherwise, there is an omitted variable bias. Candidates can argue using the OVB formula.

(b) Since MLR.5 was not given, we must use an OLS estimation of (1.2) with robust standard errors.

The regressors are in log and at level, which means a 1% increase in family income is associated with a $0.01\beta_1$ change in the outcome, whereas the estimate of β_2 is the change in the outcome given one additional year of paternal education. Therefore, the null hypothesis is $\mathrm{H}_0 : 0.2\beta_1 = \beta_2$. Let $\mathrm{H}_1 : 0.2\beta_1 \neq \beta_2$ be the two-sided alternative and set the significance level at 5%.

The t statistic to test the null hypothesis H_0 is written as:

$$t = \frac{0.2\widehat{\beta}_1 - \widehat{\beta}_2}{se(0.2\widehat{\beta}_1 - \widehat{\beta}_2)} \quad \text{under } \mathrm{H}_0.$$

Obtaining the estimates, we know $0.2\widehat{\beta}_1 - \widehat{\beta}_2$. The standard error of $0.2\widehat{\beta}_1 - \widehat{\beta}_2$ is given by:

$$se(0.2\widehat{\beta}_1 - \widehat{\beta}_2) = \sqrt{0.04 \cdot se(\widehat{\beta}_1)^2 + se(\widehat{\beta}_2)^2 - 0.4 \cdot s_{12}}$$

where s_{12} is the estimate of $\mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_2)$, and $se(\widehat{\beta}_1), se(\widehat{\beta}_2)$ are the robust standard errors of the respective estimates. To implement the above t test, we need to know s_{12}.

Under the MLR.1–4, and robust standard errors, we can use the asymptotic version of the t distribution (standard normal distribution) under H_0. Once we get s_{12}, we can implement the t test. The rejection rule is $|t| > c$, where c is the 97.5th percentile of the standard normal distribution.

Question 2

State whether the following claims are True or False. Briefly explain why (two to three sentences).

(a) **Since x^2 is an (exact) function of x, using both x and x^2 as regressors in a regression will lead to the multicollinearity problem. *Note: In your answer, clearly define what multicollinearity is.***

(2 marks)

(b) **In a multivariate regression with 3 regressors as follows,**

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + u_i, \tag{2.1}$$

the R-squared of the OLS regression of the restricted model will be larger than that of the unrestricted model if we cannot reject the restriction (or constraint) that $\alpha_2 = \alpha_3$, and smaller if we can reject the constraint.

(2 marks)

(c) **Adding an omitted variable in a regression will sometimes result in a change in the statistical significance of the estimates of the regressors.**

(2 marks)

(d) **Autocorrelation and heteroskedasticity are exactly the same problem concerning the error term of a regression model. The only difference is that the former is a time-series concept and not applicable for cross-sectional data. *Note: In your answer, clearly define what autocorrelation is.***

(2 marks)

Reading for this question

Subject guide (2022), Chapters 3.2.3; 5.2.5; 4.2.3; 8.2.1, 14.2.1.

Wooldrige, Sections 2.3, 4.5 and Appendix B.5f, 3.3, 3.2g, 8.1, 12.1–12.2.

Approaching the question

Part (a) is a routine question about multicollinearity. The best answers distinguished between perfect and imperfect collinearity and explained well the consequences. Part (b) is about R-squared and its link to the F test. Most candidates attempted this question well by writing down the R-squared version of the F-stat formula. Part (c) is a tricky question on OVB and its effect on the estimate (and also estimator). Most candidates correctly thought the statement to be true, but only a few offered the right explanation. Part (d) discusses two seemingly similar, but distinct, concepts. Most candidates answered the question very well.

(a) **False.** Multicollinearity is a problem if the regressors are *linearly related.* x are x^2 are not linearly correlated.

(b) **False.** The R^2 from the restricted model will always be smaller (fewer regressors) regardless of whether the constraint holds or not. If the null is rejected, the decrease in R^2 is larger than that when the null is not rejected.

(c) **True.** For example, if the true effect is zero but because of the OVB, say positive, we may observe a positive and statistically significant estimate. By controlling for this OVB, the estimate in the longer regression will now be closer to zero and statistically insignificant.

(d) **False.** The two problems are different: heteroskedasticity is the relationship between the variance of the error terms and the regressors; autocorrelation occurs if the error term is correlated with itself at a different time period.

Question 3

A researcher is interested in the relationship between the quarterly unemployment rate $unemploy_t$ and the quarterly rate of vacancies $vacancy_t$ in Britain since 2000. She estimates the following regression model using observations from T quarters:

$$unemploy_t = \beta_0 + \beta_1 vacancy_t + \lambda t + \gamma unemploy_{t-1} + \varepsilon_t, \ |\rho| < 1, \ t = 2, \ldots, T, \quad (3.1)$$

where ε_t is the white noise error term. The researcher observes that unemployment rates are weakly dependent.

(a) Discuss why it is important to establish the weak dependence of $unemploy_t$. In your answer, clearly define the term and its conditions.

(2 marks)

(b) The researcher includes the time trend t to account for the nonstationarity of $unemploy_t$. Define what stationarity means and distinguish between time-stationarity and difference-stationarity. Briefly explain the inclusion of the time trend t and how it helps the OLS estimation of the model.

(4 marks)

(c) Is the OLS estimator for β_1 unbiased? Explain. In your answer, clearly indicate if and when the Gauss–Markov assumptions are potentially violated, and the difference between contemporaneous exogeneity and strict exogeneity.

(2 marks)

Reading for this question

Subject guide (2022), Chapters 13.2.4, 13.2.5, 11.2.1, 15.2.2.

Wooldrige, Sections 11.1, 11.2, 16.1.

Approaching the question

Part (a) pertains to a textbook question regarding the definition of weak dependence. Some candidates did not demonstrate a clear understanding of the significance of the assumption in the context of time series analysis. Best answers illustrated the connection between weak dependence and the law of large numbers (LLN), despite this not being requested.

Part (b) requires the definition of stationarity. Only weak stationarity was expected. Best candidates made clear the difference between difference- and trend-stationarity. Note that a time trend should be included in the analysis. Part (c) posed a challenge. Most candidates did not recognise the set-up of an ADL with a lagged outcome variable. Best candidates explained clearly the weak and strict exogeneity assumptions.

(a) A process is weakly dependent if the correlation between two random temporal values of the process will disappear as the time distance between them increases: $\text{Corr}(x_t, x_{t+h}) \to 0$ as $h \to \infty$. This assumption is important because it replaces the assumption of independence so that the LLN and the CLT hold: asymptotic independence of the process.

(b) A process is stationary if, intuitively, its stochastic properties (such as the mean, and variance) and its temporal dependence structure (such as covariance over time) do not depend on time. It is a stronger version of the identically distributed assumption. It is important because we know that the nature of any correlation between adjacent terms is the same across all time periods.

A trend stationary process if by detrending, we can make it stationary. A difference stationary process is one where by differencing we can make it stationary. It is important to distinguish the two types as they have implications for the long-run behaviour of a process after a shock to the process.

We include the time trend so that we can still use standard robust inference procedures because trend stationary processes can also satisfy TS.1'.

(c) Candidates are expected to show that because of the presence of $unemploy_{t-1}$, the strict exogeneity assumption is violated: that is the error term is not strictly independent of the regressors at all times. Thus the OLS estimator is biased. The difference is that the contemporaneous exogeneity only requires the error term to be independent of the current regressors $\mathrm{E}(\varepsilon_t \mid x_t) = 0$.

Question 4

Consider the following bivariate regression model:

$$\boldsymbol{y_i = \beta_0 + \beta_1 x_i + u_i,} \tag{4.1}$$

for a given random sample of n observations $\{(y_i, x_i)\}_{i=1}^n$ under the standard Assumptions SLR.1–SLR.4 (linearity in parameters, random sampling, non-zero sample variation in the regressor, and zero conditional mean). Let:

$$\boldsymbol{z_i = 20 + \sqrt[23]{x_i}, \quad \bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i,}$$

and define the following slope estimator:

$$\boldsymbol{\tilde{\beta}_1 = \frac{\sum_{i=1}^n (z_i - \bar{z}) y_i}{\sum_{i=1}^n (z_i - \bar{z}) x_i}.}$$

(a) Carefully show that $\tilde{\beta}_1$ is an unbiased estimator of β_1.

(4 marks)

(b) Assume the homoskedasticity Assumption, SLR.5, derive the conditional variance $\mathrm{Var}(\tilde{\beta}_1 \mid x_i)$. Without technical details, discuss whether $\tilde{\beta}_1$ is the BLUE of β_1. In your answer, clearly define what BLUE means and SLR.5 involves. Discuss any consequences if Assumption SLR.5 fails.

(4 marks)

Reading for this question

Subject guide (2022): Chapters 3.2 and 7.2.1.

Wooldridge Section 2.2, 2.5, Appendix 2A and Appendices B.4a, B.4b, B.4c.

Approaching the question

Part (a) is a popular set-up with a tweak to involve the academic year 2022–23 in the question. The best answer would simplify the derivation by using w_i instead of the longer term, and correctly proceed through the derivation. The key step in getting unbiasedness is to invoke the LIE, which some candidates forgot to explicitly mention and lost marks on. Better performance must involve clear and correct annotations and explanations of the steps.

Part (b) focuses on the derivation of the variance of an OLS estimator, and a discussion of the BLUE property under the Gauss–Markow theorem. Correct algebra was expected as well as clear explanations with references to the SLRs. Similar to part (a), some candidates forgot the use of LIE, or not specifying the list of assumptions under which OLS becomes BLUE. Overall, candidates did well on this question. Some did not indicate what would happen if SLR.5 fails.

(a) Define $w_i = \frac{\sum_{i=1}^n (z_i - \bar{z})}{\sum_{i=1}^n (z_i - \bar{z}) x_i}$ and replace (4.1) in the slope estimator. It follows:

$$\begin{aligned}\tilde{\beta}_1 = \sum_{i=1}^n w_i y_i &= \sum_{i=1}^n w_i(\beta_0 + \beta_1 x_i + u_i) \\ &= \beta_0 \left[\sum_{i=1}^n \frac{(z_i - \bar{z})}{\sum_{i=1}^n (z_i - \bar{z}) x_i}\right] + \beta_1 \left[\sum_{i=1}^n \frac{(z_i - \bar{z}) x_i}{\sum_{i=1}^n (z_i - \bar{z}) x_i}\right] + \sum_{i=1}^n w_i u_i.\end{aligned}$$

Because $\sum_{i=1}^{n}(z_i - \bar{z}) = 0$ and $\sum_{i=1}^{n} \frac{(z_i-\bar{z})x_i}{\sum_{i=1}^{n}(z_i-\bar{z})x_i} = 1$, the estimator can be rewritten as:

$$\tilde{\beta}_1 = \beta_1 + \sum_{i=1}^{n} w_i u_i. \tag{4.2}$$

Take the conditional expectation on $X = \{x_i\}_{i=1}^{n}$ of $\tilde{\beta}_1$, so that:

$$\mathrm{E}(\tilde{\beta}_1 \mid X) = \beta_1 + \mathrm{E}\left(\sum_{i=1}^{n} w_i u_i \mid X\right) = \beta_1 + \sum_{i=1}^{n} E(w_i u_i \mid X) = \beta_1 + \sum_{i=1}^{n} w_i \mathrm{E}(u_i \mid X) = \beta_1.$$

The first equality follows from (4.2), the second equality follows SLR.2, the third equality follows the properties of conditional expectation and the fact that w_i is a function of x_i. The last equation follows SLR.4.

(b) From Equation (4.2) and taking the conditional variance on $X = \{x_i\}_{i=1}^{n}$ of $\tilde{\beta}_1$, we have:

$$\mathrm{Var}(\tilde{\beta}_1 \mid X) = \mathrm{Var}\left(\beta_1 + \sum_{i=1}^{n} w_i u_i \mid X\right) = \sum_{i=1}^{n} \mathrm{Var}(w_i u_i \mid X) = \sum_{i=1}^{n} w_i^2 \mathrm{Var}(u_i \mid X).$$

Using SLR.5, which assumes that $\mathrm{Var}(u_i \mid X) = \delta^2$, a constant, the variance of interest is:

$$\mathrm{Var}(\tilde{\beta}_1 \mid X) = \sigma^2 \frac{\sum_{i=1}^{n}(z_i - \bar{z})^2}{\left[\sum_{i=1}^{n}(z_i - \bar{z})x_i\right]^2}. \tag{4.3}$$

The BLUE (Best Linear Unbiased Estimator) is the best linear unbiased estimator, which implies that the estimator is unbiased and has the smallest variance. Under the Gauss–Markov theorem (SLR.1–SLR.5), our estimator is not the BLUE. If SLR.5 fails, we cannot obtain the final formula in (4.3), and the conventional t test based on this variance formula becomes invalid.

Question 5

A researcher is interested in examining the time series behaviour for the quantity of cement sold in a particular region in Japan for a given month, denoted as $\log(Q_t)$, and the quantity of new building permits, $\log(Permit_t)$. She obtains estimates using the following model over a period of T months in the Japanese region:

$$\log(Q_t) = \beta_0 + \beta_1 \log(Permit_t) + \beta_2 \log(Permit_{t-1}) + \varepsilon_t, \quad t = 2, \ldots, T,$$

where the error term ε_t follows an AR(1) model, while the other processes are weakly dependent.

(a) **Which of the Gauss–Markov assumptions is most likely to be violated? Clearly indicate the consequences of this violation to the OLS estimation of the regression.**

(2 marks)

(b) **Discuss how you can test for the violation of the Gauss–Markov assumption identified in (a). Be clear about your null hypothesis and the assumptions needed. What should we do if we reject the test's null hypothesis?**

(4 marks)

(c) **A colleague is concerned that the monthly quantity of cement is often measured with some random human mistakes when collecting the data. Briefly discuss the implications if this concern is true.**

(2 marks)

Reading for this question

Subject guide (2022), Chapters 14.2.1, 10.2.1.

Wooldrige, Sections 12.1–12.2, 12.3 (skip 12.3b), 9.4a.

Approaching the question

Part (a) asks about the implications of autocorrelation, which many candidates did well. The violation of TS.5 and TS.5' was also well discussed. Best answers discussed and made clear the definition of autocorrelation and the consequences.

Part (b) is about an autocorrelation test with an AR(1) error term. Most candidates did well, some lost marks for missing details of the testing procedure. Others did not answer the remedial measures once we reject the null hypothesis. Some discussed the Durbin–Watson test, which is not introduced in the current subject guide (2022). The Durbin–Watson test (discussed in Section 12.3, Wooldridge) is a finite sample test that permits testing for no autocorrelation against AR(1) dependence. It requires strict exogeneity and the full set of CLM assumptions (including normality). As these assumptions are deemed to be quite restrictive in time series models, we will not discuss this test in the course. You may also ignore the discussion of the LM form of the test for autocorrelation (called the Breusch–Godfrey test).

Part (c) is a discussion of measurement error (classical errors in variables) in the outcome variable. Only a few recognised this problem and discussed it correctly. The key intuition is that the errors would just add noise to the estimation, resulting in a loss of efficiency but not bias. Best candidates described the problem well.

(a) Because the error term follows an AR(1) model, the two consecutive error terms are correlated and we potentially have the problem of autocorrelation (TS.5 is violated but TS.5' is also likely). Candidates are not expected to discuss the other TS assumptions. The presence of autocorrelation itself does not affect the unbiasedness or consistency of our OLS estimator. However, autocorrelation invalidates the standard errors, meaning that tests using the conventional standard errors are invalid. The OLS estimator is also no longer BLUE (loss of efficiency) even if the other TS assumptions hold.

(b) Candidates are expected to perform an autocorrelation test. Assume that ε_t can be expressed as follows:

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t, \quad \text{with } |\rho| < 1$$

and v_t is a white noise generated randomly from a normal distribution with mean zero and constant variance. We can also assume TS.4 (or TS.4') so there is no need for robust standard errors. To test for autocorrelation, the null hypothesis is $\mathrm{H}_0 : \rho = 0$ against $\mathrm{H}_1 : \rho \neq 0$. The testing equation is obtained from using the OLS residuals:

$$\widehat{\varepsilon}_t = \log(Q_t) - (\widehat{\beta}_0 + \widehat{\beta}_1 \log(Permit_t) + \widehat{\beta}_2 \log(Permit_{t-1})).$$

The testing equation is:

$$\widehat{\varepsilon}_t = \gamma_0 + \gamma_1 \log(Permit_t) + \gamma_2 \log(Permit_{t-1}) + \rho\widehat{\varepsilon}_{t-1} + u_t.$$

We must use the asymptotic t statistic $\widehat{\rho}/se(\widehat{\rho})$ for a large sample.

At the 5% level of significance, we should reject if $|\widehat{\rho}/se(\widehat{\rho})| > 1.96$, which follows a $N(0,1)$ distribution.

Candidates can instead use the LM test statistic $(T-1)R^2 \sim \chi^2_1$.

If we reject the null hypothesis, we should be concerned about the presence of autocorrelation. We can use the HAC robust standard errors to fix the problem.

(c) Candidates are only expected to discuss the random measurement errors in the dependent variable, which will lead to an increase in the standard errors (but not invalidating). The intuition is that we are including more noise into the regression, which will be absorbed in the error term, causing a lower precision.

Section B

Answer three questions from this section. Each question carries 20 marks.

Question 6

An economist is interested in a regression model for the (logarithm of) output-per-capita y_t in a given country and at a time period t as in:

$$y_t = \alpha + \rho y_{t-1} + \beta_1 x_t + \beta_2 h_t + \gamma t + \varepsilon_t, \tag{6.1}$$

where x_t is the (logarithm of) investment-per-capita and h_t is the (logarithm of) research and development expense per capita at time t.

(a) Assume that y_t is stationary ($|\rho| < 1$) and $\varepsilon_t + \lambda\varepsilon_{t-1} = u_t$, where u_t is a white noise. Discuss the implications of this assumption for the OLS estimators for β_1.

(5 marks)

Assume for subsequent parts (b), (c), (d) that the error term (ε_t) is not serially correlated.

(b) Another economist claims that the investment rates are also affected by the current output-per-capita that can be summarised as follows:

$$x_t = \delta_0 + \delta_1 y_t + \delta_2 h_t + \eta_t. \tag{6.2}$$

Assume that η_t does not suffer from autocorrelation and $\text{Cov}(\varepsilon_t, \eta_t) = 0$. Discuss the properties of the OLS estimator of β_1 in Equation (6.1) if the claim is correct. In your answer, define the terms 'endogenous' variable and 'exogenous' variable.

(5 marks)

(c) Derive the reduced form functions for both x_t and y_t. Explain the direction of bias of the OLS estimator for δ_1 in (6.2). *Hint:* For simplicity, you can assume $\delta_2 = 0$ and investment is thought to be beneficial to output.

(5 marks)

(d) Discuss the identification of Equation (6.2). Describe a consistent estimator for δ_1 using the available information. In your answer, be clear about the assumptions needed.

(5 marks)

Reading for this question

Subject guide (2022), Chapters 13.2.4, 14.2.1, 11.2.1–3, 9.2.5.

Wooldrige, Sections 11.1, 12.3, 16.1–16.3, 15.3a-15.3d.

Approaching the question

Question 6 combines the materials on the IV method and time series analysis. It was popular and very well attempted. Part (a) is a familiar setting in which autocorrelated error terms and the presence of a lagged outcome would result in the OLS estimator being inconsistent and biased. The concern of inefficiency is of secondary importance. The best candidates showed the why TS.3' is violated by showing the correlation between y_{t-1} and the error terms via the MA(1) structure. Some candidates neglected the discussion part.

Part (b) is another familiar set-up for simultaneity (or reverse causality) bias. Some verbal explanations were expected, including which TS assumptions were violated and their consequences. Most candidates recognised the endogenous and exogenous variables in the system of equations, and explained their choices well.

Part (c) requires a derivation of the reduced forms using substitutions. Most candidates did well here. The best candidates would check for the existence condition of the equations (denominator different from zero). Most correctly identified the sign of the bias using the hint on the relationship between investment and output.

Part (d) is another routine question on the TSLS. Most correctly stated the two steps. The best answers discussed how the standard errors are to be calculated, as well as offered critical evaluations of the two assumptions for a valid IV. Only a few candidates mentioned that y_{t-1} may not satisfy the exogeneity assumption (potential correlation between the instruments and y_t being dependent on time).

(a) Candidates are expected to note that because of autocorrelation and the presence of lagged outcome, the OLS estimators are not only inefficient and biased but also inconsistent. Candidates are expected to show that $\text{Cov}(y_{t-1}, \varepsilon_t) \neq 0$ hence TS.3' is not satisfied. One can show that y_{t-1} depends on ε_{t-1} (by lagging one period), which in turn is related to u_t by the MA(1) structure of ε_t.

(b) The problem is simultaneity bias. Candidates are expected to verbally explain why. The consequence is that the OLS estimators in (6.1) are unbiased and inconsistent. y_t and x_t are determined within the structural models hence endogenous, y_{t-1} and t are determined outside of the model and hence exogenous. The exogeneity of y_{t-1} can be debated and candidates can discuss this point here as well.

(c) Candidates are expected to show the substitution and solution to the reduced forms, overall the results are:

$$y_t = \pi_{10} + \pi_{11} y_{t-1} + \pi_{12} t + \pi_{13} h_t + v_1$$

$$x_t = \pi_{20} + \pi_{21} y_{t-1} + \pi_{22} t + \pi_{23} h_t + v_2$$

where:

$$\pi_{10} = \frac{\alpha + \beta\delta_0}{1 - \beta\delta_1}, \quad \pi_{11} = \frac{\rho}{1 - \beta\delta_1}, \quad \pi_{12} = \frac{\gamma}{1 - \beta\delta_1}, \quad \pi_{13} = \frac{\beta_2}{1 - \beta\delta_1}$$

$$\pi_{20} = \frac{\alpha\delta_1 + \beta\delta_0\delta_1}{1 - \beta\delta_1}, \quad \pi_{21} = \frac{\rho\delta_1}{1 - \beta\delta_1}, \quad \pi_{22} = \frac{\gamma\delta_1}{1 - \beta\delta_1}, \quad \pi_{23} = \frac{\beta_2\delta_1}{1 - \beta\delta_1}$$

$$v_2 = \frac{\alpha\varepsilon_t + \eta_t}{1 - \beta\delta_1}, \quad v_1 = \frac{\varepsilon_t + \delta_1\eta_t}{1 - \beta\delta_1}.$$

With the reduced forms, we can clearly see that y_t and η_t are correlated. To show the direction of the bias, candidates can show that the OLS estimator of $\widehat{\delta}_1$ can be written as:

$$\text{plim}(\widehat{\delta}_1) = \delta_1 + \frac{\text{plim SampleCov}(y_t, \eta_t)}{\text{plim SampleVar}(y_t)} = \delta_1 + \frac{\text{Cov}(y_t, \eta_t)}{\text{Var}(y_t)}.$$

The sign of the bias depends on the sign of $\text{Cov}(y_t, \eta_t)$, which in turn is dependent on the sign of $\text{Cov}(v_1, \eta_t)$, which is $\frac{\delta_1}{1-\beta\delta_1}$. Normally this would be positive because we should expect $\delta_1 > 0$ (more investment, more output) and we have a positive bias.

(d) There are two available exogenous variables as instruments (t, y_{t-1}) for y_t so the Equation is over-identified. We can only obtain a consistent estimator for α_1, by using the Two Stage Least Squares:

1. Estimate the reduce form of y_t, and get the following fitted value, $\widehat{y}_t$, where:

$$\widehat{y}_t = \widehat{\pi}_0 + \widehat{\pi}_1 y_{t-1} + \widehat{\pi}_2 t + \widehat{\pi}_3 h_t.$$

2. Plug $\widehat{y}_t$ instead of y_t into the reduced form for x_t, and run an OLS estimation to obtain the consistent estimators:

$$x_t = \text{constant} + \delta_1 \widehat{y}_t + \delta_2 h_t + \text{error}.$$

Now candidates are expected to mention the relevance (how to test it by running an F test in the first stage) and the exogeneity assumption (whether we can treat y_{t-1} and particularly time as exogenous).

Question 7

A team examines the determinants of movie ratings on IMDb, the world's most popular and authoritative source for movie, TV, and celebrity content. They have the data on all 645 movies released in 2022 with information on the number of positive ratings in the first week after release $Ratings_i$, the total cost of production $cost_i$ in millions of dollars, the number of super-stars in the cast $SuperStars_i$, and a binary variable to indicate whether the director has ever won an Oscar, regarded by many as the most prestigious and significant award in the entertainment industry worldwide, $Oscars_i$. They obtain the following estimation, using the logarithms of $Ratings_i$ and $cost_i$:

$$\log(\widehat{Ratings_i}) = \underset{(.114)}{.123} + \underset{(.231)}{1.51}\log(cost_i) + \underset{(.156)}{.507}\,SuperStars_i + \underset{(.123)}{.501}\,Oscars_i, \qquad (7.1)$$

$$n = 645,\ R^2 = .458.$$

The robust standard errors are reported in parentheses.

(a) **What is the interpretation of the coefficient on $Oscars_i$? Are the sign and size of the coefficient as you would expect?**

(2 marks)

(b) **Should we also include another binary variable $nonOscars_i = 1 - Oscars_i$ to indicate whether the director has not been awarded an Oscar? Explain what happens to the regression model in (7.1) if we do.**

(3 marks)

(c) **A colleague suggests that the effect of one additional super-star is more pronounced on the rating for those movies with an Oscar-winning director. Discuss how you can test this claim. In your answer, clearly define the regression estimation or any additional information needed.**

(5 marks)

(d) **Explain how you can test whether neither having super-star actors in the casting nor production cost have any effect on the movie ratings. Clearly state any additional information you may require.**

(5 marks)

(e) **The research manager suggests that the ratings are likely to be more or less variable for movies with a large production cost than those with lower expenses. Discuss a statistical test for this claim. Discuss whether such a test is needed given the specification in (7.1). *Hint: Think about the relationship between the variance of the population error term and* $\log(cost_i)$.**

(5 marks)

Reading for this question

Subject guide (2022), Chapters 4.2.2, 5.2.2–5, 6.2.1, 6.2.4, 8.2.3.

Wooldrige, Sections 3.2–3, 4.2, 4.4-5, 6.2c, 7.3, 7.4, 9.4, 8.3–4, Appendix A.1, A.5, B.5e, B.5f.

Approaching the question

This was another popular and relatively well-attempted question, requiring real-life applications. Part (a) is a routine interpretation task. The key is to interpret the coefficient from a log-linear/level regression using the exact interpretation (see subject guide (2022)). Candidates can use an approximate interpretation to get an increase of 50.1%, but as can be seen, the difference is substantial. Most answered the part on the expectation of the sign well. Part (b) is a short question on the dummy variable trap and was well attempted by most candidates. Some mistakenly interpreted the set-up as the case of heteroskedasticity or had a convoluted explanation of the issue.

Part (c) requires the use of an interaction term, which many candidates realised and proceeded well with the test. Note that the wording of the question suggested a one-sided test for the coefficient of the interaction term ('more pronounced'). A concerning number of candidates did not recognise the interaction term being needed and directly performed a t test of the two coefficients, which can be done by splitting the sample into two. This approach is not recommended nor explained well by these candidates. In part (d) most candidates realised the need to use an F test to test for the joint significance of all coefficients. Only a few realised that the restricted R-squared can be easily obtained from the given specification. Again, some lost marks for not discussing the assumptions nor testing equation.

Part (e) posed a considerable challenge for numerous candidates. A significant number failed to identify the issue of heteroskedasticity based on the wording of the question. Even among those who recognised it, the majority described the Goldfeld–Quandt version of the test for heteroskedasticity, which is no longer covered in the subject guide (2022). The examiners did award marks for these attempts, but advise candidates to closely study the latest subject guide (2022) for the most current approach to the materials. As noted in the General Remarks, the other tests mentioned in the solution have several advantages over the Goldfeld–Quandt test and should have been utilised. While many candidates left the answer to this question blank, others attempted to do a t test on the joint significance of some coefficients, which was also incorrect. Most who performed the Breusch–Pagan test did well and specified all the testing steps.

(a) Having an Oscar-winning director is associated with an increase of $(\mathrm{e}^{0.501} - 1) \times 100 = 65\%$ in the average rating of a movie, holding other factors constant. The sign is exactly what we should expect.

(b) No, we should not as the two variables will be perfectly correlated and we will have perfect multicollinearity. If we include the constant, one of the Oscars variables will be dropped.

(c) Candidates are expected to discuss the inclusion of an interaction term $SuperStars \times Oscars$ into the regression model. The null hypothesis is $\beta_{interaction} = 0$, against a one-sided alternative hypothesis of $\beta_{interaction} > 0$. The usual t test using the estimates is expected, with the t statistic following the standard normal distribution instead of the usual t distribution.

Candidates need to clearly indicate the use of CLM assumptions, Gauss–Markov + normality, as well as the rejection rule.

(d) Candidates are expected to describe an F test of the null: $\beta_{superstars} = \beta_{oscars} = 0$, against the alternative hypothesis of at least one of them is non-zero. Estimate the restricted model by OLS when the restriction is imposed. Obtain the R_r^2. The test statistic is an F-stat:

$$F = \frac{(R_{ur}^2 - R_r^2)/q}{(1 - R_{ur}^2)(n - k - 1)}$$

where $q = 2$, $k = 3$ and $n = 645$. Under the null and the Gauss–Markow theorem assumptions, the F-stat follows an F distribution $F_{q,\, n-k-1}$. The robust standard errors should be used so candidates can discuss the approximate F-stat and the chi-squared distribution instead. If the statistic exceeds the critical value from the respective distribution, we reject the claim.

(e) Candidates are expected to realise that we need to test for heteroskedasticity: either the Breusch–Pagan version or the White version of the test.

For the Breusch-Pagan test:

- Save the OLS residuals as $\widehat{u}_i$. Compute the squared residuals $\widehat{u}_i^2$.
- Regress $\widehat{u}_i^2$ on all the regressors, with an intercept, and then compute the F test of joint significance of the regressors:

$$\widehat{u}_i^2 = \delta_0 + \delta_1 \log cost_i + \delta_2 SuperStars_i + \delta_3 Oscars_i + v_i.$$

  The null hypothesis is: $\mathrm{H}_0 : \delta_1 = \delta_2 = \delta_3 = 0$. Discuss the R-squared version or the SSR version of the F statistic in this case.
- We do not need to assume MLR.6 in this version of the test. Under the null hypothesis, the F-stat will follow the asymptotic F distribution.

For the White test:

- Save the OLS residuals as $\widehat{u}_i$. Compute the squared residuals $\widehat{u}_i^2$.
- Regress $\widehat{u}_i^2$ on all the regressors, with an intercept, and the squares and interactions of all the regressors:

$$\begin{aligned}\widehat{u}_i^2 = \delta_0 &+ \delta_1 \log cost_i + \delta_2 SuperStars_i + \delta_3 Oscars_i + \delta_4 \log cost_i^2 + \cdots \\ &+ \delta_7 \log cost_i \times SuperStars_i + \cdots + v_i.\end{aligned}$$

  The null hypothesis is: $\mathrm{H}_0 : \delta_1 = \delta_2 = \delta_3 = \cdots = \delta_9 = 0$. Obtain the F test of the joint significance of the regressors. Similarly, discuss the R-squared version or the SSR version of the F-statistic.
- We do not need to assume MLR.6 in this version of the test. Under the null hypothesis, the F-stat will follow the asymptotic F distribution.

If we reject the null hypothesis, there is evidence of heteroskedasticity. However, since we already use the logarithm in the outcome variable and also the robust standard errors, the heteroskedasticity is already accounted for. Better answers would suggest that such a test is still needed to determine whether we need robust standard errors, as this usage may cause larger standard errors.

Question 8

Intergenerational mobility is concerned with the transmission of attitudes from parents to children. A researcher uses a time-series dataset on the risk attitude for successive generations in Germany for $T = 120$ years. You estimate the following regression:

$$R_{G+1} = \alpha_0 + \alpha_1 R_G + \alpha_2 R_{G-1} + \beta_1 V_{G+1} + \beta_2 V_G + u_G, \quad 0 \le \alpha_1 \le 1, \quad G = 2, \ldots, T \tag{8.1}$$

where R_G denotes the measured risk attitude, and V_G denotes the income level of the country at time G, and u_G is a homoskedastic error term and independent of $(R_G, V_{G+1}, V_G, R_{G-1})$. Both $\{R_G\}_{G=1}^T$ and $\{V_G\}_{G=1}^T$ are $I(1)$ processes.

(a) **Explain the concept of an $I(1)$ process. Propose a suitable test to establish whether the risk attitude process $\{R_G\}_{G=1}^T$ is indeed $I(1)$, knowing that risk attitudes are often also time dependent. In your answer, be clear about why acknowledging the time dependence of risk attitudes is crucial in your test.**
Hint: **For simplicity, you can assume the β coefficients are zero in your test.**

(5 marks)

(b) **Derive the long-run relationship between $\{R_G\}_{G=1}^T$ and $\{V_G\}_{G=1}^T$, and discuss any potential problems with using OLS to estimate the relationship. In your answer, clearly discuss the concepts of co-integration and spurious relationships.**

(5 marks)

(c) We use OLS to estimate the long-run relationship model from (b) and obtain the following results:

$$\boldsymbol{R_G = \underset{(.164)}{1.351} + \underset{(.231)}{1.51}\, V_G + \widehat{e}_G} \tag{8.2}$$

$$\boldsymbol{n = 123,\ R^2 = .94.}$$

$$\boldsymbol{\Delta\widehat{e}_G = \underset{(.051)}{-0.412}\widehat{e}_{G-1} + \underset{(.156)}{.507}\,\Delta\widehat{e}_{G-1} + \Delta\widehat{e}_{G-2} + \widehat{v}_G} \tag{8.3}$$

$$\boldsymbol{R^2 = .138,\ \text{the critical value at } 5\% \text{ is } -3.34}$$

where $\widehat{u}_G$ and $\widehat{e}_G$ are the OLS residuals, and the standard errors are in parentheses. Using the available information, show that there exists an Error Correction Model (ECM) between R_G and V_G. Be clear about any statistical test that you would require and perform them if possible.

(5 marks)

(d) Describe an estimator to estimate the parameters of the ECM from (c). Describe how you can interpret the coefficients from the estimated model. *Hint:* You are not asked to derive the ECM.

(5 marks)

Reading for this question

Subject guide (2022), Chapters 13.2.2, 15.2.3–4.

Wooldrige, Sections 10.2, 10.4, 18.2–4.

Approaching the question

In line with previous years, this particular question received the least attention from candidates. However, for those who did attempt it, their efforts were mostly commendable.

Part (a) focuses on the definition and testing of stationarity for $I(1)$. The ADF test was well described by most candidates, even though some mistakenly described the DF test instead. The key of an excellent answer is the comprehensive description of the testing procedure, noting that a DF test distribution is needed. Most candidates realised that the test is without a time trend and proceeded accordingly. Part (b) is a textbook question on the calculation of the short- and long-run effects. Some candidates forgot to also account for the lagged outcome term in the regressors, resulting in an incorrect long-run calculation. The second part of the question focuses on the definition and recognition of spurious and co-integrating relationships. Best answers highlighted the differences between the two and how they affect the long-term interpretation of the results. Candidates mostly did well in this part.

Part (c) is another textbook question on the construction of ECM. The best answers centred around the realisation that the error terms must be $I(0)$ and the stationarity test presented in the question uses the residual from regression (8.3). Most candidates realised this and described the ADF test, concluding that the residuals are $I(0)$ and we have evidence for co-integration. Some forgot a few details of the ADF test, losing a few marks. Part (d) was also well attempted by most candidates. The two stages of the ECM were well discussed and the interpretations of the coefficients were presented clearly.

(a) $I(1)$ is a unit root process, or is integrated of order one. This means that the first difference of the process is weakly dependent and often stationary. For series that have clear time trends, we need to modify the ADF test for unit roots. It is important because the trend-stationary process will have a linear trend in its mean but is $I(0)$ once the trend is removed. As such, the ADF test ignoring the time trend will have little power for rejecting a unit root and we may mistake the process as $I(0)$. To allow for such a deterministic trend,

we can run the following equation:

$$\Delta R_G = \alpha + \delta t + \theta R_{G-1} + \gamma_1 \Delta R_{G-1} + e_G.$$

The null hypothesis is: $\mathrm{H}_0 : \theta = 0$ and the hypothesis is $\mathrm{H}_1 : \theta < 0$. If we reject the null, R_G will be a trend-stationary process. Candidates can add more lags to clear up any serial correlation in ΔR_G. Candidates can also discuss a joint significance test of δ and θ but this is not necessary. Candidates should discuss the ADF instead of the DF test.

The test statistic is simply the t-stat of $\widehat{\theta}/se(\widehat{\theta})$, which follows the Dickey–Fuller distribution instead of the standard normal distribution. The critical value for 5% in this test with a linear time trend is -3.41 instead of -2.86 in the case without a trend. To reject the null, the t-stat needs to be less than -3.41.

(b) Taking the equation at equilibrium (R^*, V^*) and using the assumption that $\mathrm{E}(u_G) = 0$, we have:

$$R^* = \alpha_0 + \alpha_1 R^* + (\beta_1 + \beta_2 + \beta_3)V^*$$

$$= \frac{\alpha_0}{1-\alpha_1} + \frac{(\beta_1 + \beta_2 + \beta_3)}{1-\alpha_1} V^*.$$

This is the long-run relationship between R_G and V_G. Using OLS to estimate this equation is problematic because of the potential spurious relationship between the two $I(1)$ processes. A cointegrating relationship indicates that there is a long-run relation between the two processes – that is, there is a linear combination of processes in which the order of integration is lower than that of the original processes (in this case, $I(0)$). If the only reason that there is an apparent relationship is due to the persistence of the two processes we have a spurious, meaningless relationship. OLS estimators in this case are meaningless.

(c) To have an ECM, the two processes must be cointegrating. By definition, we must show that the error term is $I(0)$. The second regression estimate allows us to test whether the residual is $I(0)$. The OLS estimation gives us an ADF without a time trend. Denote the coefficient of ε_{t-1} as π. The null hypothesis is that $\pi = 0$ (or the residual is $I(I)$ against the alternative: $\mathrm{H}_1 : \pi < 0$ (the residual is $I(0)$). The ADF statistic is practically the t-stat: $\widehat{\pi}/se(\widehat{\pi}) = -8.07$.

The ADF statistic is larger (in absolute value) than the critical value, hence we reject the null hypothesis, in favour of the alternative that the residual is $I(0)$. Thus the two processes are cointegrating and there exists an ECM between them.

(d) To estimate the ECM, we use the Engle–Granger two-step procedure.

- Step 1: Estimate the co-integrating relationship by OLS. We already obtained them, in Equation (1), and obtain the residual $\widehat{e}_{G-1}$. This term is disequilibrium.
- Step 2: Estimate the following model by OLS:

$$\Delta R_G = \phi_1 \Delta V_G - \phi_2 \Delta R_{G-1} - \theta_1 \widehat{e}_{G-1} + w_G.$$

For interpretation, ϕ_1 and ϕ_2 are the direct effects of V_G and ΔR_{G-1} on the current risk attitude. θ_1 is an indicator of how quickly the risk attitude will respond to the long-run disequilibrium: if past risk attitude is larger than the equilibrium value, then $\Delta R_{G-1} < 0$ so that the risk attitude will go back to equilibrium.

Question 9

A researcher examines the probability of a random sample of British firms that use Facebook Ad Platform ($FB_i = 1$ if yes, and $= 0$ if otherwise) for their marketing campaigns, where i indicates an individual firm. She uses OLS with robust standard errors reported in the parentheses to obtain the following regression results:

$$\widehat{FB}_i = \underset{(0.014)}{0.498} + \underset{(0.003)}{0.016}\, growth_i + \underset{(0.015)}{0.025} \log(profit_i) + \underset{(0.12)}{0.54} \log(cost_i) \qquad (9.1)$$

$$n = 1889, \ R^2 = 0.09$$

where $growth_i$, $profit_i$, $cost_i$ are respectively the average growth rate of revenues in percentage point, the average profit after tax in millions of GBP, and the average expenses (measured in millions of GBP) in the previous three years.

(a) What is the interpretation of the coefficient on $growth_i$? Is the sign of the coefficient as you would expect? Explain your answer.

(2 marks)

(b) A colleague suggests that a heteroskedasticity test is needed before employing the robust standard errors in this OLS regression. Explain, using appropriate derivations, whether you agree with this claim. In your answer, highlight the advantages and drawbacks of using an OLS to estimate such a model of Facebook usage.

(5 marks)

The researcher decides to re-estimate the model using Logit, and obtain the following regression results:

$$Pr(\widehat{FB}_i = 1) = \Lambda\left(\underset{(0.036)}{0.0254} + \underset{(0.0009)}{0.0027}growth_i - \underset{(0.03)}{0.0035}\log(profit_t) + \underset{(0.012)}{0.034}\log(cost_t)\right) \tag{9.2}$$

$$n = 1889, \text{ Pseudo } R^2 = 0.034$$

where $\Lambda(\cdot)$ is the logistic cumulative density function and $\Lambda(z) = \exp(z)/(1+\exp(z))$, and the conventional standard errors are reported in the parentheses.

(c) Why do the coefficient estimates differ between the results presented for the OLS in (9.1) and for the Logit model in (9.2)? In your answer, you are expected to derive a formula for the marginal effect of $growth_i$ for each model. Are the signs of the estimates in (9.2) as you would expect? *Hint:* The logistic probability density function is given $g(z) = \Lambda'(z) = \exp(z)/(1+\exp(z))^2$.

(5 marks)

(d) Why do we report the conventional standard errors in (9.2)? Explain intuitively the estimation method behind the Logit model and its statistical properties. Explain why the properties are important.

(5 marks)

(e) A colleague suggests that from the estimates in (9.2) the marginal effect of revenue growth on the usage of Facebook Ad is certainly dependent on the level of marketing expenses. Discuss, using appropriate technical details, whether you agree with this claim. In your answer, clearly state how to obtain the marginal effect of revenue growth for an *average* firm.

(3 marks)

Reading for this question

Subject guide (2022), Chapter 12.2 (all).

Wooldrige, Sections 7.1–5, 17.1–3, Appendix C.4b.

Approaching the question

This was another popular and well-attempted question. The question setting is familiar and most of the answers are textbook like. The best candidates could distinguish themselves in parts (a) and (d), where critical evaluation and interpretation are required.

Part (a) is a classic interpretation question on the LPM coefficients. noting the distinction between percentage and percentage points. Some candidates incorrectly used percentages. The

expected signs were well identified and discussed. Part (b), particularly the derivation of heteroskedasticity part, perhaps took some candidates by surprise. Candidates were expected to produce the correct proof of heteroskedasticity noting that the expected outcome is the same as the probability of a company using Facebook. Most candidates did well on the drawbacks and advantages of LPM.

Part (c) focuses on the key difference in interpreting the marginal effect using Logit estimations. It is key to remember that coefficients from Probit or Logit estimations are not directly interpretable, at least not the magnitudes of the estimates. The best candidates explained this statement clearly with relevant annotations. To comment on the sign of the marginal effect, best answers carefully derived the marginal effect of a Probit estimator, making use of the given logistic pdf hint. For those who directly interpreted the given coefficients, a few marks were awarded. Only a few candidates scored the maximum marks on this question.

Part (d) focuses on the properties of Logit and the mechanical process behind an MLE. The best candidates provided not only the correct likelihood function (not asked) for the estimation (and then the log-likelihood), but also an intuitive way to obtain the function and interpret it. Some candidates did not annotate their answers well, losing a few marks; others failed to provide a good intuitive understanding of the estimation or the computational approach to calculate the estimate in practice. The properties of the MLE were well recited.

Part (e) was a popular question, asking candidates to derive the marginal effect of a Logit estimation and to show that the marginal effect has the same sign and statistical significance with the Logit estimate. The better candidates explained the calculation of the APE or the PAE (partial effect of the average person). The best candidates carefully explained the steps to obtain such calculations and noted on how the averages of the log values are being treated.

(a) A 1-percentage point increase in the growth rate of a firm is on average associated with an increase of 0.016 percentage points in the probability of using Facebook Ads, holding other factors constant (or 10 percentage points in growth is with a 0.16 percentage point increase in using Facebook). The result has the expected sign (positive).

(b) Candidates are expected to prove that heteroskedasticity is inherent when OLS is used to estimate a model with a binary choice outcome. With $Pr(y = 1 \mid x)$ as the response probability, we know that $Pr(y = 0 \mid x) = 1 - Pr(y = 1 \mid x)$. To derive the variance of the error term we can examine (candidates are expected to explain the steps).

$$\mathrm{Var}(u \mid x) = \mathrm{Var}(y \mid x) = \mathrm{E}(y^2 \mid x) - (\mathrm{E}(y \mid x))^2 = Pr(y = 1 \mid x)(1 - Pr(y = 1 \mid x)).$$

This conditional variance is dependent on the regressors x, which will violate the homoskedasticity assumption. Thus the use of a robust standard error by default is appropriate. LPM is simply the OLS estimation of the model:

- The OLS model has various drawbacks relative to the Logit: Predicted probabilities are restricted to lie in $[0, 1]$, the MLE underlying Logit is (asymptotically) efficient whereas OLS (LPM) will be inefficient.
- The main advantage of the OLS model is that the coefficients can be directly interpreted as the marginal effects of the regressor(s) of interest, so that the computation and interpretation with OLS would be easier.

(c) The coefficients in the OLS model can be directly interpreted as marginal effects; whereas those in the Logit model cannot. The main reason is that the regression models are specified using a non-linear function, in this case $\Lambda(\cdot)$. Only when the relationship is linear as in OLS, we can argue that (candidates are expected to elaborate using the particular variables in question):

$$\Delta Pr(y = 1 \mid x) = \beta_j \Delta x_j.$$

To obtain the marginal effect in Logit, we need to use the chain rule and:

$$\frac{\partial Pr(y = 1 \mid x)}{\partial x_j} = \frac{\partial \Lambda(x\beta)}{\partial x_j} = \beta_j g(x\beta)$$

where $g(\cdot)$ is the logistic probability function. Because $g(\cdot) > 0$, the sign of the marginal effect will be exactly the same as that of the estimates. So in this case the Logit estimates have the signs as expected.

(d) Candidates are expected to discuss the advantages of Logit models: The MLE automatically deals with heteroskedasticity, which is a problem under the LPM, so that robust standard errors are no longer needed as default. Finally, the marginal effect is allowed to depend on individual characteristics, unlike being constant as in the LPM.

To estimate the parameters, we can use the Maximum Likelihood Estimator that chooses $\{\beta_j\}$ to maximise the joint probability of the outcomes. The intuition is that we try to choose the $\widehat{\beta}$ such that the $\widehat{Pr}(Facebook = 1)$ is high for firms using Facebook, and low for the non-users.

There are no mathematical formulae for the estimates, the popular method is to adopt an iterative trial-and-error process. In large samples, the MLE are consistent, asymptotically normally distributed, and asymptotically efficient. It is important for statistical tests as we must now use asymptotic tests. The efficiency also ensures the precision of the MLE estimators.

(e) Because the marginal effect from Logit is given as $\beta_j g(x\beta)$ and we can see that β_j is clearly statistically significant, the marginal effect is dependent on other variables via $g(x\beta)$.

The marginal effect of the average firms can be obtained by taking the averages of the regressors (characteristics of the average firm), and then substitute these averages into the marginal effect formula.

Candidates should mention whether they would take the averages of the log values or take the log of the average values.

The APE reports the average marginal effect over all individuals (here the sample size is RnR), which can be calculated as follows:

$$\frac{1}{n}\sum_{i=1}^{n}\frac{\partial Pr(Facebook = 1 \mid x)}{\partial growth} = \frac{1}{n}\sum_{i=1}^{n} g(z_i) \times 0.0027.$$

Examiners' commentaries 2023 (May)

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2022–23. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE). The examinable materials are based on the new subject guide (Schafgans and Komarova, 2022, henceforth referred to as subject guide (2022)) and the new textbook:

- Wooldridge, J. *Introductory Econometrics: A Modern Approach.* (Cengage, 2020), seventh edition [ISBN 9781337558860].

Detailed reading references in this subject guide refer to the edition of the set textbook listed above. A new edition of this textbook may have been published by the time you study this course. You can, however, use a more recent edition; use the detailed chapter and section headings to identify relevant readings. Also check the virtual learning environment (VLE) regularly for updated guidance on readings.

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2022). The subject guide is intended for you to use in conjunction with the main textbook. It will act to highlight important aspects of the content in the chapters, provide further clarifications where deemed useful and also to extend it in certain places. For the final examination, you will need to be familiar with the material in both the textbook and the subject guide. You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone B

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from **Section A** (8 marks each) and **THREE** from **Section B** (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Section A

Answer all questions from this section. Each question carries 8 marks.

Question 1

Consider the following bivariate regression model:

$$y_i = \beta_1 x_i + u_i, \tag{1.1}$$

for a given random sample of n observations $\{(y_i, x_i)\}_{i=1}^n$ under the standard Assumptions SLR.1–SLR.4 (linear in parameters, random sampling, non-zero sample variation in the regressor, and zero conditional mean). Let:

$$z_i = 20 + x_i^{23},$$

and define the following slope estimator:

$$\tilde{\beta}_1 = \frac{\sum_{i=1}^n z_i y_i}{\sum_{i=1}^n z_i x_i}.$$

(a) Carefully show that $\tilde{\beta}_1$ is a consistent estimator of β_1.

(4 marks)

(b) Suppose the data is in fact generated by the following model:

$$y_i = \beta_0 + \beta_1 x_i + u_i, \quad \beta_0 \neq 0. \tag{1.2}$$

Discuss whether $\tilde{\beta}_1$ is a biased estimator of β_1. In your answer, derive the bias of the estimator, if any, or state the assumptions needed for $\tilde{\beta}_1$ to be unbiased.

(4 marks)

Reading for this question

Subject guide (2022): Chapters 3.2 and 7.2.1.

Wooldridge Section 2.2, Appendix 2A and Appendices B.4a, B.4b, B.4c.

Approaching the question

Part (a) is a popular set-up with a tweak to involve the academic year 2022–23 in the question. The best answer would simplify the derivation by using w_i instead of the longer term, and correctly proceed through the derivation. Two key steps are taking the probability limits of the estimator and invoking the LLN and LIE, which some candidates forgot to explicitly mention and lost marks on. Better performance must involve clear and correct annotations and explanations of the steps.

Part (b) focuses on the derivation and proof of unbiasedness/biasedness of the given estimator. Correct algebra was expected as well as clear explanations with references to the SLRs. Similar to part (a), some candidates forgot the use of LIE, or getting an incorrect conclusion regarding the condition for unbiasedness. Overall, candidates did well on this question.

(a) Define $w_i = \frac{z_i}{\sum_{i=1}^n z_i x_i}$ and replace (1.1) to the slope estimator, it follows:

$$\tilde{\beta}_1 = \sum_{i=1}^n w_i y_i = \sum_{i=1}^n w_i(\beta_1 x_1 + u_i) = \beta_1 \left[\sum_{i=1}^n \frac{z_i x_i}{\sum_{i=1}^n z_i x_i}\right] + \sum_{i=1}^n w_i u_i.$$

Because $\sum_{i=1}^n \frac{z_i x_i}{\sum_{i=1}^n z_i x_i} = 1$, the estimator can be rewritten as:

$$\tilde{\beta}_1 = \beta_1 + \sum_{i=1}^n w_i u_i = \beta_1 + \frac{\sum_{i=1}^n z_i u_i}{\sum_{i=1}^n z_i x_i}.$$

Take the probability limit of $\tilde{\beta}_1$:

$$\text{plim } \tilde{\beta}_1 = \beta_1 + \frac{\text{plim } n^{-1} \sum_{i=1}^n z_i u_i}{\text{plim } n^{-1} \sum_{i=1}^n z_i x_i}.$$

Because of SLR.4 and the LIE, $\mathrm{E}(u_i z_i) = \mathrm{E}(\mathrm{E}(u_i z_i \mid X)) = 0$, we obtain:

$$\text{plim } n^{-1} \sum_{i=1}^n z_i u_i = \mathrm{E}(u_i z_i) = 0.$$

The probability limit exists because plim $n^{-1}\sum_{i=1}^{n} z_i x_i \neq 0$.

As such, plim $\tilde{\beta}_1 = \beta_1$. Therefore, by definition, $\tilde{\beta}_1$ is a consistent estimator of β_1.

(b) To discuss the unbiasedness of $\tilde{\beta}_1$, we rewrite $\tilde{\beta}_1$ with the new true relationship as:

$$\tilde{\beta}_1 = \sum_{i=1}^{n} w_i y_i = \sum_{i=1}^{n} w_i(\beta_0 + \beta_1 x_i + u_i) = \beta_0 \frac{\sum_{i=1}^{n} z_i}{\sum_{i=1}^{n} z_i x_i} + \beta_1 + \sum_{i=1}^{n} w_i u_i. \qquad (*)$$

Take the conditional expectation on $X = \{x_i\}_{i=1}^{n}$ of $\tilde{\beta}_1$ to get:

$$\begin{aligned}
\mathrm{E}(\tilde{\beta}_1 \mid X) &= \beta_0 \frac{\sum_{i=1}^{n} z_i}{\sum_{i=1}^{n} z_i x_i} + \beta_1 + \mathrm{E}\left(\sum_{i=1}^{n} w_i u_i \mid X\right) \\
&= \beta_0 \frac{\sum_{i=1}^{n} z_i}{\sum_{i=1}^{n} z_i x_i} + \beta_1 + \sum_{i=1}^{n} \mathrm{E}(w_i u_i \mid X) \\
&= \beta_0 \frac{\sum_{i=1}^{n} z_i}{\sum_{i=1}^{n} z_i x_i} + \beta_1 + \sum_{i=1}^{n} w_i \mathrm{E}(u_i \mid X) \\
&= \beta_0 \frac{\sum_{i=1}^{n} z_i}{\sum_{i=1}^{n} z_i x_i} + \beta_1.
\end{aligned}$$

The first equality follows from (*), the second equality follows SLR.2, the third equality follows the properties of conditional expectation and the fact that w_i is a function of x_i. The last equation follows SLR.4.

Following the LIE, we have:

$$\mathrm{E}(\tilde{\beta}_1) = \mathrm{E}(\mathrm{E}(\tilde{\beta}_1 \mid X)) = \beta_0 \frac{\sum_{i=1}^{n} z_i}{\sum_{i=1}^{n} z_i x_i} + \beta_1.$$

Because $\beta_0 \neq 0$, the estimator is unbiased if and only if $\sum_{i=1}^{n} z_i = 20n + \sum_{i=1}^{n} x_i^{23} = 0$.

Question 2

A researcher is interested in the relationship between (the natural logarithm of) family income ($\log \mathit{famincome}_i$) and academic performance of a student measured by GPA_i, which takes a value from 0 to 100. Using a random sample of 230 students enrolled in a UK university, she estimates the following regression:

$$\mathit{GPA}_i = \beta_0 + \beta_1 \log \mathit{famincome}_i + \beta_2 \mathit{selfstudyhours}_i + \varepsilon_i, \qquad (2.1)$$

where $\mathit{selfstudyhours}_i$ is the total hours of self-study in a week. Assume that the model given by (2.1) satisfies Assumptions MLR.1–MLR.4 (linearity in parameters, random sampling, no perfect multicollinearity and zero conditional mean).

(a) A colleague claims that the values of GPA are more or less variable for students in lower-income families than in more affluent families, holding other factors fixed. Discuss a statistical test for this claim. What is the main consequence if we cannot reject the colleague's claim? *Hint: Think about the relationship between the variance of the error term ε and $\log \mathit{famincome}_i$.*

(4 marks)

(b) Another colleague wants to test whether increasing family income by 10% has the same effect on GPA as 5 additional hours of self-study per week. Discuss how we can test this claim, be clear on the null hypothesis and any additional information needed.

(4 marks)

Reading for this question

Subject guide (2022), Chapters, 3,2,4, 8.2.3, 5.2.4.

Wooldridge, Sections 8.3, 4.4, 2.4 and Appendices A.3–A.4.

Approaching the question

Part (a) posed a considerable challenge for numerous candidates. A significant number failed to identify the issue of heteroskedasticity based on the wording of the question. Even among those who recognised it, the majority described the Goldfeld–Quandt version of the test for heteroskedasticity, which is no longer covered in the subject guide (2022). The examiners did award marks for these attempts, but advise candidates to closely study the latest subject guide for the most current approach to the materials. As noted in the General Remarks, the other tests mentioned in the solution have several advantages over the Goldfeld–Quandt test and should have been utilised. While many candidates left the answer to this question blank, others attempted to do a t test on the joint significance of some coefficients, which was also incorrect. Most who performed the Breusch–Pagan test did well and specified all the testing steps.

Part (b) discusses two important subjects: the interpretation of a level-log regression and the testing of multiple coefficients. Some candidates misinterpreted the coefficients and, as a result, arrived at an incorrect null hypothesis of $1.1\beta_1 = 5\beta_2$. The best answers elaborated on how to calculate the standard error of the specified difference in the denominator of the test statistic. A comprehensive explanation of the testing procedure was also expected, and marks were deducted for omitting certain details.

(a) Candidates are expected to realise that we need to test for heteroskedasticity: either the Breusch–Pagan version or the White version of the test.

For the Breusch–Pagan test:

We should estimate equation (2.1) using OLS, save the OLS residuals as $\widehat{u}_i$. Compute the squared residuals $\widehat{u}_i^2$. Regress $\widehat{u_i^2}$ on all the regressors, with an intercept, and then compute the F test of the joint significance of the regressors.

$$\widehat{u}_i^2 = \delta_0 + \delta_1 \log \mathit{famincome}_i + \delta_2 \mathit{selfstudyhours}_i + v_i.$$

The null hypothesis is: $\mathrm{H}_0 : \delta_1 = \delta_2 = 0$. Candidates are expected to discuss the R-squared version or the SSR version of the F statistic in this case. $k = 2$, $n - k - 1 = 230 - 2 - 1 = 227$. We do not need to assume MLR.6 in this version of the test, under the null hypothesis, the F-stat will follow the asymptotic F distribution. If we reject the null hypothesis, there is evidence of heteroskedasticity.

For the White test:

We should estimate equation (2.1) using OLS, save the OLS residuals as $\widehat{u}_i$. Compute the squared residuals $\widehat{u}_i^2$. Regress $\widehat{u}_i^2$ on all the regressors, with an intercept, and the squares and interactions of all the regressors:

$$\widehat{u}_i^2 = \delta_0 + \delta_1 \log \mathit{famincome}_i + \delta_2 \mathit{selfstudyhours}_i + \delta_3 \log \mathit{famincome}_i^2$$
$$+ \delta_4 \mathit{selfstudyhours}_i^2 + \delta_5 \log \mathit{famincome}_i \times \mathit{selfstudy}_i + v_i.$$

The null hypothesis is: $\mathrm{H}_0 : \delta_1 = \delta_2 = \delta_3 = \delta_4 = \delta_5 = 0$. We can simply obtain the F test of the joint significance of the regressors. Similarly, candidates are expected to discuss the R-squared version or the SSR version of the F statistic. $k = 5$, $n - k - 1 = 230 - 5 - 1 = 224$. We do not need to assume MLR.6 in this version of the test, under the null hypothesis, the F-stat will follow the asymptotic F distribution. If we reject the null hypothesis, there is evidence of heteroskedasticity.

In both cases, the failure to reject the claim would suggest the presence of heteroskedasticity. The main consequence is that the conventional standard errors are invalid and any statistical tests using these standard errors are invalid.

(b) Since MLR.5 was not given, we must use an OLS estimation of (2.2) with robust standard errors. The regressors are in log and at level, which means a 1% increase in family income is associated with a $0.01\beta_1$ change in the outcome, whereas the estimate of β_2 is the change in the outcome given one additional hour of self-study. Therefore the null hypothesis is $\mathrm{H}_0 : 0.1\beta_1 = 5\beta_2$. Let $\mathrm{H}_1 : 0.1\beta_1 \neq 5\beta_2$ be the two-sided alternative and set the significance level at 5%.

The t statistic to test the null hypothesis H_0 is written as:

$$t = \frac{0.1\widehat{\beta}_1 - 5\widehat{\beta}_2}{se(0.1\widehat{\beta}_1 - 5\widehat{\beta}_2)} \quad \text{under } \mathrm{H}_0.$$

Obtaining the estimates, we know $0.1\widehat{\beta}_1 - 5\widehat{\beta}_2$. The standard error of $0.1\widehat{\beta}_1 - 5\widehat{\beta}_2$ is given by:

$$se(0.1\widehat{\beta}_1 - 5\widehat{\beta}_2) = \sqrt{0.01 \cdot se(\widehat{\beta}_1)^2 + 25se(\widehat{\beta}_2)^2 - s_{12}}$$

where s_{12} is the estimate of $\mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_2)$, and $se(\widehat{\beta}_1)$, $se(\widehat{\beta}_2)$ are the robust standard errors of the respective estimates. To implement the above t test, we need to know s_{12}. Under the MLR.1–4, and robust standard errors, we can use the asymptotic version of the t distribution (standard normal distribution) under H_0. Once we get s_{12}, we can implement the t test. The rejection rule is $|t| > c$ where c is the 97.5th percentile of the standard normal distribution.

Question 3

State whether the following claims are True or False. Briefly explain why (two to three sentences).

(a) To test the exogeneity of a regressor, x, we can run an OLS of y on x and obtain the residual $u = y - \widehat{y}$, where $\widehat{y}$ is the predicted value of y. Only when x is exogenous, $\mathrm{Cov}(x, u) = 0$.

(2 marks)

(b) In a multivariate regression with 4 regressors as follows,

$$y = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 + u_i,$$

the adjusted R-squared of the OLS regression of the restricted model will be smaller than that of the unrestricted model if we cannot reject the restriction (or constraint) that $\alpha_2 = \alpha_3 = \alpha_4$, and larger if we can reject the constraint.

(2 marks)

(c) Adding an omitted variable in a regression will result in a change in the coefficients of all the regressors.

(2 marks)

(d) No serial correlation in the error terms would imply no serial correlation in the regressors. *Note: In your answer, clearly define what serial correlation is.*

(2 marks)

Reading for this question

Subject guide (2022), Chapters 3.2.3, 5.2.5, 4.2.3, 14.2.1.

Wooldrige, Sections 2.3, 4.5 and Appendix B.5f, 3.3, 3.2g, 12.1–12.2.

Approaching the question

Part (a) is a tricky question. Many candidates correctly recognised that the statement was false but explained that a Durbin–Hausman test for endogeneity would be required. Candidates failed to recall the algebraic properties of OLS on sample data (residuals) instead of the population level's error terms. Part (b) was well attempted by most candidates. Some lost marks for unclear explanations of the restricted models and its relevance to R-squared. Part (c) is another tricky question. Theoretically it should not result in any changes in the other coefficients, while in practice an absolute zero correlation with other factors are rarely found. As long as candidates explained their rationales, marks were given. Part (d) is a textbook question and involves the definition of serial correlation and the understanding of the differences between error terms and regressors. Most did well and explained the concept clearly.

(a) **False.** The residual is always uncorrelated with the regressor regardless of its exogeneity.

(b) **False.** The adjusted R-squared can fall or increase regardless of the number of variables used in a model.

(c) **False.** Theoretically, adding an omitted variable will only change the coefficients of the regressors that are correlated with the omitted variable. Other coefficients should remain unchanged. In practice, we may see different estimates. Answer 'True' to this statement is accepted if relevant explanations are present.

(d) **False.** Serial correlation occurs if the series is correlated with itself at a different time period. However, the assumption of no serial correlation in the error term has nothing to do with the serial relation in the independent variable.

Question 4

A researcher is interested in the relationship between the monthly unemployment rate $unemploy_t$ and the monthly rate of vacancies $vacancy_t$ in Britain since 2000. She estimates the following regression model using observations from T months:

$$unemploy_t = \beta_0 + \beta_1 vacancy_t + \varepsilon_t, t = 2, \ldots, T, \tag{4.1}$$

where ε_t is the error term. The researcher observes that unemployment rates are weakly dependent but tend to be sticky: the adjacent rates of unemployment are fairly correlated, $\text{Corr}(unemploy_t, unemploy_{t+1}) \neq 0$, for all time periods. Assume that ε_t is independent of $vacancy_t$ and $vacancy_t$ is not serially correlated.

(a) The researcher obtains the 95% confidence interval for β_1 after an OLS estimation. Explain why it is problematic to use the interval for any statistical inference. Be clear about any potential violation of the Gauss–Markov assumptions. *Hint: Think about how the error term is affected by the stickiness of the unemployment rate.*

(2 marks)

(b) Discuss a suitable test to verify the violation of the Gauss–Markov assumption identified in (a). Be clear about the null hypothesis and the assumptions needed. What should we do if we reject the test's null hypothesis?

(4 marks)

(c) A colleague argues that unemployment is seasonal. Discuss one method to account for the seasonality. Be clear about what seasonality means.

(2 marks)

Reading for this question

Subject guide (2022), Chapters 14.2.1, 15.2.1.

Wooldrige, Sections 12.1–12.2, 12.3 (skip 12.3b), 10.5.

Approaching the question

Part (a) asks about the implications of autocorrelation, which many candidates did well. The violation of TS.5 and TS.5' was also well discussed. Best answers discussed and made clear the definition of autocorrelation.

Part (b) is about an autocorrelation test with an AR(1) error term. Most candidates did well, some lost marks for missing details of the testing procedure. Others did not answer the remedial measures once we reject the null hypothesis. Some discussed the Durbin–Watson test, which is not introduced in the current subject guide (2022). The Durbin–Watson test (discussed in Section 12.3, Wooldridge) is a finite sample test that permits testing for no autocorrelation against AR(1) dependence. It requires strict exogeneity and the full set of CLM assumptions (including normality). As these assumptions are deemed to be quite restrictive in time series models, we will not discuss this test in the course. You may also ignore the discussion of the LM form of the test for autocorrelation (called the Breusch–Godfrey test).

Part (c) is a discussion of seasonality and its remedies. A good number of candidates skipped the question. Best answers provided a clear real-life example and remained mindful of the dummy variable trap by only including 11 dummies instead of 12. Some excellent answers discussed the deseasonalising process.

(a) Because of the correlation and the weak dependence of $unemploy_t$, the error terms are also correlated with their adjacent values $\text{Corr}(\varepsilon_t, \varepsilon_{t+1}) \neq 0$. Because of this correlation, the model suffers from autocorrelation so that the standard errors are invalid (TS.5 and TS.5' are mostly violated as there is no lagged variables). As such the construction of confidence intervals is invalid.

(b) Candidates are expected to perform an autocorrelation test. Assume that ε_t can be expressed as follows:

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t, \quad \text{with } |\rho| < 1$$

and v_t is a white noise generated randomly from a normal distribution with mean zero and constant variance. We can also assume TS.4 (or TS.4') so there is no need for a robust standard error. To test for autocorrelation, the null hypothesis is $\text{H}_0 : \rho = 0$ against $\text{H}_1 : \rho \neq 0$. The testing equation is obtained from using the OLS residuals:

$$\widehat{\varepsilon}_t = unemploy_t - (\widehat{\beta}_0 + \widehat{\beta}_1 vacancy_t).$$

The testing equation is:

$$\widehat{\varepsilon}_t = \gamma_0 + \gamma_1 vacancy_t + \rho\widehat{\varepsilon}_{t-1} + u_t.$$

We must use the asymptotic t statistic $\widehat{\rho}/se(\widehat{\rho})$ for a large sample. At the 5% level of significance, we should reject if $—\widehat{\rho}/se(\widehat{\rho})| > 1.96$, which follows a $N(0, 1)$ distribution.

Candidates can instead use the LM test statistic $(T - 1)R^2 \sim \chi^2_1$.

If we reject the null hypothesis, we should be concerned about the presence of autocorrelation. We can use the HAC robust standard errors to fix the standard errors.

(c) Because the data is observed at monthly intervals, we may be concerned of seasonality: for example, unemployment may be lower in the early months as people look for jobs after the winter holiday. We can account for seasonality by assuming that the seasonal patter within a year is roughly constant across time. We can include a set of monthly dummy variables (11 dummies instead of 12). Candidates can also discuss the deseasonalised regression models as well as an alternative.

Question 5

A researcher is interested in examining the time series behaviour for the quantity of cement sold in a particular region in Japan for a given quarter, denoted as $\log(Q_t)$, and the number of new building permits, $\log(Permit_t)$. She obtains estimates using

the following autoregressive model over a period of T quarters in the Japanese region:

$$\log(Q_t) = \alpha_0 + \alpha_1 \log(Q_{t-1}) + \beta_1 \log(Permit_t) + \beta_2 \log(Permit_{t-1}) + \varepsilon_t, \qquad (5.1)$$

where $|\alpha_1| < 1$, $t = 2, \ldots, T$ and ε_t is an i.i.d. error term.

(a) A colleague is interested in establishing whether the time series $\log(Q_t)$ is stationary and weakly dependent. Clearly define the two terms and explain why these two properties are important.

(4 marks)

(b) Is the OLS estimator for α_1 unbiased? Explain. In your answer, clearly indicate the Gauss–Markov assumptions that are potentially violated, and the difference between contemporaneous exogeneity and strict exogeneity.

(2 marks)

(c) Now suppose that $\alpha_1 = 0$ and the researcher is told that permits for new buildings are granted more often when there is a higher demand in construction. Are the OLS estimators for (β_1, β_2) unbiased? How about consistency? Be clear about the violation of the Gauss–Markov assumptions, if any.

(2 marks)

Reading for this question

Subject guide (2022), Chapters 13.2.4, 13.2.5, 11.2.1.

Wooldrige, Sections 11.1, 11.2, 16.1.

Approaching the question

Part (a) pertains to a textbook question regarding the definitions of stationarity and weak dependence. It was anticipated that only the definition of covariance stationarity would be provided. Some candidates did not demonstrate a clear understanding of the significance of these two assumptions in the context of time series analysis, particularly with respect to weak dependence. Best answers illustrated the connection between weak dependence and the law of large numbers (LLN), despite this not being requested. Part (b) requires a mathematical demonstration of endogeneity caused by the inclusion of a lagged outcome in a time series model. Most recognised the OLS estimator now became biased. Most explained well the difference between weak and strict exogeneity. Part (c) poses a challenge. Most candidates did not recognise the set-up for simultaneity bias. Best answers showed the bias mathematically, but it was not necessary for a two-mark question.

(a) A process is stationary if its stochastic properties (such as the mean, and variance) and its temporal dependence structure (such as covariance over time) do not depend on time. It is a stronger version of the identically distributed assumption. It is important because we know that the nature of any correlation between adjacent terms is the same across all time periods.

A process is weakly dependent if the correlation between two random temporal values of the process will disappear as the time distance between them increases: $\mathrm{Corr}(x_t, x_{t+h}) \to 0$ as $h \to \infty$. This assumption is important because it replaces the assumption of independence so that the LLN and the CLT hold: asymptotic independence of the process.

(b) Candidates are expected to show that because of the presence of $\log(Q_{t-1})$ the strict exogeneity assumption is violated: that is the error term is not strictly independent of the regressors at all times. Thus the OLS estimator is biased. The difference is that the contemporaneous exogeneity only requires the error term to be independent of the current regressors $\mathrm{E}(\varepsilon_t \mid x_t) = 0$.

(c) Candidates are expected to realise that this is a set-up for a simultaneity bias and omitted variable bias problem (only one is required). The consequence is that $Permit_t$ and ε_t are going to be correlated, violating TS.3' and also TS.3. The consequence is that the estimators are neither unbiased nor consistent.

Section B

Answer three questions from this section. Each question carries 20 marks.

Question 6

An economist is interested in a regression model for the (logarithm of) output-per-capita y_t in a given country and at a time period t as in:

$$y_t = \alpha + \rho y_{t-1} + \beta x_t + \gamma t + \varepsilon_t, \tag{6.1}$$

where x_t is the (logarithm of) investment-per-capita at time t.

(a) Assume that y_t is stationary ($|\rho| < 1$) and $\varepsilon_t = \lambda\varepsilon_{t-1} + u_t$, where u_t is a white noise. Discuss the implications of this assumption for the OLS estimators for β.

(5 marks)

Assume for sub-questions (b), (c), (d) that the error term (ε_t) is not serially correlated.

(b) The economist discovers that the investment rates are also affected by the current output-per-capita. The relationship is described as follows:

$$x_t = \delta_0 + \delta_1 y_t + \eta_t. \tag{6.2}$$

Assume that η_t also is not serially correlated and $\text{Cov}(\varepsilon_t, \eta_t) = 0$. Discuss how this discovery will affect the OLS estimation of Equation (6.1)? Be clear about the statistical consequences for the OLS estimation and what the terms 'endogenous' variable and 'exogenous' variable mean.

(5 marks)

(c) Derive the reduced form functions for both x_t and y_t. Explain the direction of bias of the OLS estimator for δ_1 in (6.2) if investment is thought to be detrimental to output in the country.

(5 marks)

(d) A colleague suggests that we can use y_{t-1} and the time trend t as instrumental variables for y_t in (6.2). Describe how you would implement this estimator. Discuss the validity of this suggestion.

(5 marks)

Reading for this question

Subject guide (2022), Chapters 13.2.4, 14.2.1, 11.2.1–3, 9.2.5.

Wooldrige, Sections 11.1, 12.3, 16.1–16.3, 15.3a–15.3d.

Approaching the question

Part (a) discusses the problem of endogeneity in an ADL model with autocorrelated error terms. Most candidates recognised the problem and discussed well. Better performance formally explained, using derivations, the bias due to the AR(1) error term structure. Candidates often referred to the violation of TS.3' hence inconsistency. Part (b) requires a fuller discussion on the simultaneity bias resulted from the relationship between x_t and y_t. Most candidates recognised

this well and clearly discussed the terms exogenous and endogenous. Some candidates mistakenly commented that standard errors become invalid or the estimator suffers from inefficiency.

Part (c) was attempted well by most candidates, showing all necessary steps of derivations. The best candidates would explain the direction of bias with clear annotations, as well as checking for the existence condition of the reduced forms. Part (d) requires the knowledge of the instrumental variable implementation and critical discussions of the y_{t-1} and t as potential instruments. Only a few excellent candidates discussed the potential correlation between the instruments and y_t (dependent on time). Most explained the two IV conditions well.

(a) Candidates are expected to note that because of autocorrelation and the presence of lagged outcome, the OLS estimators are not only inefficient and biased but also inconsistent. Candidates are expected to show that $\text{Cov}(y_{t-1}, \varepsilon_t) \neq 0$ hence TS.3' is not satisfied. One can show that y_{t-1} depends on ε_{t-1} (by lagging one period), which in turn is related to u_t by the AR(1) structure of ε_t.

(b) The problem is simultaneity bias. Candidates are expected to verbally explain why. The consequence is that the OLS estimators in (6.1) are biased and inconsistent. y_t and x_t are determined within the structural models hence endogenous, y_{t-1} and t are determined outside of the model and hence exogenous. The exogeneity of y_{t-1} can be debated and candidates can discuss this point here as well.

(c) Candidates are expected to show the substitution and solution to the reduced forms; overall, the results are:

$$y_t = \pi_{10} + \pi_{11} y_{t-1} + \pi_{12} t + v_1$$

$$x_t = \pi_{20} + \pi_{21} y_{t-1} + \pi_{22} t + v_2$$

where:

$$\pi_{10} = \frac{\alpha + \beta\delta_0}{1 - \beta\delta_1}, \quad \pi_{11} = \frac{\rho}{1 - \beta\delta_1}, \quad \pi_{12} = \frac{\gamma}{1 - \beta\delta_1},$$

$$\pi_{20} = \frac{\delta_0 + \alpha\delta_1}{1 - \beta\delta_1}, \quad \pi_{21} = \frac{\rho\delta_1}{1 - \beta\delta_1}, \quad \pi_{22} = \frac{\gamma\delta_1}{1 - \beta\delta_1},$$

$$v_1 = \frac{\varepsilon_t + \beta\eta_t}{1 - \beta\delta_1}, \quad v_2 = \frac{\delta_1\varepsilon_t + \eta_t}{1 - \beta\delta_1}.$$

With the reduced forms, we can clearly see that y_t and η_t are correlated. To show the direction of the bias, candidates can show that the OLS estimator of $\widehat{\delta}_1$ can be written as:

$$\text{plim}(\widehat{\delta}_1) = \delta_1 + \frac{\text{plim SampleCov}(y_t, \eta_t)}{\text{plim SampleVar}(y_t)} = \delta_1 + \frac{\text{Cov}(y_t, \eta_t)}{\text{Var}(y_t)}.$$

The sign of the bias depends on the sign of $\text{Cov}(y_t, \eta_t)$, which in turn is dependent on the sign of $\text{Cov}(v_1, \eta_t)$, which is $\delta_1/(1 - \beta\delta_1)$. Normally, this would be negative because we should expect $\delta_1 < 0$ (more investment, less output), and we have a negative bias.

(d) We can only obtain a consistent estimator for α_1, by using the Two Stage Least Squares:

- Estimate the reduced form of y_t, and get the following fitted value: $\widehat{y}_t$, where:

$$\widehat{y}_t = \widehat{\pi}_0 + \widehat{\pi}_1 y_{t-1} + \widehat{\pi}_2 t.$$

- Plug $\widehat{y}_t$ instead of y_t into the reduced form for x_t, and run an OLS estimation to obtain consistent estimators, the standard errors need correcting too:

$$x_t = \text{constant} + \delta_1 \widehat{y}_t + \text{error}.$$

Now candidates are expected to discuss the relevance (how to test it by running an F test in the first stage) and the exogeneity assumption (whether we can treat y_{t-1} and particularly time as exogenous).

Question 7

The University of Michigan Consumer Sentiment Index, C_t, is a consumer confidence index published monthly to gauge how US consumers feel about the economy. An economist is interested in its relationship with the monthly inflation rate, inf_t measured in percentage point. The economist considers an ADL(1, 1) model as below:

$$C_t = \beta_0 + \beta_1 inf_t + \beta_2 inf_{t-1} + \phi_1 C_{t-1} + v_t. \tag{7.1}$$

Using a dataset of $T = 240$ months, she obtains the following OLS regression:

$$C_t = \underset{(2.164)}{7.83} + \underset{(.231)}{0.51} C_{t-1} - \underset{(2.11)}{5.41} inf_t + \underset{(4.31)}{4.41} inf_{t-1}, \tag{7.2}$$

$$n = 240, \ R^2 = .094$$

where the standard errors are reported in parentheses. Both $\{C_t\}_{t=1}^T$ and $\{inf_t\}_{t=1}^T$ are assumed to be $I(1)$ processes.

(a) **A colleague runs the following regression to examine the assumptions we made about the inflation process?**

$$\Delta \widehat{inf}_t = 1.67 - \underset{(.251)}{0.312} inf_{t-1} + \underset{(.036)}{.107} \Delta inf_{t-1}, \tag{7.3}$$

$$T = 240, \ R^2 = .282.$$

What is the name of the test? Using the provided information, perform the test and conclude whether you believe the assumptions about the inflation process hold. In your answer, be clear about the concept of an $I(1)$ process and comment on the significance of the estimate for Δinf_{t-1}.

(5 marks)

(b) **What are the estimates of the short-run and long-run effects of inflation rates on consumer sentiment? Interpret the results. Given the information, discuss whether the results are meaningful. Hint: In your answer, discuss the concepts of spurious relationship and co-integration.**

(5 marks)

(c) **Discuss the conditions under which we can rewrite the relationship between C_t and inf_t into an Error Correction Model form. Describe a test based on the available information to examine whether such an ECM form is possible, knowing that the consumer sentiment is often dependent on the month of the year.**

(5 marks)

(d) **Describe an estimator to estimate the parameters of the following Error Correction Model.**

$$\Delta C_t = \phi_1 \Delta inf_t + \theta_1 (C_{t-1} - \alpha - \beta inf_{t-1}) + u_t. \tag{7.4}$$

Describe how you can interpret the coefficients from the estimated model.

(5 marks)

Reading for this question

Subject guide (2022), Chapters 13.2.2, 15.2.3–4.

Wooldrige, Sections 10.2, 10.4, 18.2–4.

Approaching the question

In line with previous years, this particular question received the least attention from candidates. However, for those who did attempt it, their efforts were mostly commendable.

Part (a) focuses on the definition and testing of stationarity $I(1)$. The ADF test was well described by most candidates, even though some mistakenly described the DF test instead. The key of an excellent answer is the comprehensive description of the testing procedure, noting that a DF test distribution is needed. Most candidates realised that the test is without a time trend and proceeded accordingly.

Part (b) is a textbook question on the calculation of the short- and long-run effects. Some candidates forgot to also account for the lagged outcome term in the regressors, resulting in an incorrect long-run calculation. The second part of the question focuses on the definition and recognition of spurious and cointegrating relationship. Best answers highlighted the differences between the two and how they affect the long-term interpretation of the results. Candidates mostly did well in this part.

Part (c) is another textbook question on the construction of ECM. The best answers centred around the realisation that the error terms must be $I(0)$ and the stationarity test presented in the question uses the residual from regression (7.3). Most candidates realised this and described the ADF test, concluding that the residuals are $I(0)$ and we have evidence for co-integration. Some forgot a few details of the ADF test, losing a few marks.

Part (d) was also well attempted by most candidates. The two stages were well discussed and the interpretations of the coefficients were presented clearly.

(a) $I(1)$ is a unit root process or is integrated of order one. This means that the first difference in the process is weakly dependent and often stationary. The regression provides us with the ADF regression for a unit root test for inflation. The OLS estimation gives us an ADF without time trend but with a drift. Denote the coefficient of inf_{t-1} as π. The null hypothesis is that $\pi = 0$ (or the process is $I(I)$) against the alternative: $\mathrm{H}_1 : \pi < 0$ (the residual is $I(0)$). The ADF statistic is practically the t-stat: $\widehat{\pi}/se(\widehat{\pi}) = -1.24$.

The ADF statistic is smaller (in absolute value) than the critical value of the DF distribution, in this case, -2.86. Hence we fail to reject the null hypothesis of unit root, so that we support the assumption that inflation is $I(1)$.

The estimate for Δinf_{t-1} is statistically significant so we should keep it in this ADF test, suggesting that the ADF is more appropriate than the DF version.

(b) The short-run effect is -5.41, the long-run effect is $(-5.41 + 4.41)/(1 - 0.51) = -2.04$, that is a one-percentage point increase in the inflation rate reduces the sentiment index by 2 units (consumers are less excited about the economy). Because the two processes are $I(1)$, it is likely that the relationship is spurious.

If the only reason that there is an apparent relationship is due to the persistence of the two processes we have a spurious, meaningless relationship. OLS estimators in this case are meaningless.

A cointegrating relationship indicates that there is a long-run relation between the two processes – that is, there is a linear combination of processes in which the order of integration is lower than that of the original processes (in this case, $I(0)$).

(c) To have an ECM, the two processes must be cointegrating. By definition, we must show that the error term from a linear regression between the two regressors is $I(0)$. To test this cointegrating relationship, we can test whether the residual obtain from (7.3) is $I(0)$. We obtain the residual, $\widehat{v}_t$, and run the following regression, with a time trend to account for the monthly dependence:

$$\Delta\widehat{v}_t = \alpha + \delta t + \theta\widehat{v}_{t-1} + \gamma_1\Delta\widehat{v}_{t-1} + e_t.$$

The null hypothesis is: $\mathrm{H}_0 : \theta = 0$ and the hypothesis is $\mathrm{H}_1 : \theta < 0$. If we reject the null, it will be a trend-stationary process. Candidates can add more lags to clear up any serial

correlation. Candidates can also discuss a joint significance test of δ and θ but this is not necessary. Candidates should discuss the ADF instead of the DF test.

The test statistic is simply the t-stat of $\widehat{\theta}/se(\widehat{\theta})$, which follows the Dickey–Fuller distribution instead of the standard normal distribution. The critical value for 5% in this test with a linear time trend is -3.41 instead of -2.86 in the case without a trend. To reject the null, the t-stat needs to be less than -3.41. If we reject the null hypothesis, we have evidence for a cointegrating relationship!

(d) To estimate the ECM, we use the Engle–Granger two-step procedure.

- Step 1: Estimate the cointegrating relationship by OLS. We already obtained them, in Equation (1), and obtain the residual $\widehat{e}_{t-1}$. This term is the disequilibrium.
- Step 2: Estimate the following model by OLS:

$$\Delta C_t = \phi_1 \Delta inf_t - \theta_1 \widehat{e}_{t-1} + w_t.$$

For interpretation, ϕ_1 is the direct effect of Δinf_t on the current consumer sentiment. θ_1 is an indicator of how quickly the consumer sentiment will respond to the long-run disequilibrium: if past sentiment is larger than the equilibrium value, then $\Delta C_{t-1} < 0$ so that the sentiment will go back to the equilibrium.

Question 8

A researcher is interested in explaining why people voted in the 2010 general election in Britain. Using a Probit model, she obtains the following regression results, with standard errors reported in the parentheses:

$$\Pr(\widehat{Vote}_i = 1) = \Phi\left(\underset{(0.136)}{0.544} + \underset{(0.309)}{0.727}\, income_i + \underset{(0.003)}{0.05}\, white_i + \underset{(0.122)}{0.644}\, age_i\right) \tag{8.1}$$

$$n = 1689, \text{ Pseudo } R^2 = 0.134$$

where $\Phi(\cdot)$ is the standard normal distribution, $Vote_i$ is a binary variable equal to 1 if person i voted, and 0 otherwise; $income_i$ is the total income of the person in 2020 (in thousands of GBP), $white_i$ is a binary variable of whether i is white, and age_i is the person's age.

(a) Carefully derive an economic interpretation of the Probit estimate on $income_i$. Is the sign of the coefficient as you would expect? *Hint: Students are expected to derive the marginal effect corresponded to $income_i$.*

(5 marks)

(b) Explain the estimator underlying the Probit model. In your answer, provide the relevant log-likelihood functions and discuss any properties of the estimator and their consequences on the statistical properties of hypothesis testing using the model's estimates.

(5 marks)

(c) Does being white statistically significantly increase the probability of voting in 2010? Provide an appropriate test, if any, to support your answer. Does the effect of being white on voting depend on the person's income? Explain using the results reported in (8.1).

(5 marks)

(d) Discuss two measures of goodness of fit in a probit estimation. In your answer, explain clearly the advantages and disadvantages of doing so. No technical details are expected.

(5 marks)

Reading for this question

Subject guide (2022), Chapter 12.2 (all).

Wooldrige, Sections 7.5, 17.1, Appendix C.4b.

Approaching the question

Part (a) perhaps took some candidates by surprise. Coefficients from Probit or Logit estimations are not directly interpretable, at least not the magnitudes of the estimates. The best candidates explained this statement clearly with relevant annotations. To comment on the sign of the marginal effect, best answers carefully derived the marginal effect of a Probit estimator, making use of the given standard normal distribution hint. For those who directly interpreted the given coefficients, a few marks were awarded.

Part (b) focuses on the properties of Probit and the mechanical process behind an MLE. The best candidates provided not only the correct likelihood function for the estimation (and then the log-likelihood), but also an intuitive way to obtain the function and interpret it. Some candidates did not annotate the functions well, losing a few marks, others failed to provide a good intuitive understanding of the estimation or the computational approach to calculate the estimate in practice. The properties of the MLE were well recited.

Part (c) is a tricky question and well linked to part (a). One approach is to perform a test for a single linear restriction, noting that the sign of the estimate is the same as that of the marginal effect. In general, candidates perform the z-test well, some incorrectly described a t-test even though the MLE is only asymptotically normal. Some lost marks on missing details of the testing procedures. The best candidates verbally describe the process and conclusion well.

Part (d) requires critical evaluations of two popular methods to evaluate the performance of a Probit estimation. Many discussed the pseudo-R-squared method, probably because it was given in the specification even though the topic was not examinable this year. Some candidates while discussing the joint significance of the coefficients incorrectly discussed an F test. Note that for MLEs, there will be no R-squared nor SSR, hence F tests are not applicable. We perform LR tests instead.

(a) Candidates must make clear that we cannot interpret the reported estimates as the marginal effects directly. The main reason is that the regression models are specified using a non-linear function, in this case, the $\Phi(\cdot)$. Only when the relationship is linear as in OLS, we can argue that (candidates are expected to elaborate using the particular variables in question):

$$\Delta Pr(y = 1 \,|\, x) = \beta_j \Delta x_j.$$

To obtain the marginal effect in Logit, we need to use the chain rule and:

$$\frac{\partial Pr(y = 1 \,|\, x)}{\partial x_j} = \frac{\partial \Phi(x\beta)}{\partial x_j} = \beta_j \phi(x\beta)$$

where $\phi(\cdot)$ is the probability distribution of the standard normal distribution. However, since $\phi(\cdot) > 0$, we know that people with a higher income are, on average, more likely to vote in 2010, holding other factors constant.

(b) Candidates are expected to discuss the MLE. To estimate the parameters, we can use the Maximum Likelihood Estimator that chooses $\{\beta_j\}$ to maximise the joint probability of the outcomes. The intuition is that we try to choose the $\widehat{\beta}$ such that the $\widehat{Pr}(Vote_i = 1)$ is high for voters, and low for the non-voters.

The likelihood function for this sample is:

$$\begin{aligned} L(\beta \,|\, \text{data}) &= \prod_{voters} Pr(\,Vote = 1 \,|\, x) \prod_{non-voters} Pr(\,Vote = 0 \,|\, x) \\ &= \prod_{voters} \Phi(\beta_0 + \beta_1 income_i + \beta_2 white_i + \beta_2 age_i) \\ &\quad \times \prod_{non-voters} [1 - \Phi(\beta_0 + \beta_1 income_i + \beta_2 white_i + \beta_2 age_i)]. \end{aligned}$$

So the log-likelihood function is:

$$\begin{aligned} \log L(\beta \,|\, \text{data}) &= \sum_{voters} \Phi(\beta_0 + \beta_1 income_i + \beta_2 white_i + \beta_2 age_i) \\ &\quad + \sum_{non-voters} [1 - \Phi(\beta_0 + \beta_1 income_i + \beta_2 white_i + \beta_2 age_i)]. \end{aligned}$$

There are no mathematical formulae for the estimates, the popular method was to adopt an iterative trial-and-error process. In large samples, the MLE are consistent, asymptotically normally distributed, and asymptotically efficient. It is important for statistical tests as we must now use asymptotic tests. The efficiency also ensures the precision of the MLE estimators.

(c) To test the claim, we can perform an asymptotic t test using the estimates for $white_i$. The procedure is the same as before, except we use the standard normal distribution (so technically the Wald test). The null hypothesis is $\mathrm{H}_0 : \beta_2 = 0$, the alternative is a one-sided test of $\beta_2 > 0$.

The test statistic is:

$$z = \frac{\widehat{\beta}_2}{se(\widehat{\beta}_2)} = \frac{0.05}{0.003} \quad \text{follows } N(0,1) \text{ under } \mathrm{H}_0.$$

At any conventional level of significance (one-sided), we should always reject the null hypothesis, in favour of the claim that White people are more likely to vote. Because of the statistical significance and the marginal effect of White being given as $\beta_j \phi(x\beta)$ we know that the marginal effect will be dependent on income without having to run any test.

(d) Candidates are not expected to discuss the pseudo-R-squared (not examinable). Instead, the examiners expected some discussions on:

- The percentage of observations that are correctly predicted. We can obtain the predicted response probability for each individual and then check the percentage of the predicted responses being correct. For example, a common rule is that for $y = 1$ and $Pr(y = 1) \geq 0.5$; and $y = 0$ and $Pr(y = 0) \leq 0.5$. The problem is that it does not tell us the quality of the prediction (such as for the predicted response of 0.50001 and 0.9999.
- Alternatively, we can do a joint significance of the slopes. A short discussion of the likelihood ratio test is expected (null, test statistic, and rejection rule). This method is more practical but the interpretation remains quite vague. Candidates should not discuss an F test here.

Question 9

A team examines the determinants of restaurant ratings on TripAdvisor, the world's most popular and authoritative source for restaurants, and hospitality services. They have the data on all 4562 restaurants from 50 US states being reviewed in 2022 with information on the percentage of positive reviews for restaurant i, $Review_i$; the total running expenses and overhead cost $expenses_i$ in millions of dollars; the number of new menus developed in a year, $Menu_i$; and a binary of

whether the head chef has ever worked in a Michelin-star restaurant before $Michelin_i$ (an award for a few selective establishments with exceptional services). They obtain the following estimation, using the logarithm of $expenses_i$.

$$\widehat{Review}_i = \underset{(.054)}{15.23} + \underset{(.131)}{1.72}\log(expenses_i) + \underset{(1.15)}{0.507}Menus_i + \underset{(10.23)}{25.01}Michelin_i \tag{9.1}$$

$$n = 4562, R^2 = .628.$$

The robust standard errors are reported in parentheses.

(a) **What is the interpretation of the coefficient on $Michelin_i$? Are the sign and size of the coefficient as you would expect? Can you interpret the estimate as a causal effect? Explain your answer. Be specific about any source and direction of bias if any of the MLR assumptions fail.**

(5 marks)

(b) **A colleague obtains a set of 50 dummy variables for the 50 states. Should the team include them all in the regression (9.1)? Briefly explain.**

(2 marks)

(c) **Explain how you can test the claim that expenses, the quality of the menu, and having a Michelin head chef would jointly have no effect on positive reviews. Be clear about the null hypothesis and if you require any further information.**

(4 marks)

(d) **Describe a statistical test to examine whether a Michelin-starred head-chef is more able to utilise the restaurant's expenses to improve the restaurant's reviews. Describe any necessary new regressions or information.**

(4 marks)

(e) **The team was told by an auditor that the expense details from the restaurants are likely to contain many errors. Discuss the implication of this piece of information on the properties of the OLS estimators for the parameters in (9.1). Be clear about the direction of bias and discuss a potential remedy. *Hint:* You can assume that the errors are entirely random.**

(5 marks)

Reading for this question

Subject guide (2022), Chapters 4.2.2, 5.2.2–5, 6.2.1, 6.2.4, 10.2.1–2.

Wooldrige, Sections 3.2–3, 4.2, 4.4–5, 6.2c, 7.3, 7.4, 9.4, 15.4, Appendix A.1, A.5, B.5e, B.5f.

Approaching the question

Part (a) is a routine interpretation task, noting the distinction between percentage and percentage points. Some candidates incorrectly used percentage. The expected signs and the causal inference story were the weaker part of the answers, highlighting the needed emphasis on real-life applications and practical aspects of the course. Some candidates presented plausible OVB stories and employed the OVB very well to explain the bias. Many forgot to discuss the direction of bias with a presence of the OVB.

Part (b) is a short question on the dummy variable trap and was well attempted by most candidates. Some mistakenly interpreted the set-up as the case of heteroskedasticity or had a convoluted explanation of the issue.

Part (c) was a popular and very well-attempted sub-question. Most candidates realised the need to use an F-test to test for the joint significance of all coefficients. Only a few realised that the restricted R-squared can be easily obtained from the given specification. Again, some lost marks for not discussing the assumptions nor testing equation.

Part (d) requires the use of an interaction term, which many candidates realised and proceeded well with the test. Note that the wording of the question suggested a one-sided test for the coefficient of the interaction term ('improve'). A concerning number of candidates did not recognise the interaction term being needed and directly performed a t test of the two coefficients, which can be done by splitting the sample into two. This approach is not recommended nor explained well by these candidates.

Part (e) focuses on the measurement error issue, particularly on the classical errors in variables. Like previous years, it is a challenging topic and only a small number of candidates did well in this question. Best candidates explained attenuation bias and its remedial measures (using an IV or a new dataset). Only a very few described in details the conditions of such IVs, and were handsomely rewarded.

(a) Having a Michelin-winning head chef is on average associated with an increase of 25 percentage points in the positive review, holding other factors constant. It is exactly what we should expect in terms of sign.

Candidates are expected to discuss the omitted variable bias here (the measurement error is another issue, but not expected). Any reasonable omitted variable is accepted. For example, the location of the restaurant should be positively correlated with Michelin head chef and higher rating. We have a positive bias. Candidates can discuss using the OVB formula or any derivation.

(b) No, we should not as including all the dummies and the constant will cause perfect multicollinearity. If we include them all, one of the dummies will be dropped.

(c) Candidates are expected to describe an F-test of the null of joint significance: $\beta_{expenses} = \beta_{Menus} = \beta_{Michelin} = 0$, against the alternative hypothesis of at least one of them is non-zero. Estimate the restricted model by OLS when the restriction is imposed. It is when we regress against the constant so $R_r^2 = 0$. The test statistic is F-stat:

$$F = \frac{(R_{ur}^2 - R_r^2)/k}{(1 - R_{ur}^2)/(n-k-1)} = 989.407$$

where $k = 3$, $n = 4562$. Under the null and the Gauss–Markov theorem assumptions, the F-stat follows an F distribution $F_{3,\,4558}$. The robust standard errors should be used so candidates can discuss the approximate F-stat and the chi-squared distribution instead.

The F-stat is very large so we should reject the null at any level of significance. We reject the claim.

(d) Candidates are expected to discuss the inclusion of an interaction term $\mathit{Michelin} \times \log(\mathit{expenses})$ into the regression model. The null hypothesis is $\beta_{interaction} = 0$, against a one-sided alternative hypothesis of $\beta_{interaction} > 0$. The usual t test using the estimates is expected, with the t-stat following the standard normal distribution instead of the usual t distribution.

Candidates need to clearly indicate the use of CLM assumptions, Gauss–Markov + normality, as well as the rejection rule.

(e) Candidates are expected to discuss the measurement error in explanatory variables, the classical version. Let:

$$\log(\mathit{expenses}) = x^* + v$$

where x^* is the true expenses detail, and v is the random error which has zero mean and uncorrelated with the other factors. Candidates are expected to show that in general $\log(\mathit{expenses})$ becomes correlated with the error terms (some derivation is needed).

Under the Classical Errors-in-Variables assumption, the OLS estimator will exhibit attenuation bias, that is the estimate will be closer to zero than the true effect. In this case we are likely to underestimate the effect.

To solve this problem, we may need an IV: perhaps obtain another source of information for such expenses (for example, from tax returns). Candidates are expected to briefly discuss the details here.

Another remedy is to collect a better set of data.

EC2020 Elements of Econometrics

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from Section A (8 marks each) and **THREE** from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Please find the questions on the following page.

© University of London 2024

SECTION A

Answer all questions from this section. Each question carries 8 marks.

1. Consider the following simple regression model between Y_i and X_i:

$$Y_i = \beta_0 + \beta_1 \log(X_i) + u_i, \qquad i = 1, ..., n. \tag{1.1}$$

The error term u_i is assumed to be i.i.d. with zero mean and variance equal to σ^2 and independent of the regressor $X_i > 0$. You are given given a random sample of observations $\{(Y_i, X_i)\}_{i=1}^n$.

(a) **(4 marks)** Derive the OLS estimator of β_1. Provide an economic interpretation of the OLS estimate for β_1.

(b) **(4 marks)** You are told that Equation (1.1) suffers from the classical errors-in-variables (CEV) problem of the outcome variable. Discuss what this problem means and the assumptions behind the problem. Be clear about the consequences of the problem on the OLS estimator in (a), in terms of unbiasedness, consistency, standard errors, and efficiency. You are not expected to provide technical derivation. Clear intuitions suffice.

2. Consider the following regression model:

$$y_t = \alpha + \rho y_{t-1} + \gamma_1 x_t + \varepsilon_t, \quad t = 1, ..., T, \tag{2.1}$$

where $|\rho| < 1$ and the error term ε_t is a white noise process generated randomly from a normal distribution with mean zero and constant variance. ε_t is also independent of $y_{t-1}, y_{t-2,}, ..., z_t, z_{t-1}, ...$ You are told that y_t and x_t are I(1) variables.

(a) **(4 marks)** What are the short and long-term effect of x on y?

(b) **(4 marks)** It is important to see if y_t and x_t are co-integrated. Clearly define what it means by co-integration? What are the consequences to the OLS estimation in (2.1) if y_t and x_t are not co-integrated?

3. An economist is interested in the bed and breakfast industry in the UK. Let $size$ denote the room size in squared meters m^2 in a town, $avginc$ the average town residents' income in £1000, and $tourist$ the tourist population as a percentage of the total town residents. The economist obtains the following equation with the total revenue from the hospitality industry per week in each town, $Revenue$ as the outcome variable.

$$\log(Revenue) = \beta_0 + \beta_1 \log(size) + \beta_2 \log(avginc) + \beta_3 tourist + u.$$

Using data from 216 towns, she obtains the following estimation, with the standard errors reported in parenthesis:

$$\begin{aligned} \log(\widehat{Revenues}) &= \underset{(.134)}{0.113} + \underset{(0.34)}{0.211} \log(size) + \underset{(0.081)}{1.07} \log(avginc) + \underset{(0.07)}{0.16} tourist \\ & \quad n = 216, \qquad R^2 = 0.528. \end{aligned}$$

(a) **(4 marks)** You are interested in testing the following hypothesis $H_0 : 5\beta_1 = \beta_2$. Discuss how to implement the test for H_0, clearly describing any additional information you would require to perform the test.

(b) **(4 marks)** A colleague observes that the standard error for $\log(size)$ is unreasonably large and claims that this is evidence for heteroscedasticity. Discuss whether the colleague is correct or not. In your answer, clarify what heteroscedasticity means and indicate how you should proceed with your test of H_0 from part (a) in light of the presence of heteroscedasticity (of unknown form) if the colleague is correct.

4. Consider the following time series model for $\{y_t\}_{t=1}^T$:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t + \lambda\varepsilon_{t-1}, \tag{4.1}$$

where t is a deterministic trend, ε_t is i.i.d with mean zero and variance σ^2, for $t = 1, ..., T$, and $\beta_1 \neq 0$.

(a) **(4 marks)** Show that y_t is a trend stationary. In your answer, clearly define what a trend stationary process is and how it is different from a difference-stationary process.

(b) **(4 marks)** Show that the auto-covariance $Cov(y_{t+h}, y_t) = 0$ for $h > 1$, which ensures that y_t is weakly dependent. In your answer, define what weak dependence means and why it is important.

5. Let us consider the following system of simultaneous equations that describe the equilibrium condition of beer consumption and production in a country:

$$beer = \alpha_1 price + \alpha_2 GDP + u_1, \quad \text{demand} \tag{5.1}$$

$$beer = \beta_1 price + \beta_2 sunshine + \beta_3 sunshine^2 + \beta_4 flood + u_2, \quad \text{supply} \tag{5.2}$$

where (u_1, u_2) are i.i.d. errors with zero mean and constant variances, and $Cov(u_1, u_2) \neq 0$. Denote $beer$ per capita consumption/production of beer (in gallons) at the country aggregate level, $price$ the domestic price per gallon of beer, GDP per capita income, $sunshine$ the level of sunshine hours, and $flood$ the number of severe flooding events during the last wheat-growing season, an important ingredient for the beer industry. You are told that $GDP, sunshine, flood$ are exogenous variables.

(a) **(4 marks)** Explain the concepts of endogenous and exogenous variables. Do you think it is reasonable to assume that $GDP, sunshine, flood$ are exogenous?

(b) **(4 marks)** Derive the reduced form equation for $price$. Explain why the problems with estimating (5.2) by OLS.

SECTION B

Answer three questions from this section. Each question carries 20 marks.

6. Consider the following OLS estimation of a model to explain the total compensation package, $comp$ (in \$100,000) of the CEO/managers of the 500 largest companies in the US. Using a data on $profit$, the profit before tax in millions of USD, $research$ is the expenditure on research and development in \$100,000, and $news$ is the number of newsletters issued by the company, we obtain the following estimation.

$$\widehat{comp} = \underset{(1.02)}{0.96} + \underset{(0.27)}{0.74} \log(\textit{profit}) - \underset{(0.32)}{0.79} \log(\textit{research}) + \underset{(0.12)}{0.34}\,\textit{news}, \tag{6.1}$$

$$n = 120,\ R^2 = 0.19, SSR = 2.59.$$

The standard errors are in the parentheses, SSR is the Sum of Squared Residuals.

(a) **(5 marks)** What is the interpretation of the coefficient on $\log(profit_t)$? Is the sign of the coefficient as you would expect? Explain your answer. Can you interpret the results causally? Explain.

(b) **(5 marks)** A colleague claims that the variable $news$ seems to be irrelevant. Explain what it means by "irrelevant variable" and the consequences of including such a variable in the estimation. Clearly support your answers with intuitions or appropriate technical derivations.

(c) **(5 marks)** Looking at the estimates, a colleague claims that the effect of profit before tax $\log(profit)$ on the CEO's compensation is exactly offset by the effect of research and development expenses $\log(research)$. Describe a suitable test to examine this claim. Clearly specify the null and the alternative hypothesis and the assumptions you need to make. Indicate additional information, if necessary, to conduct such a test.

(d) **(5 marks)** The research manager is concerned that the regression model is likely to suffer from heteroscedasticity. Discuss a suitable test for the presence of heteroscedasticity when the form of heteroscedasticity is unknown. What are the consequences if we reject the null of the heteroscedasticity test?

7. Job hopping is an increasingly popular trend in the labour market. It is defined as the situation when a worker moves job frequently and voluntarily. A labour economist is interested in understanding the motivation behind changing job within the first year of employment, $hopping$ (= 1 if yes, and = 0 otherwise). She obtains data from 1000 workers in Britain on $bonus_i$ (= 1 if worker i works in a company eligible for the bonus, = 0 otherwise); age_i of the worker, $gender_i$ (1 = female, 0 = otherwise), and $graduate_i$ (= 1 if graduated from a university, = 0 otherwise). She aims to use the following Probit estimation:

$$\text{Pr}(hopping_i = 1) = \Phi\left(\beta_0 + \beta_1 bonus_i + \beta_2 age_i + \beta_3 gender_i + \beta_4 graduate_i\right), \tag{7.1}$$

where $\Phi(.)$ is the cumulative distribution function of the standard normal distribution.

(a) **(5 marks)** A colleague argues that using the Probit regression model is superior to using an OLS when explaining the decision to job hopping. Discuss whether you would agree with him, highlighting both the drawbacks and advantages of Probit compared to the OLS.

(b) **(5 marks)** Describe how the Probit estimates are calculated. In your answer, specify the log likelihood function involved, the intuition of the estimator, and its statistical properties.

(c) **(5 marks)** Describe how you can obtain the effect of having a bonus on the probability of job hopping for a 45-year-old female worker without a university degree. Discuss how you can test whether the effect is statistically significant and negative. *Hint:* You are not expected to give a number: clarity of the computations would suffices.

(d) **(5 marks)** Discuss how you can test the joint significance of the workers' characteristics (age, gender, education) using the Probit estimation. Be clear about the test statistic, its distribution, the rejection rule, and any additional information you would require.

8. Consider the following regression model that describes the demand for Money in the USA:

$$\log(M_t^*) = \beta_0 + \beta_1 \log(R_t) + \beta_2 \log(Y_t) + u_t, \tag{8.1}$$

where, M_t^* is the desired (unobserved) or long run demand for Money (the real cash balances), R_t is the long-term interest (%, measured by the 90-day corporate rate of interest), and Y_t is the aggregate real national income (measured by the GDP), and u_t is an i.i.d error term, and independent of $M_t, M_{t-1}, ..., R_t, R_{t-1}, ..., Y_t, Y_{t-1}$.

The unobserved desired demand variable $\log(M_t^*)$ can also be written as an adaptive process as follows:

$$\log(M_t) - \log(M_{t-1}) = \delta \left[\log(M_t^*) - \log(M_{t-1})\right], \tag{8.2}$$

where M_t is the actual demand for money at time t.

(a) **(5 marks)** What is the interpretation of δ in (8.2)? Show that we can transform the models (8.1) and (8.2) into following short-run model of the demand function for money.

$$\log(M_t) = \gamma_0 + \gamma_1 \log(R_t) + \gamma_2 \log(Y_t) + \gamma_3 \log(M_{t-1}) + v_t, \tag{8.3}$$

where v_t is a error term.

(b) **(5 marks)** Carefully discuss the properties of the OLS estimators for (8.3) in terms of unbiasedness and consistency. Comment on the validity of the OLS standard errors. Provide suitable arguments for your claims. No formal derivations are expected.

(c) **(5 marks)** Discuss how we can test for the presence of autocorrelation in (8.3). We can assume the alternative v_t can be approximated by a stationary AR(1).

(d) **(5 marks)** Instead of fitting model (8.3), a researcher estimates the following model without $\log(M_{t-1})$.

$$\log(M_t) = \alpha_0 + \alpha_1 \log(R_t) + \alpha_2 \log(Y_t) + e_t, \tag{8.4}$$

Discuss the properties of the OLS estimators for model (8.4).
Hint: Prove that the error term in model (8.4) exhibits autocorrelation.

9. Let $X = \{X_1, X_2, ..., X_n\}$ be a random sample of size n from an exponential distribution with parameter θ. The probability density function of the distribution is given as:

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}},\ x \geq 0 \text{ and } \theta > 0, \tag{9.1}$$

You are told that $E(X_i) = \theta$ and $Var(X_i) = \theta^2$. Answer the following questions with different estimators of θ. Let us first consider a sequence of estimators given by

$$\widehat{\theta}_c = c\overline{X},$$

where c is a positive constant and $\overline{X} = n^{-1} \sum_{i=1}^{n} X_i$. For example, $\widehat{\theta}_1 = \overline{X}$ and $\widehat{\theta}_{0.5} = 0.5\bar{X}$.

(a) **(6 marks)** For any given c, determine the bias, variance, and mean squared error (MSE) of the estimator $\widehat{\theta}_c$. *Hints:* $MSE(\widehat{\theta}) = E(\widehat{\theta} - \theta)^2 = Bias^2(\widehat{\theta}) + Var(\widehat{\theta})$.

(b) **(3 marks)** Demonstrate that when we choose $c^* = \dfrac{n}{n+1}$, it minimizes the MSE of $\widehat{\theta}_c$ obtained in (i).
Hint: You will need to look at the derivative of $MSE(\widehat{\theta})$ with respect to c.

(c) **(5 marks)** Is the estimator $\widehat{\theta}_{c^*} = (n+1)^{-1} \sum_{i=1}^{n} X_i$ unbiased? Is it consistent? Support your answers with appropriate derivations. In your answer, be clear about the intuitive difference between unbiasedness and consistency.

(d) **(6 marks)** Derive the MLE estimator for θ, $\hat{\theta}_{MLE}$. Briefly discuss what desirable properties this estimator has.

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	**2·55**	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	**2·60**	0·99534	**2·75**	0·99702	**2·90**	0·99813	**3·05**	0·99886	**3·20**	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	**2·65**	0·99598	**2·80**	0·99744	**2·95**	0·99841	**3·10**	0·99903	**3·25**	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918	**3·30**	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	**3·263**	0·9994	**3·731**	0·99990	**3·916**	0·99995
3·105	0·9991	**3·320**	0·9995	**3·759**	0·99991	**3·976**	0·99996
3·138	0·9992	**3·389**	0·9996	**3·791**	0·99992	**4·055**	0·99997
3·174	0·9993	**3·480**	0·9997	**3·826**	0·99993	**4·173**	0·99998
3·215	0·9994	**3·615**	0·9998	**3·867**	0·99994	**4·417**	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	**5·0**	1·6449	**3·0**	1·8808	**2·0**	2·0537	**1·0**	2·3263	**0·10**	3·0902
45	0·1257	**4·8**	1·6646	**2·9**	1·8957	**1·9**	2·0749	**0·9**	2·3656	**0·09**	3·1214
40	0·2533	**4·6**	1·6849	**2·8**	1·9110	**1·8**	2·0969	**0·8**	2·4089	**0·08**	3·1559
35	0·3853	**4·4**	1·7060	**2·7**	1·9268	**1·7**	2·1201	**0·7**	2·4573	**0·07**	3·1947
30	0·5244	**4·2**	1·7279	**2·6**	1·9431	**1·6**	2·1444	**0·6**	2·5121	**0·06**	3·2389
25	0·6745	**4·0**	1·7507	**2·5**	1·9600	**1·5**	2·1701	**0·5**	2·5758	**0·05**	3·2905
20	0·8416	**3·8**	1·7744	**2·4**	1·9774	**1·4**	2·1973	**0·4**	2·6521	**0·01**	3·7190
15	1·0364	**3·6**	1·7991	**2·3**	1·9954	**1·3**	2·2262	**0·3**	2·7478	**0·005**	3·8906
10	1·2816	**3·4**	1·8250	**2·2**	2·0141	**1·2**	2·2571	**0·2**	2·8782	**0·001**	4·2649
5	1·6449	**3·2**	1·8522	**2·1**	2·0335	**1·1**	2·2904	**0·1**	3·0902	**0·0005**	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

Examiners' commentaries 2024 (October)

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2023–24. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE). The examinable materials are based on the new subject guide (Schafgans and Komarova, 2022, henceforth referred to as subject guide (2022)) and the new textbook:

- Wooldridge, J. *Introductory Econometrics: A Modern Approach.* (Cengage, 2020), seventh edition [ISBN 9781337558860].

Detailed reading references in this subject guide refer to the edition of the set textbook listed above. A new edition of this textbook may have been published by the time you study this course. You can, however, use a more recent edition; use the detailed chapter and section headings to identify relevant readings. Also check the virtual learning environment (VLE) regularly for updated guidance on readings.

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2022). The subject guide is intended for you to use in conjunction with the main textbook. It will act to highlight important aspects of the content in the chapters, provide further clarifications where deemed useful and also to extend it in certain places. For the final examination, you will need to be familiar with the material in both the textbook and the subject guide. You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

General remarks

Learning outcomes

At the end of the course, and having completed the Essential reading and activities, you should be able to:

- describe and apply the classical regression model and its application to cross-section data
- describe and apply the:
  - Gauss–Markov conditions and other assumptions required in the application of the classical regression model
  - reasons for expecting violations of these assumptions in certain circumstances
  - tests for violations

  - potential remedial measures, including, where appropriate, the use of instrumental variables
- recognise and apply the advantages of logit, probit and similar models over regression analysis when fitting binary choice models
- competently use regression, logit and probit analysis to quantify economic relationships using R
- describe and explain the principles underlying the use of maximum likelihood estimation
- apply regression analysis to time-series models using stationary time series, with awareness of some of the econometric problems specific to time series applications (for example, autocorrelation) and remedial measures
- recognise the difficulties that arise in the application of regression analysis to nonstationary time series, know how to test for unit roots, and know what is meant by cointegration.

Common mistakes committed by candidates

In the following discussion, we will address the common errors made by candidates, many of which reoccur on an annual basis despite changes to the subject guide. Subsequently, we will offer specific comments on each question in the examination paper.

Similar to the previous year, considerable effort was dedicated to ensuring that the examination structure and expected level of effort were consistent with those of previous years. It was expected that candidates would take approximately three hours to complete the examination. As the examinations return to in-person, we have used again questions based on derivations and proofs *with textbook settings*. An equal emphasis was placed on questions that required candidates to adopt learned materials to a new setting, or evaluate and discuss the properties of estimators and test procedures. Furthermore, candidates were asked to explain their answers in a more intuitive manner and provide real-life examples.

Outstanding answers would demonstrate a discernible understanding of econometrics principles and accurate implementation in the given context. Additionally, we anticipate critical analysis of the methods or estimators employed, along with clear real-world examples. Attention to details and a thorough adherence to testing or estimation procedures are also expected. We commend the excellent performance of many candidates this year. Several candidates displayed a lack of effort and preparation in their submitted attempts, even though the number of such insufficient performance decreased already considerably this year.

The examination made frequent use of the Gauss–Markov conditions, particularly in relation to discussing the BLUE property of an estimator or in conducting hypotheses tests. It is necessary to explicitly state the Gauss–Markov conditions rather than simply asserting their validity. A recommended approach, often followed by *some* candidates, is to begin by explicitly stating the Gauss–Markov conditions and then refer back to them throughout the analysis. However, many candidates neglected to mention either the Gauss–Markov conditions or the normality assumption when conducting the t test.

Candidates should realise that $\frac{1}{n}\sum_{i=1}^{n} u_i$ is not the same as $\mathrm{E}(u_i)$. So, while we typically assume $\mathrm{E}(u_i) = 0$, this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} u_i = 0$. Also, while we may be happy to assume $\mathrm{E}(x_i u_i) = 0$ (uncorrelatedness between the errors and regressors), this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} x_i u_i = 0$. Note that:

- both $\frac{1}{n}\sum_{i=1}^{n} u_i$ and $\frac{1}{n}\sum_{i=1}^{n} x_i u_i$ are random variables, which take the value 0 with probability 0 (continuous random variables)!

- $\mathrm{E}(u_i) = 0$ and $\mathrm{E}(x_i u_i) = 0$ are fixed, not stochastic!
- If we treat x_i as a random variable, we must use the Law of Iterated Expectations to obtain $\mathrm{E}(\widehat{\beta}) = \mathrm{E}(\mathrm{E}(\widehat{\beta} \mid X))$. Candidates often skipped this part in their proof.

Relatedly, a good number of candidates seem to be put off by the use of polynomial nature of the regressor in the derivation questions (such as x^{2023} instead of the usual x). A trick in this sort of question is to first think about x^{2023} as a regressor in its own right (think about it as $z = x^{2023}$) and do the proof and derivation as such. The power of stochastic regressors is that once we arrive at the final steps and given the zero conditional mean hold, we can use the Law of Iterated Expectation to relate the polynomial term to the base level. Perhaps missing the last step of the LIE impedes completing this sort of questions for some.

The proofs for unbiasedness and consistency often require clarification of the differences between sample and population moments, which seems missing for a few. In both cases, we first simplify our estimator (plug in the true model) to obtain:

$$\widehat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \beta + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \quad \text{with } x_i = X_i - \bar{X}.$$

- Note that this method applies for a wide range of estimator set-ups, not only for OLS. So in this case x_i can take any form of a function of a regressor.
- For *unbiasedness*, clearly indicate that you want to show that $\mathrm{E}(\widehat{\beta}) = \beta$. Unbiasedness does not follow from $\sum_{i=1}^{n} x_i u_i = 0$, instead it follows from $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) = 0$.

  If we treat x_i as fixed:

  $$\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) \equiv \mathrm{E}\left(\sum_{i=1}^{n} d_i u_i\right) = \sum_{i=1}^{n} d_i \mathrm{E}(u_i)$$

  and then unbiasedness follows as $\mathrm{E}(u_i) = 0$.
- For *consistency*, clearly indicate that you want to show that $\mathrm{plim}(\widehat{\beta}) = \beta$. Using the plim properties, we show:

  $$\begin{aligned}
  \mathrm{plim}\ \widehat{\beta} = \beta + \mathrm{plim}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) &= \beta + \frac{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i u_i\right)}{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)} \\
  &\equiv \beta + \frac{\mathrm{plim}(\text{Sample Cov}(x,u))}{\mathrm{plim}(\text{Sample Var}(x))} \\
  &= \beta + \frac{\mathrm{Cov}(x,u)}{\mathrm{Var}(x)} \quad \text{using the law of large numbers}
  \end{aligned}$$

  where $\mathrm{Cov}(x,u) = 0$ and $\mathrm{Var}(x) > 0$, ensuring we get consistency.
- Remember, the Law of Large Numbers guarantees that sample averages approach their corresponding population values. In order to maintain consistency in proofs, it is crucial to utilise the LLN to finalise the necessary steps. Likewise, the Law of Iterated Expectations (LIE) plays a significant role in establishing the connection between the conditional expectations and the unconditional expectations of an estimator.

A large number of candidates are not able to clearly distinguish between sample variance and covariance, and population variance and covariance.

The use of $\text{Cov}(X, Y)$ and $\text{Var}(X)$ should be restricted to describing the population covariance and variances, respectively, with definitions:

$$\text{Cov}(X, Y) = \text{E}((X - \text{E}(X))(Y - \text{E}(Y))) = \text{E}(XY) - \text{E}(X)\,\text{E}(Y)$$

and:

$$\text{Var}(X) = \text{E}((X - \text{E}(X))^2) = \text{E}(X^2) - (\text{E}(X))^2$$

(you also may denote $\text{Cov}(X, Y) = \sigma_{XY}$ and $\text{Var}(X) = \sigma_X^2$). They are typically unknown, but fixed, quantities.

The sample covariance and variance are estimators of the population covariance and variance, respectively. They are defined as:

$$\text{Sample Cov}(X, Y) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

and:

$$\text{Sample Var}(X) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

(you also may use $\widehat{\sigma}_{XY}$ and $\widehat{\sigma}_X^2$). You can compute them given the data.

With a slight abuse of notation, we often divide by n instead, which is irrelevant if we let n be large. The division by $n - 1$ is a finite sample issue only (unbiasedness).

The sample covariance and variance show up in our definition of the OLS estimator of the slope in the simple linear regression model, not the population covariance and variance, as:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{\text{Sample Cov}(X, Y)}{\text{Sample Var}(X)} \neq \frac{\text{Cov}(X, Y)}{\text{Var}(X)}.$$

Treating them as being the same results in incorrect analyses and candidates losing significant marks.

Recall that:

$$\widehat{\beta}_{IV} = \frac{\sum_{i=1}^{n}(Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(Z_i - \bar{Z})(X_i - \bar{X})} = \frac{\text{SampleCov}(Z_i, Y_i)}{\text{SampleCov}(Z_i, X_i)} \neq \frac{\text{Cov}(Z_i, Y_i)}{\text{Cov}(Z_i, X_i)}.$$

It was very common to misunderstand the central limit theorem. While it provides the distribution of the *estimator* $\widehat{\beta}$, no matter what the true distribution of the errors is, it does *not* say that the distribution of the error term is approximately normal in large samples. A good number of candidates failed to connect the concepts of stationarity and weak dependence in allowing for the LLN and CLT to be used in our time series setting.

Some candidates did not provide satisfactory answers when it came to interpreting empirical results. When engaging in the interpretation of an empirical result, it is essential to analyse the significance, magnitude, and sign of the coefficients. Moreover, it is crucial to establish the principle of *ceteris paribus* or holding all other factors constant. This principle ensures that the estimation can be properly understood as the marginal change in the dependent variable in response to a change in the independent variable of interest. Furthermore, a significant number of candidates failed to address the issue of omitting important and relevant explanatory variables, as well as irrelevant explanatory variables. By neglecting the former, there exists a risk of introducing bias, known as 'omitted variable bias'. On the other hand, omitting the latter would merely result in a loss of efficiency. The concepts of omitted variable bias and the associated OVB formula are significant subjects in the curricula, particularly in the latest versions of the course. It would be advantageous for candidates to engage in more extensive revisions on these topics.

Two other sources of endogeneity, which are often overlooked by candidates, are reverse causality (simultaneity bias) and measurement errors. Only a few candidates allude to them when discussing potential biases and are rewarded.

It is also important to understand that unbiasedness or consistency is a property of an estimator (not an estimate). It may be tempting to say an 'unbiased estimate'. Technically this statement is incorrect. It is the estimator (or the rule of which estimates are calculated) is biased or inconsistent, not the estimate itself.

Numerous candidates appeared to encounter difficulties when interpreting coefficients in the context of a non-linear regression, such as log-linear/linear-log or log-log relationships. One helpful strategy for such interpretations is to bear in mind that coefficients pertaining to variables measured in logarithmic form should be interpreted in a relative sense. We have already provided necessary notes in the subject guide for the interpretation. Another trick to think about the interpretation is that changes in the variables in logs must be interpreted in a relative term (in percentage).

In the case of dummy variables or categorical variables in general as regressors, candidates often missed that the coefficients have to be interpreted relative to the base case which is the left out dummy variable (see appropriate questions in the examinations for details)

Candidates particularly struggled with the difference between percentage points and % changes when it comes to the interpretation of coefficients in binary choice models. A coefficient in such a model gives the percentage point difference in the outcome variable when the regressor changes by 1 unit. The % change is a *relative* change that depends on the baseline value of the outcome variable (for example, the realisation when a regressor equals zero in the case of a dummy variable regressor). For example, if we have initially 50% of a sample buying ecological apples, a 10% increase will mean now $(1 + 0.1) \times 50 = 55\%$ of the sample buying ecological apples. A 10% point increase instead will result in $50 + 10 = 60\%$ of the sample choosing to buy ecological apples.

When conducting hypothesis testing, it is important to ensure that you clearly indicate the null and alternative hypotheses in terms of the true parameters, denoted as β_1). Additionally, provide the test statistic in terms of the parameter estimator (t-stat, τ-stat, F-stat, LR). Specify the distribution of the test statistic under the null hypothesis and appropriate assumptions with the corresponding degrees of freedom. Clearly state the rejection rule, whether it is one-sided or two-sided, for a given significance level, typically at 5%, using appropriate critical values. Finally, provide a comprehensive interpretation of the result. It should be noted that missing any of these steps may result in a loss of marks. Many candidates fail to obtain the full marks for these questions because of the insufficient details.

Another important consideration is selecting the critical value that is appropriate for the sample size. Using a critical value of 1.96, which is suitable for a large degree of freedom, in a small sample is an error that should be avoided. Furthermore, it is crucial to understand when to use an asymptotic version of a test, such as when normality of the errors is not assumed or when robust standard errors are employed.

When using confidence intervals in a hypothesis test, candidates should decide if the hypothesised value for the population parameter of interest β lies within the confidence interval. Many candidates used the default value of zero when discussing their confidence interval regardless of the population parameter in question. Another popular mistake is to conduct a two-sided test when the statement was worded with a clear direction of the relationship.

Several candidates inaccurately defined multicollinearity as a correlation between independent variables. While we acknowledge that a high correlation among the independent variables may serve as an indicator of multicollinearity, it is incorrect to solely rely on the correlation between the regressors to define multicollinearity. Additionally, numerous candidates mistakenly claimed that multicollinearity leads to biasedness and/or inconsistency.

Many candidates made the classic error of not recognising that:

$$\mathrm{E}\left(\frac{X}{Y}\right) \neq \frac{\mathrm{E}(X)}{\mathrm{E}(Y)}$$

in general. We do have:

$$\text{plim}\left(\frac{X}{Y}\right) = \frac{\text{plim}(X)}{\text{plim}(Y)}.$$

The concepts of autocorrelation and heteroskedasticity proved to be challenging for many candidates. A considerable number of candidates struggled to adequately explain the rationale behind employing robust standard errors or asymptotic methods for testing in the presence of autocorrelation and/or heteroskedasticity. Furthermore, many candidates received deductions in their grades when discussing the properties of an estimator that can effectively address these issues. It should be noted that some candidates erroneously proposed that addressing autocorrelation would automatically resolve issues related to inconsistency and biasedness.

Testing and recognising the problem of heteroskedasticity posed challenges, particularly when the scenarios were presented in a more real-life manner. One popular wording was that the variability of outcome variables differed across different groups of one or some of the regressors. A significant number of candidates failed to identify this as the classic problem of heteroskedasticity. Among those who did recognise it, many employed the Goldfeld–Quandt version of the heteroskedasticity test. However, in the new subject guide, the Goldfeld–Quandt test is no longer introduced due to its limitations compared to alternative options such as the Breusch–Pagan and the White tests. These alternatives offer advantages such as relaxing the assumption that the error term variance to be a monotonic function of a regressor. For future iterations of the course, it is recommended that candidates thoroughly study the latest version of the subject guide, as it presents the most up-to-date approach to the discussed materials.

When discussing the significance of differentiating between trend stationary and difference stationary processes, there are two essential aspects to consider: the implications for the long-term behaviour of the processes and the statistical implications, specifically the consistency of the estimators. While the majority of candidates discussed the former aspect effectively, only a few briefly addressed the latter.

The treatment of non-stationary time series often proves a challenge to many cohorts. Candidates can often verbally explain the differences between spurious and cointegrating, however details related to its testing or consequences of the relationship often seem more difficult.

Candidates frequently received lower marks for their interpretation of the Error Correction Model (ECM), especially when discussing the characteristics of the estimator in the context of stationary or non-stationary variables. Many candidates encountered difficulties in deriving an ECM model. A significant portion of the candidates either chose to skip this question entirely or performed poorly when attempting it. In general, these questions have always been unpopular to candidates, even though the answers are often straightforward.

Some candidates appeared to be uncertain about the distinction between the Dickey–Fuller test and the Augmented Dickey–Fuller test (ADF). The discrepancy lies in the formulation of the test equation. In the augmented variant, additional lagged differences of the residuals are incorporated to ensure the absence of remaining autocorrelation in the regression of the test equation. Consequently, some individuals mistakenly utilised the ADF test equation when the DF test equation was intended. Relatedly, others forgot that even though we can use a very similar test statistic to the t statistic in the DF stationarity test (we call it τ statistic), it does not follow a t distribution but the Dickey–Fuller distribution instead.

When discussing the marginal effect after a Probit/Logit estimation, candidates often did not specify the term z in the pdf $\phi(z)$ or $\Lambda(z)$. Some used the coefficients from the LPM regressions to calculate the marginal effects after a Probit/Logit estimation. A popular, related mistake was that candidates did not realise the inherent heteroscedasticity of LPM.

A number of candidates seemed to be confused between the conditions for a valid instrument (exclusion restriction and relevance) and factors related to obtaining a more precise IV estimate (a higher sample size, a higher sample variability of X_i, a lower variance of the error term, and a stronger correlation between X_i and Z_i). More seriously, most candidates did not realise the need to use an IV estimator when the regression has a lagged dependent variable as a regressor in the

presence of autocorrelation.

Note that the IV framework works for both a cross-sectional setting and a time series analysis.

Exogeneity and endogeneity are often very simply explained as variables 'outside of the model' or 'inside the model', respectively. While this approach is not wrong; it is not a particularly clear or helpful definition in general contexts that are not related to simultaneous equations models. Defining these concepts in terms of the relationship between regressor and the error term is much clearer and naturally leads to a way to show exogeneity or endogeneity by considering the covariance between regressors and error terms.

Candidates often missed that all exogenous regressors and instruments have to be included in the first stage when performing a TSLS estimation. Some candidates use the fitted values of the endogenous variable obtained from the first stage as the dependent variable in the second stage. This mistake is a serious misunderstanding. It is fine to indicate that this fitted value can be used as an instrument of the endogenous variable, but is not really the way the second stage of TSLS is typically understood.

When asked about identification of structural form equations, candidates often compared the number of excluded regressors to the number of endogenous variables in general, without pointing out which variables are the excluded regressors and could therefore be used as an instrument. Candidates were also often unclear of the definitions of over-identification, under-identification, and exact identification. The best approach to this type of questions is to discuss the endogenous and exogenous variables, and then think about the available exogenous variables that can be used as instruments.

When deriving the reduced forms, candidates should check conditions for their existence, in particular, whether the denominator of the expression is different from zero.

The question on deriving the maximum likelihood estimator seems to be unpopular to many. However, it is actually a straightforward one. Those who chose the question, did very well.

Just as in previous years, many candidates did not answer all parts of a question. Make sure you read the questions properly and provide all details that are requested. Not answering a question will automatically earn you a zero. Some candidates probably did not have a good examination plan.

Key steps to improvement

Essential reading for this course includes the new subject guide (2022) and the following:

- Wooldridge, J. *Introductory Econometrics: A Modern Approach.* (Cengage, 2020) seventh edition [ISBN 9781337558860].

Please note that as long as you read the Essential reading you are then free to read around the subject area in any text, paper or online resource. You will need to support your learning by reading as widely as possible and by thinking about how these principles apply in the real world. To help you read extensively, you have free access to the virtual learning environment (VLE) and University of London Online Library (see below).

- Dougherty, C. *Introduction to econometrics.* (Oxford: Oxford University Press, 2016) 5th edition [ISBN 9780199676828].
- Gujarati, D.N. and D.C. Porter *Basic econometrics.* (McGraw–Hill, 2009, International edition) 5th edition [ISBN 9780071276252].
- Maddala, G.S. and K. Lahiri *Introduction to econometrics.* (Wiley, 2009) 4th edition [ISBN 9780470015124].
- Stock, J.H. and M.W. Watson *Introduction to econometrics.* (Pearson, 2012) 3rd edition [ISBN 9781408264331].

To understand the subject it is important to supplement the reading of the textbook Wooldridge with the subject guide (Schafgans and Komarova, 2022). The subject guide is intended for you to use in conjunction with the main textbook. It will act to highlight important aspects of the content in the chapters, provide further clarifications where deemed useful and also to extend it in certain places. For the final examination, you will need to be familiar with the material in both the textbook and the subject guide.

The subject guide provides a range of activities that will enable you to test your understanding of the basic ideas and concepts. We want to encourage you to try the exercises you encounter throughout the material before working through the solutions provided on the virtual learning environment (VLE). With econometrics the motto has to be 'practise, practise, practise...'. It is the best way to learn the material and prepare for examinations. The course is rigorous and demanding, but the skills you will be developing will be rewarding and well recognised by future employers.

At the end of each chapter you will find an overview of the material covered, a list of key terms and concepts, and a reminder of the learning outcomes. It concludes with additional examination based exercises that will enable you to test your knowledge and understanding with worked-out solutions. Worked-out solutions for the exercises can be found on the VLE (downloadable pdf files). Candidates should utilise datasets using standard regression programmes (R in our course, but Stata or Python are other popular choices). This will help the understanding of the subject.

Examination revision strategy

Many candidates are disappointed to find that their examination performance is poorer than they expected. This may be due to a number of reasons, but one particular failing is '**question spotting**', that is, confining your examination preparation to a few questions and/or topics which have come up in past papers for the course. This can have serious consequences.

We recognise that candidates might not cover all topics in the syllabus in the same depth, but you need to be aware that examiners are free to set questions on **any aspect** of the syllabus. This means that you need to study enough of the syllabus to enable you to answer the required number of examination questions.

The syllabus can be found in the Course information sheet available on the VLE. You should read the syllabus carefully and ensure that you cover sufficient material in preparation for the examination. Examiners will vary the topics and questions from year to year and may well set questions that have not appeared in past papers. Examination papers may legitimately include questions on any topic in the syllabus. So, although past papers can be helpful during your revision, you cannot assume that topics or specific questions that have come up in past examinations will occur again.

If you rely on a question-spotting strategy, it is likely you will find yourself in difficulties when you sit the examination. We strongly advise you not to adopt this strategy.

Examiners' commentaries 2024 (October)

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2023–24. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE). The examinable materials are based on the new subject guide (Schafgans and Komarova, 2022, henceforth referred to as subject guide (2022)) and the new textbook:

- Wooldridge, J. *Introductory Econometrics: A Modern Approach.* (Cengage, 2020), seventh edition [ISBN 9781337558860].

Detailed reading references in this subject guide refer to the edition of the set textbook listed above. A new edition of this textbook may have been published by the time you study this course. You can, however, use a more recent edition; use the detailed chapter and section headings to identify relevant readings. Also check the virtual learning environment (VLE) regularly for updated guidance on readings.

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2022). The subject guide is intended for you to use in conjunction with the main textbook. It will act to highlight important aspects of the content in the chapters, provide further clarifications where deemed useful and also to extend it in certain places. For the final examination, you will need to be familiar with the material in both the textbook and the subject guide. You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from **Section A** (8 marks each) and **THREE** from **Section B** (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Section A

Answer all questions from this section. Each question carries 8 marks.

Question 1

Consider the following simple regression model between Y_i and X_i:

$$Y_i = \beta_0 + \beta_1 \log(X_i) + u_i, \quad i = 1, \ldots, n. \tag{1.1}$$

The error term u_i is assumed to be i.i.d. with zero mean and variance equal to σ^2 and independent of the regressor $X_i > 0$. You are given given a random sample of observations $\{(Y_i, X_i)\}_{i=1}^n$.

(a) Derive the OLS estimator of β_1. Provide an economic interpretation of the OLS estimate for β_1.

(4 marks)

(b) You are told that Equation (1.1) suffers from the classical errors-in-variables (CEV) problem of the outcome variable. Discuss what this problem means and the assumptions behind the problem. Be clear about the consequences of the problem on the OLS estimator in (a), in terms of unbiasedness, consistency, standard errors, and efficiency. You are not expected to provide technical derivation. Clear intuitions suffice.

(4 marks)

Reading for this question

Subject guide (2022), Chapter 10, Sections 9.4a and 9.4b.

Wooldridge, Sections 9.4a, 9.4b and 15.4.

Approaching the question

(a) To facilitate our discussion, let $Z_i = \log(X_i)$. To obtain the OLS estimator we should solve:

$$\min_{b_0, b_1} \sum_{i=1}^n (Y_i - b_0 - b_1 Z_i)^2.$$

This requires us to solve the FOC:

$$-2\sum_{i=1}^n (Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 Z_i) = 0$$

$$-2\sum_{i=1}^n (Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 Z_i) Z_i = 0.$$

Solving these equations yields:

$$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1 \bar{Z} = \bar{Y} - \widehat{\beta}_1 \overline{\log(X)}$$

$$\widehat{\beta}_1 = \frac{\sum_{i=1}^n (Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum_{i=1}^n (Z_i - \bar{Z})^2} = \frac{\sum_{i=1}^n (\log(X_i) - \overline{\log(X)})(Y_i - \bar{Y})}{\sum_{i=1}^n (\log(X_i) - \overline{\log(X)})^2}.$$

Observe $\beta_1 = \Delta Y / \Delta \log X = \Delta Y / \% \Delta X$, for a small change in X. Therefore, if you change X by 1%, we expect Y to change by $\beta_1/100$ units.

(b) The classical CEV in outcome means there is a random measurement error in the outcome variable, we can write the observed Y_i^* as $Y_i^* = Y_i + e$ where the measurement error e has zero means and most importantly independent of X_i. We do not need to have e to be uncorrelated with Y_i. Under these assumptions, the OLS estimator remains unbiased, consistent. The standard errors are valid, but will be larger than in the case without CEV. The key intuition is that we are adding the noise into the estimation, which will not affect the other statistical properties but the precision of the estimator.

In general, Question 1 was well attempted by most candidates.

- Part (a) requires candidates to derive the OLS estimators and then interpret the results using the approximation of the log ratio. Most did well. Some went even further to provide the intuition, perhaps unnecessary as it is only 4 marks.
- Part (b) is a classical question on the CEV problem. Candidates need to discuss the consequences. This is all textbook-information and most did well. The key intuition was not discussed well by many, costing some marks.

Question 2

Consider the following regression model:

$$y_t = \alpha + \rho y_{t-1} + \gamma_1 x_t + \varepsilon_t, \quad t = 1, \ldots, T, \tag{2.1}$$

where $|\rho| < 1$ and the error term ε_t is a white noise process generated randomly from a normal distribution with mean zero and constant variance. ε_t is also independent of $y_{t-1}, y_{t-2}, \ldots, z_t, z_{t-1}, \ldots$. You are told that y_t and x_t are $I(1)$ variables.

(a) What are the short and long-term effect of x on y?

(4 marks)

(b) It is important to see if y_t and x_t are co-integrated. Clearly define what it means by co-integration? What are the consequences to the OLS estimation in (2.1) if y_t and x_t are not co-integrated?

(4 marks)

Reading for this question

Subject guide (2022), Chapters 14.2.1, 11, 13.2.

Wooldrige, Sections 12.1–12.2, 12.3 (skip 12.3b), 10.5.

Approaching the question

(a) The short-run effect is simply γ_1. To obtain the long-run relationship, we can set the variables at equilibrium:

$$y = \frac{\gamma_1}{(1-\rho)} z$$

so that the long-run effect is $\gamma_1/(1-\rho)$.

(b) We have co-integration so we do have a long-run relationship between our $I(1)$ variables x_t and y_t. We have a linear combination of $I(1)$ variables that becomes stationary $I(0)$. For example, if x_t and y_t are $I(1)$ processes, but $y_t - \beta x_t$ is stationary. Then we can say X and Y are co-integrated, meaning there is a long-term relationship between these two. In that case, there will be an ECM representation. In contrast, two $I(1)$ variables can appear to have a meaningful relationship but in fact there is none. That's the spurious relationship: where two unrelated series give the impression that they are related, purely because of the fact that both are trending. It is important to confirm because otherwise the statistical inference is invalid and regression is meaningless (OLS not consistent)!

Part (a) is a popular setup. Candidates should pay attention to the condition for the long-run: we can set y_t and y_{t-1} to its long-run equilibrium. Most candidates got this correctly.

Part (b) is a discussion question for co-integration. Best candidates realised the non-cointegrated cases refer to spurious relationships and discussed them well. Best candidates also referred to the

invalid statistical inference, but only a few did so. Overall, candidates did quite well on this question.

Question 3

An economist is interested in the bed and breakfast industry in the UK. Let *size* denote the room size in squared meters m^2 in a town, *avginc* the average town residents' income in £1000, and *tourist* the tourist population as a percentage of the total town residents. The economist obtains the following equation with the total revenue from the hospitality industry per week in each town, *Revenue* as the outcome variable.

$$\log(Revenue) = \beta_0 + \beta_1 \log(size) + \beta_2 \log(avginc) + \beta_3 tourist + u.$$

Using data from 216 towns, she obtains the following estimation, with the standard errors reported in parenthesis:

$$\widehat{\log(Revenues)} = \underset{(0.134)}{0.113} + \underset{(0.34)}{0.211} \log(size) + \underset{(0.081)}{1.07} \log(avginc) + \underset{(0.07)}{0.16} tourist$$

$$n = 216, \ R^2 = 0.528.$$

(a) **You are interested in testing the following hypothesis $H_0 : 5\beta_1 = \beta_2$. Discuss how to implement the test for H_0, clearly describing any additional information you would require to perform the test.**

(4 marks)

(b) **A colleague observes that the standard error for $\log(size)$ is unreasonably large and claims that this is evidence for heteroscedasticity. Discuss whether the colleague is correct or not. In your answer, clarify what heteroscedasticity means and indicate how you should proceed with your test of H_0 from part (a) in light of the presence of heteroscedasticity (of unknown form) if the colleague is correct.**

(4 marks)

Reading for this question

Subject guide (2022), Chapters 5, 6.2.1, 7, 8.

Wooldrige, Sections 4.1–4.5; Section 6.2b and Appendices A.4a, A.5; Sections 5.1–5.3 and Appendices C.3a and C3.b 8.1–8.4.

Approaching the question

(a) The t statistic to test the null hypothesis H_0 is written as:

$$t = \frac{5\widehat{\beta}_1 - \widehat{\beta}_2}{se(5\widehat{\beta}_1 - \widehat{\beta}_2)}.$$

From the above output, we know $5\widehat{\beta}_1 - \widehat{\beta}_2$. On the other hand, the standard error of $5\widehat{\beta}_1 - \widehat{\beta}_2$ is given by:

$$se(5\widehat{\beta}_1 - \widehat{\beta}_2) = \sqrt{25 \times se(\widehat{\beta}_1)^2 + se(\widehat{\beta}_2)^2 - 10 \times s_{12}}$$

where s_{12} is the estimate of $\text{Cov}(\widehat{\beta}_1, \widehat{\beta}_2)$. To implement the above t test, we need s_{12}. Under the Classical Linear Regression Assumptions (Gauss–Markov assumptions and normality) (candidates must state these somewhere in the script), we know that $t \sim t_{62}$ under H_0. Once we get s_{12}, we can implement the t test.

Let $H_1 : 50\beta_1 \neq \beta_2$ be the two-sided alternative and set the significance level at 5%. The rejection rule is $|t| > c$, where c is the 97.5th percentile of the t_{212} distribution.

(b) An unreasonably large standard error might to a little extent give some informal sign of heteroscedasticity. We cannot rely on this as there are many other reasons for a large standard error (mis-specification, multicollinearity, or low variations in the regressors). Instead, one must use classical tests such as the White test, or the Breusch–Pagan, or Goldfeld–Quandt tests. Heteroscedasticity means the variance of the error terms u depend on some of the regressors, invalidating the standard errors.

If there is indeed heteroscedasticity, we should use Newey–West–White's heteroscedasticity robust standard errors. Using the robust standard errors we then apply an asymptotic t test (does not require the assumption of normality of the errors), and the critical values must be obtained from the $N(0,1)$ distribution instead of the t distribution.

Part (a) looks complicated but has been a popular question. The key is to calculate the estimate of the covariance of the two estimated coefficients. Some candidates were confused with the null hypothesis as well. The main mistake is often missing some details of the testing procedure.

Part (b) discusses the symptoms of having a large standard errors. Most candidates referred, correctly, to the variance formula to discuss the various reasons. The discussion on the robust standard errors is lacking though.

Question 4

Consider the following time series model for $\{y_t\}_{t=1}^T$:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t + \lambda\varepsilon_{t-1}, \tag{4.1}$$

where t is a deterministic trend, ε_t is i.i.d. with mean zero and variance σ^2, for $t = 1, \ldots, T$, and $\beta_1 \neq 0$.

(a) Show that y_t is trend stationary. In your answer, clearly define what a trend stationary process is and how it is different from a difference-stationary process.

(4 marks)

(b) Show that the auto-covariance $\text{Cov}(y_{t+h}, y_t) = 0$ for $h > 1$, which ensures that y_t is weakly dependent. In your answer, define what weak dependence means and why it is important.

(4 marks)

Reading for this question

Subject guide (2022), Sections 13.2.1, 13.3.4, 14.2.1, 15.2.1.

Wooldrige, Sections 10.1–10.3, 11.1–11.2, 12.1–12.2, 12.3 (skip 12.3b).

Approaching the question

(a) A process is trend stationary if, by detrending, we can make it stationary. A process is difference stationary if, by differencing, we can make it stationary. It is important to distinguish the two types as they have implications for the long-run behaviours of a process after a shock to the processes. After a shock, trend stationary processes will always revert to the trend in the long run (they would exhibit weak dependence). Difference stationary processes, however, never recover from shocks to the system. Shocks to the system will have permanent effects on the processes and the processes exhibit strong dependence. To show y_t is trend stationary, it is clear that $y_t - t = \varepsilon_t + \lambda\varepsilon_{t-1}$, a clearly stationary process as it is an MA(1) structure.

(b) We check the covariance $\text{Cov}(y_{t+h}, y_t)$, where $h > 1$:

$$\begin{aligned}\text{Cov}(y_{t+h}, y_t) &= \text{Cov}(\beta_0 + \beta_1(t+h) + \varepsilon_{t+h} + \lambda\varepsilon_{t+h-1}, \beta_0 + \beta_1 t + \varepsilon_t + \lambda\varepsilon_{t-1}) \\ &= \text{Cov}(\varepsilon_{t+h} + \lambda\varepsilon_{t+h-1}, \varepsilon_t + \lambda\varepsilon_{t-1}).\end{aligned}$$

By definition of the covariance the additive non-stochastic terms vanish. Since $\text{Cov}(\varepsilon_i, \varepsilon_j) = 0$, $i \neq j$, then we should note that the covariance must be equal to zero for $h > 1$. In particular, with $h = 1$, we get:

$$\text{Cov}(y_{t+1}, y_t) = \text{Cov}(\varepsilon_{t+1} + \lambda\varepsilon_t, \varepsilon_t + \lambda\varepsilon_{t-1}) = \lambda\text{Var}(\varepsilon_t) \neq 0$$

but for $h = 2$, we get:

$$\text{Cov}(y_{t+2}, y_t) = \text{Cov}(\varepsilon_{t+2} + \lambda\varepsilon_{t+1}, \varepsilon_t + \lambda\varepsilon_{t-1}) = 0$$

for $h > 2$ we do not get any corresponding error terms either.

A time series is weakly dependent if the correlation between y_t and future periods y_{t+h} tends to zero asymptotically; with the periods increasing to infinity, the dependency between the two values tends to zero. It is important because it allows us to use LTN and CLT (replacing the i.i.d. assumption).

Part (a) is a standard question. Candidates are expected to show stationarity by taking away the trend, and then define correct what difference-stationarity is. Most lost marks for the explanation of why the distinction is important.

Part (b) asks candidates to be a bit more involved by deriving the covariance for a time distance of 2 ($h = 2$). Only a few got that far. Not many recognised (or rather recited) the fact that stationarity allows us to use LTN and CLT in lieu of the i.i.d. assumption.

Question 5

Let us consider the following system of simultaneous equations that describe the equilibrium condition of beer consumption and production in a country:

$$\boldsymbol{beer = \alpha_1 price + \alpha_2 GDP + u_1,} \qquad \textbf{(demand (5.1))}$$

$$\boldsymbol{beer = \beta_1 price + \beta_2 sunshine + \beta_3 sunshine^2 + \beta_4 flood + u_2,} \qquad \textbf{(supply (5.2))}$$

where (u_1, u_2) are i.i.d. errors with zero mean and constant variances, and $\text{Cov}(u_1, u_2) \neq 0$. Denote *beer* per capita consumption/production of beer (in gallons) at the country aggregate level, *price* the domestic price per gallon of beer, *GDP* per capita income, *sunshine* the level of sunshine hours, and *flood* the number of severe flooding events during the last wheat-growing season, an important ingredient for the beer industry. You are told that *GDP*, *sunshine*, *flood* are exogenous variables.

(a) Explain the concepts of endogenous and exogenous variables. Do you think it is reasonable to assume that *GDP*, *sunshine*, *flood* are exogenous?

(4 marks)

(b) Derive the reduced form equation for *price*. Explain why the problems with estimating (5.2) by OLS.

(4 marks)

Reading for this question

Subject guide (2022), Sections 9.2.3–4, 9.2.5.

Wooldrige, Sections 15.2, 15.1 remainder, 15.1a–15.1c.

Approaching the question

(a) Exogenous variables are those whose values are determined outside of the model. *GDP*, *sunshine* and *flood* are such variables. It is reasonable as it is unlikely that beer consumption and production can affect these variables, mostly natural ones.

Endogenous variables are those whose values are determined interactively within the model (the system of equations). In our case, *wheat* and *price* are endogenous because their values are determined through the models.

(b) The key problem is that *price* is endogenous in (5.1). To demonstrate the endogeneity of *price* we can express *price* in terms of the exogenous variable by equating (5.1) and (5.2):

$$\alpha_1 price + \alpha_2 GDP + u_1 = \beta_1 price + \beta_2 sunshine + \beta_3 sunshine^2 + \beta_4 flood + u_2$$

$$\rightarrow \quad price = \frac{\beta_2 sunshine + \beta_3 sunshine^2 + \beta_4 flood - \alpha_2 GDP + u_2 - u_1}{\alpha_1 - \beta_1}.$$

This reduced form exists because it is reasonable to assume $\alpha_1 \neq \beta_1$. We can clearly see that now $\text{Cov}(price, u_1) \neq 0$ because in the reduced form of *price* there is a component of u_1. Thus MLR4 that the error term is independent of the explanatory variable is violated, OLS estimator would be inconsistent.

Part (a) is a short question on the definitions of endogenous and exogenous variables. Candidates can approach either way: using the zero conditional mean or using the within and outside of the models. The identification of the equation then is rather straightforward.

Part (b) seems tedious to work out the reduced form but candidates rather did well in this question. Only a few missed some annotations or explanations of some steps. Most got the answers right.

This is, however, the least popular short question in Section A. Perhaps some time management can be of help.

Section B

Answer three questions from this section. Each question carries 20 marks.

Question 6

Consider the following OLS estimation of a model to explain the total compensation package, *comp* (in \$100,000) of the CEO/managers of the 500 largest companies in the US. Using data on *profit*, the profit before tax in millions of USD, *research* is the expenditure on research and development in \$100,000, and *news* is the number of newsletters issued by the company, we obtain the following estimation.

$$\widehat{comp} = \underset{(1.02)}{0.96} + \underset{(0.27)}{0.74} \log(profit) - \underset{(0.32)}{0.79} \log(research) + \underset{(0.12)}{0.34}\, news, \qquad (6.1)$$

$$n = 120, \; R^2 = 0.19, \; SSR = 2.59.$$

The standard errors are in the parentheses, *SSR* is the Sum of Squared Residuals.

(a) What is the interpretation of the coefficient on $\log(profit_t)$? Is the sign of the coefficient as you would expect? Explain your answer. Can you interpret the results causally? Explain.

(5 marks)

(b) A colleague claims that the variable *news* seems to be irrelevant. Explain what it means by 'irrelevant variable' and the consequences of including such a variable in the estimation. Clearly support your answers with intuitions or appropriate technical derivations.

(5 marks)

(c) Looking at the estimates, a colleague claims that the effect of profit before tax log(*profit*) on the CEO's compensation is exactly offset by the effect of research and development expenses log(*research*). Describe a suitable test to examine this claim. Clearly specify the null and the alternative hypothesis and the assumptions you need to make. Indicate additional information, if necessary, to conduct such a test.

(5 marks)

(d) The research manager is concerned that the regression model is likely to suffer from heteroscedasticity. Discuss a suitable test for the presence of heteroscedasticity when the form of heteroscedasticity is unknown. What are the consequences if we reject the null of the heteroscedasticity test?

(5 marks)

Reading for this question

Subject guide (2022), Sections 3.2.4, 8.2.3, 5.2.4, 4.2.2, 5.2.2–5, 6.2.4, 10.2.

Wooldrige, Sections 3.2, 6.2, 8.3, 4.4, Appendices A1–4, B.5e; 9.4a, 9.4b and 15.4.

Approaching the question

(a) A 1% increase in profit is on average associated with an increase of $0.74 \times 100{,}000/100 = 7{,}400$ USD increase in the CEO compensation. This is expected as higher profits can allow the company to further expand. It's unlikely to be causal. Candidates are expected to discuss one example of OVB (for example, company's core performance or pay structure, or quality of the CEO), with a clear direction of bias. Best candidates would also mention reverse causality (higher compensation motivates more profit).

(b) The colleague is concerned about including a variable that has no correlation with the outcome variable. In this case, it seems to be true that the number of news issues seems to be irrelevant to the CEO compensation. Including this irrelevant variable is not harmful if *news* is also uncorrelated with log(*profit*). But if *news* is correlated with *profit*, which seems to be likely as firms may issue more news if they are doing well, then including this irrelevant variable will decrease the precision of the estimate (increasing standard errors). The intuition is that we are introducing some noise in the error terms. Candidates can cite the variance of an estimator formula to show why this happens.

(c) The colleague claims that $\beta_1 + \beta_2 = 0$, where β_1, β_2 are the population parameters for log(*profit*) and log(*research*). The null hypothesis is $\mathrm{H}_0 : \beta_2 = -\beta_1$. The restriction is $\beta_2 = -\beta_1$, and the alternative is otherwise. At this point, there are three approaches:

- Candidates can do a t test, in which the covariance between $\mathrm{Cov}(\widehat{\beta}_1, \widehat{\beta}_2)$ is needed.

$$t = \frac{\widehat{\beta}_1 + \widehat{\beta}_2}{s.e.(\widehat{\beta}_1 + \widehat{\beta}_2)}.$$

  The t test statistic follows a t distribution and details of the test are expected. Candidates must explain how to obtain the standard errors for the t test.

- Alternatively, candidates can create a new variable $X = \log(profit) - \log(research)$ and run the following model:

$$hire_i = \alpha + \theta \log(profit) + \beta_2 X + \beta_3 news_t + error$$

  then do a t test on $\theta = 0$.

- Candidates can also describe an F-test with the restriction of $\beta_2 = -\beta_1$. Candidates need to clearly indicate the use of CLM assumptions, Gauss–Markov + normality.

(d) Heteroscedasticity refers to the situation when the variance of the error terms (or the outcome variable) is dependent on the regressors. Heteroscedasticity in itself does not affect the unbiasedness nor consistency of an OLS estimator. The standard errors, however, now become invalid and the conventional testing inference will become invalid. Candidates are expected to do the following steps:

- Estimate the original model by OLS, as usual. Obtain the OLS residuals $\widehat{u}_i$. Compute the squared residuals $\widehat{u}_i^2$.
- Run the following regression and keep its R-squared $R^2_{\widehat{u}^2}$:

$$\begin{aligned}\widehat{u}_i^2 = \delta_0 &+ \delta_1 \log(\textit{profit}) + \delta_2 \log(\textit{research}) + \delta_3 \textit{news} + \beta_4 \log(\textit{profit})^2 \\ &+ \delta_5 \log(\textit{research})^2 + \delta_6 \textit{news}^2 + \delta_7 \log(\textit{profit}) \times \log(\textit{research}) \\ &+ \delta_8 \log(\textit{profit}) \times \textit{news} + \delta_9 \log(\textit{research}) \times \textit{news} + \textit{error}.\end{aligned}$$

- The null hypothesis is $\mathrm{H}_0 : \delta_1 = \cdots = \delta_9 = 0$ and H_1 : otherwise. It is equivalent to a test of $k = 9$ restrictions.
- The test statistic is either:

$$F\text{-stat} \quad F = \frac{R^2_{\widehat{u}^2}/k}{(1 - R^2_{\widehat{u}^2})(n-k-1)} \qquad \text{or} \qquad LM\text{-statistic} \quad LM = nR^2_{\widehat{u}^2}$$

  with $n = 130$.
- Under the null, the F-stat follows the $F_{k,\,n-k-1}$ distribution or the χ^2_9 distribution. If the statistic exceeds the critical value from the respective distribution, we reject the null hypothesis of homoscedasticity. If that is the case, all of the standard errors reported are invalid.

Part (a) is a typical question regarding the interpretation of an OLS estimation, which is a key learning outcome of the course. Most candidates correctly interpreted the estimates, with some good explanation of the expected sign. Candidates can be rather liberal in choosing the example of the OVB. The key here is to get a reasonable discussion for the direction of bias. Only a few mentioned the potential of a reverse causality. Most recognised the linear-log regression and interpreted accordingly.

Part (b) is opposite of part (a) in which candidates are asked to discuss the irrelevant variables. Best answers gave also some examples and discussed their properties.

Part (c) is a standard t test question and most did reasonably well. The key is to correctly follow the testing procedure.

Part (d) requires the knowledge of heteroscedasticity tests, either the Breusch–Pagan or the White test. Only a very few did this well, even though the question was repeated from previous years (and part of an important course material). Best candidates also noticed the use of robust standard errors under heteroscedasticity.

Question 7

Job hopping is an increasingly popular trend in the labour market. It is defined as the situation when a worker moves job frequently and voluntarily. A labour economist is interested in understanding the motivation behind changing job within the first year of employment, *hopping* (= 1 if yes, and = 0 otherwise). She obtains data from 1000 workers in Britain on $bonus_i$ (= 1 if worker i works in a company eligible for the bonus, = 0 otherwise); age_i of the worker, $gender_i$ (1 = female, 0 = otherwise), and $graduate_i$ (= 1 if graduated from a university, = 0 otherwise). She aims to use the following Probit estimation:

$$\Pr(hopping_i = 1) = \Phi(\beta_0 + \beta_1 bonus_i + \beta_2 age_i + \beta_3 gender_i + \beta_4 graduate_i), \qquad (7.1)$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

(a) **A colleague argues that using the Probit regression model is superior to using an OLS when explaining the decision to job hopping. Discuss whether you would agree with him, highlighting both the drawbacks and advantages of Probit compared to the OLS.**

(5 marks)

(b) **Describe how the Probit estimates are calculated. In your answer, specify the log likelihood function involved, the intuition of the estimator, and its statistical properties.**

(5 marks)

(c) **Describe how you can obtain the effect of having a bonus on the probability of job hopping for a 45-year-old female worker without a university degree. Discuss how you can test whether the effect is statistically significant and negative. *Hint:* You are not expected to give a number: clarity of the computations would suffices.**

(5 marks)

(d) **Discuss how you can test the joint significance of the workers' characteristics (age, gender, education) using the Probit estimation. Be clear about the test statistic, its distribution, the rejection rule, and any additional information you would require.**

(5 marks)

Reading for this question

Subject guide (2022), Chapters 12 and 6.

Wooldrige, Sections 7.5 and 17.1.

Approaching the question

(a) Candidates are expected to make a discussion, the final decision is not relevant.

- The Probit model has three main advantages over the LPM: predicted probabilities are restricted to lie in $[0, 1]$, MLE is (asymptotically) efficient whereas OLS (LPM) will be inefficient, and the MLE automatically deals with heteroscedasticity.
- The main drawback(s) of the Probit model relative to the linear probability model is that the coefficients cannot be directly interpreted as the marginal effects of the regressor(s) of interest, we need to compute predicted probabilities using the Probit specification and it is also computationally more complicated.

(b) The parameters are estimated by MLE, where the log-likelihood is given by:

$$\begin{aligned}\log L(\beta) &= \sum_{i=1}^{n}\{hopping_i \log(\Pr(hopping_i = 1 \mid X)) + (1 - hopping_i)\log(\Pr(hopping_i = 0 \mid X))\} \\ &= \sum_{i=1}^{n}\{hopping_i \log(\Phi(\beta_0 + \beta_1 bonus_i + \beta_2 age_i + \cdots + \beta_4 graduate_i)) \\ &\quad + (1 - hopping_i)\log(1 - \Phi(\beta_0 + \beta_1 bonus_i + \beta_2 age_i + \cdots + \beta_4 graduate_i))\}.\end{aligned}$$

To obtain the parameter estimates the first-order conditions are solved (numerically as no explicit formulae exist). The approach is iterative computation. Under suitable regularity conditions, the estimates are consistent, asymptotically normally, and asymptotically efficient.

(c) The effect is the difference in predicted probabilities between a female, non-graduate worker the characteristics given, with $bonus = 1$ versus $bonus = 0$:

$$\begin{aligned}\widehat{\Pr}[y_i = 1 \mid X, bonus = 1] - \widehat{\Pr}[y_i = 1 \mid X, bonus = 0] &= \Phi\left(\sum_k x_{ki}\widehat{\beta}_k + \widehat{\beta}_1\right) - \Phi\left(\sum_k x_{ki}\widehat{\beta}_k\right) \\ &= \Phi(\beta_0 + \beta_1 + \beta_2 \times 45 + \beta_3 + 0 \times \beta_4) \\ &\quad - \Phi(\beta_0 + \beta_1 \times 0 + \beta_2 \times 45 + \beta_3 + 0 \times \beta_4).\end{aligned}$$

(d) We would like to use the LR test, to test:

$$\mathrm{H}_0 : \beta_j = 0 \quad \forall\, j \in \{\text{Worker Characteristics}\}$$

$$\mathrm{H}_1 : \beta_j \neq 0 \quad \text{for at least one } j \in \{\text{Worker Characteristics}\}.$$

Estimate the restricted (R) and the unrestricted Probit model (U). The restricted model imposes H_0. The unrestricted model is just the originally estimated model above. The LR test makes use of the difference between the two estimated log-likelihood functions (log ratio of the likelihood functions). The test statistic is:

$$LR = 2(\log L_U - \log L_R) \sim \chi^2_J$$

where J denotes the degrees of freedom which is equal to the number of restrictions, that is, $J = 3$ here. For a given level of significance, we reject if its realisation exceeds the critical value given by the χ^2_3 distribution, which equals 7.815 for the 5% level of significance.

Part (a) asks candidates to discuss the Probit versus an LPM model, which was a standard task and well attempted. Most recognised the inherent heteroscedasticity.

Part (b) focuses on the properties of Probit and the mechanical process behind an MLE. The best candidates provided not only the correct likelihood function for the estimation (and then the log-likelihood), but also an intuitive way to obtain the function and interpret it. Some candidates did not annotate the functions well, losing a few marks, others failed to provide a good intuitive understanding of the estimation or the computational approach to calculate the estimate in practice. The properties of the MLE were well recited.

Part (c) is another popular question. Most actually obtained the marginal effect by substituting the characteristics into the pdf of the standard normal distribution function. Some, similar to the solution, substitute directly to the probability functions. Both are correct.

Part (d) asks to conduct a hypothesis test on the coefficient. Candidates should realise the use of the LR test, instead of the normal F test, and discuss as well as recite the correct formula and the null hypothesis. This is the least well done subquestion. Only a few got the details right.

Question 8

Consider the following regression model that describes the demand for Money in the USA:

$$\log(M_t^*) = \beta_0 + \beta_1 \log(R_t) + \beta_2 \log(Y_t) + u_t, \tag{8.1}$$

where, M_t^* is the desired (unobserved) or long-run demand for Money (the real cash balances), R_t is the long-term interest (%, measured by the 90-day corporate rate of interest), and Y_t is the aggregate real national income (measured by the GDP), and u_t is an i.i.d. error term, and independent of $M_t, M_{t-1}, \ldots, R_t, R_{t-1}, \ldots, Y_t, Y_{t-1}$.

The unobserved desired demand variable $\log(M_t^*)$ can also be written as an adaptive process as follows:

$$\log(M_t) - \log(M_{t-1}) = \delta[\log(M_t^*) - \log(M_{t-1})], \tag{8.2}$$

where M_t is the actual demand for money at time t.

(a) **What is the interpretation of δ in (8.2)? Show that we can transform the models (8.1) and (8.2) into following short-run model of the demand function for money.**

$$\log(M_t) = \gamma_0 + \gamma_1 \log(R_t) + \gamma_2 \log(Y_t) + \gamma_3 \log(M_{t-1}) + v_t, \tag{8.3}$$

where v_t is an error term.

(5 marks)

(b) **Carefully discuss the properties of the OLS estimators for (8.3) in terms of unbiasedness and consistency. Comment on the validity of the OLS standard errors. Provide suitable arguments for your claims. No formal derivations are expected.**

(5 marks)

(c) **Discuss how we can test for the presence of autocorrelation in (8.3). We can assume the alternative v_t can be approximated by a stationary AR(1).**

(5 marks)

(d) **Instead of fitting model (8.3), a researcher estimates the following model without $\log(M_{t-1})$.**

$$\log(M_t) = \alpha_0 + \alpha_1 \log(R_t) + \alpha_2 \log(Y_t) + e_t. \tag{8.4}$$

Discuss the properties of the OLS estimators for model (8.4).

Hint: **Prove that the error term in model (8.4) exhibits autocorrelation.**

(5 marks)

Reading for this question

Subject guide (2022), Chapters 1, 13, 15.2, 14.2.

Wooldrige, Sections 10.2, 10.4, 11.1–11.2, 12.1, Appendix B, C.

Approaching the question

(a) Rearrange (8.2), we have:

$$\begin{aligned}\log(M_t) &= \delta \log(M_t^*) + (1-\delta)\log(M_{t-1}) \\ &= \delta(\beta_0 + \beta_1 \log(R_t) + \beta_2 \log(Y_t) + u_t) + (1 - \delta \log(M_{t-1})) \\ &= \delta\beta_0 + \delta\beta_1 \log(R_t) + \delta\beta_2 \log(Y_t) + (1-\delta)\log(M_{t-1}) + \delta u_t.\end{aligned}$$

δ captures the percentage of the differences between the actual and desired real cash balances (demand for money) is eliminated within a single period (here a year). It tells us how quickly the actual and desired demand for money converge.

(b) Because of the presence of $\log(M_{t-1})$ in the above model, we cannot establish unbiasedness.

The assumptions for consistency are satisfied because for this we only need to ensure that $v_t = \delta u_t$ is uncorrelated with the regressors $\log(M_{t-1})$, $\log(R_t)$, $\log(Y_t)$ and X_{t-1}. We have the assumption for the independent of the errors so they are uncorrelated. As long as there is no autocorrelation, we will also not have any correlation between $\log(M_{t-1})$ and u_t.

We do not need to use robust standard errors as there is no heteroscedasticity or autocorrelation.

(c) As the alternative v_t follows an AR(1) process, it can be written as $v_t = \rho v_{t-1} + \varepsilon_t$, where ε_t is a white noise and $|\rho| < 1$. The null hypothesis is $\mathrm{H}_0 : \rho = 0$ against $\mathrm{H}_1 : \rho \neq 0$. We will proceed as follows:

1. Run OLS on the original equation to obtain the residuals $\widehat{v}_t$.

$$\widehat{v}_t = Y_t - \widehat{\gamma}_0 - \widehat{\gamma}_1 \log(R_t) - \widehat{\gamma}_2 \log(Y_t) - \widehat{\beta}_3 \log(M_{t-1}).$$

2. By the assumption that the alternative is a stationary AR(1) process, the test equation is:

$$\widehat{v}_t = \alpha_0^* + \alpha_1^* \log(R_t) + \alpha_2^* \log(Y_t) + \alpha_3^* \log(M_{t-1}) + \rho \widehat{v}_{t-1} + \varepsilon_t.$$

3. Finally, using the sample size, n, and its R^2 (goodness of fit measure), we obtain our test statistic $nR^2 \overset{a}{\sim} \chi_1^2$. We should reject the null of no autocorrelation if nR^2 exceeds the critical value of the χ_1^2 distribution at a given level of significance. Or, we use the asymptotic t test, given by $\widehat{\rho}/s.e.(\widehat{\rho}) \overset{a}{\sim} N(0,1)$. We should reject the null if the absolute value of the t test exceeds the critical value of the $N(0,1)$ at a given level of significance.

(d) The errors in model (8.4) that excluded the relevant variable $\log(M_{t-1})$ is:

$$e_t = \gamma_3 \log(M_{t-1}) + v_t$$

where $v_t = \delta u_t$ is assumed as a white noise. The omission will lead to autocorrelation. Indeed, we can show that, for example, $\mathrm{Cov}(e_{t+1}, e_t) \neq 0$. Using our definition of e_t and e_{t+1}, we have:

$$\begin{aligned}
\mathrm{Cov}(e_{t+1}, e_t) &= \mathrm{Cov}(\gamma_3 \log(M_t) + v_{t+1}, \gamma_3 \log(M_{t-1}) + v_t) \\
&= \gamma_3^2 \mathrm{Cov}(\log(M_t), \log(M_{t-1})) + \gamma_3 \mathrm{Cov}(\log(M_t), v_t) \\
&\quad + \gamma_3 \mathrm{Cov}(v_{t+1}, \log(M_{t-1})) + \mathrm{Cov}(v_{t+1}, v_t) \\
&= \gamma_3 \mathrm{Cov}(\log(M_t), v_t).
\end{aligned}$$

The last equality holds because v_t is white noise (i.i.d. errors unrelated to anything in the past) hence $\mathrm{Cov}(v_{t+1}, v_t) = 0$ and $\mathrm{Cov}(u_{t+1}, \log(M_{t-1})) = 0$. $\mathrm{Cov}(\log(M_t), v_t) \neq 0$ because of the relationship between the two variables from (8.3). Therefore, we have:

$$\mathrm{Cov}(e_{t+1}, e_t) \neq 0$$

or the error term exhibits autocorrelation. With the presence of autocorrelation, the weak exogeneity assumption will not hold in this model with a lagged dependent variable, the OLS estimators are no longer consistent. The standard errors are now invalid too (but really irrelevant).

Part (a) is simply a transformation exercise to really interpret the coefficient δ. Only a few can meaningfully interpret the coefficient as the speed for the convergence of the money demands. This subquestion may put off candidates to attend the whole question.

Part (b) is a standard question of the consequence of including lagged dependent variables in the regression. Candidates can simply provide intuition for biasedness and the condition for consistency. The best recognise the irrelevance of robust standard errors here.

Part (c) is a question on testing for autocorrelation. It is a textbook exercise with the AR(1) model. Those who attempted did well, but very few did.

Part (d) is an extension of part (b) on the consequence of excluding a relevant variable. Candidates should recognise the presence now of autocorrelation and show mathematically the biases. Only very few candidates went this far, making this whole question the least popular and least well attempted.

Question 9

Let $X = \{X_1, X_2, \ldots, X_n\}$ be a random sample of size n from an exponential distribution with parameter θ. The probability density function of the distribution is given as:

$$f(x) = \frac{1}{\theta} e^{-x/\theta}, \quad x \geq 0 \text{ and } \theta > 0. \tag{9.1}$$

You are told that $\mathrm{E}(X_i) = \theta$ and $\mathrm{Var}(X_i) = \theta^2$. Answer the following questions with different estimators of θ. Let us first consider a sequence of estimators given by

$$\widehat{\theta}_c = c\bar{X},$$

where c is a positive constant and $\bar{X} = n^{-1} \sum_{i=1}^{n} X_i$. For example, $\widehat{\theta}_1 = \bar{X}$ and $\widehat{\theta}_{0.5} = 0.5\bar{X}$.

(a) **For any given c, determine the bias, variance, and mean squared error (MSE) of the estimator $\widehat{\theta}_c$. *Hints:* $MSE(\widehat{\theta}) = \mathrm{E}(\widehat{\theta} - \theta)^2 = Bias^2(\widehat{\theta}) + \mathrm{Var}(\widehat{\theta})$.**

(6 marks)

(b) **Demonstrate that when we choose $c^* = \dfrac{n}{n+1}$, it minimises the MSE of $\widehat{\theta}_c$ obtained in (i).**

***Hint:* You will need to look at the derivative of $MSE(\widehat{\theta})$ with respect to c.**

(3 marks)

(c) **Is the estimator $\widehat{\theta}_{c^*} = (n+1)^{-1} \sum_{i=1}^{n} X_i$ unbiased? Is it consistent? Support your answers with appropriate derivations. In your answer, be clear about the intuitive difference between unbiasedness and consistency.**

(5 marks)

(d) **Derive the MLE estimator for θ, $\widehat{\theta}_{MLE}$. Briefly discuss what desirable properties this estimator has.**

(6 marks)

Reading for this question

Subject guide (2022), Section 12.2 (all).

Wooldrige, Sections 7.5, 17.1, Appendix C.4b.

Approaching the question

(a) Candidates are expected to derive the following terms:

Mean: Using the fact that c is a constant:

$$\mathrm{E}(\widehat{\theta}_c) = \mathrm{E}(c\bar{X}) = c\mathrm{E}(\bar{X}) = c\theta$$

as:

$$\mathrm{E}(\bar{X}) = \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}(X_i) = \theta.$$

So the bias is: $c\theta - \theta = (c-1)\theta$.

Variance: Using the fact that c is a constant:

$$\mathrm{Var}(\widehat{\theta}_c) = \mathrm{Var}(c\bar{X}) = c^2\mathrm{Var}(\bar{X}) = \frac{c^2\theta^2}{n}$$

as:

$$\mathrm{Var}(\bar{X}) = \frac{1}{n^2}\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i)$$

by independence and $\mathrm{Var}(X_i) = \theta^2$.

Finally, we can find the MSE as follows:

$$\mathrm{MSE}(\bar{X}) = \mathrm{Var}(\bar{X}) + (\mathrm{E}(\widehat{\theta}_c) - \theta)^2 = \frac{c^2\theta^2}{n} + ((c-1)\theta)^2 = \frac{c^2\theta^2}{n} + (c-1)^2\theta^2.$$

(b) We should look at the value of c that will set the first derivative of $\mathrm{MSE}(\bar{X})$ with respect to c equal to zero:

$$\frac{\mathrm{d}}{\mathrm{d}c}\left(\frac{c^2\theta^2}{n} + (c-1)^2\theta^2\right) = \frac{2c\theta^2}{n} + 2(c-1)\theta^2$$

$$2c^*\frac{\theta^2}{n} + 2(c^*-1)\theta^2 = 0$$

$$2\theta^2\left(\frac{c^*}{n} + c^* - 1\right) = 0$$

$$c^*\left(1 + \frac{1}{n}\right) = 1 \quad\Longrightarrow\quad c^*\left(\frac{n+1}{n}\right) = 1 \quad\Longrightarrow\quad c^* = \frac{n}{n+1}.$$

The SOC is obviously positive, hence c^* minimises the MSE.

(c) We want to look at the property of:

$$\widehat{\theta}_{c^*} = \frac{n}{n+1}\bar{X} = \frac{1}{n+1}\sum_{i=1}^{n} X_i.$$

The estimator is biased:

$$\mathrm{E}(\widehat{\theta}_{c^*}) = \frac{1}{n+1}\sum_{i=1}^{n}\mathrm{E}(X_i) = \frac{n}{n+1}\theta \neq \theta.$$

Candidates can use sufficient conditions to prove consistency. That is:

$$\lim_{n\to\infty}\mathrm{E}(\widehat{\theta}_{c^*}) = \theta \quad\text{and}\quad \lim_{n\to\infty}\mathrm{Var}(\widehat{\theta}_{c^*}) = \lim_{n\to\infty}\left(\frac{n}{n+1}\right)^2\frac{\theta^2}{n} = \lim_{n\to\infty}\left(\frac{n}{(n+1)^2}\theta^2\right) = 0$$

or alternatively show:

$$\mathrm{plim}\ \widehat{\theta}_{c^*} = \mathrm{plim}\left(\frac{n}{n+1}\right) \times \mathrm{plim}(\bar{X}) = 1\cdot\theta = \theta$$

where LLN ensures that $\mathrm{plim}(\bar{X}) = \mathrm{E}(X_i) = \theta$.

Candidates must discuss the two concepts: unbiasedness (repeated samples, so that the expected value of the estimate from the estimator would be the parameter) and consistency (increasing sample size, the estimate would approach in probability the parameter).

(d) To obtain the maximum likelihood estimator, we first define the joint density (defines the likelihood function). Given independence we have:

$$f(X_1, \ldots, X_n) = \prod_{i=1}^{n} f(X_i; \theta) = \frac{1}{\theta}\mathrm{e}^{-X_1/\theta}.\frac{1}{\theta}\mathrm{e}^{-X_2/\theta} \cdots \frac{1}{\theta}\mathrm{e}^{-X_n/\theta} = \theta^{-n}\mathrm{e}^{-\sum X_i/\theta} = L(\theta; X_1, \ldots, X_n).$$

We then specify the log-likelihood function:

$$\log L(\theta; X_1, \ldots, X_n) = -n\log(\theta) - \frac{1}{\theta}\sum_{i=1}^{n} X_i.$$

The maximum likelihood estimator requires us to solve the FOC:

$$\frac{\mathrm{d}\log L(\theta)}{\mathrm{d}\theta} = -\frac{n}{\theta} + \frac{1}{\theta^2}\sum_{i=1}^{n} X_i.$$

For $\widehat{\theta}_{\mathrm{MLE}}$, we have:

$$-\frac{n}{\widehat{\theta}_{\mathrm{MLE}}} + \frac{1}{\widehat{\theta}^2_{\mathrm{MLE}}}\sum_{i=1}^{n} X_i = 0.$$

Multiplying this equation by $\widehat{\theta}^2_{\mathrm{MLE}}$ yields:

$$-n\widehat{\theta}_{\mathrm{MLE}} + \sum_{i=1}^{n} X_i = 0 \quad \Longrightarrow \quad \widehat{\theta}_{\mathrm{MLE}} = \frac{\sum_{i=1}^{n} X_i}{n} = \bar{X}.$$

The MLE is consistent, asymptotically normal and asymptotically efficient.

This question, expectedly, was the second least attempted long question, but those who chose it did extremely well. It perhaps has a less economic setup and contains pure statistics flavours. It fits, however, well with the MLE section of the syllabus.

Part (a) asks candidates to derive the mean, MSE, and variance of the estimator. If candidates are comfortable with working through the expected values and variances forms, the derivation is straightforward. The more challenging part is the MSE, but the formula was given in the hint.

Part (b) focuses on the minimisation of the MSE, which is a simple substitution of the given optimal c^*.

Part (c) asks to prove the unbiasedness and consistency of the given estimator, together with the distinctions between the two concepts. Most got the answers right.

Part (d) is a procedural question on maximum likelihood estimation. Candidates were expected to derive the maximum likelihood estimator from the log-likelihood. The subject guide provided the recipe for this process. Most did extremely well.

EC2020 Elements of Econometrics

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from Section A (8 marks each) and **THREE** from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Please find the questions on the following page.

© University of London 2024

SECTION A

Answer all questions from this section. Each question carries 8 marks.

1. Consider the following bivariate regression model

$$y_i = \alpha + \beta x_i + u_i, \tag{1.1}$$

for $i = 1, ..., n$. The regressor x_i is non-negative, *stochastic*, and has a non-zero variance, i.e., $\sum_{i=1}^{n}(x_i - \bar{x})^2 \neq 0$. The error term u_i is i.i.d. with mean zero and constant variance of σ^2, and independent of x_i. Given a random sample of observations $\{(y_{i,}, x_i)\}_{i=1}^{n}$ from the population of interest, consider the following estimator:

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} x_i^{2023} y_i}{\sum_{i=1}^{n} x_i^{2024}}.$$

   (a) **(3 marks)** Derive the condition under which $\tilde{\beta}$ is an unbiased estimator of β? Can $\tilde{\beta}$ be the BLUE estimator of (1.1)?

   (b) **(2 marks)** Discuss the difference between unbiasedness and consistency.

   (c) **(3 marks)** Assume that the condition in (a) is satisfied. Will this condition also ensure that $\tilde{\beta}$ is consistent?

2. Consider the following multivariate model with two regressors X_i and Z_i:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \varepsilon_i, \tag{2.1}$$

with OLS parameter estimates $\hat{\beta}_0, \hat{\beta}_1$ and $\hat{\beta}_2$. Under the Gauss Markov assumptions, you are given the following formula:

$$Var\left(\hat{\beta}_2\right) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n}\left(Z_i - \bar{Z}\right)^2} \times \frac{1}{1 - r_{XZ}^2},$$

where σ_ε^2 is the variance of ε and r_{XZ} is the sample correlation between Z and X, and n is the sample size.

   (a) **(4 marks)** Discuss four factors that can improve the precision of the parameter estimate for $\hat{\beta}_2$.

   (b) **(4 marks)** You are told that X_i is an irrelevant variable, that is the true value of $\beta_1 = 0$. Discuss the effect of including X_i in (2.1) on the unbiasedness and precision of the OLS estimator for β_2.

3. Consider a linear regression model:

$$y_t = \beta x_t + u_t, \qquad t = 1, ..., T \tag{3.1}$$

where the error terms u_t are independent of x_t and exhibit autocorrelation of an unknown form, y_t and x_t are covariance stationary and weakly dependent processes.

(a) **(5 marks)** Explain the concepts covariance stationarity and weak dependence, using appropriate technical terms. Intuitively discuss why it is desirable for the processes y_t and x_t to have these properties.

(b) **(3 marks)** Discuss how you would conduct a statistical test of the hypothesis $H_0 : \beta = 0.5$ against the alternative $H_1 : \beta > 0.5$ using the OLS estimate for β. Be specific about the steps and statistic that you use.

4. Consider the following simplified production function for a firm i:

$$output_i = \alpha_0 + \alpha_1 capital_i + \alpha_2 labour_i + \alpha_3 labour_i^2 + u_i, \qquad i = 1, ..., n \tag{4.1}$$

where $output$ denotes the total products of firm i, and $capital$ and $labour$ denote the amount of capital (machines) and workers employed by the firm. Assume that we obtain a random sample of n firms and the error terms u_i are i.i.d with zero mean and variance δ^2. We also assume that u_i and the regressors are independent.

(a) **(3 marks)** Discuss the economic and econometric rationales for including both $labour$ and $labour^2$ in this model. What are the expected signs for α_2 and α_3?

(b) **(2 marks)** Derive the formula for the effect of employing on additional worker on the total output for a firm with 100 workers.

(c) **(3 marks)** Discuss how you would test if the number of workers, $labour$ in general, has a statistically significant effect on $output$. Be clear about your null hypothesis, the test statistic, and any additional information required.

5. Consider the following model for the gin and tonic industry in Britain:

$$Q_t = \alpha_0 + \alpha_1 P_t^g + \alpha_2 P_t^b + \alpha_3 Y_t + \alpha_4 A_t + u_t \qquad \text{demand} \tag{5.1}$$

$$Q_t = \beta_0 + \beta_1 P_t^g + \beta_2 S_t + v_t \qquad \text{supply} \tag{5.2}$$

where Q_t is the real per capita consumption of gin and tonic in year t, P_t^g is the average price of 200ml of gin and tonic, P_t^b is the average price of other alcoholic beverages, Y_t is the real per capita disposable income, A_t is the real per capita expense on related advertising, and S_t is the price of sugar, an ingredient to make gin and tonic. P_t^b, Y_t, S_t, A_t are assumed to be exogenous. The error terms u_t and v_t do not exhibit any correlation over time and have constant variances.

(a) **(4 marks)** Derive the reduced form for P_t^g.

(b) **(4 marks)** Use the answer in (a) to explain why an OLS estimator for β_1 in the supply equation is likely to be biased and inconsistent.

SECTION B

Answer three questions from this section. Each question carries 20 marks.

6. A football team executive is evaluating the team's transfer activities. A sport economist is hired to perform the analysis. Using the data from 204 football teams from the top 10 leagues in Europe, she runs the following regression of the team final points after the 2023-2024 season, $Points$ (the value ranges from 20 to 100), on the net transfer expenses (in log (USD)), $\log(transfer)$, the number of players playing for their national team, $NatPlayer$, and the number of injuries in the season ($Injuries$).

$$\begin{aligned} \widehat{Points}_i &= \underset{(4.14)}{55.3} + \underset{(.231)}{0.51} \log(transfer_i) + \underset{(2.56)}{2.07} NatPlayer_i - \underset{(1.23)}{2.01} Injuries_i, \\ & n = 204, \qquad R^2 = .658. \end{aligned} \tag{6.1}$$

The robust standard errors are reported in parentheses.

(a) **(5 marks)** What is the interpretation of the coefficient on $\log(transfer)$? Are the sign and size of the coefficient as you would expect? Do you think the effect is qualitatively large or small? Explain using the available information in (6.1)

(b) **(3 marks)** Can we interpret the results in (a) causally? Explain intuitively.

(c) **(4 marks)** The economist believes that the final points for the teams are likely to be more or less variable for teams with more injuries than average than for teams with much less injuries. Discuss a statistical test for this claim. What are the consequences to the regression (6.1) if she is correct?

(d) **(4 marks)** The executive wants to see if the effect of the transfer activities can be improved by having more players who play for their national team. Discuss how you can test this claim. In your answer, clearly define the regression estimation, the assumptions, test statistics, or any additional information needed.

(e) **(4 marks)** The executive wants to see if having more national team players can offset the effect of injuries on the final points. Discuss how you can test this claim. In your answer, clearly define the regression estimation, the assumptions, test statistics, or any additional information needed.

7. Central Banks are paying attention to economic indicators in their fight against inflation. An economist uses OLS to estimate the relationship between the monthly inflation rate inf_t and the monthly employment growth rate emp_t and obtains the following results, with the standard errors are reported in the parentheses:

$$\begin{aligned} \widehat{inf_t} &= \underset{(0.014)}{0.003} + \underset{(.051)}{1.12} emp_t - \underset{(0.016)}{0.907} emp_{t-1}, \\ & T = 222, \qquad R^2 = 0.12. \end{aligned} \tag{7.1}$$

(a) **(4 marks)** What are the estimates of the short-run and long-run effect of emp_t on inf_t? Interpret the economic meaning of these estimates.

(b) **(4 marks)** The economist believes employment growth is a leading indicator of inflation. Conduct a statistical test that a unit increase in last month's employment growth results in a one unit reduction in the current month inflation. Be specific about the null hypothesis and the assumptions for the test to be valid.

Let e_t be the OLS residuals from the regression (7.1). The economist obtains the following OLS regression of e_t on e_{t-1}:

$$\begin{aligned} e_t &= \underset{(2.21)}{0.53} + \underset{(.18)}{0.41} e_{t-1} + \underset{(.118)}{0.541} emp_t - \underset{(0.126)}{1,07}\, emp_{t-1}, \\ & T = 221, \qquad R^2 = 0.42. \end{aligned} \tag{7.2}$$

(c) **(5 marks)** The economist uses (7.2) to perform a test. What is the name of the test? Clearly indicate the null and alternative hypothesis, the test statistic, rejection rule, and the *underlying assumptions*. Be clear about the outcome of the test.

(d) **(2 marks)** In light of the result in (c), the economist wants to see the impact of the previous inflation inf_{t-1} on the current inflation by including inf_{t-1} into regression (7.1). Briefly explain if it is a good idea.

(e) **(5 marks)** The economist wants to test whether the long-run effect of emp_t on inf_t is statistically significant. Discuss how you can reparameterise (7.1) to make sure your regression results can provide you with appropriate information for the test. Be clear about the new regression that you would run, and how you would proceed with the test.

8. A researcher is interested in how sales assistants in a large retailer are being promoted. Collecting a sample of 550 sales assistants who recently applied for promotion, she has data on whether the promotion is approved, $Promoted = 1$ if Yes, 0 otherwise; the amount in 100,000 of dollars sales generated by each assistant, $Sales$; the expenses in 10,000 of dollars incurred for the sales, $expenses$; years of experience, Exp; gender $female = 1$ if female, 0 otherwise; the highest qualification $educ$; and the age of the assistant, age. The researcher uses OLS and Logit to obtain the following results, with the dependent variable being $Promoted$. The heteroskedasticity robust standard errors are reported in squared brackets and the (asymptotic) standard errors are reported in parentheses.

	OLS A	Logit A	Logit B
constant	0.890 [.068]	1.088 (.206)	−0.244 (.474)
Sales	0.735 [.132]	2.029 (.378)	2.030 (.381)
expenses	−0.845 [.106]	−2.344 (.318)	−2.267 (.321)
Exp			0.14 (.015)
female			0.069 (.037)
educ			0.071 (.024)
age			−0.001 (.004)
R^2	.086		
$\log L$		−407.60	−399.04

(a) **(5 marks)** Interpret the coefficient on $Sales$ reported in Column "OLS A" and discuss whether the effect is statistically significant. Why is it important to use robust standard errors? You are expected to support your answer with technical derivations.

(b) **(5 marks)** It is argued that using the Logit model is better than using the OLS when explaining the promotion decision $Promoted$. Discuss the benefits and drawbacks of using the Logit model when trying to explain a binary variable. In your answer explain what the linear probability model refers to.

(c) **(5 marks)** Using the results in Column Logit B, indicate how you can obtain the marginal effect of one additional year of experience Exp on promotion. *Hints:* You are given $\Lambda(z) = \exp(z)/(1 + \exp(z))$, the logistic cumulative density function. Is the marginal effect positive or negative? Is it statistically significant? Explain.

(d) **(5 marks)** The researcher wants to test whether the personal attributes (Exp, $female$, $educ$, and age) have a jointly significant effect on promotion. Using an appropriate test with the given information, indicate the hypothesis, the test statistic, and the outcome of the test.

9. We are interested in the relationship between the quarterly price level CPI_t and the quarterly employment level Emp_t in the following ADL(1,1) model:

$$CPI_t = \beta_0 + \beta_1 CPI_{t-1} + \beta_2 Emp_t + \beta_3 Emp_{t-1} + u_t, \quad t = 1, ..., T, \tag{9.1}$$

where Emp_t and CPI_t are both $I(1)$ variables, and $|\beta_1| < 1$. Assume also that the error term has a zero mean and is independent of CPI_{t-1}, Emp_t and Emp_{t-1}.

(a) **(10 marks)** Discuss how you can test whether the above specification is spurious or cointegrating. Clearly indicate the test equation, test statistic, the the rejection rule. In your answer, clearly distinguish the concepts and explain why it is so important that we have to test for cointegration.

(b) **(6 marks)** A colleague ascertains that there exists an Error Correction Model (ECM) that describes the short-run and long-run dynamics between CPI_t and Emp_t. Derive the Error Correction Model and interpret the various components of the ECM. Discuss how does this claim relate to your answer in (a). *Hints:* Subtract CPI_{t-1} from both sides of (9.1) to obtain the following version:

$$\Delta CPI_t = \delta_0 + \delta_1 \Delta Emp_t + \gamma s_{t-1} + \epsilon_t,$$

with s_{t-1} is the error correction term.

(c) **(4 marks)** Discuss the Engle-Granger two-step procedure to estimate the ECM obtained in (c).

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	**2·55**	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	**2·60**	0·99534	**2·75**	0·99702	**2·90**	0·99813	**3·05**	0·99886	**3·20**	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	**2·65**	0·99598	**2·80**	0·99744	**2·95**	0·99841	**3·10**	0·99903	**3·25**	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918	**3·30**	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	**3·263**	0·9994	**3·731**	0·99990	**3·916**	0·99995
3·105	0·9991	**3·320**	0·9995	**3·759**	0·99991	**3·976**	0·99996
3·138	0·9992	**3·389**	0·9996	**3·791**	0·99992	**4·055**	0·99997
3·174	0·9993	**3·480**	0·9997	**3·826**	0·99993	**4·173**	0·99998
3·215	0·9994	**3·615**	0·9998	**3·867**	0·99994	**4·417**	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	**5·0**	1·6449	**3·0**	1·8808	**2·0**	2·0537	**1·0**	2·3263	**0·10**	3·0902
45	0·1257	**4·8**	1·6646	**2·9**	1·8957	**1·9**	2·0749	**0·9**	2·3656	**0·09**	3·1214
40	0·2533	**4·6**	1·6849	**2·8**	1·9110	**1·8**	2·0969	**0·8**	2·4089	**0·08**	3·1559
35	0·3853	**4·4**	1·7060	**2·7**	1·9268	**1·7**	2·1201	**0·7**	2·4573	**0·07**	3·1947
30	0·5244	**4·2**	1·7279	**2·6**	1·9431	**1·6**	2·1444	**0·6**	2·5121	**0·06**	3·2389
25	0·6745	**4·0**	1·7507	**2·5**	1·9600	**1·5**	2·1701	**0·5**	2·5758	**0·05**	3·2905
20	0·8416	**3·8**	1·7744	**2·4**	1·9774	**1·4**	2·1973	**0·4**	2·6521	**0·01**	3·7190
15	1·0364	**3·6**	1·7991	**2·3**	1·9954	**1·3**	2·2262	**0·3**	2·7478	**0·005**	3·8906
10	1·2816	**3·4**	1·8250	**2·2**	2·0141	**1·2**	2·2571	**0·2**	2·8782	**0·001**	4·2649
5	1·6449	**3·2**	1·8522	**2·1**	2·0335	**1·1**	2·2904	**0·1**	3·0902	**0·0005**	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

EC2020 Elements of Econometrics

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from Section A (8 marks each) and **THREE** from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Please find the questions on the following page.

© University of London 2024

SECTION A

Answer all questions from this section. Each question carries 8 marks.

1. Consider the following regression model of a time series Y_t:

$$Y_t = \beta Y_{t-1} + u_t, \qquad t = 1, ..., T \tag{1.1}$$

where the error terms u_t have a zero mean.

(a) **(3 marks)** Discuss whether the OLS estimator $\hat{\beta}$ is an unbiased estimator for β. Justify your answers with suitable technical derivations. *Note:* You are not expected to derive the OLS estimator.

You are now told that u_t follows an MA(1) process.

(b) **(2 marks)** Discuss the consequences of making the error terms an MA(1) process. What is the name given to this problem?

(c) **(3 marks)** Discuss whether the OLS estimator $\hat{\beta}$ is a consistent estimator for β. Justify your answers with suitable technical derivations.

2. Consider the following bivariate regression model

$$y_i = \alpha + \beta z_i + u_i, \tag{2.1}$$

for $i = 1, ..., n$. The regressor z_i is non-negative, *stochastic*, and has a non-zero variance, i.e., $\sum_{i=1}^{n}(z_i - \bar{z})^2 \neq 0$. The error term u_i is i.i.d. with mean zero and constant variance of σ^2, and independent of z_i. Given a random sample of observations $\{(y_{i,}, z_i)\}_{i=1}^{n}$ from the population of interest, consider the following estimator:

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} z_i^{2023} y_i}{\sum_{i=1}^{n} z_i^{2024}}.$$

(a) **(3 marks)** Derive the condition under which $\tilde{\beta}$ is a consistent estimator of β?

(b) **(5 marks)** Let us assume that the condition in (a) is satisfied so that $\tilde{\beta}$ is consistent. Will this condition also ensure that $\tilde{\beta}$ is unbiased? In your answer, discuss the difference between unbiasedness and consistency.

3. Consider the following model for the energy drink industry in Britain:

$$Q_t = \alpha_0 + \alpha_1 P_t^e + \alpha_2 P_t^c + \alpha_3 Y_t + u_t \qquad \text{demand} \tag{3.1}$$

$$Q_t = \beta_0 + \beta_1 P_t^e + \beta_2 P_t^S + v_t \qquad \text{supply} \tag{3.2}$$

where Q_t is the real per capita consumption of energy drink in year t, P_t^e is the average price of a can of energy drink, P_t^c is the average price of other caffeine drink, Y_t is the real per capita disposable income, and P_t^S is the price of sugar, an ingredient to make energy drink. P_t^c, Y_t, P_t^S are assumed to be exogenous. The error terms u_t and v_t do not exhibit any correlation over time and have constant variances.

(a) **(4 marks)** Derive the reduced form for Q_t.

(b) **(4 marks)** Use the answer in (a) to explain why an OLS estimator for α_1 in the demand equation is likely to be biased and inconsistent.

4. Consider a linear regression model:

$$y_t = \beta_1 x_t + \beta_2 x_{t-1} + u_t, \qquad t = 1, ..., T \tag{4.1}$$

where the error terms u_t are independent of x_t and x_{t-1}, y_t and x_t are covariance stationary and weakly dependent. You are told that y_t and y_{t-1} are correlated so that u_t exhibits autocorrelation.

(a) **(4 marks)** Discuss the concepts of weak dependence and autocorrelation. Discuss the consequences of these having weak dependence and autocorrelation on the OLS estimation of (4.1)

(b) **(4 marks)** Discuss a suitable test to test for autocorrelation. Be clear about the null hypothesis and the assumptions needed. What should we do if we reject the test's null hypothesis?

5. Consider the following OLS estimation results. The outcome variable is the average math scores of a primary school, math$_i$ (ranging from 0 to 10), in a school i in Britain. The regressors are $\log expend_i$, the spending per student of a school in log(USD); $\log enroll_i$, the number of registered students in logs, and $poverty$, the percentage of students live in poverty during the school year. The OLS results with and without $poverty_i$ as an explanatory variable are given in (5.1) and (5.2), with the standard errors are in parentheses:

$$\widehat{\text{math}_i} = \underset{(26.72)}{-69.24} + \underset{(3.30)}{11.13} \log(\text{expend}_i) + \underset{(0.615)}{0.022} \log(\text{enroll}_i), \tag{5.1}$$
$$N = 428, R^2 = 0.0297$$

$$\widehat{\text{math}_i} = \underset{(24.99)}{-23.14} + \underset{(3.04)}{7.75} \log(\text{expend}_i) - \underset{(0.58)}{1.26} \log(\text{enroll}_i) - \underset{(0.036)}{0.324} \text{ poverty}_i. \tag{5.2}$$
$$N = 428, R^2 = 0.1893$$

(a) **(3 marks)** Interpret the coefficient on $poverty_i$ in (5.2). Is the sign of the coefficient estimate as what you would expect? Explain.

(b) **(5 marks)** Carefully explain why the effect of expenditures on math$_i$ is lower in Equation (5.2) where $poverty_i$ is included than in Equation (5.1) where it is excluded. What is the name given to this difference?

SECTION B

Answer three questions from this section. Each question carries 20 marks.

6. Central Banks are paying attention to economic indicators in their fight against inflation. An economist is interested in the following ADL(1,1) relationship between the monthly inflation rate inf_t and the monthly employment growth rate emp_t

$$infl_t = \beta_0 + \beta_1 infl_{t-1} + \beta_2 emp_t + \beta_3 emp_{t-1} + u_t, \quad |\beta_1| < 1 \tag{6.1}$$

where emp_t and $infl_t$ are covariance stationary variables. The error terms u_t are an i.i.d. $(0, \sigma_u^2)$ and independent of $infl_{t-1}, infl_{t-2}, infl_{t-3}$..... We also assume that u_t is independent of emp_t for all t.

(a) **(4 marks)** What does it mean by "covariance stationary"? How is it different from "difference stationary" and "trend stationary"? Briefly explain the concepts and why it is important to distinguish them.

(b) **(5 marks)** Provide the short-run and long-run effect of emp on Y. Provide an economic interpretation of the effects.

(c) **(5 marks)** Discuss the statistical properties your OLS estimators for the ADL(1,1) parameters. In particular: (i) are the estimators unbiased and consistent?, and (ii) should we use robust standard errors? Support your answers with appropriate technical derivations.

(d) **(6 marks)** Show that when we exclude the relevant variable $infl_{t-1}$ from (6.1), the new model will suffer from autocorrelation. Explain your answer. Propose a solution to the problem(s) resulted from the exclusion $infl_{t-1}$.
Hint: You are expected to reformulate your model as

$$infl_t = \beta_0 + \beta_2 emp_t + \beta_3 emp_{t-1} + v_t,$$

where the new error term is $v_t = \beta_1 infl_{t-1} + u_t$.

7. A consultancy team executive is evaluating the students' performance in an online course. An education economist is hired to perform the analysis. Using the data from surveying 894 students participating in the course, she runs the following regression of the team final exam performance in the 2023-2024 academic year (marks range from 0 - 100), $Grade_i$, on the self-reported number of self-study hours per week in logs, $\log selfstudy$; whether the student also worked a job, $work_i = 1$ if Yes, and 0 otherwise; and the number of concurrent courses the student took ($othercourses$).

$$\widehat{Grade_i} = \underset{(7.14)}{48.3} + \underset{(.23)}{0.51} \log(selfstudy_i) - \underset{(4.56)}{12.07} work_i - \underset{(2.23)}{3.01} othercourses_i, \qquad (7.1)$$
$$n = 894, \qquad R^2 = .345.$$

The robust standard errors are reported in parentheses.

(a) **(4 marks)** What is the interpretation of the coefficient on $\log(selfstudy)$? Are the sign and size of the coefficient as you would expect? Do you think the effect is qualitatively large or small? Explain using the available information in (7.1).

(b) **(4 marks)** A colleague believes the regression (7.1) may suffer from heteroskedasticity. Discuss a statistical test for this claim. What are the consequences to the regression (7.1) if she is correct?

(c) **(4 marks)** A colleague wants to see if the effect of self-study on the course performance is lower when a student also had to work during the study. Discuss how you can test this claim. In your answer, clearly define the regression estimation, the assumptions, test statistics, or any additional information needed.

(d) **(4 marks)** Another colleague says that since the coefficient of $othercourses_i$ is insignificant, including the variable is harmful and the team should remove the variable from the regression. Critically discuss the suggestion.

(e) **(4 marks)** The colleague also suggests that students tend to misreport the self-study hours to the survey. Intuitively discuss the consequences on the interpretation of the regression results in (7.1) if the colleague is correct.

8. We are interested in the relationship between the quarterly price level CPI_t and the quarterly employment level Emp_t in the following model:

$$CPI_t = \alpha_0 + \alpha_1 Emp_t + u_t, \quad t = 1, ..., T, \tag{8.1}$$

where Emp_t and CPI_t are both $I(1)$ variables. Assume that the error terms exhibit autocorrelation and follow an AR(1) structure as:

$$u_t = \rho u_{t-1} + v_t, |\rho| < 1, \tag{8.2}$$

where v_t is an i.i.d. innovation with zero mean and constant variance, and is independent of e_{t-1}.

(a) **(5 marks)** Explain what it means to say CPI_t is an $I(1)$ variable. Discuss how you would test for it. In your answer, specify the test equation, test statistic, and assumptions.

(b) **(6 marks)** A colleague suggests to test for co-integration before proceeding with analysing the data. Discuss the key idea of how you can test for evidence of a co-integrating relationship between CPI and Emp. In your answer, make clear the distinction between co-integration and spurious relationship. *Hints:* we can use (8.2) to get rid of the autocorrelation in (8.1).

(c) **(5 marks)** A colleague ascertains that there exists an Error Correction Model (ECM) that describes the short-run and long-run dynamics between CPI_t and Emp_t. Derive the Error Correction Model and interpret the various components of the ECM. Discuss how does this claim relate to your answer in (b). *Hints:* You would need to rely on the model in (a) after getting rid of autocorrelation to obtain the following version:

$$\Delta CPI_t = \delta_0 + \delta_1 \Delta Emp_t + \gamma s_{t-1} + \epsilon_t, \tag{8.3}$$

with s_{t-1} is the error correction term.

(d) **(4 marks)** Using the notations in (8.3), discuss what happens to CPI_t in equilibrium in the following situations:

(i) $s_{t-1} > 0$?

(ii) $\gamma > 0$?

9. A education economist is interested in the decision to undertake an MBA course of a group of 456 employees with at least 2 years of experience. Among the 456 employees, 25% of whom belong to a minority ethnicity group, $minority_i = 1$; 126 went on to do an MBA within the next two years$MBA_i = 1$; and the other 300 employees did not have a plan to do so $MBA_i = 0$. The annual salary, measured in £1,000, of the employees pre-MBA is denoted as $salary_i$. The employees have an average salary of £43,000 and on average 4.5 years of work experience, denoted as exp_i. Using OLS and Logit, the economist obtains the following results.

$$\widehat{MBA}_i = \Lambda\left(\underset{(0.011)}{0.32} - \underset{(0.032)}{0.123}\, salary_i + \underset{(0.354)}{0.840}\, minority_i - \underset{(0.344)}{1.195}\, exp_i\right) \quad \text{(Logit)}$$

$$\widehat{MBA}_i = \underset{(0.04)}{0.113} - \underset{(0.042)}{0.223}\, salary_i + \underset{(1.124)}{2.62}\, minority_i - \underset{(1.144)}{3.295}\, exp_i \quad \text{(OLS)}$$

where $\Lambda(z) = {}^{\exp(z)}/_{(1+\exp(z))}$ is the logistic function. The usual standard errors for the Logit model and the robust standard errors for the LPM are reported in parentheses.

(a) **(5 marks)** Interpret the coefficient on $salary_i$ reported in the OLS estimation and discuss whether the effect is statistically significant. Why is it important to use robust standard errors? You are expected to support your answer with appropriate technical derivations.

(b) **(5 marks)** Describe the estimator underlying the Logit model. **Note:** clearly indicate the likelihood function used, but a detailed derivation of the estimator is not expected. What are the key advantages of using Logit over OLS to model the MBA study decision of the employees.

(c) **(5 marks)** Discuss how you can test the null hypothesis that the ceteris paribus effect of $salary$ on the probability of MBA is equal to -0.25. Clearly state the alternative hypothesis, the test statistic, and rejection rule.

(d) **(5 marks)** Using the logit model, what is the marginal effect of one additional year of work experience on the decision to do an MBA evaluated at the mean values of the explanatory variables? Discuss the limitation(s) of the marginal effect obtained and propose an alternative that overcomes such limitation(s). **Notes:** there is no need to calculate the marginal effect using the alternative approach. You are not expected to use a calculator, only the clarity of the steps is required.

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	**2·55**	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	**2·60**	0·99534	**2·75**	0·99702	**2·90**	0·99813	**3·05**	0·99886	**3·20**	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	**2·65**	0·99598	**2·80**	0·99744	**2·95**	0·99841	**3·10**	0·99903	**3·25**	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918	**3·30**	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	**3·263**	0·9994	**3·731**	0·99990	**3·916**	0·99995
3·105	0·9991	**3·320**	0·9995	**3·759**	0·99991	**3·976**	0·99996
3·138	0·9992	**3·389**	0·9996	**3·791**	0·99992	**4·055**	0·99997
3·174	0·9993	**3·480**	0·9997	**3·826**	0·99993	**4·173**	0·99998
3·215	0·9994	**3·615**	0·9998	**3·867**	0·99994	**4·417**	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	**5·0**	1·6449	**3·0**	1·8808	**2·0**	2·0537	**1·0**	2·3263	**0·10**	3·0902
45	0·1257	**4·8**	1·6646	**2·9**	1·8957	**1·9**	2·0749	**0·9**	2·3656	**0·09**	3·1214
40	0·2533	**4·6**	1·6849	**2·8**	1·9110	**1·8**	2·0969	**0·8**	2·4089	**0·08**	3·1559
35	0·3853	**4·4**	1·7060	**2·7**	1·9268	**1·7**	2·1201	**0·7**	2·4573	**0·07**	3·1947
30	0·5244	**4·2**	1·7279	**2·6**	1·9431	**1·6**	2·1444	**0·6**	2·5121	**0·06**	3·2389
25	0·6745	**4·0**	1·7507	**2·5**	1·9600	**1·5**	2·1701	**0·5**	2·5758	**0·05**	3·2905
20	0·8416	**3·8**	1·7744	**2·4**	1·9774	**1·4**	2·1973	**0·4**	2·6521	**0·01**	3·7190
15	1·0364	**3·6**	1·7991	**2·3**	1·9954	**1·3**	2·2262	**0·3**	2·7478	**0·005**	3·8906
10	1·2816	**3·4**	1·8250	**2·2**	2·0141	**1·2**	2·2571	**0·2**	2·8782	**0·001**	4·2649
5	1·6449	**3·2**	1·8522	**2·1**	2·0335	**1·1**	2·2904	**0·1**	3·0902	**0·0005**	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

Examiners' commentaries 2024 (May)

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2023–24. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE). The examinable materials are based on the new subject guide (Schafgans and Komarova, 2022, henceforth referred to as subject guide (2022)) and the new textbook:

- Wooldridge, J. *Introductory Econometrics: A Modern Approach.* (Cengage, 2020), seventh edition [ISBN 9781337558860].

Detailed reading references in this subject guide refer to the edition of the set textbook listed above. A new edition of this textbook may have been published by the time you study this course. You can, however, use a more recent edition; use the detailed chapter and section headings to identify relevant readings. Also check the virtual learning environment (VLE) regularly for updated guidance on readings.

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2022). The subject guide is intended for you to use in conjunction with the main textbook. It will act to highlight important aspects of the content in the chapters, provide further clarifications where deemed useful and also to extend it in certain places. For the final examination, you will need to be familiar with the material in both the textbook and the subject guide. You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

General remarks

Learning outcomes

At the end of the course, and having completed the Essential reading and activities, you should be able to:

- describe and apply the classical regression model and its application to cross-section data
- describe and apply the:
  - Gauss–Markov conditions and other assumptions required in the application of the classical regression model
  - reasons for expecting violations of these assumptions in certain circumstances
  - tests for violations

  - potential remedial measures, including, where appropriate, the use of instrumental variables
- recognise and apply the advantages of logit, probit and similar models over regression analysis when fitting binary choice models
- competently use regression, logit and probit analysis to quantify economic relationships using R
- describe and explain the principles underlying the use of maximum likelihood estimation
- apply regression analysis to time-series models using stationary time series, with awareness of some of the econometric problems specific to time series applications (for example, autocorrelation) and remedial measures
- recognise the difficulties that arise in the application of regression analysis to nonstationary time series, know how to test for unit roots, and know what is meant by cointegration.

Common mistakes committed by candidates

Different from the online adopted over the past four years due to the pandemic, the examination was conducted in person. Two examination papers were administered, one for Zone A and another for Zone B. In the following discussion, we will address the common errors made by candidates in both zones, many of which reoccur on an annual basis despite changes to the subject guide. Subsequently, we will offer specific comments on each question in the examination papers for both zones.

Similar to the previous year, considerable effort was dedicated to ensuring that the examination structure and expected level of effort were consistent with those of previous years. Questions based on derivations and proofs with textbook settings that could easily be found in reference materials were limited. Greater emphasis was placed on questions that required candidates to adopt learned materials to a new setting, or evaluate and discuss the properties of estimators and test procedures. Furthermore, candidates were asked to explain their answers in a more intuitive manner and provide real-life examples.

Outstanding responses would demonstrate a discernible understanding of econometrics principles and accurate implementation in the given context. Additionally, we anticipate critical analysis of the methods or estimators employed, along with clear real-world examples. Attention to details and a thorough adherence to testing or estimation procedures are also expected. We commend the excellent performance of many candidates this year. However, a number of candidates displayed a lack of effort and preparation in their submitted attempt.

The examination made frequent use of the Gauss–Markov conditions, particularly in relation to discussing the BLUE property of an estimator or in conducting hypotheses tests. It is necessary to explicitly state the Gauss–Markov conditions rather than simply asserting their validity. A recommended approach, often followed by some candidates, is to begin by explicitly stating the Gauss–Markov conditions and then refer back to them throughout the analysis. However, many candidates neglected to mention either the Gauss–Markov conditions or the normality assumption when conducting the t test.

Candidates should realise that $\frac{1}{n}\sum_{i=1}^{n} u_i$ is not the same as $\mathrm{E}(u_i)$. So, while we typically assume $\mathrm{E}(u_i) = 0$, this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} u_i = 0$. Also, while we may be happy to assume $\mathrm{E}(x_i u_i) = 0$ (uncorrelatedness between the errors and regressors), this does not guarantee that $\frac{1}{n}\sum_{i=1}^{n} x_i u_i = 0$. Note that:

- both $\frac{1}{n}\sum_{i=1}^{n} u_i$ and $\frac{1}{n}\sum_{i=1}^{n} x_i u_i$ are random variables, which take the value 0 with probability 0 (continuous random variables)!

- $\mathrm{E}(u_i) = 0$ and $\mathrm{E}(x_i u_i) = 0$ are fixed, not stochastic!
- If we treat x_i as a random variable, we must use the Law of Iterated Expectations to obtain $\mathrm{E}(\widehat{\beta}) = \mathrm{E}(\mathrm{E}(\widehat{\beta} \mid X))$. Candidates often skipped this part in their proof.

Relatedly, a good number of candidates seem to be put off by the use of polynomial nature of the regressor in the derivation questions (such as x^{2023} instead of the usual x). A trick in this sort of question is to first think about x^{2023} as a regressor in its own right (think about it as $z = x^{2023}$) and do the proof and derivation as such. The power of stochastic regressors is that once we arrive at the final steps and given the zero conditional mean hold, we can use the Law of Iterated Expectation to relate the polynomial term to the base level. Perhaps missing the last step of the LIE impedes completing this sort of questions for some.

The proofs for unbiasedness and consistency often require clarification of the differences between sample and population moments, which seems missing for a few. In both cases, we first simplify our estimator (plug in the true model) to obtain:

$$\widehat{\beta} = \beta + \frac{\sum_{i=1}^{n}(X_i - \bar{X})u_i}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \beta + \frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2} \quad \text{with } x_i = X_i - \bar{X}.$$

- Note that this method applies for a wide range of estimator set-ups, not only for OLS. So in this case x_i can take any form of a function of a regressor.
- For *unbiasedness*, clearly indicate that you want to show that $\mathrm{E}(\widehat{\beta}) = \beta$. Unbiasedness does not follow from $\sum_{i=1}^{n} x_i u_i = 0$, instead it follows from $\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) = 0$.

  If we treat x_i as fixed:

  $$\mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) \equiv \mathrm{E}\left(\sum_{i=1}^{n} d_i u_i\right) = \sum_{i=1}^{n} d_i \mathrm{E}(u_i)$$

  and then unbiasedness follows as $\mathrm{E}(u_i) = 0$.
- For *consistency*, clearly indicate that you want to show that $\mathrm{plim}(\widehat{\beta}) = \beta$. Using the plim properties, we show:

  $$\begin{aligned}\mathrm{plim}\ \widehat{\beta} = \beta + \mathrm{plim}\left(\frac{\sum_{i=1}^{n} x_i u_i}{\sum_{i=1}^{n} x_i^2}\right) &= \beta + \frac{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i u_i\right)}{\mathrm{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)} \\ &\equiv \beta + \frac{\mathrm{plim}(\text{Sample Cov}(x,u))}{\mathrm{plim}(\text{Sample Var}(x))} \\ &= \beta + \frac{\mathrm{Cov}(x,u)}{\mathrm{Var}(x)} \quad \text{using the law of large numbers}\end{aligned}$$

  where $\mathrm{Cov}(x,u) = 0$ and $\mathrm{Var}(x) > 0$, ensuring we get consistency.
- Remember, the Law of Large Numbers guarantees that sample averages approach their corresponding population values. In order to maintain consistency in proofs, it is crucial to utilise the LLN to finalise the necessary steps. Likewise, the Law of Iterated Expectations (LIE) plays a significant role in establishing the connection between the conditional expectations and the unconditional expectations of an estimator.

A large number of candidates are not able to clearly distinguish between sample variance and covariance, and population variance and covariance.

The use of $\text{Cov}(X, Y)$ and $\text{Var}(X)$ should be restricted to describing the population covariance and variances, respectively, with definitions:

$$\text{Cov}(X, Y) = \text{E}((X - \text{E}(X))(Y - \text{E}(Y))) = \text{E}(XY) - \text{E}(X)\,\text{E}(Y)$$

and:

$$\text{Var}(X) = \text{E}((X - \text{E}(X))^2) = \text{E}(X^2) - (\text{E}(X))^2$$

(you also may denote $\text{Cov}(X, Y) = \sigma_{XY}$ and $\text{Var}(X) = \sigma_X^2$). They are typically unknown, but fixed, quantities.

The sample covariance and variance are estimators of the population covariance and variance, respectively. They are defined as:

$$\text{Sample Cov}(X, Y) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})$$

and:

$$\text{Sample Var}(X) = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

(you also may use $\widehat{\sigma}_{XY}$ and $\widehat{\sigma}_X^2$). You can compute them given the data.

With a slight abuse of notation, we often divide by n instead, which is irrelevant if we let n be large. The division by $n - 1$ is a finite sample issue only (unbiasedness).

The sample covariance and variance show up in our definition of the OLS estimator of the slope in the simple linear regression model, not the population covariance and variance, as:

$$\widehat{\beta} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} = \frac{\text{Sample Cov}(X, Y)}{\text{Sample Var}(X)} \neq \frac{\text{Cov}(X, Y)}{\text{Var}(X)}.$$

Treating them as being the same results in incorrect analyses and candidates losing significant marks.

Recall that:

$$\widehat{\beta}_{IV} = \frac{\sum_{i=1}^{n}(Z_i - \bar{Z})(Y_i - \bar{Y})}{\sum_{i=1}^{n}(Z_i - \bar{Z})(X_i - \bar{X})} = \frac{\text{SampleCov}(Z_i, Y_i)}{\text{SampleCov}(Z_i, X_i)} \neq \frac{\text{Cov}(Z_i, Y_i)}{\text{Cov}(Z_i, X_i)}.$$

It was very common to misunderstand the central limit theorem. While it provides the distribution of the *estimator* $\widehat{\beta}$, no matter what the true distribution of the errors is, it does *not* say that the distribution of the error term is approximately normal in large samples. A good number of candidates failed to connect the concepts of stationarity and weak dependence in allowing for the LLN and CLT to be used in our time series setting.

Some candidates did not provide satisfactory answers when it came to interpreting empirical results. When engaging in the interpretation of an empirical result, it is essential to analyse the significance, magnitude, and sign of the coefficients. Moreover, it is crucial to establish the principle of *ceteris paribus* or holding all other factors constant. This principle ensures that the estimation can be properly understood as the marginal change in the dependent variable in response to a change in the independent variable of interest. Furthermore, a significant number of candidates failed to address the issue of omitting important and relevant explanatory variables, as well as irrelevant explanatory variables. By neglecting the former, there exists a risk of introducing bias, known as 'omitted variable bias'. On the other hand, omitting the latter would merely result in a loss of efficiency. The concepts of omitted variable bias and the associated OVB formula are significant subjects in the curricula, particularly in the latest versions of the course. It would be advantageous for candidates to engage in more extensive revisions on these topics.

Two other sources of endogeneity, which are often overlooked by candidates, are reverse causality (simultaneity bias) and measurement errors. Only a few candidates allude to them when discussing potential biases and are rewarded.

It is also important to understand that unbiasedness or consistency is a property of an estimator (not an estimate). It may be tempting to say an 'unbiased estimate'. Technically this statement is incorrect. It is the estimator (or the rule of which estimates are calculated) is biased or inconsistent, not the estimate itself.

Numerous candidates appeared to encounter difficulties when interpreting coefficients in the context of a non-linear regression, such as log-linear/linear-log or log-log relationships. One helpful strategy for such interpretations is to bear in mind that coefficients pertaining to variables measured in logarithmic form should be interpreted in a relative sense. We have already provided necessary notes in the subject guide for the interpretation. Another trick to think about the interpretation is that changes in the variables in logs must be interpreted in a relative term (in percentage).

In the case of dummy variables or categorical variables in general as regressors, candidates often missed that the coefficients have to be interpreted relative to the base case which is the left out dummy variable (see appropriate questions in the examinations for details)

Candidates particularly struggled with the difference between percentage points and % changes when it comes to the interpretation of coefficients in binary choice models. A coefficient in such a model gives the percentage point difference in the outcome variable when the regressor changes by 1 unit. The % change is a *relative* change that depends on the baseline value of the outcome variable (for example, the realisation when a regressor equals zero in the case of a dummy variable regressor). For example, if we have initially 50% of a sample buying ecological apples, a 10% increase will mean now $(1 + 0.1) \times 50 = 55\%$ of the sample buying ecological apples. A 10% point increase instead will result in $50 + 10 = 60\%$ of the sample choosing to buy ecological apples.

When conducting hypothesis testing, it is important to ensure that you clearly indicate the null and alternative hypotheses in terms of the true parameters, denoted as β_1). Additionally, provide the test statistic in terms of the parameter estimator (t-stat, τ-stat, F-stat, LR). Specify the distribution of the test statistic under the null hypothesis and appropriate assumptions with the corresponding degrees of freedom. Clearly state the rejection rule, whether it is one-sided or two-sided, for a given significance level, typically at 5%, using appropriate critical values. Finally, provide a comprehensive interpretation of the result. It should be noted that missing any of these steps may result in a loss of marks. Many candidates fail to obtain the full marks for these questions because of the insufficient details.

Another important consideration is selecting the critical value that is appropriate for the sample size. Using a critical value of 1.96, which is suitable for a large degree of freedom, in a small sample is an error that should be avoided. Furthermore, it is crucial to understand when to use an asymptotic version of a test, such as when normality of the errors is not assumed or when robust standard errors are employed.

When using confidence intervals in a hypothesis test, candidates should decide if the hypothesised value for the population parameter of interest β lies within the confidence interval. Many candidates used the default value of zero when discussing their confidence interval regardless of the population parameter in question. Another popular mistake is to conduct a two-sided test when the statement was worded with a clear direction of the relationship.

Several candidates inaccurately defined multicollinearity as a correlation between independent variables. While we acknowledge that a high correlation among the independent variables may serve as an indicator of multicollinearity, it is incorrect to solely rely on the correlation between the regressors to define multicollinearity. Additionally, numerous candidates mistakenly claimed that multicollinearity leads to biasedness and/or inconsistency.

Many candidates made the classic error of not recognising that:

$$\mathrm{E}\left(\frac{X}{Y}\right) \neq \frac{\mathrm{E}(X)}{\mathrm{E}(Y)}$$

in general. We do have:

$$\text{plim}\left(\frac{X}{Y}\right) = \frac{\text{plim}(X)}{\text{plim}(Y)}.$$

The concepts of autocorrelation and heteroskedasticity proved to be challenging for many candidates. A considerable number of candidates struggled to adequately explain the rationale behind employing robust standard errors or asymptotic methods for testing in the presence of autocorrelation and/or heteroskedasticity. Furthermore, many candidates received deductions in their grades when discussing the properties of an estimator that can effectively address these issues. It should be noted that some candidates erroneously proposed that addressing autocorrelation would automatically resolve issues related to inconsistency and biasedness.

Testing and recognising the problem of heteroskedasticity posed challenges, particularly when the scenarios were presented in a more real-life manner. One popular wording was that the variability of outcome variables differed across different groups of one or some of the regressors. A significant number of candidates failed to identify this as the classic problem of heteroskedasticity. Among those who did recognise it, many employed the Goldfeld–Quandt version of the heteroskedasticity test. However, in the new subject guide, the Goldfeld–Quandt test is no longer introduced due to its limitations compared to alternative options such as the Breusch–Pagan and the White tests. These alternatives offer advantages such as relaxing the assumption that the error term variance to be a monotonic function of a regressor. For future iterations of the course, it is recommended that candidates thoroughly study the latest version of the subject guide, as it presents the most up-to-date approach to the discussed materials.

When discussing the significance of differentiating between trend stationary and difference stationary processes, there are two essential aspects to consider: the implications for the long-term behaviour of the processes and the statistical implications, specifically the consistency of the estimators. While the majority of candidates discussed the former aspect effectively, only a few briefly addressed the latter.

The treatment of non-stationary time series often proves a challenge to many cohorts. Candidates can often verbally explain the differences between spurious and cointegrating, however details related to its testing or consequences of the relationship often seem more difficult.

Candidates frequently received lower marks for their interpretation of the Error Correction Model (ECM), especially when discussing the characteristics of the estimator in the context of stationary or non-stationary variables. Many candidates encountered difficulties in deriving an ECM model. A significant portion of the candidates either chose to skip this question entirely or performed poorly when attempting it. In general, these questions have always been unpopular to candidates, even though the answers are often straightforward.

Some candidates appeared to be uncertain about the distinction between the Dickey–Fuller test and the Augmented Dickey–Fuller test (ADF). The discrepancy lies in the formulation of the test equation. In the augmented variant, additional lagged differences of the residuals are incorporated to ensure the absence of remaining autocorrelation in the regression of the test equation. Consequently, some individuals mistakenly utilised the ADF test equation when the DF test equation was intended. Relatedly, others forgot that even though we can use a very similar test statistic to the t statistic in the DF stationarity test (we call it τ statistic), it does not follow a t distribution but the Dickey–Fuller distribution instead.

When discussing the marginal effect after a Probit/Logit estimation, candidates often did not specify the term z in the pdf $\phi(z)$ or $\Lambda(z)$. Some used the coefficients from the LPM regressions to calculate the marginal effects after a Probit/Logit estimation. A popular, related mistake was that candidates did not realise the inherent heteroscedasticity of LPM.

A number of candidates seemed to be confused between the conditions for a valid instrument (exclusion restriction and relevance) and factors related to obtaining a more precise IV estimate (a higher sample size, a higher sample variability of X_i, a lower variance of the error term, and a stronger correlation between X_i and Z_i). More seriously, most candidates did not realise the need to use an IV estimator when the regression has a lagged dependent variable as a regressor in the

presence of autocorrelation.

Note that the IV framework works for both a cross-sectional setting and a time series analysis.

Exogeneity and endogeneity are often very simply explained as variables 'outside of the model' or 'inside the model', respectively. While this approach is not wrong; it is not a particularly clear or helpful definition in general contexts that are not related to simultaneous equations models. Defining these concepts in terms of the relationship between regressor and the error term is much clearer and naturally leads to a way to show exogeneity or endogeneity by considering the covariance between regressors and error terms.

Candidates often missed that all exogenous regressors and instruments have to be included in the first stage when performing a TSLS estimation. Some candidates use the fitted values of the endogenous variable obtained from the first stage as the dependent variable in the second stage. This mistake is a serious misunderstanding. It is fine to indicate that this fitted value can be used as an instrument of the endogenous variable, but is not really the way the second stage of TSLS is typically understood.

When asked about identification of structural form equations, candidates often compared the number of excluded regressors to the number of endogenous variables in general, without pointing out which variables are the excluded regressors and could therefore be used as an instrument. Candidates were also often unclear of the definitions of over-identification, under-identification, and exact identification. The best approach to this type of questions is to discuss the endogenous and exogenous variables, and then think about the available exogenous variables that can be used as instruments.

When deriving the reduced forms, candidates should check conditions for their existence, in particular, whether the denominator of the expression is different from zero.

Just as in previous years, many candidates did not answer all parts of a question. Make sure you read the questions properly and provide all details that are requested. Not answering a question will automatically earn you a zero. Some candidates probably did not have a good examination plan.

Key steps to improvement

Essential reading for this course includes the new subject guide (2022) and the following:

- Wooldridge, J. *Introductory Econometrics: A Modern Approach.* (Cengage, 2020) seventh edition [ISBN 9781337558860].

Please note that as long as you read the Essential reading you are then free to read around the subject area in any text, paper or online resource. You will need to support your learning by reading as widely as possible and by thinking about how these principles apply in the real world. To help you read extensively, you have free access to the virtual learning environment (VLE) and University of London Online Library (see below).

- Dougherty, C. *Introduction to econometrics.* (Oxford: Oxford University Press, 2016) 5th edition [ISBN 9780199676828].
- Gujarati, D.N. and D.C. Porter *Basic econometrics.* (McGraw–Hill, 2009, International edition) 5th edition [ISBN 9780071276252].
- Maddala, G.S. and K. Lahiri *Introduction to econometrics.* (Wiley, 2009) 4th edition [ISBN 9780470015124].
- Stock, J.H. and M.W. Watson *Introduction to econometrics.* (Pearson, 2012) 3rd edition [ISBN 9781408264331].

To understand the subject it is important to supplement the reading of the textbook Wooldridge with the subject guide (Schafgans and Komarova, 2022). The subject guide is intended for you to

use in conjunction with the main textbook. It will act to highlight important aspects of the content in the chapters, provide further clarifications where deemed useful and also to extend it in certain places. For the final examination, you will need to be familiar with the material in both the textbook and the subject guide.

The subject guide provides a range of activities that will enable you to test your understanding of the basic ideas and concepts. We want to encourage you to try the exercises you encounter throughout the material before working through the solutions provided on the virtual learning environment (VLE). With econometrics the motto has to be 'practise, practise, practise...'. It is the best way to learn the material and prepare for examinations. The course is rigorous and demanding, but the skills you will be developing will be rewarding and well recognised by future employers.

At the end of each chapter you will find an overview of the material covered, a list of key terms and concepts, and a reminder of the learning outcomes. It concludes with additional examination based exercises that will enable you to test your knowledge and understanding with worked-out solutions. Worked-out solutions for the exercises can be found on the VLE (downloadable pdf files). Candidates should utilise datasets using standard regression programmes (R in our course, but Stata or Python are other popular choices). This will help the understanding of the subject.

Examination revision strategy

Many candidates are disappointed to find that their examination performance is poorer than they expected. This may be due to a number of reasons, but one particular failing is '**question spotting**', that is, confining your examination preparation to a few questions and/or topics which have come up in past papers for the course. This can have serious consequences.

We recognise that candidates might not cover all topics in the syllabus in the same depth, but you need to be aware that examiners are free to set questions on **any aspect** of the syllabus. This means that you need to study enough of the syllabus to enable you to answer the required number of examination questions.

The syllabus can be found in the Course information sheet available on the VLE. You should read the syllabus carefully and ensure that you cover sufficient material in preparation for the examination. Examiners will vary the topics and questions from year to year and may well set questions that have not appeared in past papers. Examination papers may legitimately include questions on any topic in the syllabus. So, although past papers can be helpful during your revision, you cannot assume that topics or specific questions that have come up in past examinations will occur again.

If you rely on a question-spotting strategy, it is likely you will find yourself in difficulties when you sit the examination. We strongly advise you not to adopt this strategy.

Examiners' commentaries 2024 (May)

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2023–24. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE). The examinable materials are based on the new subject guide (Schafgans and Komarova, 2022, henceforth referred to as subject guide (2022)) and the new textbook:

- Wooldridge, J. *Introductory Econometrics: A Modern Approach.* (Cengage, 2020), seventh edition [ISBN 9781337558860].

Detailed reading references in this subject guide refer to the edition of the set textbook listed above. A new edition of this textbook may have been published by the time you study this course. You can, however, use a more recent edition; use the detailed chapter and section headings to identify relevant readings. Also check the virtual learning environment (VLE) regularly for updated guidance on readings.

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2022). The subject guide is intended for you to use in conjunction with the main textbook. It will act to highlight important aspects of the content in the chapters, provide further clarifications where deemed useful and also to extend it in certain places. For the final examination, you will need to be familiar with the material in both the textbook and the subject guide. You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone A

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from **Section A** (8 marks each) and **THREE** from **Section B** (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Section A

Answer all questions from this section. Each question carries 8 marks.

Question 1

Consider the following bivariate regression model

$$y_i = \alpha + \beta x_i + u_i, \tag{1.1}$$

for $i = 1, \ldots, n$. The regressor x_i is non-negative, *stochastic*, and has a non-zero variance, i.e. $\sum_{i=1}^{n}(x_i - \bar{x})^2 \neq 0$. The error term u_i is i.i.d. with mean zero and constant variance of σ^2, and independent of x_i. Given a random sample of observations $\{(y_i, x_i)\}_{i=1}^{n}$ from the population of interest, consider the following estimator:

$$\widetilde{\beta} = \frac{\sum_{i=1}^{n} x_i^{2023} y_i}{\sum_{i=1}^{n} x_i^{2024}}.$$

(a) Derive the condition under which $\widetilde{\beta}$ is an unbiased estimator of β. Can $\widetilde{\beta}$ be the BLUE estimator of (1.1)?

(3 marks)

(b) Discuss the difference between unbiasedness and consistency.

(2 marks)

(c) Assume that the condition in (a) is satisfied. Will this condition also ensure that $\widetilde{\beta}$ is consistent?

(3 marks)

Reading for this question

Subject guide (2022), Chapters 3.2 and 7.2.1.

Wooldridge, Section 2.2, Appendix 2A and Appendices B.4a, B.4b, B.4c.

Approaching the question

Parts (a) and (c) are a popular setup with a tweak to involve the academic year 2024 in the question. It is actually a similar setting like one in the 2023 summer examination, albeit in a simplified form. The best answer carefully derived the steps with the use of the assumptions and a clear distinction between unbiasedness and consistency in (b).

Two key steps are taking the probability limits of the estimator and invoking the LLN and LIE, which some candidates forgot to explicitly mention and lost marks on. Better performance must involve clear and correct annotations and explanations of the steps. Some candidates mistakenly in (c) claimed that unbiasedness implies consistency. Even though the conclusion is correct for (c), that line of reasoning is not and candidates are penalised for that. Some candidates forgot to refer to the LIE and LLN in their answers and were penalised.

(a) For unbiasedness, we need to show that $\mathrm{E}(\widetilde{\beta}) = \beta$. We need to take the conditional expectation on $X = \{x_i\}_{i=1}^{n}$, $\mathrm{E}[\widetilde{\beta} \,|\, X]$. The conditional expectation of $\widetilde{\beta}$ is given by:

$$\begin{aligned}
\mathrm{E}(\widetilde{\beta} \,|\, X) &= \mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i^{2023} y_i}{\sum_{i=1}^{n} x_i^{2024}} \,\Big|\, X \right) \\
&= \mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i^{2023} (\alpha + \beta x_i + u_i)}{\sum_{i=1}^{n} x_i^{2024}} \,\Big|\, X \right) \quad \text{replacing } y_i \text{ from the model} \\
&= \mathrm{E}\left[\frac{\alpha(\sum_{i=1}^{n} x_i^{2023}) + \beta(\sum_{i=1}^{n} x_i^{2024}) + \sum_{i=1}^{n} x_i^{2023} u_i}{\sum_{i=1}^{n} x_i^{2024}} \,\Big|\, X \right] \\
&= \beta + \alpha\, \mathrm{E}\left(\frac{\sum_{i=1}^{n} x_i^{2023}}{\sum_{i=1}^{n} x_i^{2024}} \,\Big|\, X \right) + \frac{\sum_{i=1}^{n} x_i^{2023} \mathrm{E}(u_i \,|\, X)}{\sum_{i=1}^{n} x_i^{2024}} \quad \text{because of i.i.d.} \\
&= \beta + \alpha \left(\frac{\sum_{i=1}^{n} x_i^{2023}}{\sum_{i=1}^{n} x_i^{2024}} \right) \quad \text{because } \mathrm{E}(u_i \,|\, X) = 0.
\end{aligned}$$

Thus, in general $\widetilde{\beta}$ is a biased estimator unless the population intercept $\alpha = 0$. The best candidates would realise that when $\sum_{i=1}^{n} x_i^{2023}$ cannot be zero because of x_i being non-negative and sample variability. $\widetilde{\beta}$ is not the BLUE because the OLS in this case would be the BLUE (as the Gauss–Markov assumptions hold).

(b) Consistency requires that as the sample size increases, the estimator gets arbitrarily close to the true parameter (or it converges in probability to the true parameter). It is a large sample property, and we need to show that $\text{plim}(\widetilde{\beta}) = \beta$. Unbiasedness is a finite sample property, which requires the estimator on average to be the true parameter when we repeatedly draw an infinite number of samples of size n.

(c) We need to look at the condition for consistency.

- When the estimator is unbiased, we must have $\alpha = 0$. From (a), we can rewrite the estimator as:

$$\widetilde{\beta} = \beta + \alpha \left(\frac{\sum_{i=1}^{n} x_i^{2023}}{\sum_{i=1}^{n} x_i^{2024}} \right) + \frac{\sum_{i=1}^{n} x_i^{2023} u_i}{\sum_{i=1}^{n} x_i^{2024}} = \beta + \frac{\sum_{i=1}^{n} x_i^{2023} u_i}{\sum_{i=1}^{n} x_i^{2024}}.$$

- The probability limit of $\widetilde{\beta}$ is:

$$\begin{aligned}
\text{plim}(\widetilde{\beta}) &= \beta + \text{plim} \left(\frac{\sum_{i=1}^{n} x_i^{2023} u_i}{\sum_{i=1}^{n} x_i^{2024}} \right) \\
&= \beta + \frac{\text{plim} \left(\frac{1}{n} \sum_{i=1}^{n} x_i^{2023} u_i \right)}{\text{plim} \left(\frac{1}{n} \sum_{i=1}^{n} x_i^{2024} \right)} \quad \text{properties of plim} \\
&= \beta + \frac{\text{E}(x_i^{2023} u_i)}{\text{E}(x_i^{2024})} \quad \text{from the Law of Large Numbers} \\
&= \beta \quad \text{as } \text{E}(xu) = 0.
\end{aligned}$$

- Therefore, when $\widetilde{\beta}$ is an unbiased estimator, it is also consistent.

Question 2

Consider the following multivariate model with two regressors X_i and Z_i:

$$Y_i = \beta_0 + \beta_1 X_i + \beta_2 Z_i + \varepsilon_i, \tag{2.1}$$

with OLS parameter estimates $\widehat{\beta}_0$, $\widehat{\beta}_1$ and $\widehat{\beta}_2$. Under the Gauss–Markov assumptions, you are given the following formula:

$$\text{Var}\left(\widehat{\beta}_2\right) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^{n} (Z_i - \bar{Z})^2} \times \frac{1}{1 - r_{XZ}^2},$$

where σ_ε^2 is the variance of ε and r_{XZ} is the sample correlation between Z and X, and n is the sample size.

(a) Discuss four factors that can improve the precision of the parameter estimate for $\widehat{\beta}_2$.

(4 marks)

(b) You are told that X_i is an irrelevant variable, that is the true value of $\beta_1 = 0$. Discuss the effect of including X_i in (2.1) on the unbiasedness and precision of the OLS estimator for β_2.

(4 marks)

Reading for this question

Subject guide (2022), Chapters 4.2 and 4.4.

Wooldridge, Section 3.3, 3.2g, 3.4 and 3.5.

Approaching the question

Part (a) is a classic question on the factors related to the size of the standard error (here, candidates are given the variance formula). Some candidates forgot to rewrite $\sum(Z_i - \bar{Z})^2$, which is the sample variance of Z to obtain the sample size n from the formula. Overall, it is another well-done question.

Part (b)'s challenge is in the recognition that omitting an irrelevant variable does not affect unbiasedness nor consistency of the other coefficient's estimator. The best candidates realised that the consequence on the precision of the estimator now lies on r_{XZ}. Only a few did this and got marks.

(a) Candidates are expected to discuss how to get more precise estimates: (i) larger sample size, n, (ii) smaller variance of ε_i, (iii) larger variance of Z, and (iv) smaller sample correlation between Z and X.

(b) Candidates are expected to explain that the OLS estimator will remain unbiased as the Gauss–Markov assumptions still hold: β_1 is still consistent with $\mathrm{E}(\varepsilon_i \mid X, Z) = 0$. To discuss precision, candidates are expected to realise that the consequence depends on whether X and Z are correlated or not. Indeed, if $r_{XZ} = 0$ then there is no loss of efficiency (precision), but if $r_{XZ} > 0$ then including the irrelevant variable will decrease the precision (increasing the variance) of the estimator.

$$\mathrm{Var}(\widehat{\beta}_2) = \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (Z_i - \bar{Z})^2} \times \frac{1}{1 - r_{XZ}^2} > \frac{\sigma_\varepsilon^2}{\sum_{i=1}^n (Z_i - \bar{Z})^2}.$$

Question 3

Consider a linear regression model:

$$\boldsymbol{y_t = \beta x_t + u_t, \quad t = 1, \ldots, T} \tag{3.1}$$

where the error terms u_t are independent of x_t and exhibit autocorrelation of an unknown form, y_t and x_t are covariance stationary and weakly dependent processes.

(a) Explain the concepts covariance stationarity and weak dependence, using appropriate technical terms. Intuitively discuss why it is desirable for the processes y_t and x_t to have these properties.

(5 marks)

(b) Discuss how you would conduct a statistical test of the hypothesis $\mathrm{H_0} : \beta = 0.5$ against the alternative $\mathrm{H_1} : \beta > 0.5$ using the OLS estimate for β. Be specific about the steps and statistic that you use.

(3 marks)

Reading for this question

Subject guide (2022), Chapters 13.2.1, 13.3.4, 14.2.1, 15.2.1.

Wooldridge, Sections 10.1–10.3, 11.1–11.2, 12.1–12.2, 12.3 (skip 12.3b).

Approaching the question

Part (a) asks a definition question of the two important assumptions in time series. The definitions are discussed well in the subject guide so, as expected, most candidates did this part well. What was less well done is the explanation on the desirability of the assumptions. Most lightly alluded to the replacement of the i.i.d. assumptions, but only a few really expanded their answers.

Part (b) was a rather popular question as most candidates recognised the need to use the Newey–West standard error (not the White robust standard errors). The then one-sided z test was well-discussed by a good number of candidates. Some lost marks for not using a one-sided hypothesis, or forgetting procedural details of the hypothesis testing, or simply the use of HAC standard errors.

(a) Candidates are expected to discuss the following:

- A process is weakly dependent if the correlation between two random temporal values of the process will disappear as the time distance between them increases: i.e. $\text{Corr}(x_t, x_{t+h}) \to 0$ as $h \to \infty$. This assumption is important because it replaces the assumption of independence of the sample distribution so that the LLN and the CLT hold: asymptotic independence of the process.
- A process is stationary if, intuitively, its stochastic properties (such as the mean and variance) and its temporal dependence structure (such as covariance over time) do not depend on time. It is a stronger version of the identically distributed assumption. A more formal definition is that for every collection of time indices, the joint distribution of the temporal values of the process is identical to the temporal values when we shift the sequence ahead h time periods. Candidates can discuss the covariance stationarity instead of this strict stationarity. It is important because we know that the nature of any correlation between adjacent terms is the same across all time periods.

(b) Because of the autocorrelation, we cannot use the usual standard errors of the OLS estimators. We must use the Newey–West (or HAC) robust standard errors. Using this HAC standard error, we can obtain the t statistic:

$$\frac{\widehat{\beta} - 0.5}{HACse(\widehat{\beta})}.$$

Using the 5% significance level, we will reject H_0 if the t statistic is smaller than -1.645, which is the one-sided asymptotic z critical value (candidates are expected to discuss this part).

Question 4

Consider the following simplified production function for a firm i:

$$\boldsymbol{output_i = \alpha_0 + \alpha_1 capital_i + \alpha_2 labour_i + \alpha_3 labour_i^2 + u_i, \quad i = 1, \ldots, n} \tag{4.1}$$

where *output* denotes the total products of firm i, and *capital* and *labour* denote the amount of capital (machines) and workers employed by the firm. Assume that we obtain a random sample of n firms and the error terms u_i are i.i.d. with zero mean and variance δ^2. We also assume that u_i and the regressors are independent.

(a) Discuss the economic and econometric rationales for including both *labour* and *labour*^{2} in this model. What are the expected signs for α_2 and α_3?

(3 marks)

(b) Derive the formula for the effect of employing on additional worker on the total output for a firm with 100 workers.

(2 marks)

(c) Discuss how you would test if the number of workers, *labour* in general, has a statistically significant effect on *output*. Be clear about your null hypothesis, the test statistic, and any additional information required.

(3 marks)

Reading for this question

Subject guide (2022), Chapters 6.2.1; 5.2.5.

Wooldridge, Section 6.2b and Appendices A.4a, A.5; Section 4.5 and Appendix B.5f.

Approaching the question

Part (a) is a short question on the economics side of the course. Candidates can use any reasonable stories to back up their claims. We also expected candidates to explain the need to use the squared term for a non-linear relationship of the variables. Most did well in this question.

Part (b) is another simple question, most got full marks for the derivation! Part (c) relies directly on the derivation in (b). An F test is needed (as in the solution below). A few candidates did not recognise this link and discussed the significance of α_2 instead, which is not correct.

(a) Candidates are expected to discuss the diminishing marginal returns to labour. For the econometric rationale, because of the economic rationale, we need to account for the non-linear relation between output and labour, hence omitting the squared term will cause omitted variable bias. We expect $\alpha_2 > 0$ and $\alpha_3 < 0$.

(b) Candidates are expected to take the first derivative with respect of labour. We will obtain the marginal effect for a 100-worker firm since:

$$\alpha_2 + 2\alpha_3 labour_i = \alpha_2 + 200\alpha_3.$$

(c) From (b), the marginal effect of labour on output is $\alpha_2 + 2\alpha_3 labour_i$. For this to be insignificant (zero) we would need $\alpha_2 = \alpha_3 = 0$. Candidates are then expected to discuss an F test using either R-squared or SSR. Candidates are expected to write down the restricted equation as well.

Question 5

Consider the following model for the gin and tonic industry in Britain:

$$\boldsymbol{Q_t = \alpha_0 + \alpha_1 P_t^g + \alpha_2 P_t^b + \alpha_3 Y_t + \alpha_4 A_t + u_t \quad \text{demand}} \tag{5.1}$$

$$\boldsymbol{Q_t = \beta_0 + \beta_1 P_t^g + \beta_2 S_t + v_t \quad \text{supply}} \tag{5.2}$$

where Q_t is the real per capita consumption of gin and tonic in year t, P_t^g is the average price of 200ml of gin and tonic, P_t^b is the average price of other alcoholic beverages, Y_t is the real per capita disposable income, A_t is the real per capita expense on related advertising, and S_t is the price of sugar, an ingredient to make gin and tonic. P_t^b, Y_t, S_t, A_t are assumed to be exogenous. The error terms u_t and v_t do not exhibit any correlation over time and have constant variances.

(a) Derive the reduced form for P_t^g.

(4 marks)

(b) Use the answer in (a) to explain why an OLS estimator for β_1 in the supply equation is likely to be biased and inconsistent.

(4 marks)

Reading for this question

Subject guide (2022), Chapters 9.2.3-4, 9.2.5.

Wooldridge, Sections 15.2, 15.1 remainder, 15.1a–15.1c.

Approaching the question

Part (a) seems tedious with a combination of reduced form equations and time series, but candidates did rather well in this question. Only a few missed some annotations or explanations of some steps. Most got the answers right. Note that the question is slightly different from the equivalent question in Zone B (a different reduced form was asked).

Part (b) follows directly from the derivation in (a) and most candidates recognised the bias which is exactly $\text{Cov}(P_t^g, u_t)$. Some opted to use intuitive explanations to explain the endogeneity – do note that it was clearly asked in the question to 'use the answer in (a)' for explanations.

(a) To derive the reduced form for P_t^g, candidates are expected to combine the demand curve and supply curve and set quantity of demand equal to quantity of supply. Hence:

$$\alpha_0 + \alpha_1 P_t^g + \alpha_2 P_t^b + \alpha_3 Y_t + \alpha_4 A_t + u_t = \beta_0 + \beta_1 P_t^g + \beta_3 S_t + v_t.$$

We then rewrite to get an expression of P_t^g in terms of the exogenous regressors and errors only:

$$\begin{aligned} P_t^g &= \frac{\beta_0 + \beta_3 S_t + v_t - \alpha_0 - \alpha_2 P_t^b - \alpha_3 Y_t - \alpha_4 A_t - u_t}{\alpha_1 - \beta_1} \\ &= \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \frac{\beta_3}{\alpha_1 - \beta_1} S_t - \frac{\alpha_2}{\alpha_1 - \beta_1} P_t^b - \frac{\alpha_3}{\alpha_1 - \beta_1} Y_t - \frac{\alpha_4}{\alpha_1 - \beta_1} A_t + \frac{v_t - u_t}{\alpha_1 - \beta_1} \\ P_t^g &= \pi_0 + \pi_1 S_t + \pi_2 P_t^b + \pi_3 Y_t + \pi_4 A_t + V_t \end{aligned}$$

where π_j are our reduced-form parameters.

(b) Our parameter estimators suffer from simultaneity bias. Q_t and P_t^g are jointly determined in the system of simultaneous equations, and this results in an endogeneity problem as $\text{Cov}(P_t^g, u_t) \neq 0$. Using the result from (a), we have:

$$\text{Cov}(P_t^g, u_t) \neq 0 = \text{Cov}(\pi_0 + \pi_1 S_t + \pi_2 P_t^b + \pi_3 Y_t + \pi_4 A_t + V_t, u_t)$$

where $V_t = \dfrac{v_t - u_t}{\alpha_1 - \beta_1}$.

As we are told that S_t, P_t^b, Y_t and A_t are exogenous, we get:

$$\text{Cov}(P_t^g, u_t) \neq 0 = \text{Cov}\left(\frac{v_t - u_t}{\alpha_1 - \beta_1}, u_t\right) = \frac{1}{\alpha_1 - \beta_1}(\sigma_{vu} - \sigma_u^2) \neq 0.$$

Section B

Answer three questions from this section. Each question carries 20 marks.

Question 6

A football team executive is evaluating the team's transfer activities. A sport economist is hired to perform the analysis. Using the data from 204 football teams from the top 10 leagues in Europe, she runs the following regression of the team final points after the 2023–2024 season, *Points* (the value ranges from 20 to 100), on the net transfer expenses (in log (USD)), log(*transfer*), the number of players playing for their national team, *NatPlayer*, and the number of injuries in the season (*Injuries*).

$$\widehat{Points}_i = \underset{(4.14)}{55.3} + \underset{(0.231)}{0.51}\log(transfer_i) + \underset{(2.56)}{2.07}\,NatPlayer_i - \underset{(1.23)}{2.01}\,Injuries_i, \qquad (6.1)$$

$$n = 204, \;\; R^2 = 0.658.$$

The robust standard errors are reported in parentheses.

(a) **What is the interpretation of the coefficient on log(*transfer*)? Are the sign and size of the coefficient as you would expect? Do you think the effect is qualitatively large or small? Explain using the available information in (6.1)**

(5 marks)

(b) Can we interpret the results in (a) causally? Explain intuitively.

(3 marks)

(c) The economist believes that the final points for the teams are likely to be more or less variable for teams with more injuries than average than for teams with much less injuries. Discuss a statistical test for this claim. What are the consequences to the regression (6.1) if she is correct?

(4 marks)

(d) The executive wants to see if the effect of the transfer activities can be improved by having more players who play for their national team. Discuss how you can test this claim. In your answer, clearly define the regression estimation, the assumptions, test statistics, or any additional information needed.

(4 marks)

(e) The executive wants to see if having more national team players can offset the effect of injuries on the final points. Discuss how you can test this claim. In your answer, clearly define the regression estimation, the assumptions, test statistics, or any additional information needed.

(4 marks)

Reading for this question

Subject guide (2022), Chapters 3.2.4, 8.2.3, 5.2.4, 4.2.2, 5.2.2-5, 6.2.4.

Wooldridge, Sections 3.2, 4.2, 6.2, 8.3, 4.4, 2.4, Appendices A1, A.3, A.4, B.5e.

Approaching the question

Part (a) is a typical question regarding the interpretation of an OLS estimation, which is a key learning outcome of the course. Most candidates correctly interpreted the estimates, with some good explanation of the expected sign. Candidates can be rather liberal in making a judgement on the qualitative magnitude of the results but the key is to get the linear-log interpretation correct. Most candidates agreed that it was a rather small effect.

Answers in part (b) were weaker when discussing the reasons for the estimates not being causal: we expected examples and discussions around OVB, with the use of the OVB formula. Some better candidates attempted stories around reverse causality, or measurement errors, but these are of less importance in this setting.

Part (c) requires knowledge of heteroskedasticity tests, either the Breusch–Pagan test or White's test. Only a very few did this well, even though the question was repeated from previous years (and part of important course material). The best candidates also noticed the use of robust standard errors under heteroskedasticity.

Part (d) requires candidates to include an interaction between *NatPlayer* and $\log(\textit{transfer})$, which most recognised. Another tricky part of the question is the use of a one-sided t test since the expected direction of the alternative hypothesis was clearly specified. Some again forgot the assumption set for the t test, or failed to recognise the use of a one-sided test.

Part (e) is a question on a multivariate hypothesis. Candidates can use either an F test or a reparameterisation of the regression. The reparameterisation approach seemed to be more popular and was done well.

(a) Candidates are expected to provide economic interpretations: for example, a 20-percent increase in the transfer expenses is associated with a $20 \times 0.5 \times 1/100$ or 0.1 points increase in the final position. This is perhaps an expected result. To evaluate the magnitude, for a team with minimal transfer, no national team players, and no injuries, the expected points is 55.3, so increasing the transfer budget by 20 percent to gain 0.1 additional points, seems to be not a good investment. It's a small effect. Candidates can have different answers on the size as long as the reasoning is clear.

(b) Candidates are expected to discuss the omitted variable bias here (the measurement error should not be a concern). Any reasonable omitted variable is accepted. For example, the quality of the team: it should be positively correlated with a larger transfer budget and a higher points tally, hence omitting this factor will create a positive OVB. Candidates can discuss using the OVB formula or any derivation.

(c) Candidates are expected to realise that we need to test for heteroskedasticity: either the Breusch–Pagan version or the White version of the test.

For the Breusch–Pagan test, we should:

- Save the OLS residuals as $\widehat{u}_i$. Compute the squared residuals $\widehat{u}_i^2$.
- Regress $\widehat{u}_i^2$ on all the regressors, with an intercept, and then compute the F test of joint significance of the regressors. We have:

$$\widehat{u}_i^2 = \delta_0 + \delta_1 \log \mathit{transfer}_i + \delta_2 \mathit{NatPlayer}_i + \delta_3 \mathit{Injuries} + v_i.$$

- The null hypothesis is: $\mathrm{H}_0 : \delta_1 = \delta_2 = \delta_3 = 0$. Candidates are expected to discuss the R-squared version or the SSR version of the F statistic. In this case, $k = 3$, $n - k - 1$.
- We do not need to assume MLR.6 in this version of the test, under the null hypothesis, the F statistic will follow the asymptotic F distribution.

For the White test, we should:

- Save the OLS residuals as $\widehat{u}_i$. Compute the squared residuals $\widehat{u}_i^2$.
- Regress $\widehat{u}_i^2$ on all the regressors, with an intercept, and the squares and interactions of all the regressors:

$$\begin{aligned}\widehat{u}_i^2 = {} & \delta_0 + \delta_1 \log \mathit{transfer}_i + \delta_2 \mathit{NatPlayer}_i + \delta_3 \mathit{Injuries}_i \\ & + \delta_4 \log \mathit{transfer}_i^2 + \cdots + \delta_7 \log \mathit{transfer}_i \times \mathit{NatPlayer}_i + \cdots + v_i.\end{aligned}$$

  The null hypothesis is: $\mathrm{H}_0 : \delta_1 = \delta_2 = \delta_3 = \cdots = \delta_9 = 0$. We can simply obtain the F test of the joint significance of the regressors. Similarly, candidates are expected to discuss the R-squared version or the SSR version of the F statistic. $k = 9$, $n - k - 1$.
- We do not need to assume MLR.6 in this version of the test, under the null hypothesis, the F statistic will follow the asymptotic F distribution.

If we reject the null hypothesis, there is evidence of heteroskedasticity. The normal standard errors are invalid. But we already use robust standard errors so there is no concern.

(d) Candidates are expected to discuss the inclusion of an interaction term $\log(\mathit{transfer}) \times \mathit{NatPlayer}$ into the regression model. The null hypothesis is $\beta_{interaction} = 0$, against a one-sided alternative hypothesis of $\beta_{interaction} > 0$. The usual t test using the estimates is expected, with the t statistic following the standard normal distribution instead of the usual t distribution. Robust standard errors should also be used. If candidates performed a two-sided test, some marks were deducted.

Candidates need to clearly indicate the use of CLM assumptions, Gauss–Markov + normality, as well as the rejection rule.

(e) Candidates are expected to describe an F test of the null: $\beta_{NatPlayer} = -\beta_{Injuries} = 0$, or $\beta_{NatPlayer} + \beta_{Injuries} = 0$ against the alternative hypothesis of non-zero. Candidates are expected to discuss either of the following approaches:

Approach 1: Perform a t test, with $n - 4 = 200$ degrees of freedom (4 parameters), we obtain the t statistic:

$$t = \frac{\widehat{\beta}_2 + \widehat{\beta}_3}{se(\widehat{\beta}_2 + \widehat{\beta}_3)}$$

which follows a t_{df} distribution under H_0. To obtain the test statistic, we must obtain $se(\widehat{\beta}_2 + \widehat{\beta}_3)$ which involves the calculation of the estimate of $\mathrm{Cov}(\widehat{\beta}_2, \widehat{\beta}_3)$.

Approach 2: Alternatively, candidates can reformulate or reparameterise the model by considering $\theta = \beta_2 + \beta_3$ and $\beta_2 = \theta - \beta_3$, and then run a new regression of:

$$Poins_i = \beta_0 + \beta_1 \log(transfer_i) + \theta NatPlayer_i + \beta_3(Injuries_i - NatPlayer_i) + u_t.$$

We can then perform a t test on the significance of θ. If the statistic exceeds the critical value from the respective distribution, we reject the claim.

Question 7

Central Banks are paying attention to economic indicators in their fight against inflation. An economist uses OLS to estimate the relationship between the monthly inflation rate inf_t and the monthly employment growth rate emp_t and obtains the following results, with the standard errors are reported in the parentheses:

$$\widehat{inf}_t = \underset{(0.014)}{0.003} + \underset{(0.051)}{1.12}\, emp_t - \underset{(0.016)}{0.907}\, emp_{t-1}, \tag{7.1}$$

$$T = 222, \;\; R^2 = 0.12.$$

(a) **What are the estimates of the short-run and long-run effect of emp_t on inf_t? Interpret the economic meaning of these estimates.**

(4 marks)

(b) **The economist believes employment growth is a leading indicator of inflation. Conduct a statistical test that a unit increase in last month's employment growth results in a one unit reduction in the current month inflation. Be specific about the null hypothesis and the assumptions for the test to be valid.**

(4 marks)

Let e_t be the OLS residuals from the regression (7.1). The economist obtains the following OLS regression of e_t on e_{t-1}:

$$e_t = \underset{(2.21)}{0.53} + \underset{(0.18)}{0.41}\, e_{t-1} + \underset{(0.118)}{0.541}\, emp_t - \underset{(0.126)}{1.07}\, emp_{t-1}, \tag{7.2}$$

$$T = 221, \;\; R^2 = 0.42.$$

(c) **The economist uses (7.2) to perform a test. What is the name of the test? Clearly indicate the null and alternative hypothesis, the test statistic, rejection rule, and the *underlying assumptions*. Be clear about the outcome of the test.**

(5 marks)

(d) **In light of the result in (c), the economist wants to see the impact of the previous inflation inf_{t-1} on the current inflation by including inf_{t-1} into regression (7.1). Briefly explain if it is a good idea.**

(2 marks)

(e) **The economist wants to test whether the long-run effect of emp_t on inf_t is statistically significant. Discuss how you can reparameterise (7.1) to make sure your regression results can provide you with appropriate information for the test. Be clear about the new regression that you would run, and how you would proceed with the test.**

(5 marks)

Reading for this question

Subject guide (2022), Chapters 14.2.1, 11, 13.2, 5.2.

Wooldridge, Sections 12.1–12.2, 12.3 (skip 12.3b), 10.5; 11.2, 12, 4.4, 10.2.

Approaching the question

Part (a) asks for the calculation of the short- and long-term effects from the regression estimates, together with their interpretations. Some candidates forgot the interpretation of 'percentage points' instead of percentage (which is incorrect). Other than that, most did well. Part (b) is a typical t test hypothesis procedure with the hypothesised value of -1. Candidates were expected to discuss or list the TS assumptions instead (TS' assumptions also work fine), with a clear note for the use of an asymptotic test. Most did well and rejected the null.

Part (c) is an autocorrelation test for the error term. Candidates were also told the expected structure is AR(1). Most candidates did well, some lost marks for missing details of the testing procedure. Others did not answer the remedial measures once we reject the null hypothesis. Some discussed the Durbin–Watson test, which is not in the current subject guide (2022). The Durbin–Watson test (discussed in Section 12.3, Wooldridge) is a finite sample test that permits testing for no autocorrelation against AR(1) dependence. It requires strict exogeneity and the full set of CLM assumptions (including normality). As these assumptions are deemed to be quite restrictive in time series models, we will not discuss this test in the course. You may also ignore the discussion of the LM form of the test for autocorrelation (called the Breusch–Godfrey test).

Part (d) is a question on the consequence of autocorrelation in the error term coupled with the inclusion of the lagged outcome. Candidates often correctly stated the inconsistency and biasedness. Weaker answers were on the explanations: either intuition or mathematical derivation would suffice. Most did reasonably well. Part (e) is a question on a multivariate hypothesis. Candidates can use either an F test or a reparameterisation of the regression, quite similar to part (e) in Question 6. Most also focused on reparameterising the regression equation. The key difference is that we must use the HAC standard errors, which some candidates missed.

(a) The short-term effect is 1.12, that is, a one-percentage point increase in the current employment growth rate leads to (or is associated with) a 1.12 percentage-point increase in the inflation rate.

The long-term effect is $1.12 - 0.907 = 0.213$, so a one-percentage point increase in the equilibrium employment growth rate leads to (or is associated with) a 0.213 percentage-point increase in the long-term inflation rate. Candidates are expected to interpret the results in percentage points.

(b) Candidates are expected to perform a test that the coefficient of emp_{t-1} is -1, that is $\beta_{emp_{t-1}} = -1$. The TS1–TS5 assumptions (particularly no autocorrelation so the standard errors are valid) are to be listed for a full mark. Assumptions TS1'–TS5' are also fine, but the test must be asymptotic. The test statistic is:

$$t = \frac{\widehat{\beta}_{emp_{t-1}} - (-1)}{se(\widehat{\beta}_{emp_{t-1}})} = \frac{-0.907 + 1}{0.016} = \frac{0.093}{0.016} = 5.81.$$

Under the assumptions (TS1–TS5, or TS1'–TS5'), we compare the t statistic against the critical value (at any conventional level), two-sided, and we can reject the null of emp_{t-1} is a leading indicator.

(c) This is a test for autocorrelation in the error term of the main equation of (7.1). The key assumption here is that the error terms ε_t follow an AR(1) structure:

$$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$$

where u_t is a white noise.

The null hypothesis is that the coefficient of e_{t-1} in (7.2) is zero. Now, if candidates assume strict exogeneity in part (b), which is absolutely fine, we do not need the inclusion of the regressors. But we do in this case, so the weak exogeneity is assumed and we should perform an asymptotic t test using the t statistic:

$$\frac{\widehat{\rho}}{se(\widehat{\rho})} = \frac{0.41}{0.18} = 2.28$$

under the asymptotic distribution $N(0,1)$.

In this case, we reject the null hypothesis of no autocorrelation, in favour of the alternative that we must be dealing with autocorrelation in the error term here.

(d) Candidates are expected to note that because of autocorrelation and the presence of a lagged outcome, the OLS estimators are not only inefficient and biased but also inconsistent. Candidates are expected to show that $\text{Cov}(inf_{t-1}, \varepsilon_t) \neq 0$ hence TS.3' is not satisfied. One can show that inf_{t-1} depends on ε_{t-1} (by lagging one period), which in turn is related to ε_t by the AR(1) structure of ε_t. So it is not a good idea.

(e) Essentially candidates are expected to test whether the coefficients for emp_t and emp_{t-1}, say β_1 and β_2 satisfy: $\theta = \beta_1 + \beta_2 = 0$. We can rewrite $\beta_1 = \theta - \beta_2$. Plug that into (7.1) and we can run the following regression:

$$\begin{aligned} inf_t &= \beta_0 + (\theta - \beta_2) emp_t + \beta_2 emp_{t-1} + \varepsilon_t \\ &= \beta_0 + \theta emp_t + \beta_2 (emp_t + emp_{t-1}) + \varepsilon_t. \end{aligned}$$

That is, we run a regression of inf_t on emp_t and $emp_t + emp_{t-1}$ and then perform a t test on the significance of θ. Note that, however, because of the presence of autocorrelation in ε_t, we must use the HAC standard errors and also the asymptotic version of the t test for the significance test (essentially an asymptotic t test). Candidates must discuss these points for full marks.

Question 8

A researcher is interested in how sales assistants in a large retailer are being promoted. Collecting a sample of 550 sales assistants who recently applied for promotion, she has data on whether the promotion is approved, *Promoted* = 1 if Yes, 0 otherwise; the amount in 100,000 of dollars sales generated by each assistant, *Sales*; the expenses in 10,000 of dollars incurred for the sales, *expenses*; years of experience, *Exp*; gender *female* = 1 if female, 0 otherwise; the highest qualification *educ*; and the age of the assistant, *age*. The researcher uses OLS and Logit to obtain the following results, with the dependent variable being *Promoted*. The heteroskedasticity robust standard errors are reported in squared brackets and the (asymptotic) standard errors are reported in parentheses.

	OLS A	Logit A	Logit B
constant	0.890 [0.068]	1.088 (0.206)	−0.244 (0.474)
Sales	0.735 [0.132]	2.029 (0.378)	2.030 (0.381)
expenses	−0.845 [0.106]	−2.344 (0.318)	−2.267 (0.321)
Exp			0.14 (0.015)
female			0.069 (0.037)
educ			0.071 (0.024)
age			−0.001 (0.004)
R^2	0.086		
$\log L$		−407.60	−399.04

(a) Interpret the coefficient on *Sales* reported in Column 'OLS A' and discuss whether the effect is statistically significant. Why is it important to use robust standard errors? You are expected to support your answer with technical derivations.

(5 marks)

(b) It is argued that using the Logit model is better than using the OLS when explaining the promotion decision *Promoted*. Discuss the benefits and drawbacks of using the Logit model when trying to explain a binary variable. In your answer explain what the linear probability model refers to.

(5 marks)

(c) Using the results in Column Logit B, indicate how you can obtain the marginal effect of one additional year of experience *Exp* on promotion. *Hints:* You are given $\Lambda(z) = \exp(z)/(1 + \exp(z))$, the logistic cumulative density function. Is the marginal effect positive or negative? Is it statistically significant? Explain.

(5 marks)

(d) The researcher wants to test whether the personal attributes (*Exp*, *female*, *educ*, and *age*) have a jointly significant effect on promotion. Using an appropriate test with the given information, indicate the hypothesis, the test statistic, and the outcome of the test.

(5 marks)

Reading for this question

Subject guide (2022), Chapter 12.2 (all).

Wooldridge, Sections 7.5, 17.1, Appendix C.4b.

Approaching the question

Part (a) asks candidates to interpret an LPM model, which was a standard task and well attempted. The less attempted part was why there is a need for robust standard errors in an LPM. Most recognised the inherent heteroskedasticity, but only a few provided the correct technical discussions. Part (b) focuses on the properties of Logit and the mechanical process behind an MLE. The best candidates provided not only the correct likelihood function for the estimation (and then the log-likelihood), but also an intuitive way to obtain the function and interpret it. Some candidates did not annotate the functions well, losing a few marks, others failed to provide a good intuitive understanding of the estimation or the computational approach to calculate the estimate in practice. The properties of the MLE were well-recited.

Part (c) asks the calculation of the marginal effect of one regressor. Most did this question well. Part (d) was another popular question. One approach is to perform a test for a joint hypothesis. Most candidates realised the need to use an LR test. Some lost marks on missing details of the testing procedures. The best candidates verbally described the process and conclusion well.

(a) An increase of 100,000 USD in sales is associated with an increase of 73.5 percentage points in the probability of being promoted, holding every other factor constant. Using the t-statistic $\widehat{\beta}_{Sales}/se(\widehat{\beta}_{Sales})$, with some necessary information, we should expect a statistically significant effect.

It is important to use robust standard errors because heteroskedasticity is inherent in the LPM. With $P(y = 1 \,|\, x)$ as the response probability, we know that $P(y = 0 \,|\, x) = 1 - P(y = 1 \,|\, x)$. To derive the variance of the error term we can examine (candidates are expected to explain the steps).

$$\mathrm{Var}(u \,|\, x) = \mathrm{Var}(y \,|\, x) = \mathrm{E}(y^2 \,|\, x) - (\mathrm{E}(y \,|\, x))^2 = P(y = 1 \,|\, x)(1 - P(y = 1 \,|\, x)).$$

This conditional variance is dependent on the regressors x, which will violate the homoskedasticity assumption. Thus the use of robust standard errors by default is appropriate.

(b) The Logit model has three main advantages over the LPM: (i) predicted probabilities are restricted to lie in $[0, 1]$, (ii) MLE is (asymptotically) efficient whereas OLS (LPM) will be inefficient, and (iii) the MLE automatically deals with heteroskedasticity.

The main drawback(s) of the Logit model relative to the linear probability model is that the coefficients cannot be directly interpreted as the marginal effects of the regressor(s) of interest, we need to compute predicted probabilities using the Logit specification and it is also computationally more complicated.

(c) In this setting, candidates are expected to first calculate the z-values by plugging the estimates from Logit B and the average sales and expenses to obtain the value of z. The marginal effect of a one-unit change is given by $\Lambda(z) \times 0.14$.

Alternatively, candidates can indicate that they could compute:

$$\Lambda(z + 0.14) - \Lambda(z).$$

(d) Candidates are expected to use the likelihood ratio test to test the joint significance of the *personal* variables.

We are asked to test $\mathrm{H}_0 : \beta_3 = \beta_4 = \beta_5 = \beta_6 = 0$ against the alternative $\mathrm{H}_1 : \exists\ i \in [3, 4, 5, 6]$, $\beta_i \neq 0$. We need to use the LR test for this, which requires us to estimate the restricted Logit (where only *Sales* and *expenses* are used, i.e. Logit A) and unrestricted Logit (where all variables are included, i.e. Logit B). The test statistic is given by:

$$LR = 2(\log L^U - \log L^R) \overset{a}{\sim} \chi^2_4 \text{ under } \mathrm{H}_0.$$

The χ^2 has four degrees of freedom because we are testing 4 restrictions. We should reject at the 5% level of significance if $LR > \chi^2_{0.05,\,4} = 9.488$. Above we note that our test statistic value is $2 \times (-399.04 - (-407.60)) = 17.12$ which exceeds the critical value, hence we reject the joint significance of non-price variables. The test uses the property that our MLE estimators are asymptotically normally distributed.

Question 9

We are interested in the relationship between the quarterly price level CPI_t and the quarterly employment level Emp_t in the following ADL(1, 1) model:

$$CPI_t = \beta_0 + \beta_1 CPI_{t-1} + \beta_2 Emp_t + \beta_3 Emp_{t-1} + u_t, \quad t = 1, \ldots, T, \tag{9.1}$$

where Emp_t and CPI_t are both $I(1)$ variables, and $|\beta_1| < 1$. Assume also that the error term has a zero mean and is independent of CPI_{t-1}, Emp_t and Emp_{t-1}.

(a) Discuss how you can test whether the above specification is spurious or cointegrating. Clearly indicate the test equation, test statistic, the the rejection rule. In your answer, clearly distinguish the concepts and explain why it is so important that we have to test for cointegration.

(10 marks)

(b) A colleague ascertains that there exists an Error Correction Model (ECM) that describes the short-run and long-run dynamics between CPI_t and Emp_t. Derive the Error Correction Model and interpret the various components of the ECM. Discuss how does this claim relate to your answer in (a). *Hints:* Subtract CPI_{t-1} from both sides of (9.1) to obtain the following version:

$$\Delta CPI_t = \delta_0 + \delta_1 \Delta Emp_t + \gamma s_{t-1} + \epsilon_t,$$

with s_{t-1} is the error correction term.

(6 marks)

(c) Discuss the Engle–Granger two-step procedure to estimate the ECM obtained in (c).

(4 marks)

Reading for this question

Subject guide (2022), Chapters 13.2, 15.2.

Wooldridge, Sections 10.2, 10.4, 18.2–4.

Approaching the question

Part (a) was a short and standard question about the concepts of spurious regression and cointegration, which should be recited well by candidates. The statistical properties should also be explained well. The test for cointegration was also well-addressed, with details and procedures provided. It's a 10-mark question so candidates seemed to pay much attention to it (even though only a very few chose this question).

Part (b) perhaps caught most off-guard as candidates must recognise this as a derivation of the ECM. The question is actually artificially complex and may have dissuaded candidates from attempting the question. It was actually a simple definition and discussion of the ECM. Marks were given to any attempt to derive the models and the standard interpretation. Part (c) was a conceptual question regarding the theoretical advantages and drawbacks of the two-step Engle–Granger estimation of the ECM.

(a) Candidates are expected to discuss the two concepts, and then talk about the ADF test of the residual for cointegration (in general it's called the Engle–Granger test).

Spurious regression refers to the scenario where two unrelated series give the impression that they are related, purely because of the fact that both are trending. A classic example is where CPI_t and Emp_t are both unrelated $I(1)$ series (random walk or random walk with drift). Running a regression of CPI_t on Emp_t (spurious) has a large R^2 and a seemingly significant slope parameter. The problem is that here the error term is $I(1)$ as well and statistical inference is invalid. The regression is meaningless!

Another example of a spurious regression is where two completely unrelated trend-stationary series may give the impression that they are related purely because of the omission of the 'time' variable (OVB). Once account is taken of the deterministic trend, the regression will be able to detect that there is no relation between X and Y. By omitting the time trend in the regression, the parameter on X will capture their joint trending behaviour.

If, on the other hand, we have cointegration than we do have a long-run relationship between our $I(1)$ variables Emp_t and CPI_t. We have a linear combination of $I(1)$ variables that becomes stationary $I(0)$. For example, if Emp_t and CPI_t are $I(1)$ processes, but $CPI_t - \beta Emp_t$ is stationary. Then we can say X and Y are cointegrated, meaning there is a long-term relationship between these two. In that case, there will be an ECM representation as in part (b).

To test for cointegration, we can test for a unit root in the residual. We run the above equation to obtain the residuals:

$$\widehat{u}_t = CPI_t - \widehat{\beta}_0 - \widehat{\beta}_1 CPI_{t-1} - \widehat{\beta}_2 Emp_t - \widehat{\beta}_3 Emp_{t-1}.$$

Next we use an ADF specification, given by:

$$\Delta\widehat{u}_t = \rho\widehat{u}_{t-1} + \delta\Delta\widehat{u}_{t-1} + e_t$$

where by having added sufficient lagged differences we ensure that e_t does not exhibit autocorrelation. We now test $\mathrm{H}_0 : \rho = 0$ (spurious, nonstationary) against $\mathrm{H}_1 : \rho < 0$ (cointegrating, stationary). The test statistic is $\widehat{\rho}/se(\widehat{\rho})$, and we reject at the 5% level of significance if this value is smaller than the critical value of the Engle–Granger test (so not exactly equal to those in the DF or ADF tests).

(b) If there is an ECM model for X and Y, then X and Y must be cointegrated and the test conducted in part (b) should have rejected the null. Candidates are expected to use the hints to derive the ECM, which can be written as:

$$\Delta CPI_t = (\beta_1 - 1)\left(CPI_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta_1}\,Emp_{t-1}\right) + \beta_2\Delta Emp_t + u_t$$

where:

- $CPI_{t-1} - \dfrac{\beta_0}{1-\beta_1} - \dfrac{\beta_2+\beta_3}{1-\beta_1} Emp_{t-1}$ is the error correction term, measuring the disequilibrium at time $t-1$
- $(\beta_1 - 1)$ indicates how fast the two processes are coming back to the steady state
- β_2 is the direct effect a change in *Emp* has on *CPI*.

(c) The error correction model can be estimated by two steps.

Firstly, we estimate the cointegating relationship between X and Y by OLS to obtain (need the residuals!):

$$\widehat{diseq}_{t-1} = CPI_{t-1} - \widehat{\gamma}_1 - \widehat{\gamma} Emp_{t-1}.$$

Next, using $\widehat{diseq}_{t-1}$ we run OLS on the error correction equation:

$$\Delta CPI_t = (\beta_1 - 1)\widehat{diseq}_{t-1} + \beta_2 \Delta Emp_t + u_t.$$

Examiners' commentaries 2024 (May)

EC2020 Elements of econometrics

Important note

This commentary reflects the examination and assessment arrangements for this course in the academic year 2023–24. The format and structure of the examination may change in future years, and any such changes will be publicised on the virtual learning environment (VLE). The examinable materials are based on the new subject guide (Schafgans and Komarova, 2022, henceforth referred to as subject guide (2022)) and the new textbook:

- Wooldridge, J. *Introductory Econometrics: A Modern Approach.* (Cengage, 2020), seventh edition [ISBN 9781337558860].

Detailed reading references in this subject guide refer to the edition of the set textbook listed above. A new edition of this textbook may have been published by the time you study this course. You can, however, use a more recent edition; use the detailed chapter and section headings to identify relevant readings. Also check the virtual learning environment (VLE) regularly for updated guidance on readings.

Information about the subject guide and the Essential reading references

Unless otherwise stated, all cross-references will be to the latest version of the subject guide (2022). The subject guide is intended for you to use in conjunction with the main textbook. It will act to highlight important aspects of the content in the chapters, provide further clarifications where deemed useful and also to extend it in certain places. For the final examination, you will need to be familiar with the material in both the textbook and the subject guide. You should always attempt to use the most recent edition of any Essential reading textbook, even if the commentary and/or online reading list and/or subject guide refer to an earlier edition. If different editions of Essential reading are listed, please check the VLE for reading supplements – if none are available, please use the contents list and index of the new edition to find the relevant section.

Comments on specific questions – Zone B

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from **Section A** (8 marks each) and **THREE** from **Section B** (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Section A

Answer all questions from this section. Each question carries 8 marks.

Question 1

Consider the following regression model of a time series Y_t:

$$Y_t = \beta Y_{t-1} + u_t, \quad t = 1, \ldots, T \tag{1.1}$$

where the error terms u_t have a zero mean.

(a) Discuss whether the OLS estimator $\widehat{\beta}$ is an unbiased estimator for β. Justify your answers with suitable technical derivations. *Note:* You are not expected to derive the OLS estimator.

(3 marks)

You are now told that u_t follows an MA(1) process.

(b) Discuss the consequences of making the error terms an MA(1) process. What is the name given to this problem?

(2 marks)

(c) Discuss whether the OLS estimator $\widehat{\beta}$ is a consistent estimator for β. Justify your answers with suitable technical derivations.

(3 marks)

Reading for this question

Subject guide (2022), Chapters 13.2.4, 13.2.5, 11.2.1.

Wooldridge, Sections 11.1, 11.2, 16.1.

Approaching the question

Part (a) is a standard discussion on the violation of the strict exogeneity assumption. Most realised this and discussed well. Parts (b) and (c) further assume the error term is autocorrelated and asks about its consequences. It's another standard discussion and candidates often recited the Gauss–Markov assumptions (for time series) to discuss the inconsistency. Candidates did well in naming the problem 'autocorrelation' and the best ones specified also the inclusion of a lagged outcome variable.

(a) It is a classical ADL model with a lagged dependent variable. Because u_t is correlated with Y_t which is a future value of Y_{t-1}, the strict exogeneity assumption does not hold, hence $\widehat{\beta}$ is not unbiased. The best candidates showed this by showing $\mathrm{E}(\widehat{\beta}) \neq \beta$.

(b) Candidates are expected to name and explain the 'autocorrelation' concept, in which a time series is correlated with some of its values in the past. This autocorrelation problem causes the standard errors to be invalid. Also, as shown in part (b), the estimator is no longer consistent (and already biased).

(c) Now u_t follows an MA(1) process (say $u_t = e_t + \theta e_{t-1}$ where e_t is white noise), candidates are expected to show that u_t and u_{t-1} are correlated (through e_{t-1}). Candidates are expected to show that now u_t and Y_{t-1} are correlated (through u_{t-1}). As such, the OLS estimator is not even consistent. The best candidates showed this by showing $\operatorname{plim} \widehat{\beta} = \beta$.

Question 2

Consider the following bivariate regression model

$$y_i = \alpha + \beta z_i + u_i, \tag{2.1}$$

for $i = 1, \ldots, n$. The regressor z_i is non-negative, *stochastic*, and has a non-zero variance, i.e. $\sum_{i=1}^{n}(z_i - \bar{z})^2 \neq 0$. The error term u_i is i.i.d. with mean zero and constant variance of σ^2, and independent of z_i. Given a random sample of observations $\{(y_i, z_i)\}_{i=1}^{n}$ from the population of interest, consider the following estimator:

$$\widetilde{\beta} = \frac{\sum_{i=1}^{n} z_i^{2023} y_i}{\sum_{i=1}^{n} z_i^{2024}}.$$

(a) Derive the condition under which $\widetilde{\beta}$ is a consistent estimator of β.

(3 marks)

(b) Let us assume that the condition in (a) is satisfied so that $\widetilde{\beta}$ is consistent. Will this condition also ensure that $\widetilde{\beta}$ is unbiased? In your answer, discuss the difference between unbiasedness and consistency.

(5 marks)

Reading for this question

Subject guide (2022), Chapters 3.2 and 7.2.1.

Wooldridge, Section 2.2, Appendix 2A and Appendices B.4a, B.4b, B.4c.

Approaching the question

Part (a) is a popular setup with a tweak to involve the academic year 2024 in the question. It is actually a similar setting like one in the 2023 Summer examination, albeit in a simplified form. The best answers carefully derived the steps with the use of the assumptions and a clear distinction between unbiasedness and consistency in part (b).

Two key steps are taking the probability limits of the estimator and invoking the LLN and LIE, which some candidates forgot to explicitly mention and lost marks on. Better performance must involve clear and correct annotations and explanations of the steps. Some candidates mistakenly in claimed that unbiasedness implies consistency. Even though the conclusion is correct for (b), that line of reasoning is not and candidates are penalised for that. Some candidates forgot to refer to the LIE and LLN in their answers and were penalised.

(a) For consistency, we need to show that $\text{plim}(\widetilde{\beta}) = \beta$. We first simplify $\widetilde{\beta}$ as follows:

$$\begin{aligned}
\widetilde{\beta} &= \frac{\sum_{i=1}^n z_i^{2023} y_i}{\sum_{i=1}^n z_i^{2024}} \\
&= \frac{\sum_{i=1}^n z_i^{2023}(\alpha + \beta z_i + u_i)}{\sum_{i=1}^n z_i^{2024}} \quad \text{replacing } y_i \text{ from the model} \\
&= \beta + \alpha \left(\frac{\sum_{i=1}^n z_i^{2023}}{\sum_{i=1}^n z_i^{2024}} \right) + \frac{\sum_{i=1}^n z_i^{2023} u_i}{\sum_{i=1}^n z_i^{2024}}.
\end{aligned}$$

Taking the plim, we get:

$$\begin{aligned}
\text{plim}(\widetilde{\beta}) &= \beta + \text{plim } \alpha \left(\frac{\sum_{i=1}^n z_i^{2023}}{\sum_{i=1}^n z_i^{2024}} \right) + \text{plim} \left(\frac{\sum_{i=1}^n z_i^{2023} u_i}{\sum_{i=1}^n z_i^{2024}} \right) \\
&= \beta + \alpha \left(\frac{\text{plim } 1/n \sum_{i=1}^n z_i^{2023}}{\text{plim } 1/n \sum_{i=1}^n z_i^{2024}} \right) + \frac{\text{plim} \left(\frac{1}{n} \sum_{i=1}^n z_i^{2023} u_i \right)}{\text{plim} \left(\frac{1}{n} \sum_{i=1}^n z_i^{2024} \right)} \quad \text{properties of plim} \\
&= \beta + \alpha \frac{\text{E}(z^{2023})}{\text{E}(z^{2024})} + \frac{\text{E}(z_i^{2023} u_i)}{\text{E}(z_i^{2024})} \quad \text{from the Law of Large Numbers} \\
&= \beta + \alpha \frac{\text{E}(z^{2023})}{\text{E}(z^{2024})} \quad \text{as } \text{E}(z u_i) = 0.
\end{aligned}$$

At this point, for the estimator to be consistent, we would need $\alpha = 0$ because $\text{E}(z^{2023}) = 0$ is only when all of the zs are zero, contradicting the assumption. So, the condition is $\alpha = 0$.

(b) Consistency requires that as the sample size increases, the estimator gets arbitrarily close to the true parameter (or it converges in probability to the true parameter). It is a large sample property and we need to show that $\text{plim}(\widetilde{\beta}) = \beta$. Unbiasedness is a finite sample property, which requires the estimator on average is the true parameter when we repeatedly draw an infinite number of samples of size n.

When the estimator is consistent, we must have $\alpha = 0$. From (a), unbiasedness, we need to show that $\mathrm{E}(\widetilde{\beta}) = \beta$. The expectation of $\widetilde{\beta}$ is given by:

$$\begin{aligned}
\mathrm{E}(\widetilde{\beta}) &= \mathrm{E}\left(\frac{\sum_{i=1}^{n} z_i^{2023} y_i}{\sum_{i=1}^{n} z_i^{2024}}\right) \\
&= \mathrm{E}\left(\frac{\sum_{i=1}^{n} z_i^{2023}(\alpha + \beta z_i + u_i)}{\sum_{i=1}^{n} z_i^{2024}}\right) \\
&= \mathrm{E}\left[\frac{\alpha(\sum_{i=1}^{n} z_i^{2023}) + \beta(\sum_{i=1}^{n} z_i^{2024}) + \sum_{i=1}^{n} z_i^{2023} u_i}{\sum_{i=1}^{n} z_i^{2024}}\right] \\
&= \beta + \alpha\,\mathrm{E}\left(\frac{\sum_{i=1}^{n} z_i^{2023}}{\sum_{i=1}^{n} z_i^{2024}}\right) + \frac{\sum_{i=1}^{n} z_i^{2023}\mathrm{E}(u_i)}{\sum_{i=1}^{n} z_i^{2024}} \\
&= \beta + \alpha\left(\frac{\sum_{i=1}^{n} z_i^{2023}}{\sum_{i=1}^{n} z_i^{2024}}\right) \quad \text{because } \mathrm{E}(z_i u_i) = 0.
\end{aligned}$$

Thus, in general $\widetilde{\beta}$ is a biased estimator unless the population intercept $\alpha = 0$. The best candidates would realise that when $\sum_{i=1}^{n} z_i^{2023}$ cannot be zero because of z_i being non-negative and sample variability. Therefore, when $\widetilde{\beta}$ is a consistent estimator, it is also unbiased.

Question 3

Consider the following model for the energy drink industry in Britain:

$$\boldsymbol{Q_t = \alpha_0 + \alpha_1 P_t^e + \alpha_2 P_t^c + \alpha_3 Y_t + u_t \quad \text{demand}} \tag{3.1}$$

$$\boldsymbol{Q_t = \beta_0 + \beta_1 P_t^e + \beta_2 P_t^S + v_t \quad \text{supply}} \tag{3.2}$$

where Q_t is the real per capita consumption of energy drink in year t, P_t^e is the average price of a can of energy drink, P_t^c is the average price of other caffeine drink, Y_t is the real per capita disposable income, and P_t^S is the price of sugar, an ingredient to make energy drink. P_t^c, Y_t, P_t^S are assumed to be exogenous. The error terms u_t and v_t do not exhibit any correlation over time and have constant variances.

(a) Derive the reduced form for Q_t.

(4 marks)

(b) Use the answer in (a) to explain why an OLS estimator for α_1 in the demand equation is likely to be biased and inconsistent.

(4 marks)

Reading for this question

Subject guide (2022), Chapters 9.2.3–4, 9.2.5.

Wooldridge, Sections 15.2, 15.1 remainder, 15.1a–15.1c.

Approaching the question

Part (a) seems tedious with a combination of reduced form equations and time series, but candidates rather did well in this question. Only a few missed some annotations or explanations of some steps. Most got the answers right. Note that the question is slightly different from the equivalent question in Zone A (a different reduced form was asked). Part (b) follows directly from the derivation in (a) and most candidates recognised the bias which is exactly $\mathrm{Cov}(P_t^g, u_t)$. Some opted to use intuitive explanations to explain the endogeneity – do note that it was clearly asked in the question to 'use the answer in (a)' for explanations.

(a) To derive the reduced form for Q_t, we can first obtain the reduced form for P_t^e (not necessarily, but it follows from the setup), candidates are expected to combine the demand curve and supply curve and set quantity of demand equal to quantity of supply:

$$\alpha_0 + \alpha_1 P_t^e + \alpha_2 P_t^c + \alpha_3 Y_t + u_t = \beta_0 + \beta_1 P_t^e + \beta_3 P_t^S + v_t.$$

We then rewrite to get an expression of P_t^e in terms of the exogenous regressors and errors only:

$$\begin{aligned} P_t^e &= \frac{\beta_0 + \beta_3 P_t^S + v_t - \alpha_0 - \alpha_2 P_t^c - \alpha_3 Y_t - u_t}{\alpha_1 - \beta_1} \\ &= \frac{\beta_0 - \alpha_0}{\alpha_1 - \beta_1} + \frac{\beta_3}{\alpha_1 - \beta_1} P_t^S - \frac{\alpha_2}{\alpha_1 - \beta_1} P_t^c - \frac{\alpha_3}{\alpha_1 - \beta_1} Y_t + \frac{v_t - u_t}{\alpha_1 - \beta_1} \\ P_t^e &= \pi_0 + \pi_1 P_t^S + \pi_2 P_t^c + \pi_3 Y_t + V_t \end{aligned}$$

where π_j are our reduced-form parameters. From here, we can easily obtain the reduced form for Q_t as:

$$\begin{aligned} Q_t &= \beta_0 + \beta_1 P_t^e + \beta_2 P_t^S + v_t \\ &= \pi_0 + \beta_0 + (\pi_1 + \beta_1) P_t^e + (\beta_2 + \pi_2) P_t^S + \pi_3 Y_t + V_t + v_t. \end{aligned}$$

(b) Candidates would need also to obtain the reduced form of P_t^e for this question. Our parameter estimators suffer from simultaneity bias. Q_t and P_t^e are jointly determined in the system of simultaneous equations, and this results in an endogeneity problem as $\text{Cov}(P_t^e, u_t) \neq 0$. Using the result from (a), we have:

$$\text{Cov}(P_t^e, u_t) \neq 0 = \text{Cov}(\pi_0 + \pi_1 P_t^S + \pi_2 P_t^c + \pi_3 Y_t + \pi_4 A_t + V_t, u_t)$$

where $V_t = (v_t - u_t)/(\alpha_1 - \beta_1)$. As we are told that P_t^S, P_t^c, Y_t and A_t are exogenous, we get:

$$\text{Cov}(P_t^e, u_t) \neq 0 = \text{Cov}\left(\frac{v_t - u_t}{\alpha_1 - \beta_1}, u_t\right) = \frac{1}{\alpha_1 - \beta_1}(\sigma_{vu} - \sigma_u^2) \neq 0.$$

Question 4

Consider a linear regression model:

$$\boldsymbol{y_t = \beta_1 x_t + \beta_2 x_{t-1} + u_t, \quad t = 1, \ldots, T} \tag{4.1}$$

where the error terms u_t are independent of x_t and x_{t-1}, y_t and x_t are covariance stationary and weakly dependent. You are told that y_t and y_{t-1} are correlated so that u_t exhibits autocorrelation.

(a) Discuss the concepts of weak dependence and autocorrelation. Discuss the consequences of these having weak dependence and autocorrelation on the OLS estimation of (4.1).

(4 marks)

(b) Discuss a suitable test to test for autocorrelation. Be clear about the null hypothesis and the assumptions needed. What should we do if we reject the test's null hypothesis?

(4 marks)

Reading for this question

Subject guide (2022), Chapters 13.2.1, 13.3.4, 14.2.1, 15.2.1.

Wooldridge, Sections 10.1–10.3, 11.1–11.2, 12.1–12.2, 12.3 (skip 12.3b).

Approaching the question

Part (a) asks a definition question of the two important assumptions in time series. The definitions are discussed well in the subject guide so, as expected, most candidates did this part well. What was less well done was the explanation on the desirability of the assumptions. Most lightly alluded to the replacement of the i.i.d. assumptions, but only a few really expanded their answers. Part (b) was a rather popular question as most candidates recognised the need to use the Newey–West standard error (not the White robust standard errors). The then one-sided z test was well discussed by a good number of candidates. Some lost marks for not using a one-sided hypothesis, or forgetting procedural details of the hypothesis testing, or simply the use of the HAC standard errors.

(a) A process is weakly dependent if the correlation between two random temporal values of the process will disappear as the time distance between them increases: $\text{Corr}(x_t, x_{t+h}) \to 0$ as $h \to \infty$. This assumption is important because together with stationarity it replaces the assumption of independence of the sample distribution so that the LLN and the CLT hold: asymptotic independence of the process. That is, it allows for correlation within the time series, but the correlation will disappear if the time distance gets large. Autocorrelation simply means there is a correlation between two time periods $\text{Corr}(x_t, x_{t+h})$, so in a sense, weakly dependent series exhibit autocorrelation (not vice versa). Autocorrelation in the error terms violates TS5 (or 5'). So the standard errors are invalid.

(b) Candidates are expected to perform an autocorrelation test. Assume that u_t can be expressed as follows:

$$u_t = \rho u_{t-1} + v_t, \quad \text{with } |\rho| < 1$$

and v_t is a white noise generated randomly from a normal distribution with mean zero and constant variance. Since we only assume weak exogeneity here, it's safe to include all the regressors in the autocorrelation test. To test for autocorrelation, the null hypothesis is $\text{H}_0 : \rho = 0$ against $\text{H}_1 : \rho \neq 0$. The testing equation is obtained from using the OLS residuals $\widehat{u}_t$ of (4.1): $\widehat{u}_t = y_t - (\widehat{\beta}_1 x_t + \widehat{\beta}_2 x_{t-1})$. The testing equation is:

$$\widehat{u}_t = \gamma_0 + \gamma_1 x_t + \gamma_2 x_{t-1} + \rho \widehat{\varepsilon}_{t-1} + u_t.$$

We must use the asymptotic t statistic $\widehat{\rho}/se(\widehat{\rho})$ for a large sample. At the 5% level of significance, we should reject if $|\widehat{\rho}/se(\widehat{\rho})| > 1.96$, which follows a $N(0,1)$ distribution. (Candidates can instead use the LM test statistic $(T-1)R^2 \sim \chi^2_1$ – only mentioned but not discussed in the subject guide.)

If we reject the null hypothesis, we should be concerned about the presence of autocorrelation. We can use the HAC robust standard errors (not robust alone) to fix the standard errors.

Question 5

Consider the following OLS estimation results. The outcome variable is the average math scores of a primary school, math_i (ranging from 0 to 10), in a school i in Britain. The regressors are $\log expend_i$, the spending per student of a school in log(USD); $\log enroll_i$, the number of registered students in logs, and *poverty*, the percentage of students live in poverty during the school year. The OLS results with and without $poverty_i$ as an explanatory variable are given in (5.1) and (5.2), with the standard errors are in parentheses:

$$\widehat{\text{math}}_i = \underset{(26.72)}{-69.24} + \underset{(3.30)}{11.13} \log(expend_i) + \underset{(0.615)}{0.022} \log(enroll_i), \tag{5.1}$$

$$N = 428, \; R^2 = 0.0297$$

$$\widehat{\text{math}}_i = \underset{(24.99)}{-23.14} + \underset{(3.04)}{7.75} \log(expend_i) - \underset{(0.58)}{1.26} \log(enroll_i) - \underset{(0.036)}{0.324} poverty_i. \tag{5.2}$$

$$N = 428, \; R^2 = 0.1893$$

(a) Interpret the coefficient on $poverty_i$ in (5.2). Is the sign of the coefficient estimate as what you would expect? Explain.

(3 marks)

(b) Carefully explain why the effect of expenditures on math_i is lower in Equation (5.2) where $poverty_i$ is included than in Equation (5.1) where it is excluded. What is the name given to this difference?

(5 marks)

Reading for this question

Subject guide (2022), Chapters 6.2.1; 5.2.5.

Wooldridge, Section 6.2b and Appendices A.4a, A.5; Section 4.5 and Appendix B.5f.

Approaching the question

Part (a) is a short question on the economics side of the course. Candidates are expected to correctly use the linear-log interpretation of the functional form. Most did very well. Part (b) is an interesting application of the OVB formula and candidates are expected to observe the difference between the two estimates and manage to attribute that to the omission of *poverty*. Only a few made relevant discussions around this topic. As similar to previous years, it seems candidates are well prepared for theoretical (technical) questions that involved derivation and proof. Questions surrounding the application of the OVB concepts seem to be less prepared.

(a) A one percentage increase in the number of students in poverty is associated with a deduction of 0.324 points in the maths score, holding other factors constant and on average. This is as expected.

(b) We are looking at the OVB here, by omitting the poverty we have a positive OVB on expenditure. Candidates are expected to use the OVB formula to show that poverty is negatively correlated with the outcome (math) and also negatively correlated with expenditure (more poverty is associated with a lower district tax revenue, hence lower expenditure). The OVB formula gives us a positive OVB.

Section B

Answer three questions from this section. Each question carries 20 marks.

Question 6

Central Banks are paying attention to economic indicators in their fight against inflation. An economist is interested in the following ADL(1, 1) relationship between the monthly inflation rate $infl_t$ and the monthly employment growth rate emp_t

$$infl_t = \beta_0 + \beta_1 infl_{t-1} + \beta_2 emp_t + \beta_3 emp_{t-1} + u_t, \quad |\beta_1| < 1 \tag{6.1}$$

where emp_t and $infl_t$ are covariance stationary variables. The error terms u_t are an i.i.d. $(0, \sigma_u^2)$ and independent of $infl_{t-1}, infl_{t-2}, infl_{t-3} \ldots$. We also assume that u_t is independent of emp_t for all t.

(a) What does it mean by 'covariance stationary'? How is it different from 'difference stationary' and 'trend stationary'? Briefly explain the concepts and why it is important to distinguish them.

(4 marks)

(b) Provide the short-run and long-run effect of emp on Y. Provide an economic interpretation of the effects.

(5 marks)

(c) Discuss the statistical properties your OLS estimators for the ADL(1, 1) parameters. In particular: (i) are the estimators unbiased and consistent?, and (ii) should we use robust standard errors? Support your answers with appropriate technical derivations.

(5 marks)

(d) Show that when we exclude the relevant variable $\boldsymbol{infl_{t-1}}$ from (6.1), the new model will suffer from autocorrelation. Explain your answer. Propose a solution to the problem(s) resulted from the exclusion $\boldsymbol{infl_{t-1}}$.

Hint: You are expected to reformulate your model as

$$\boldsymbol{infl_t = \beta_0 + \beta_2 emp_t + \beta_3 emp_{t-1} + v_t,}$$

where the new error term is $\boldsymbol{v_t = \beta_1 infl_{t-1} + u_t}$.

(6 marks)

Reading for this question

Subject guide (2022), Chapters 14.2.1, 11, 13.2, 5.2; 11.2.

Wooldrige, Sections 12.1-12.2, 12.3 (skip 12.3b), 10.5; 11.2, 12, 4.4, 10.2; 16.1-3

Approaching the question

Part (a) asks to discuss a familiar concept of 'covariance stationary' and to distinguish between trend and difference stationary. Most candidates did well as this question has appeared in past examination papers very often. Part (b) asks for the calculation and interpretation of the long- and short-term effects. Most did well. Part (c) is a question on the consequence of autocorrelation in the error term coupled with the inclusion of the lagged outcome. Candidates often correctly stated the inconsistency and biasedness. Weaker answers were on the explanations: either intuition or mathematical derivation would suffice. Most did reasonably well. Part (d) is a rather interesting question. Candidates are expected to use an IV to solve for this problem as using HAC standard errors is not sufficient (the OVB issue is not addressed). Surprisingly even though it is expected that top candidates would discuss the IV, many candidates did this very well.

(a) Covariance stationarity is when the first two moments of a series do not change over time ($\mathrm{E}(x_t)$ and $\mathrm{Var}(x_t)$) and the covariance between two values x_t and x_{t+h} depends only on the distance in time, h, but not in time. We need this assumption, together with weak dependence, to replace the i.i.d. assumption and to use LLT and CLN.

Difference stationarity is completely different, which means the series is not stationary and can only be made stationary by taking the first difference. Trend stationarity means the series can be made stationary by taking away the trend.

It is important to distinguish the two because the former is highly persistent while the latter is weakly dependent. For the former, shocks or policy changes will have a lasting/permanent effect, while for the latter the effect is transitory. Therefore it is important to distinguish the two to avoid estimating meaningless and spurious relationships due to highly persistent behaviour.

(b) Candidates are expected to clearly show the work for the following points.

The short-run effect is β_2 (gives the immediate effect that *emp* has on Y). To economically interpret: a one percentage point increase in employment growth is associated with a β_2 percentage point change in the inflation rate. The key is the percentage point interpretation.

The long-run effect is $(\beta_2 + \beta_3)/(1 - \beta_1)$. We obtain the latter by solving, at equilibrium:

$$infl^e = \beta_0 + \beta_1 infl^e + \beta_2 emp^e + \beta_3 emp^e$$

so that:

$$infl^e = \frac{\beta_0}{1-\beta_1} + \underbrace{\frac{\beta_2+\beta_3}{1-\beta_1}}_{\text{long-run effect}} emp^e.$$

The short-run effect gives the immediate effect that a change (potentially just temporary) in *emp* has on *infl* (temporary change), whereas the long-run effect indicates what in equilibrium the effect would be of a permanent change in *emp*.

(c) This question is a classic discussion on the properties of estimation with lagged dependent variables.

Because of the presence of $infl_{t-1}$ in the above model, we cannot establish unbiasedness because the strict exogeneity assumption is violated (u_t and $infl_t$ are correlated).

The assumptions for consistency are satisfied because for this we only need to ensure that u_t is uncorrelated with the regressors $infl_{t-1}$, emp_t and emp_{t-1}. We are given that the *emp* variables are independent of the errors so they are uncorrelated. As long as there is no autocorrelation, we will also not have any correlation between $infl_{t-1}$ and u_t.

We do not need to use robust standard errors as there is no heteroskedasticity or autocorrelation.

(d) Candidates must annotate the derivations clearly for full marks. The key is to show that $\text{Cov}(v_{t+1}, v_t) \neq 0$.

Using our definition of v_t and v_{t+1}, we note:

$$\begin{aligned}\text{Cov}(v_{t+1}, v_t) &= \text{Cov}(\beta_1 infl_t + u_{t+1}, \beta_1 infl_{t-1} + u_t)\\ &= \beta_1^2 \text{Cov}(infl_t, infl_{t-1}) + \beta_1 \text{Cov}(infl_t, u_t)\\ &\quad + \beta_1 \text{Cov}(u_{t+1}, infl_{t-1}) + \text{Cov}(u_{t+1}, u_t).\end{aligned}$$

Because u_t is white noise (i.i.d. errors unrelated to anything in the past) $\text{Cov}(u_{t+1}, u_t) = 0$ and $\text{Cov}(u_{t+1}, infl_{t-1}) = 0$.

To show that $\text{Cov}(v_{t+1}, v_t) \neq 0$ it is sufficient to point out that $\text{Cov}(infl_t, u_t) \neq 0$. This follows quickly when noting that as our model is given by:

$$infl_t = \beta_0 + \beta_1 infl_{t-1} + \beta_2 emp_t + \beta_3 emp_{t-1} + u_t, \quad |\beta_1| < 1$$

which shows a direct relation between $infl_t$ and u_t, and hence $\text{Cov}(infl_t, u_t) \neq 0$. (Clearly $\text{Cov}(infl_t, infl_{t-1}) \neq 0$ as $\beta_1 \neq 0$.)

Now there are two problems with the exclusion: autocorrelation, which can be solved by using HAC, but more importantly, OVB due to the exclusion of an important variable. The most potential solution is to use an IV (for example, further lags of $infl_t$ for the emp_t variables, or some other external variables (supply-side shocks)). This is a difficult part and reserved for top candidates.

Question 7

A consultancy team executive is evaluating the students' performance in an online course. An education economist is hired to perform the analysis. Using the data from surveying 894 students participating in the course, she runs the following regression of the team final exam performance in the 2023–2024 academic year (marks range from 0 – 100), $Grade_i$, on the self-reported number of self-study hours per week in logs, log *selfstudy*; whether the student also worked a job, $work_i = 1$ if Yes, and 0 otherwise; and the number of concurrent courses the student took (*othercourses*).

$$\widehat{Grade}_i = \underset{(7.14)}{48.3} + \underset{(0.23)}{0.51} \log(selfstudy_i) - \underset{(4.56)}{12.07} work_i - \underset{(2.23)}{3.01}\, othercourses_i, \qquad ((7.1)$$

$$n = 894, \ R^2 = 0.345.$$

The robust standard errors are reported in parentheses.

(a) **What is the interpretation of the coefficient on log(*selfstudy*)? Are the sign and size of the coefficient as you would expect? Do you think the effect is qualitatively large or small? Explain using the available information in (7.1).**

(4 marks)

(b) **A colleague believes the regression (7.1) may suffer from heteroskedasticity. Discuss a statistical test for this claim. What are the consequences to the regression (7.1) if she is correct?**

(4 marks)

(c) **A colleague wants to see if the effect of self-study on the course performance is lower when a student also had to work during the study. Discuss how you can test this claim. In your answer, clearly define the regression estimation, the assumptions, test statistics, or any additional information needed.**

(4 marks)

(d) **Another colleague says that since the coefficient of $othercourses_i$ is insignificant, including the variable is harmful and the team should remove the variable from the regression. Critically discuss the suggestion.**

(4 marks)

(e) **The colleague also suggests that students tend to misreport the self-study hours to the survey. Intuitively discuss the consequences on the interpretation of the regression results in (7.1) if the colleague is correct.**

(4 marks)

Reading for this question

Subject guide (2022), Chapters 3.2.4, 8.2.3, 5.2.4, 4.2.2, 5.2.2–5, 6.2.4, 10.2.

Wooldridge, Sections 3.2, 6.2, 8.3, 4.4, Appendices A1–4, B.5e; 9.4a, 9.4b and 15.4.

Approaching the question

Part (a) is a typical question regarding the interpretation of an OLS estimation, which is a key learning outcome of the course. Most candidates correctly interpreted the estimates, with some good explanation of the expected signs. Candidates can be rather liberal in making a judgement on the qualitative magnitude of the results but the key is to get the linear-log interpretation correct. Most candidates agreed that it was a rather small effect. Part (b) requires the knowledge of heteroskedasticity tests, either the Breusch–Pagan or White test. Only a very few did this well, even though the question was repeated from previous years (and part of important course material). The best candidates also noticed the use of robust standard errors under heteroskedasticity. Part (c) requires candidates to include an interaction between *work* and log(*selfstudy*), which most recognised. Another tricky part of the question is the use of a one-sided t test since we clearly specified the expected direction of the null hypothesis. Some again forgot the assumption set for the t test, or failed to recognise the use of a one-sided test.

Part (d) is a question on the interpretation of including an irrelevant variable, evidenced by the seemingly insignificant coefficient. Only the best candidates discussed the nuisances behind the inclusion of such variables by using the danger of leaving out a potentially important variable (OVB). The conclusion is to keep the variable in (in the risk of OVB and doing so does not harm the estimator). Only a very few candidates did this. Part (e) is rather challenging and only the very best candidates recognised the errors in variable issue. In fact, measurement error and ECM seem to be the challenging topics for candidates year after year.

(a) Candidates are expected to provide economic interpretations: for example, a 10 percent increase in the self-study hours is associated with a 10×0.5 or 5 points increase in the final examination. This is perhaps an expected result. To evaluate the magnitude, for a student with minimal self-study hours, no work, and no other courses, they will earn around 48 marks, so the effect seems quite small. Candidates can have different answers on the size as long as the reasoning is clear.

(b) Candidates are expected discuss clearly the technical definition of heteroskedasticity (the error term variance depends on some or all of the regressors), and also the intuition: the GPA would be more or less variable for students who do more self-study than those who do less self-study.

Candidates are expected to explain the test for heteroskedasticity: either the Breusch–Pagan version or the White version of the test.

For the Breusch–Pagan test, we should:

- Save the OLS residuals as $\widehat{u}_i$. Compute the squared residuals $\widehat{u}_i^2$.
- Regress $\widehat{u}_i^2$ on all the regressors, with an intercept, and then compute the F test of joint significance of the regressors. We have:

$$\widehat{u}_i^2 = \delta_0 + \delta_1 \log selfstudy_i + \delta_2 work_i + \delta_3 othercourses + v_i.$$

- The null hypothesis is: $\mathrm{H}_0 : \delta_1 = \delta_2 = \delta_3 = 0$. Candidates are expected to discuss the R-squared version or the SSR version of the F statistic. In this case, $k = 3$, $n - k - 1$.
- We do not need to assume MLR.6 in this version of the test, under the null hypothesis, the F statistic will follow the asymptotic F distribution.

For the White test, we should:

- Save the OLS residuals as $\widehat{u}_i$. Compute the squared residuals $\widehat{u}_i^2$.
- Regress $\widehat{u}_i^2$ on all the regressors, with an intercept, and the squares and interactions of all the regressors:

$$\begin{aligned}\widehat{u}_i^2 = \delta_0 &+ \delta_1 \log selfstudy_i + \delta_2 work_i + \delta_3 othercourses_i \\ &+ \delta_4 \log selfstudy_i^2 + \cdots + \delta_7 \log selfstudy_i \times work_i + \cdots + v_i.\end{aligned}$$

  The null hypothesis is: $\mathrm{H}_0 : \delta_1 = \delta_2 = \delta_3 = \cdots = \delta_9 = 0$. We can simply obtain the F test of the joint significance of the regressors. Similarly, candidates are expected to discuss the R-squared version or the SSR version of the F statistic. $k = 9$, $n - k - 1$.
- We do not need to assume MLR.6 in this version of the test, under the null hypothesis, the F statistic will follow the asymptotic F distribution.

If we reject the null hypothesis, there is evidence of heteroskedasticity. The normal standard errors are invalid. But we already use robust standard errors so there is no concern.

(c) Candidates are expected to discuss the inclusion of an interaction term $\log(selfstudy) \times work$ into the regression model. The null hypothesis is $\beta_{interaction} = 0$, against a one-sided alternative hypothesis of $\beta_{interaction} > 0$. The usual t test using the estimates is expected, with the t statistic following the standard normal distribution instead of the usual t distribution. Robust standard errors should also be used.

Candidates need to clearly indicate the use of CLM assumptions, Gauss–Markov + normality, as well as the rejection rule.

(d) The colleague is right that the coefficient is insignificant (candidates are expected to briefly mention the t statistic or the large standard error here). However, including an irrelevant variable is not particularly harmful, it also increases the standard errors of the other estimates, which in this case are still statistically significant at the 5% level. So the effect of such inclusion is minimal. Intuitively, excluding the variable is potentially more harmful as having other courses may actually be an important factor in student performance. So we should keep that variable in.

(e) Candidates are expected to discuss the errors-in-variables situation, at least, if the self-reported numbers are randomly misreported: the errors may come from students misremembering their self-study in a random manner. If that's the case, we will have attenuation bias, which biases the estimate towards zero. As the expected true effect is positive, in this case, we are having an underestimate – so the true effect of self-study should be larger than what we observe here. The best candidates would comment that if students tend to overreport their self-study hours, then actually the effect we observe is an overestimate of the true effect!

Question 8

We are interested in the relationship between the quarterly price level CPI_t and the quarterly employment level Emp_t in the following model:

$$CPI_t = \alpha_0 + \alpha_1 Emp_t + u_t, \quad t = 1, \ldots, T, \tag{8.1}$$

where Emp_t and CPI_t are both $I(1)$ variables. Assume that the error terms exhibit autocorrelation and follow an AR(1) structure as:

$$u_t = \rho u_{t-1} + v_t, \quad |\rho| < 1, \tag{8.2}$$

where v_t is an i.i.d. innovation with zero mean and constant variance, and is independent of e_{t-1}.

(a) **Explain what it means to say CPI_t is an $I(1)$ variable. Discuss how you would test for it. In your answer, specify the test equation, test statistic, and assumptions.**

(5 marks)

(b) **A colleague suggests to test for co-integration before proceeding with analysing the data. Discuss the key idea of how you can test for evidence of a cointegrating relationship between CPI and Emp. In your answer, make clear the distinction between cointegration and spurious relationship. *Hints:* we can use (8.2) to get rid of the autocorrelation in (8.1).**

(6 marks)

(c) **A colleague ascertains that there exists an Error Correction Model (ECM) that describes the short-run and long-run dynamics between CPI_t and Emp_t. Derive the Error Correction Model and interpret the various components of the ECM. Discuss how does this claim relate to your answer in (b). *Hints:* You would need to rely on the model in (a) after getting rid of autocorrelation to obtain the following version:**

$$\Delta CPI_t = \delta_0 + \delta_1 \Delta Emp_t + \gamma s_{t-1} + \epsilon_t, \tag{8.3}$$

where s_{t-1} is the error correction term.

(5 marks)

(d) **Using the notations in (8.3), discuss what happens to CPI_t in equilibrium in the following situations:**

i. **$s_{t-1} > 0$?**

ii. **$\gamma > 0$?**

(4 marks)

Reading for this question

Subject guide (2022), Chapters 13.2, 15.2.

Wooldridge, Sections 10.2, 10.4, 18.2–4.

Approaching the question

Part (a) was a short and standard question about the concepts of nonstationarity (here unit root of order 1). Most could discuss well the concept but rather lacked the details of the ADF (or DF) test. The key part we looked for is the testing equation of the DF (or ADF test) and the components of the testing procedure (using the τ statistics instead of the normal t statistics). The few candidates who attempted did this well.

Part (b) was a short and standard question about the concepts of spurious regression and cointegration, which should be recited well by candidates. The statistical properties should also

be explained well. The test for cointegration was well addressed, with details and procedures well provided. It's a 6-mark question with a hint so candidates seemed to pay much attention to it (even though only a very few chose this question).

Parts (c) and (d) perhaps caught most off-guard as candidates must recognise this as a derivation of the ECM. The question is actually artificially complex and may have dissuaded candidates from attempting the question. It was actually a simple definition and discussion of the ECM. Marks were given to any attempt to derive the models and the standard interpretation. Part (d) is a simple interpretation of the ECM components, which should be present after deriving the model.

(a) When we can make a process stationary by differencing a process once, we say that we have a process that is integrated of order 1, or $I(1)$. *CPI* is a classic example of an $I(1)$ process. To perform a test, we need a DF or ADF test in this case. For example, assume *CPI* follows an AR(1) structure, which is reasonable. The testing equation of the DF test requires the estimation of:

$$\Delta CPI_t = \pi_0 + \pi_1 CPI_{t-1} + e_t.$$

We will test $\mathrm{H}_0 : \pi_1 = 0$ against $\mathrm{H}_1 : \pi_1 < 0$, using the τ statistic: $\widehat{\pi}_1/se(\widehat{\pi}_1)$. At the 5% significance level, we need to reject if the τ statistic is larger than the DF critical value (not the t), which is about -2.86.

(b) The key to the rest of the answers is to realise that we can lag one period of (8.1) to obtain CPI_{t-1}, u_{t-1} and then substitute (8.2) into (8.1) to obtain the following general form:

$$CPI_t = \beta_0 + \beta_1 CPI_{t-1} + \beta_2 Emp_t + \beta_3 Emp_{t-1} + u_t, \quad t = 1, \ldots, T.$$

Candidates are then expected to discuss the two concepts, and then talk about the ADF test for cointegration (named the Engle–Granger test).

Spurious regression refers to the scenario where two unrelated series give the impression that they are related, purely because of the fact that both are trending. A classic example is where CPI_t and Emp_t are both unrelated $I(1)$ series (random walk or random walk with drift). Running a regression of CPI_t on Emp_t (spurious) has a large R^2 and a seemingly significant slope parameter. The problem is that here the error term is $I(1)$ as well and statistical inference is invalid. The regression is meaningless!

Another example of a spurious regression is where two completely unrelated trend stationary series may give the impression that they are related purely because of the omission of the 'time' variable (OVB). Once account is taken of the deterministic trend, the regression will be able to detect that there is no relation between X and Y. By omitting the time trend in the regression, the parameter on X will capture their joint trending behaviour.

If, on the other hand, we have cointegration than we do have a long-run relationship between our $I(1)$ variables Emp_t and CPI_t. We have a linear combination of $I(1)$ variables that becomes stationary $I(0)$. For example, if Emp_t and CPI_t are $I(1)$ processes, but $CPI_t - \beta Emp_t$ is stationary. Then we can say X and Y are cointegrated, meaning there is a long-term relationship between these two. In that case, there will be an ECM representation as in part (c).

To test for cointegration, we can test for a unit root in the residual. Candidates are not required to discuss the following details, only briefly.

We run the above equation to obtain the residuals:

$$\widehat{u}_t = CPI_t - \widehat{\beta}_0 - \widehat{\beta}_1 CPI_{t-1} - \widehat{\beta}_2 Emp_t - \widehat{\beta}_3 Emp_{t-1}.$$

Next we use an ADF specification, given by:

$$\Delta\widehat{u}_t = \rho\widehat{u}_{t-1} + \delta\Delta\widehat{u}_{t-1} + e_t$$

where by having added sufficient lagged differences we ensure that e_t does not exhibit autocorrelation. We now test $\mathrm{H}_0 : \rho = 0$ (spurious, nonstationary) against $\mathrm{H}_1 : \rho < 0$ (cointegrating, stationary). The test statistic is $\widehat{\rho}/se(\widehat{\rho})$, and we reject at the 5% level of significance if the test statistic value is smaller than -2.86 (the actual Engle–Granger critical value is slightly different from this DF number).

(c) If there is an ECM model for X and Y, then X and Y must be cointegrated and the test conducted in part (b) should have rejected the null. Candidates are expected to use the hints to derive the ECM, which can be written as:

$$\Delta CPI_t = (\beta_1 - 1)\left(CPI_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta_1} Emp_{t-1}\right) + \beta_2 \Delta Emp_t + u_t$$

where:

- $CPI_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta_1} Emp_{t-1}$ is the error correction term, measuring the disequilibrium at time $t-1$
- $(\beta_1 - 1)$ indicates how fast the two processes are coming back to the steady state
- β_2 is the direct effect a change in Emp has on CPI.

(d) This is a classic interpretation question: if $s_{t-1} > 0$, it means in equilibrium CPI_t is too high so that in the next period γs_{t-1} will push CPI_t back to the equilibrium level. With a smaller γ, the adjustment pushing CPI back to equilibrium is stronger (faster adjustment); if $\gamma > 0$, we will see a movement further away from the equilibrium level, which is non-sensical. If $\gamma = 0$ then essentially we do not have the error correction term and there is no ECM or no cointegration.

Question 9

An education economist is interested in the decision to undertake an MBA course of a group of 456 employees with at least 2 years of experience. Among the 456 employees, 25% of whom belong to a minority ethnicity group, $minority_i = 1$; 126 went on to do an MBA within the next two years $MBA_i = 1$; and the other 300 employees did not have a plan to do so $MBA_i = 0$. The annual salary, measured in £1,000, of the employees pre-MBA is denoted as $salary_i$. The employees have an average salary of £43,000 and on average 4.5 years of work experience, denoted as exp_i. Using OLS and Logit, the economist obtains the following results.

$$\widehat{MBA}_i = \Lambda\left(\underset{(0.011)}{0.32} - \underset{(0.032)}{0.123}\, salary_i + \underset{(0.354)}{0.840}\, minority_i - \underset{(0.344)}{1.195}\, exp_i\right) \quad \text{(Logit)}$$

$$\widehat{MBA}_i = \underset{(0.04)}{0.113} - \underset{(0.042)}{0.223}\, salary_i + \underset{(1.124)}{2.62}\, minority_i - \underset{(1.144)}{3.295}\, exp_i \quad \text{(OLS)}$$

where $\Lambda(z) = \exp(z)/(1+\exp(z))$ is the logistic function. The usual standard errors for the Logit model and the robust standard errors for the LPM are reported in parentheses.

(a) Interpret the coefficient on $salary_i$ reported in the OLS estimation and discuss whether the effect is statistically significant. Why is it important to use robust standard errors? You are expected to support your answer with appropriate technical derivations.

(5 marks)

(b) Describe the estimator underlying the Logit model. Note: clearly indicate the likelihood function used, but a detailed derivation of the estimator is not expected. What are the key advantages of using Logit over OLS to model the MBA study decision of the employees.

(5 marks)

(c) Discuss how you can test the null hypothesis that the *ceteris paribus* effect of *salary* on the probability of MBA is equal to -0.25. Clearly state the alternative hypothesis, the test statistic, and rejection rule.

(5 marks)

(d) Using the logit model, what is the marginal effect of one additional year of work experience on the decision to do an MBA evaluated at the mean values of the explanatory variables? Discuss the limitation(s) of the marginal effect obtained and propose an alternative that overcomes such limitation(s). Notes: there is no need to calculate the marginal effect using the alternative approach. You are not expected to use a calculator, only the clarity of the steps is required.

(5 marks)

Reading for this question

Subject guide (2022), Chapter 12.2 (all).

Wooldridge, Sections 7.5, 17.1, Appendix C.4b.

Approaching the question

Part (a) asks candidates to interpret an LPM model, which was a standard task and well attempted. The less well-attempted part was why there is a need for robust standard errors in an LPM. Most recognised the inherent heteroskedasticity, but only a few provided the correct technical discussions. Part (b) focuses on the properties of Logit and the mechanical process behind an MLE. The best candidates provided not only the correct likelihood function for the estimation (and then the log-likelihood), but also an intuitive way to obtain the function and interpret it. Some candidates did not annotate the functions well, losing a few marks, others failed to provide a good intuitive understanding of the estimation or the computational approach to calculate the estimate in practice. The properties of the MLE were well recited.

Part (c) asks to conduct hypothesis testing on a coefficient. Candidates could describe a Wald test, but the most popular one should simply be a z test on the coefficient of the LPM. The best candidates recognised this. In part (d) most recognised the need to calculate the average person value and plug that into the pdf of the logistic function. The discussion on the shortfalls of this method was also well done (as it's rather standard). Candidates proposed the alternatives and were well rewarded.

(a) An increase of £10,000 in salary is associated with a decrease of 2.23 percentage points in the probability of doing an MBA holding every other factor constant. Using the t statistic $\widehat{\beta}_{salary}/se(\widehat{\beta}_{salary})$, with some necessary information, we should expect a statistically significant effect. This seems to be a small effect!

It is important to use robust standard errors because heteroskedasticity is inherent in the LPM. With $P(y = 1 \,|\, x)$ as the response probability, we know that $P(y = 0 \,|\, x) = 1 - P(y = 1 \,|\, x)$. To derive the variance of the error term we can examine (candidates are expected to explain the steps):

$$\mathrm{Var}(u \,|\, x) = \mathrm{Var}(y \,|\, x) = \mathrm{E}(y^2 \,|\, x) - (\mathrm{E}(y \,|\, x))^2 = P(y = 1 \,|\, x)(1 - P(y = 1 \,|\, x)).$$

This conditional variance is dependent on the regressors x, which will violate the homoskedasticity assumption. Thus the use of robust standard errors by default is appropriate.

(b) The probability of doing an MBA for each individual is:

$$p_i(MBA_i = 1) = \widehat{MBA}_i = \Lambda(\alpha_0 + \alpha_1 salary_i + \alpha_2 minority_i + \alpha_3 exp_i)$$

and the probability of not doing an MBA is, where $\Lambda(z) = \exp(z)/(1 + \exp(z))$:

$$p_i(MBA_i = 0) = 1 - \Lambda(\alpha_0 + \alpha_1 salary_i + \alpha_2 minority_i + \alpha_3 exp_i).$$

The likelihood function for this sample is:

$$\begin{aligned} L(\alpha_1, \alpha_2, \alpha_3 \,|\, \text{data}) &= \prod_{MBA} p_i(MBA = 1) \prod_{non-MBA} p_i(MBA = 0) \\ &= \prod_{MBA} \Lambda(\alpha_0 + \alpha_1 salary_i + \alpha_2 minority_i + \alpha_3 exp_i) MBA_i \\ &\quad \times \prod_{non-MBA} [1 - \Lambda(\alpha_0 + \alpha_1 salary_i + \alpha_2 minority_i + \alpha_3 exp_i)](1 - MBA_i). \end{aligned}$$

The key problem with the LPM is that the predicted probabilities of MBA obtained from the LPM can be larger than 1 or smaller than 0, which is non-sensical. This problem is very much likely to happen, given the estimates from the LPM model.

The logit model addresses this concern as the probability is defined using the CDF, $\Lambda(z)$, which ensures the predicted probabilities are contained between 0 and 1.

The LPM implicitly assumes that the marginal effect of minority (or any variables) is constant over its entire range. Even though it is easier and straightforward to interpret, this assumption might not be intuitively appealing. Since the sample is large, we do not have to worry about the error term of the LPM not having a normal distribution.

Candidates can also discuss the easy implementation of LPM compared to Logit (which is no longer a concern due to computability nowadays, or LPM will inherently be heteroskedastic).

(c) We need to do a hypothesis test on whether the coefficient on $salary_i$ is equal to -0.25. Since estimates from the LPM and the Logit estimations have different interpretations, we must use the estimate from the LPM for this test.

The null hypothesis is $\mathrm{H}_0 : \alpha_1 = -0.25$, the alternative hypothesis is $\mathrm{H}_1 : \alpha_1 \neq -0.25$. The appropriate test for the LPM estimate is the z test (asymptotic t test). The z test is:

$$\frac{\widehat{\alpha}_1 + 0.25}{se(\widehat{\alpha}_1)} = 0.64.$$

Under the null hypothesis, the z test has a $N(0, 1)$ distribution. At the 5% significance level and for a two-tailed test, the critical value is 1.96. If the test statistic is smaller than the critical value, we cannot reject the null hypothesis. We cannot reject the hypothesis that the effect of salary of the probability of MBA is -0.25.

(d) The marginal effect of interest will tell us how the probability of MBA $p_i(MBA_i = 1)$ for the employees changes for an increase of 1 year of experience for average employees with average attributes. The marginal effect at the mean values of the explanatory variables is:

$$\frac{\partial p_i(MBA_i = 1)}{\partial exp} = -1.195 \times f(z)$$

where $f(z) = \exp^{-z} / (1 + \exp^{-z})^2$ is the PDF of the logistic function, and:

$$\begin{aligned} z &= 0.32 - 0.123\overline{salary} + 0.840\overline{minority} - 1.195\overline{exp} \\ &= 0.32 - 0.123 \times 43 + 0.84 \times 0.25 - 1.195 \times 4.5. \end{aligned}$$

Plug z into the pdf of the logistic function $f(z)$, we can obtain the marginal effect.

Alternatively, candidates can show the difference in probabilities of the two employees at the averages that differ by one year of work experience. Candidates are not expected to calculate the numbers.

The problem with this calculation is that its interpretation is non-nonsensical as there cannot be such a person that is of 25% minority. The interpretation for an individual with average characteristics might not be of any practical use. The alternative is to look at the average marginal effect over all individuals (here the sample size is $n = 1234$), which can be calculated as follows:

$$\frac{1}{n} \sum_{i=1}^{n} \frac{\partial p_i(MBA_i = 1)}{\partial exp} = \frac{1}{n} \sum_{i=1}^{n} \Lambda(z_i) \times [1 - \Lambda(z)] \times 0.240 = \frac{1}{n} \sum_{i=1}^{n} f(z_i) \times 0.840$$

where $z_i = \alpha_0 + \alpha_1 salary_i + \alpha_2 minority_i + \alpha_3 exp_i$, and $f(z)$ is the pdf.

Examiners' commentaries November 2025

EC2020 Elements of Econometrics

What the examiners are looking for

This commentary reflects the November examination and assessment arrangements for this course in the academic year 2024-25. This exam was designed to test both technical skills and economic interpretation. While many students performed well, common mistakes persist year after year. Below is a summary the popular mistakes:

- Confusing sample averages with expectations: $\frac{1}{n}\sum u_i \neq E(u_i)$.
- Misinterpreting unbiasedness and consistency proofs; forgetting to use Law of Iterated Expectations.
- Mixing up sample vs. population variance/covariance in formulas.
- Incorrectly assuming CLT implies normality of errors.
- Oversimplifying multicollinearity and its effects (it does not cause bias).
- Forgetting that $E(X/Y) \neq E(X)/E(Y)$ in general.
- Weak interpretation of coefficients, especially in log models and binary choice models (percentage points vs. % changes).
- Confusion between Dickey-Fuller and Augmented Dickey-Fuller tests.
- Poor discussion of omitted variable bias and irrelevant variables.
- Misunderstanding IV conditions and TSLS steps; not checking identification.
- Incomplete hypothesis testing steps (missing distribution, critical values, or interpretation).

In general, we are looking for the following points in a strong answer:

- Clear understanding of assumptions (e.g., strict exogeneity, Gauss-Markov conditions).
- Correct distinction between sample and population concepts.
- Logical derivations supported by interpretation and intuition.
- Accurate application of hypothesis tests (null/alternative, test statistic, distribution, rejection rule), and an economic interpretation of the decision.
- Ability to link economic and econometric theory to practice (including proofs of unbiased/consistency, and derivations of estimators) and interpret empirical results.

Suggestions for Improvement

- **Interpretation matters:** Many answers focused only on derivations. Always explain what your result means economically. For example, after computing an estimator, discuss its implications for bias, consistency, and real-world interpretation as these are often asked in the questions.
- **Master assumptions and conditions:** Review the concepts of strict (and weak) exogeneity, Gauss-Markov assumptions and extension, stationarity, weak dependence and other fundamental assumptions. Be able to explain why these assumptions matter and what happens when they fail (e.g., lagged dependent variables violate strict exogeneity).
- **Distinguish sample vs. population concepts:** Clearly separate sample moments (random variables) from population moments (fixed quantities). Misusing these concepts leads to incorrect arguments about unbiasedness and consistency.
- **Practice unbiasedness and consistency proofs:** When proving unbiasedness, condition correctly if regressors are stochastic and use the Law of Iterated Expectations. For consistency, apply probability limits and the Law of Large Numbers to show convergence to population parameters.
- **Strengthen hypothesis testing skills:** Always include all steps: state null and alternative hypotheses in terms of true parameters, specify the test statistic and its distribution under the null, choose the correct critical value for the sample size, and provide a clear interpretation. Avoid common mistakes such as conducting two-sided tests when the question specifies a direction.
- **Work on advanced topics:** Practice writing likelihood functions clearly and understand their role in estimation. Review ECM models and stationarity tests (Dickey-Fuller vs. Augmented Dickey-Fuller) and know how to interpret convergence rates and other parameters from an ECM model.
- **Interpret coefficients correctly:** Pay attention to log-linear and log-log models, percentage changes vs. percentage points, and interpretation in binary choice models. For dummy variables, always interpret coefficients relative to the omitted reference category.
- **Understand IV and TSLS procedures:** Know the conditions for valid instruments (relevance and exclusion restriction) and the steps in two-stage least squares. Avoid confusion between identification conditions and precision factors. Always include all exogenous variables in the first stage.
- **Address omitted variable bias and efficiency:** Omitting relevant variables causes bias; omitting irrelevant ones reduces efficiency. Be prepared to discuss these implications clearly.

- **Improve time management and completeness:** Many students lost marks by leaving parts unanswered. Read questions carefully, plan your time, and ensure all requested details are provided.
- **Use past papers and timed practice:** Simulate exam conditions to balance speed and depth. Focus on questions requiring interpretation and advanced setups.
- **Create summary sheets:** Include key formulas, assumptions, and interpretations for estimators and tests. This may help many students consolidate understanding and improves recall during exams.

Final Tip: Examiners reward answers that combine **technical accuracy** with **economic insight**. Show not only *how* to compute results but also *the steps and working* as well as economic interpretation when possible.

The next section discusses the specific expectations for each individual question. The revised format of this section provides brief guidance on how to approach each question and includes commentary on what constitutes a good or outstanding answer. The solutions shown are intentionally kept to a minimum and should not be considered model answers. There may be more than one valid way to tackle a question, and a well-written answer can differ in structure or style. Use these notes to understand the key ideas and expectations rather than memorising specific responses. The commentary also reflects how students performed overall and highlights common mistakes to help you improve.

SECTION A

Answer all questions from this section.

Comments on specific questions

1. Consider the following stationary time series model

$$y_t = \alpha + \rho y_{t-1} + \beta x_t + \varepsilon_t, \qquad |\rho| < 1, \qquad t = 1, ..., T. \tag{1.1}$$

The error terms ε_t have zero mean and exhibit autocorrelation represented by a stationary AR(1) process. Also assume that ε_t and x_s are independent for all t, s and $y_0 = 0$.

(a) **(4 marks)** Briefly discuss the properties of the OLS estimator for ρ (unbiasedness, consistency). Support your answers with suitable arguments (no derivations are expected).

(b) **(4 marks)** What is the purpose of heteroskedasticity and autocorrelation robust (HAC) standard errors? Discuss whether HAC standard errors can resolve the main problem associated with estimating the model using OLS.

Approaching the question

(a) When we apply OLS to the above model in the presence of autocorrelation, AR(2), we will get a biased and inconsistent estimator. The presence of a lagged dependent variable is already enough to tell us that we will get a biased estimator, it violates the strict exogeneity assumption. In addition, the errors exhibit autocorrelation we will get inconsistency as well as that will result in $Cov(y_{t-1}, \varepsilon_t) \neq 0$. Intuitively, lagging our model shows that y_{t-1} depends on ε_{t-1}. Given the AR(1) structure of our error ($\varepsilon_t = \rho_1 \varepsilon_{t-1} + e_t$), ε_t also depends on ε_{t-1}.

(b) HAC robust standard errors is required to ensure our standard errors are valid in the presence of autocorrelation and heteroskedasticity. HAC robust standard errors is used when we apply OLS in the presence of autocorrelation, it does not solve the problem of inconsistency of OLS.

Comments

- **Overall performance:** Most students reached the correct conclusion that OLS is biased and inconsistent, but only a few explained why the lagged dependent variable violates strict exogeneity. Many answers lacked clarity on the role of autocorrelation.
- **What makes a good answer: Part (a)** A good answer clearly states that strict exogeneity fails and links this to the correlation between y_{t-1} and ε_t. Outstanding answers provide intuition on why this correlation arises in dynamic models. **Part (b):** Most students correctly noted that HAC standard errors improve inference but do not

solve the main problem. A better answer mentions that consistency requires alternative estimators such as IV.

- **Common mistakes:** Some students wrote "robust SE fixes the problem" without clarifying its limitations. Others ignored autocorrelation entirely or confused robust SE with GLS.

Reading for the question: Subject Guide: Chapter 14.2.1, Wooldridge, Sections 12.1 - 12.2

2. Consider the following simple regression model between Y_i and X_i:

$$Y_i = \beta_0 + \beta_1 \log(X_i) + u_i, \qquad i = 1, ..., n. \tag{2.1}$$

The error term u_i is assumed to be i.i.d, with zero mean and variance equal to σ^2 and independent of the regressor $X_i > 0$. You are given given a random sample of observations $\{(Y_i, X_i)\}_{i=1}^n$.

(a) **(4 marks)** Discuss at least two econometric reasons for the use of $\log(X_i)$ instead of X_i in an OLS estimation.

(b) **(4 marks)** Derive the OLS estimator of β_1. Provide an economic interpretation of the OLS estimate for β_1.

Approaching the question

(a) There are three reasons. First, most importantly, when we care about the relative change in X rather than an absolute change, for example, easier comparability, then using the $\log(X)$ will facilitate the interpretation. Second, if X has a large magnitude, using a log will scale down the magnitude (for example, when X is income in thousands of GBP). Third, following the second point, the use helps reduce the problem of heteroskedasticity.

(b) To facilitate our discussion, let $Z_i = \log(X_i)$. To obtain the OLS estimator we should solve:

$$\min_{b_0, b_1} \sum_{i=1}^n (Y_i - b_0 - b_1 Z_i)^2$$

This requires us to solve the FOC:

$$\begin{aligned} -2\sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 Z_i) &= 0 \\ -2\sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 Z_i) Z_i &= 0 \end{aligned}$$

Solving these equations yields:

$$\begin{aligned} \hat{\beta}_0 &= \bar{Y} - \hat{\beta}_1 \bar{Z} = \bar{Y} - \hat{\beta}_1 \overline{\log(X)} \\ \hat{\beta}_1 &= \frac{\sum_{i=1}^n \left(Z_i - \bar{Z}\right)\left(Y_i - \bar{Y}\right)}{\sum_{i=1}^n \left(Z_i - \bar{Z}\right)^2} = \frac{\sum_{i=1}^n \left(\log(X_i) - \overline{\log(X)}\right)\left(Y_i - \bar{Y}\right)}{\sum_{i=1}^n \left(\log(X_i) - \overline{\log(X)}\right)^2} \end{aligned}$$

Observe $\beta_1 = \frac{\Delta Y}{\Delta \log X} = \frac{\Delta Y}{\%\Delta X}$, for a small change in X. Therefore, if you change X by 1%, we expect Y to change by $\beta_1/100$ units.

Comments:

- **Overall performance:** Most students did well on this question. Part (a) was generally answered correctly, with most mentioning at least two reasons. However, some answers lacked clarity or practical examples.
- **What makes a good answer for (a):** A good answer explains why logs help interpretation and heteroskedasticity reduction. Outstanding answers link these reasons to real-world examples (e.g., income scaling). **For (b):** Most derived the estimator correctly, but many skipped interpretation. A better answer explains that β_1 represents the marginal effect of a percentage change in X on Y, holding others constant.
- **Common mistakes:** Some students confused percentage changes with percentage points or gave vague interpretations without economic meaning.

Reading for the question: Subject Guide: Chapter 1.2.1, 3.2.2, Wooldridge, Section 2.2, Appendix 2A and Appendices B.4a, B.4b, B.4c, Appendix A

3. Consider the following regression model:

$$y_t = \alpha + \rho y_{t-1} + \gamma_1 x_t + \varepsilon_t, \quad t = 1, ..., T, \tag{3.1}$$

where $|\rho| < 1$ and the error term ε_t is a white noise process generated randomly from a normal distribution with mean zero and constant variance. ε_t is also independent of $y_{t-1}, y_{t-2,}, ..., z_t, z_{t-1}, ...$ You are told that y_t and x_t are I(1) variables.

(a) **(4 marks)** What are the short and long-term effect of x on y?

(b) **(4 marks)** It is important to see if y_t and x_t are co-integrated. Clearly define what co-integration means. What are the consequences to the OLS estimation in (3.1) if y_t and x_t are not co-integrated?

Approaching the question

(a) The short-run effect is simply γ_1. To obtain the long-run relationship, we can set the variables at equilibrium:

$$y = \frac{(\gamma_1)}{(1-\rho)} z.$$

so that the long-run effect is $\frac{(\gamma_1)}{(1-\rho)}$.

(b) We have co-integration than we do have a long run relationship between our I(1) variables x_t and y_t. We have a linear combination of I(1) variables that becomes stationary I(0). For example, if x_t and y_t are I(1) process, but $y_t - \beta x_t$ is stationary. Then we can say X and Y are co-integrated, meaning there is a long term relationship between these two. In that case, there will be an ECM representation. In contrast, two I(1)

variables can appear to have a meaningful relationship but in fact there is none. That's the spurious relationship: where two unrelated series give the impression that they are related, purely because of the fact that both are trending. Otherwise the statistical inference is invalid and regression is meaningless (OLS not consistent)!

Comments:

- **Overall performance:** Most students did well on part (a), correctly identifying the short-run effect as γ_1 and the long-run effect as $\gamma_1/(1-\rho)$. However, only a few provided clear reasoning or mentioned equilibrium conditions explicitly.
- **What makes a good answer for (a):** A good answer states both effects clearly. An outstanding answer explains why the long-run effect depends on equilibrium and persistence in the dynamic model.
- **For (b):** This was more challenging. A good answer defines co-integration as a linear combination of I(1) variables being stationary and explains its importance for valid inference. A better answer discusses the consequences of no co-integration—spurious regression, invalid statistical inference, and inconsistency of OLS.
- **Common mistakes:** Many students confused co-integration with correlation or gave vague definitions without mentioning stationarity. Some failed to explain why spurious regression is problematic for hypothesis testing and interpretation.

Reading for the question: Subject Guide: Chapters 13.2.2, 15.2.4; Wooldridge, Sections 18.3 and 18.4, 10.2 and 10.4.

4. An economist investigates the determinants of sold house prices in the UK, *price*, using the total area (*totsize*), the number of bedrooms (*bedrooms*), and the house location (*urban*), which takes value of 1 if in an urban area, 0 otherwise, in the following OLS estimation.

$$price = \beta_0 + \beta_1 totsize + \beta_2 bedrooms + \beta_3 urban + u. \tag{4.1}$$

The error term u has a variance $var(u) = \sigma^2 \log(totsize)$, where σ^2 is an unknown constant. The economist uses a sample of houses sold in 2024. We assume that the regressors are independent of the error terms.

(a) **(5 marks)** Discuss the problem of using estimates from an OLS to perform the hypothesis test of $H_0 : \beta_1 = 0$. Provide and briefly discuss two potential solutions to the problem.

(b) **(3 marks)** A colleague suggests including also $urban^2$. Is this suggestion sound? Explain your answer.

Approaching the question

(a) The test is no longer invalid since OLS standard error is invalid due to heteroskedasticity. With $var(u) = \sigma^2 \log(totsize) = \sigma^2 h(totsize)$, where $h(totsize)$ is the heteroskedasticity function. We first transform the equation by dividing the original equation by $\sqrt{h(totsize)} = \sqrt{\log(totsize)}$:

$$\frac{price}{\sqrt{\log(totsize)}} = \beta_0 \frac{1}{\sqrt{\log(totsize)}} + \beta_1 \frac{totsize}{\sqrt{\log(totsize)}} + \beta_2 \frac{bedrooms}{\sqrt{\log(totsize)}} + \beta_3 \frac{urban}{\sqrt{\log(totsize)}} + \frac{u}{\sqrt{\log(totsize)}}.$$

- The transformed equation now has an error term that is homoskedastic. This WLS is the BLUE, but we must have known the formula. No intercept is used here.
- Another option is to use the robust standard errors. OLS is no longer BLUE though.

(b) Including the term is to account for the potential non linearity relationship between urban and price. But $urban$ and $urban^2$ are exactly the same, so doing so would result in perfect collinearity, one of them will be dropped.

Comments:

- **Overall performance:** Most students answered part (a) reasonably well, identifying heteroskedasticity as the main issue and mentioning robust standard errors. However, only a few discussed the weighted least squares (WLS) transformation correctly.
- **What makes a good answer for (a):** A good answer explains why heteroskedasticity invalidates standard errors and suggests robust SE as a solution. An outstanding answer also mentions WLS, explains why it restores efficiency, and notes that OLS is no longer BLUE when using robust SE. **For (b):** Most recognized that including $urban^2$ adds no new information because $urban$ is binary, leading to perfect collinearity. A better answer briefly explains why this happens and concludes that the suggestion is unsound.
- **Common mistakes:** Some students only wrote "use robust SE" without explaining why, while others incorrectly suggested that adding $urban^2$ would capture non-linearity without realizing the collinearity issue.

Reading for the question: Subject Guide: Chapter 8.2.1; Wooldridge, Sections 8.1 - 8.2.

5. Consider the following time series model for $\{y_t\}_{t=1}^T$:

$$y_t = \beta_0 + \beta_1 t + \varepsilon_t + \lambda \varepsilon_{t-1}, \tag{5.1}$$

where t denotes the time period, ε_t is i.i.d with mean zero and variance σ^2, for $t = 1, ..., T$, and $\beta_1 \neq 0$.

(a) **(4 marks)** Show that y_t is trend stationary. In your answer, clearly define what a trend stationary process is and how it is different from a difference-stationary process.

(b) **(4 marks)** Define what weak dependence means and why it is important to have weak dependence. How is this concept different from covariance stationarity?

Approaching the question

(a) A trend stationary process if by detrending, we can make it stationary. A difference stationary process if by differencing, we can make it stationary. It is important to distinguish the two types as they have implications for the long-run behaviours of a process after a shock to the processes. After a shock, trend stationary processes will always revert to the trend in the long run (they would exhibit weak dependence). Difference stationary processes, however, never recover from shocks to the system. Shocks to the system will have a permanent effects on the processes and the processes exhibit strong dependence. To show y_t is trend stationary, it is clear that $y_t - t = \varepsilon_t + \lambda\varepsilon_{t-1}$, a clearly stationary process as it's an MA(1) structure.

(b) A time series is weakly dependent as the correlation between y_t and future periods y_{t+h} tends to zero asymptotically; with the periods increasing to infinity, the dependency between the two values tends to zero. It is important because it allows us to use LTN and CLT (replacing the i.i.d assumption). Covariance stationarity is rather different: it asks for three statistical properties of a time series: mean, variance, co-variance, are independent of time.

Comments:

- **Overall performance:** Most students performed well on part (a), correctly defining trend stationarity and distinguishing it from difference stationarity. Many also showed that detrending makes the process stationary. However, only a few discussed the implications for long-run behavior after shocks.
- **What makes a good answer for Part a:** A good answer provides clear definitions and demonstrates why y_t becomes stationary after removing the trend. An outstanding answer explains the economic intuition—trend stationary processes revert to trend after shocks, while difference-stationary processes do not. For Part (b): This was more challenging. A good answer defines weak dependence and explains its importance for asymptotic theory (LLN and CLT). A better answer contrasts weak dependence with covariance stationarity, noting that the latter concerns constant mean, variance, and covariance over time.
- **Common mistakes:** Some students confused weak dependence with stationarity (strange!) or gave vague definitions without mentioning its role in large-sample properties. Others failed to clearly differentiate weak dependence from the independence assumption in the cross-sectional setup.

Reading for the question: Subject Guide: Chapter 15.2, Wooldridge, Sections 10.5 (skip 10.5d), 11.3, and 18.2–18.4.

SECTION B

Answer three questions from this section.

Comments on specific questions

6. A Logit regression model and a Linear Probability Model (LPM) are used to explain the mortgage denial rates amongst a random sample of 1234 young families in Britain, 35% of whom belong to a minority ethnicity group. There are 520 successful mortgage applications and 714 denied applications, with an average salary of £33,000 and an average deposit of £50,000. The regression results are as follows:

$$\widehat{denial}_i = \Lambda\left(\underset{(0.011)}{-0.032} - \underset{(0.052)}{0.103}\ salary_i + \underset{(0.054)}{0.240}\ minority_i - \underset{(0.044)}{0.095}\ deposit_i\right) \quad \text{(Logit)}$$

$$\widehat{denial}_i = \underset{(0.022)}{-0.103} - \underset{(0.152)}{0.263}\ salary_i + \underset{(0.024)}{0.062}\ minority_i - \underset{(0.144)}{0.295}\ deposit_i \quad \text{(LPM)}$$

where $\Lambda(z) = {}^{\exp(z)}/_{(1+\exp(z))}$ is the logistic function. The usual standard errors for the Logit model and the robust standard errors for the LPM are reported in parentheses.

The variable $denial_i$ equals to 1 if the mortgage application was denied, 0 otherwise. The variables $salary_i$ and $deposit_i$ are the total annual salary and the down payment by family i in £10,000. The variable $minority_i$ equals to 1 if the applicant belongs to a minority group, 0 otherwise.

(a) **(5 marks)** Describe the estimator underlying the Logit model and discuss its properties. **Note:** clearly indicate the likelihood function used, but a detailed derivation of the estimator is not expected.

(b) **(5 marks)** Discuss the advantages and drawbacks of using the LPM, rather than the Logit model, to model the mortgage denial rates.

(c) **(5 marks)** Discuss how you can test the null hypothesis that the ceteris paribus effect of *salary* on the probability of denial is equal to -0.25. Clearly state the alternative hypothesis, the test statistic, and rejection rule.

(d) **(5 marks)** Using the Logit model, what is the marginal effect of belonging to a minority ethnicity group on mortgage denial rate at the mean values of the explanatory variables. Discuss the limitation(s) of the marginal effect obtained and propose an alternative that overcomes such limitation(s). **Note:** there is no need to calculate the marginal effect using the alternative approach.

Approaching the question

(a) The probability of denial for each individual is:

$$p_i(denial_i = 1) = \widehat{denial}_i = \Lambda\left(\alpha_0 + \alpha_1\ salary_i + \alpha_2\ minority_i + \alpha_3\ deposit_i\right),$$

and the probability of successful application is, where $\Lambda(z) = {}^{\exp(z)}/_{(1+\exp(z))}$:

$$p_i(denial_i = 0) = 1 - \Lambda\left(\alpha_0 + \alpha_1\ salary_i + \alpha_2\ minority_i + \alpha_3\ deposit_i\right),$$

The likelihood function for this sample is:

$$\begin{aligned} L(\alpha_1, \alpha_2, \alpha_3 \mid data) &= \prod_{denial} p_i(denial = 1) \prod_{success} p_i(denial = 0) \\ &= \prod_{denial} \Lambda\left(\alpha_0 + \alpha_1\ salary_i + \alpha_2\ minority_i + \alpha_3\ deposit_i\right) \times \\ &\quad \prod_{success} \Lambda(1 - \alpha_0 - \alpha_{i1}\ salary_i - \alpha_2\ minority_i - \alpha_3\ deposit_i). \end{aligned}$$

- To estimate the parameters, we can use the Maximum Likelihood Estimator that chooses $\{\alpha_j\}$ to maximise the joint probability of the outcomes.
- There are no mathematical formulae for the estimates, the popular method was to adopt an iterative trial-and-error process.
- In large samples, the MLE are consistent, asymptotically normally distributed, and asymptotically efficient.

(b) The key problem with the LPM is that the predicted probabilities of denial obtained from the LPM can be larger than 1 or smaller than 0, which is nonsensical. This problem is very much likely to happen, given the estimates from the LPM model.

- The logit model addresses this concern as the probability is defined using CDF ($\Lambda(z)$), which ensures the predicted probabilities are contained between 0 and 1.
- The LPM implicitly assumes that the marginal effect of minority (or any variables) is constant over its entire range. Even though it is easier and straightforward to interpret, this assumption might not be intuitively appealing. Since the sample is large, we do not have to worry about the error term of the LPM not having a normal distribution.
- Students can also discuss the easy implementation of LPM compared to Logit (which is no longer a concern due to computability nowadays, or LPM will inherently be heteroscedastic).

(c) We need to do a hypothesis test on whether the coefficient on $salary_i$ is equal to -0.25. Since estimates from the LPM and the Logit estimations have different interpretation, we must use the estimate from the LPM, with robust standard errors, for this test.

- The null hypothesis is $\alpha_1 = -0.25$, the alternative hypothesis is $\alpha_{1LPM} \neq 0.25$.
- The appropriate test for the LPM estimate is the z test (asymptotic t test). The z test is $\frac{\hat{\alpha}_1 + .25}{se(\hat{\alpha}_1)} = -0.09$.
- Under the null hypothesis, the z test has a N(0,1) distribution. At 5% significance level and two-tailed test, the critical value is 1.96, and using robust standard errors.
- If the test statistic is larger than the critical value, we reject the null hypothesis, in favour of statistical evidence for racial discrimination in mortgage approval process.

We cannot reject the hypothesis that the effect of salary of the probability of denial is -0.25.

(d) The marginal effect of interest will tell us how the probability of denial $p_i(denial_i = 1)$ for minority applicants differs from that of non-minority applicants. The marginal effect at the mean values of the explanatory variables is:

$$\frac{\partial p_i(denial_i = 1)}{\partial minority} = 0.240 \times f(z),$$

where $f(z) = \frac{e^{-z}}{(1+e^{-z})^2}$ is the PDF of the logistic function, and

$$\begin{aligned} z &= -0.032 - 0.103\ \overline{salary} + 0.240\ \overline{minority} - 0.095\ \overline{deposit} \\ &= -0.032 - 0.103 \times 3.3 + 0.24 \times 0.35 - 0.095 \times 5 = -0.763. \end{aligned}$$

- Plug $z = -0.763$ to pdf of the logistic function $f(z) = \frac{e^{-z}}{(1+e^{-z})^2} = 0.217$. Thus the marginal effect of interest is $0.240 \times 0.217 = 0.052$. Instead of using the pdf, candidates can use $\Lambda(z)[1 - \Lambda(z)]$. Students are not expected to calculate the numbers.
- Alternative approach simply compares the probability at averages with minority=1 versus 0.
- The problem with this calculation is that its interpretation is non-nonsensical as there cannot no person that is of 35% minority. The interpretation for an individual with average characteristics might not be of any practical use. The alternative is to look at the average marginal effect over all individuals (here the sample size is $n = 1234$), which can be calculated as follows:

$$\frac{1}{n}\sum_{i=1}^{n} \frac{\partial p_i(denial_i = 1)}{\partial minority} = \frac{1}{n}\sum_{i=1}^{n} \Lambda(z_i) \times [1 - \Lambda(z)] \times 0.240 = \frac{1}{n}\sum_{i=1}^{n} f(z_i) \times 0.240,$$

where $z_i = \alpha_0 + \alpha_1\ salary_i + \alpha_2\ minority_i + \alpha_3\ deposit_i$, and $f(z)$ is the pdf.

Comments:

- **Overall performance:** This was one of the more challenging questions. Very few students were able to write down the likelihood function clearly in part (a), and even fewer commented on the properties of the maximum likelihood estimator. Most students did better on part (b), identifying key limitations of the LPM and advantages of the Logit model. Parts (c) and (d) were attempted by many, but answers often lacked completeness and interpretation.
- **Part (a):** A good answer sets up the likelihood function correctly and explains that MLE maximizes the joint probability of observed outcomes. Outstanding answers mention asymptotic properties such as consistency and efficiency. Common mistake: omitting the likelihood structure or confusing it with the log-likelihood.

- **Part (b):** Most students recognized that LPM can predict probabilities outside [0,1] and that Logit addresses this issue. A better answer also discusses the assumption of constant marginal effects in LPM and heteroskedasticity concerns. Common mistake: focusing only on interpretability without mentioning these drawbacks.
- **Part (c):** Many students struggled with hypothesis testing. A good answer specifies the null and alternative hypotheses, uses the LPM estimate for the test, and explains the steps of the z-test. Common mistake: using Logit estimates for the test or omitting the distribution and critical value.
- **Part (d):** Few students provided a clear formula for the marginal effect or discussed its interpretation. A better answer notes the limitation of interpreting marginal effects at mean values and suggests alternatives such as average marginal effects. Common mistake: giving only a numeric result without explanation or ignoring practical interpretation issues.

Reading for this question: Subject Guide: Chapter 12.2, Wooldridge Sections 7.5, 8.5, 17.1, and Appendices 17A and C4.b.

7. Consider the following OLS estimation of a model to explain the stock prices of a FTSE100company using 120 observations from 2012m1 to 2021m12. All variables are calculated at the end of month t:

$$\widehat{\log(stock_t)} = \underset{(1.12)}{0.86} + \underset{(0.24)}{0.64} \log(profit_t) - \underset{(0.10)}{0.03} research_t - \underset{(0.12)}{1.34} \log(marketing_t) \quad (7.1)$$
$$n = 120,\ R^2 = 0.34, SSR = 1.29, F^3_{116} = 3.89.$$

where $stock_t$ is the stock price in GBP (British Sterling). $profit_t$ (the profit before tax in millions of GBP), $research_t$ (expenditure on research and development), and $marketing_t$ (expenditure on marketing) are measured in millions GBP. Standard errors are reported in parentheses, SSR is the Sum of Squared Residuals, and the F statistic for the significance of the regression is provided.

(a) **(5 marks)** What is the interpretation of the coefficient on $research_t$? Is the sign of the coefficient as you would expect? How about the magnitude? Explain your answer.

(b) **(5 marks)** Looking at the estimates, a colleague claims that the effect of (log) marketing expenses is *more* than twice as large as the effect of (log) profit on the stock price. Describe a suitable test to examine this claim. Clearly specify the null and the alternative hypothesis and assumptions underlying your test. Indicate additional information, if necessary, to conduct such a test.

(c) **(5 marks)** Another colleague is concerned that there may be some problem with the recording of the expenses on research and development activities. Explain the consequences of this concern to your answer in (a).

(d) **(5 marks)** Your research manager believes that the stock price action of the FTSE100 company behaves differently after 2016m6. The manager claims that the coefficients of the regressors in (7.1) for the months after 2016m6 are different from the coefficients in (7.1) for the months in and before 2016m6.

Discuss how you can test whether the manager is correct. Clearly specify the null and the alternative hypothesis. Indicate additional information, if necessary, to conduct such a test.

Approaching the question

(a) Everything else equal (ceteris paribus), a 1 million increase in the monthly research budget is associated with on average an increase of 3% in the monthly stock price. The interpretation will change with the new unit. Another way to explain is to mathematically derive the new estimates (not necessary). The sign is not as expected as maybe research should help stock performance. But it's not significant anyway so maybe not that important.

(b) Let β_2 and β_3 be the population parameters for $\log(profit)$ and $\log(marketing)$. We are interested in testing $H_o : \beta_3 = 2\beta_1$ against the alternative hypothesis $H_a : \beta_3 > 2\beta_1$.

The t statistic to test the null hypothesis H_0^a is written as:

$$t = \frac{\hat{\beta}_3 - 2\hat{\beta}_1}{se(\hat{\beta}_3 - 2\hat{\beta}_1)}.$$

From the above output, we know $\hat{\beta}_3 - 2\hat{\beta}_1$. On the other hand, the standard error of $\hat{\beta}_3 - 2\hat{\beta}_1$ is given by:

$$se(\hat{\beta}_3 - 2\hat{\beta}_1) = \sqrt{se(\hat{\beta}_3)^2 + 4se(\hat{\beta}_1)^2 - 4 \cdot s_{13}}.$$

where s_{12} is the estimate of $Cov(\hat{\beta}_1, \hat{\beta}_3)$. To implement the above t test, we need to know s_{13}. Under the Classical Linear Regression Assumptions (Gauss-Markow assumptions and normality), we know that $t \sim t_{120-4}$. Once we get s_{13}, we can implement the one-sided t test. With the significance level at 5%. The rejection rule is

$$t > c;$$

where c is the 95 percentile of the t_{116} distribution.

(c) There may be issues of measurement error (errors-in-variables). If the misreporting is random (human errors), then there will be attenuation bias, which makes the estimate for research closer to zero. This looks like to be the case as it's already insignificant. If the issue is more sinister, that is some systematic misreporting, the problem is far more worse. Over-reporting may lead to an overestimation of the estimate, and vice versa. Overall, it's more difficult to deal with. The other estimates are also affected by this error.

(d) The appropriate test is the Chow test. We have to split the data into two groups, from 2012m1 to 2016m6, and 2016m7 to 2021m12, then run the following regressions:

$$\log(stock_t) = \alpha_1 + \alpha_2 \log(profit_t) + \alpha_3 \log(research_t) + \alpha_4 \log(marketing_t) + u,$$

where, t = 2012m1, ..., 2016m6

$$\log(stock_t) = \gamma_1 + \gamma_2 \log(profit_t) + \gamma_3 \log(research_t) + \gamma_4 \log(marketing_t) + v,$$

where t = 2016m7, ..., 2021m12. The null hypothesis of the Chow test is $H_o : \alpha_i = \gamma_i$, with $i = 1, 2, 3, 4$, the alternative hypothesis is otherwise.

- The test-statistic is an F-statistic, that uses the sum of squared errors from the combined data SSR, from the first data group SSR_1, and from the second data group SSR_2. The numbers of observations are $n = 120$, $n_1 = 66$ and $n_2 = 54$, the total number of parameters, including the intercept, is 4. The Chow-test statistic is:

$$\text{F-stat} = \frac{SSR - (SSR_1 + SSR_2))/4}{(SSR_1 + SSR_2)/(120 - 2 \times 4)}$$

- Under CLM assumptions (Gauss-Markov + normality). The critical value is from $F_{4,120-2\times4}$.

- If the F-stat is larger than the critical F-value, we reject the null hypothesis, suggesting that the manager is correct (the model coefficients are different before and after 2016m6). If we cannot reject the null hypothesis, the manager suggestion cannot be concluded.

Comments:

- **Overall performance:** Most students attempted part (a) and provided some interpretation of the coefficient, but many missed key elements such as stating "ceteris paribus" or clarifying that the interpretation is in percentage terms. Part (b) was moderately well done, with students generally identifying the null and alternative hypotheses, but few provided the full details of the test statistic and assumptions. Parts (c) and (d) were more challenging, and many answers lacked depth or completeness.
- **Part (a):** A good answer interprets the coefficient correctly, noting that a 1 million increase in research budget is associated with a 3% increase in stock price, holding other factors constant. Outstanding answers also comment on the sign, significance, and possible reasons for unexpected results.
- **Part (b):** A good answer specifies H_0 and H_a, writes the correct t-statistic formula, and mentions the assumptions and rejection rule. An outstanding answer explains how to compute the standard error using the covariance term and may suggest reparameterizing the model. Common mistake: omitting the covariance term or failing to state the distribution and critical value.
- **Part (c):** A good answer identifies measurement error and explains attenuation bias under classical random error. A better answer discusses the implications of systematic misreporting and its effect on other coefficients. Common mistake: only mentioning "measurement error" without explaining its consequences.
- **Part (d):** Most students recognized the Chow test but did not provide all details. A good answer sets up the null and alternative hypotheses and explains the F-statistic formula. Outstanding answers include the calculation structure, degrees of freedom, and interpretation of rejecting or failing to reject the null. Common mistake: confusing the Chow test with a simple t-test or omitting the assumptions.

Reading for this question: Subject Guide: Chapter 7.2.2, 6.1.2, Wooldridge Section 5.2 and Appendix C.3b, Sections 6.2b, 6.2c, 6.3a, 6.3c, 7.1–7.4 and Appendices A.4a, A.5.

8. Consider the following regression model that describes the demand for money in the USA:

$$\log(M_t^*) = \beta_0 + \beta_1 \log(R_t) + \beta_2 \log(Y_t) + u_t, \tag{8.1}$$

where, M_t^* is the desired (unobserved) or long run demand for money (the real cash balances), R_t is the long-term interest (%, measured by the 90-day corporate rate of interest), Y_t is the aggregate real national income (measured by the GDP), and u_t is an i.i.d error term, and independent of $M_t, M_{t-1}, ..., R_t, R_{t-1}, ..., Y_t, Y_{t-1}$.
The unobserved desired demand variable $\log(M_t^*)$ can also be written as an adaptive process as follows:

$$\log(M_t) - \log(M_{t-1}) = \delta \left[\log(M_t^*) - \log(M_{t-1})\right], \tag{8.2}$$

where M_t is the actual demand for money at time t.

(a) **(5 marks)** What is the interpretation of δ in (8.2)? Show that we can transform the models (8.1) and (8.2) into the following short-run model of the demand function for money:

$$\log(M_t) = \gamma_0 + \gamma_1 \log(R_t) + \gamma_2 \log(Y_t) + \gamma_3 \log(M_{t-1}) + v_t, \tag{8.3}$$

where v_t is a error term.

(b) **(5 marks)** Carefully discuss the properties of the OLS estimators for (8.3) in terms of unbiasedness and consistency. Comment on the validity of the OLS standard errors. Provide suitable arguments for your claims. No formal derivations are expected.

(c) **(5 marks)** Discuss how we can test for the presence of autocorrelation in (8.3). We can assume the alternative v_t can be approximated by a stationary AR(1).

(d) **(5 marks)** Instead of fitting model (8.3), a researcher estimates the following model without $\log(M_{t-1})$.

$$\log(M_t) = \alpha_0 + \alpha_1 \log(R_t) + \alpha_2 \log(Y_t) + e_t, \tag{8.4}$$

Discuss the properties of the OLS estimators for model (8.4).
Hint: Prove that the error term in model (8.4) exhibits autocorrelation.

Approaching the question

(a) Rearrange (8.2), we have:

$$\begin{aligned}
\log(M_t) &= \delta \log(M_t^*) + (1-\delta)\log(M_{t-1}) \\
&= \delta(\beta_0 + \beta_1 \log(R_t) + \beta_2 \log(Y_t) + u_t) + (1 - \delta \log(M_{t-1})) \\
&= \delta\beta_0 + \delta\beta_1 \log(R_t) + \delta\beta_2 \log(Y_t) + (1-\delta)\log(M_{t-1}) + \delta u_t.
\end{aligned}$$

δ captures the percentage of the differences between the actual and desired real cash balances (demand for money) is eliminated within a single period (here a year). It tells us how quickly the actual and desired demand for money converge.

(b)
- Because of the presence of $\log(M_{t-1})$ in the above model, we can not establish unbiasedness.
- The assumptions for consistency are satisfied because for this we only need to ensure that $v_t = \delta u_t$ is uncorrelated with the regressors $\log(M_{t-1}), \log(R_t), \log(Y_t)$ and X_{t-1}. We have the assumption for the independent of the errors so they are uncorrelated. As long as there is no autocorrelation, we will also not have any correlation between $\log(M_{t-1})$ and u_t.
- We don't need to use robust standard errors as there is no heteroscedasticity or autocorrelation.

(c) As the alternative v_t follows an AR(1) process, it can be written as $v_t = \rho v_{t-1} + \varepsilon_t$, where ε_t is a white noise and $|\rho| < 1$. The null hypothesis is $H_0 : \rho = 0$ against $H_1 : \rho \neq 0$. We will proceed as follows:

- (1). Run OLS on the original equation to obtain the residuals $\hat{v}_t$

$$\hat{v}_t = Y_t - \hat{\gamma}_0 - \hat{\gamma}_1 \log(R_t) - \hat{\gamma}_2 \log(Y_t) - \hat{\beta}_3 \log(M_{t-1}).$$

- (2). By the assumption that the alternative is a stationary AR(1) process, the test equation is:

$$\hat{v}_t = \alpha_0^* + \alpha_1^* \log(R_t) + \alpha_2^* \log(Y_t) + \alpha_3^* \log(M_{t-1}) + \rho \hat{v}_{t-1} + \varepsilon_t$$

- (3). Finally, using the sample size, n, and its R^2 (goodness of fit measure), we obtain our test statistic $nR^2 \overset{a}{\sim} \chi^2(1)$. We should reject the null of no autocorrelation if nR^2 exceeds the critical value of the $\chi^2(1)$ distribution at a given level of significance. Or, we use the asymptotic t-test, given by $\hat{\rho}/se(\hat{\rho}) \overset{a}{\sim} N(0,1)$. We should reject the null if the absolute value of the t-test exceeds the critical value of the $N(0,1)$ at a given level of significance.

(d) The errors in model (8.4) that excluded the relevant variable $\log(M_{t-1})$ is

$$e_t = \gamma_3 \log(M_{t-1}) + v_t,$$

where $v_t = \delta u_t$ is assumed as a white noise. The omission will lead to autocorrelation. Indeed, we can show that e.g., $Cov(e_{t+1}, e_t) \neq 0$. Using our definition of e_t and e_{t+1}, we have:

$$\begin{aligned} Cov(e_{t+1}, e_t) &= Cov\left(\gamma_3 \log(M_t) + v_{t+1}, \gamma_3 \log(M_{t-1}) + v_t\right) \\ &= \gamma_3^2 Cov\left(\log(M_t), \log(M_{t-1})\right) + \gamma_3 Cov\left(\log(M_t), v_t\right) \\ &\quad + \gamma_3 Cov\left(v_{t+1}, \log(M_{t-1})\right) + Cov\left(v_{t+1}, v_t\right) \\ &= \gamma_3 Cov\left(\log(M_t), v_t\right). \end{aligned}$$

The last equality holds because v_t is white noise (i.i.d errors unrelated to anything in the past) hence $Cov\left(v_{t+1}, v_t\right) = 0$ and $cov\left(u_{t+1}, \log(M_{t-1})\right) = 0$. $cov\left(\log(M_t), v_t\right) \neq 0$

because of the relationship between the two variables from (8.3). Therefore, we have: $Cov(e_{t+1}, e_t) \neq 0$, or the error term exhibits autocorrelation. With the presence of autocorrelation, the weak exogeneity assumption will not hold in this model with a lagged dependent variable, the OLS estimators are no longer consistent. The standard errors are now invalid too (but really irrelevant).

Comments

- **Overall performance:** This was one of the most challenging questions in the paper. Very few students were able to interpret δ correctly as the speed of adjustment or rate of convergence in part (a). Part (b) was generally better, with most students recognizing the issue of unbiasedness and discussing consistency, but few provided a clear explanation. Parts (c) and (d) were poorly answered overall, with many students struggling to set up the correct test equation or explain the implications of omitted variables.
- **Part (a):** A good answer rearranges the equation and explains that δ measures how quickly actual and desired money demand converge. Outstanding answers provide an intuitive interpretation and link it to adjustment speed in economic terms. Common mistake: treating δ as just another coefficient without interpretation.
- **Part (b):** Most students correctly noted that unbiasedness does not hold due to the lagged dependent variable. A better answer explains why consistency can still hold under certain assumptions and why robust standard errors are unnecessary in the absence of heteroskedasticity or autocorrelation. Common mistake: confusing unbiasedness with consistency or failing to justify the assumptions.
- **Part (c):** Few students provided a complete answer. A good answer specifies the null hypothesis, sets up the test equation, and explains the use of nR^2 or the t-test for autocorrelation. Outstanding answers mention the distribution and rejection rule. Common mistake: omitting the test statistic or confusing this with a standard Durbin-Watson test.
- **Part (d):** Most students struggled to explain why omitting $\log(M_{t-1})$ leads to autocorrelation and inconsistency. A good answer shows the covariance argument and concludes that OLS is no longer consistent. Common mistake: stating "autocorrelation occurs" without proof or failing to link it to the violation of weak exogeneity.

Reading for this question: Subject Guide: Chapter 7.2.2, 6.1.2, 13.2. Wooldridge, Sections 5.1–5.3, 10.1–10.3 and 11.1–11.2, and Appendices C.3a and C3.b.

9. Consider the following system of simultaneous equations

$$y_1 = \alpha_1 y_2 + \alpha_2 x_1 + \alpha_3 x_2 + u_1 \tag{9.1}$$

$$y_2 = \beta_1 y_1 + \beta_2 x_1 + \beta_3 x_3 + \beta_4 x_4 + u_2, \tag{9.2}$$

where (u_1, u_2) are i.i.d. errors with zero mean, $Var\,(u_1) = \sigma_1^2$, $Var(u_2) = \sigma_2^2$, and let $Cov(u_1, u_2) = 0$. The endogenous variables are (y_1, y_2) and (x_1, x_2, x_3, x_4) are exogenous variables.

(a) **(5 marks)** Explain the main problem when estimating the two equations (9.1) and (9.2) using OLS. Give an economic example that fits the structural model.

(b) **(5 marks)** You are told that in the above SEM there exists a perfect linear relation between x_2 and x_3, in particular,

$$x_3 = 2x_2.$$

Obtain the reduced form equations for y_1 and y_2, recognising this exact linear relation between x_2 and x_3.

(c) **(5 marks)** Based on your answer in (b), derive the bias of the OLS estimator for β_1 in (9.2).

(d) **(5 marks)** Use the result obtained in (b) to discuss the identification of the two structural equations. Clearly state whether the equations are over identified, exact (just) identified, or under identified.

Approaching the question

(a) The problem here is simultaneity bias: both equations have their own ceteris paribus interpretation, the simultaneity gives rise to a correlation between the two endogenous variables that are jointly determined in this system which makes it that when we estimate these structural equation in isolation will give rise to inconsistency. One example is police&crime.

(b) Substitute (9.2) to (9.1) we have:

$$\begin{aligned} y_1 &= \alpha_1(\beta_1 y_1 + \beta_2 x_1 + \beta_3 x_3 + \beta_4 x_4 + u_2) + \alpha_2 x_1 + \alpha_3 x_2 + u_1 \\ (1 - \alpha_1\beta_1) y_1 &= (\alpha_1\beta_2 + \alpha_2) x_1 + \alpha_3 x_2 + \alpha_1\beta_3 x_3 + \alpha_1\beta_4 x_4 + u_1 + \alpha_1 u_2 \\ y_1 &= \frac{(\alpha_1\beta_2 + \alpha_2) x_1 + \alpha_3 x_2 + \alpha_1\beta_3 x_3 + \alpha_1\beta_4 x_4 + u_1 + \alpha_1 u_2}{(1 - \alpha_1\beta_1)}, \end{aligned} \tag{9.3}$$

and for y_2, from (9.1) to (9.2):

$$\begin{aligned} y_2 &= \beta_1 y_1 + \beta_2 x_1 + \beta_3 x_3 + \beta_4 x_4 + u_2 \\ y_2 &= \beta_1(\alpha_1 y_2 + \alpha_2 x_1 + \alpha_3 x_2 + u_1) + \beta_2 x_1 + \beta_3 x_3 + \beta_4 x_4 + u_2 \\ y_2 &= \frac{(\beta_1\alpha_2 + \alpha_2) x_1 + \beta_1\alpha_3 x_2 + \beta_3 x_3 + \beta_4 x_4 + u_2 + \beta_1 u_1}{(1 - \alpha_1\beta_1)}. \end{aligned} \tag{9.4}$$

(c) The problem is the violation of the zero conditional mean: the error terms must be independent of the explanatory variables. For example, for (9.1) the covariance of y_1 and u_1 can be obtained from (9.4), and (x_1, x_2, x_3, x_4). The biases are (details of the derivation are expected).

$$\text{plim } \hat{\alpha}_1^{OLS} - \alpha_1 = cov(y_2, u_1) = \frac{\beta_1 \sigma_1^2}{(1 - \alpha_1 \beta_1)}.$$

and:

$$\text{plim } \hat{\beta}_1^{OLS} - \beta_1 = cov(y_1, u_2) = \frac{\alpha_1 \sigma_2^2}{(1 - \alpha_1 \beta_1)}.$$

The signs of the biases depend on β_1 and α_1 because $(1 - \alpha_1 \beta_1)$ is in general positive (as a condition for the equilibrium of the model).

Because the covariance is non-zero, the OLS estimators are both biased and inconsistent.

(d) Equation(9.1) is just identified as we have x_4 as the exogenous variable to work as an instrument for the only endogenous variable (y_2). x_2 and x_3 cannot be used as instruments as they appear in both the equations. The conditions are: β_4 is statistically significant (strongly predicts y_2), and the exogeneity of x_4 holds ($cov(x_4, u_1) = 0$). Equation (9.2) is underidentified because we cannot use any of the regressors as an instrument to identify β_1 following the multicollinearity between x_2 and x_3.

Comments

- **Overall performance:** Most students recognized the presence of simultaneity bias in part (a), but only a few explained why it leads to inconsistency or provided an example. Part (b) was generally well attempted, with many students able to derive the reduced forms, though explanations were often brief. Part (c) was challenging—very few students derived the bias formula correctly or commented on inconsistency. Part (d) was mixed: some students identified the identification status of each equation, but few explained the conditions clearly.
- **Part (a):** A good answer names simultaneity bias and explains that endogenous variables are jointly determined, violating the zero conditional mean assumption. Outstanding answers provide an intuitive example (e.g., crime and policing) and link this to why OLS fails.
- **Part (b):** Most students managed to substitute and derive the reduced forms for y_1 and y_2. A better answer also comments on the interpretation and why these forms matter for estimation. Common mistake: algebraic errors or skipping explanation.
- **Part (c):** Few students derived the bias expressions correctly. A good answer shows the covariance argument and concludes that OLS is biased and inconsistent. Outstanding answers discuss how the sign of the bias depends on parameters and why this matters for interpretation. Common mistake: stating "OLS is biased" without proof or reasoning.

- **Part (d):** Many students identified that equation (9.1) is just-identified and equation (9.2) is under-identified. A better answer explains the role of instruments, the exclusion restriction, and why multicollinearity prevents identification in (9.2). Common mistake: vague statements like "needs instruments" without specifying which variables or conditions.

Reading for this question: Subject Guide: Chapter 9.2. Wooldridge, Sections 15.1, 15.3a, 15.3d, 15.5a.

END OF EXAMINERS' COMMENTARY

© University of London 2025

UNIVERSITY OF LONDON

EC2020 Elements of Econometrics

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from Section A (8 marks each) and **THREE** from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Please find the questions on the following page.

© University of London 2025

© University of London 2025

SECTION A

Answer all questions from this section. Each question carries 8 marks.

1. Consider the following bivariate regression model:

$$y_i = \alpha + \beta x_i + u_i, \tag{1.1}$$

for $i = 1, ..., n$. The regressor x_i is non-negative, *stochastic*, and has a non-zero finite variance, i.e., $\sum_{i=1}^{n}(x_i - \bar{x})^2 \neq 0$. The error term u_i is i.i.d. with mean zero and constant variance of σ^2, and independent of x_i. Given a random sample of observations $\{(y_{i,}, x_i)\}_{i=1}^{n}$ from the population of interest, consider the following estimator:

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} x_i^{2025} y_i}{\sum_{i=1}^{n} x_i^{2026}}.$$

(a) **(4 marks)** Derive the condition under which $\tilde{\beta}$ is an unbiased estimator of β? Can $\tilde{\beta}$ be the BLUE estimator of (1.1)?

(b) **(4 marks)** Assume that the condition in (a) is satisfied. Will this condition also ensure that $\tilde{\beta}$ is consistent?

Approaching the question

(a) For unbiasedness, we need to show that $E(\tilde{\beta}) = \beta$. We need to take the conditional expectation on $X = \{x_i\}_{i=1}^{n}$, $E[\tilde{\beta}|X]$. The conditional expectation of $\tilde{\beta}$ is given by:

$$\begin{aligned}
E(\tilde{\beta}|X) &= E\left(\frac{\sum_{i=1}^{n} x_i^{2025} y_i}{\sum_{i=1}^{n} x_i^{2026}}|X\right) \\
&= E\left(\frac{\sum_{i=1}^{n} x_i^{2025}(\alpha + \beta x_i + u_i)}{\sum_{i=1}^{n} x_i^{2026}}|X\right) \quad \text{replacing } y_i \text{ from the model} \\
&= E\left[\frac{\alpha\left(\sum_{i=1}^{n} x_i^{2025}\right) + \beta\left(\sum_{i=1}^{n} x_i^{2026}\right) + \sum_{i=1}^{n} x_i^{2025} u_i}{\sum_{i=1}^{n} x_i^{2026}}|X\right] \\
&= \beta + \alpha E\left(\frac{\sum_{i=1}^{n} x_i^{2025}}{\sum_{i=1}^{n} x_i^{2026}}|X\right) + \frac{\sum_{i=1}^{n} x_i^{2025} E(u_i|X)}{\sum_{i=1}^{n} x_i^{2026}} \quad \text{because of i.i.d.} \\
&= \beta + \alpha\left(\frac{\sum_{i=1}^{n} x_i^{2025}}{\sum_{i=1}^{n} x_i^{2026}}\right) \quad \text{because } E(u_i|X) = 0.
\end{aligned}$$

Thus, in general $\tilde{\beta}$ is a biased estimator unless the population intercept $\alpha = 0$. Best candidates would realise that when $\sum_{i=1}^{n} x_i^{2025}$ cannot be zero because of x_i being non-negative and sample variability. $\tilde{\beta}$ is not a BLUE because the OLS in this case would be the BLUE (as the Gauss-Markow assumptions hold).

(b) We need to look at the condition for consistency.

- When the estimator is unbiased, we must have $\alpha = 0$. From (a), we can rewrite the estimator as:

$$\tilde{\beta} = \beta + \alpha\left(\frac{\sum_{i=1}^{n} x_i^{2025}}{\sum_{i=1}^{n} x_i^{2026}}\right) + \frac{\sum_{i=1}^{n} x_i^{2025} u_i}{\sum_{i=1}^{n} x_i^{2026}} = \beta + \frac{\sum_{i=1}^{n} x_i^{2025} u_i}{\sum_{i=1}^{n} x_i^{2026}}.$$

– The probability limit of $\tilde{\beta}$ is:

$$\begin{aligned}
\text{plim}(\tilde{\beta}) &= \beta + \text{plim}\left(\frac{\sum_{i=1}^{n} x_i^{2025} u_i}{\sum_{i=1}^{n} x_i^{2026}}\right) = \beta + \frac{\text{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^{2025} u_i\right)}{\text{plim}\left(\frac{1}{n}\sum_{i=1}^{n} x_i^{2026}\right)} \quad \text{properties of plim} \\
&= \beta + \frac{E\left(x_i^{2025} u_i\right)}{E\left(x_i^{2026}\right)} \quad \text{from the Law of Large Numbers} \\
&= \beta \quad \text{as } E(xu) = 0
\end{aligned}$$

– Therefore, when $\tilde{\beta}$ is an unbiased estimator, it is also consistent.

2. Consider the following demand curve for taxi in London in year t:

$$\ln(Q_t) = \beta_0 + \beta_1 \ln(P_t) + \beta_2 \ln(Y_t) + u_t, \tag{2.1}$$

where Q_t and P_t are the quantity (number of rides) and the average price of taxi per km in London in year t, and Y_t is the average income in London in year t.

(a) **(2 marks)** What is the interpretation for β_1? Briefly explain.

(b) **(2 marks)** You are told that price and quantity in the taxi market are simultaneously determined in a market equilibrium. Briefly discuss how this piece of information poses problems to the OLS estimator for β_1.

(c) **(4 marks)** You are given the average price of electric vehicles in London, EV_t in year t. Discuss whether EV_t is useful to address the problems in (b).

Approaching the question

(a) 1% increase in price is associated with β_1 percent change in the quantity. It's the price elasticity of demand for taxi ride.

(b) There will be simultaneity bias, that is the error term is correlated with P_t. OLS estimator will be biased and inconsistent.

(c) We may want to use the variable as an IV. (1 mark), but it won't be really valid. Relevance can be weakly satisfied. But it won't be exogenous (probably related to demand for taxi rides, or not random at all). Candidates can be creative here, as long as discussing the three conditions.

3. Consider a linear regression model:

$$y_t = \beta x_t + u_t, \qquad t = 1, ..., T \tag{3.1}$$

where the error terms u_t are independent of x_t and exhibit autocorrelation of an unknown form, y_t and x_t are covariance stationary and weakly dependent processes.

(a) **(3 marks)** Explain the concept of covariance stationarity. Intuitively discuss why the property is often assumed in time series analysis.

(b) **(5 marks)** Discuss how you would conduct a statistical test of the hypothesis $H_0 : \beta = 0.5$ against the alternative $H_1 : \beta \neq 0.5$ using the OLS estimate for β. Be specific about the steps and statistic that you use.

Approaching the question

(a) Candidates are expected to discuss either of the following:

- A process is stationary if, intuitively, its stochastic properties (such as the mean, and variance) and its temporal dependence structure (such as co-variance over time) do not depend on time. It is a stronger version of the identically distributed assumption.
- A more formal definition is that for every collection of time indices, the joint distribution of the temporal values of the process is identical to the temporal values when we shift the sequence ahead an h time periods. Candidates can discuss the covariance stationarity instead of this strict stationarity. It is important because we know that the nature of any correlation between adjacent terms is the same across all time periods.

(b) Because of the autocorrelation, we cannot use the usual standard errors of the OLS estimators. We must use the Newey-West (or HAC) robust standard errors. Using this HAC standard error, we can obtain the t statistic:

$$\frac{\hat{\beta} - 0.5}{HACse(\hat{\beta})}.$$

Using the 5% significance level, we will reject of the t-statistic is smaller than -1.645, which is the one-sided asymptotic z critical value (candidates are expected to discuss this part).

4. Consider the following simplified production function for a sample of taxi firms i:

$$revenue_i = \alpha_0 + \alpha_1 vehicles_i + \alpha_2 driver_i + \alpha_3 vehicles_i^2 + u_i, \qquad i = 1, ..., n \tag{4.1}$$

where $revenue_i$ denotes the total revenue of firm i, and $vehicles_i$ and $driver_i$ denote the number of vehicles and the number of drivers employed by the firm. Assume that we obtain a random sample of n firms and the error terms u_i are i.i.d with zero mean and variance δ^2. We also assume that u_i and the regressors are independent.

(a) **(4 marks)** Discuss the economic and econometric rationales for including $vehicles_i^2$ in this model. What is the expected sign for α_3?

(b) **(4 marks)** Discuss how we can test for whether an additional purchase of vehicles has a total significant effect on the revenues. Specify the testing steps and the hypothesis.

© University of London 2025

Approaching the question

(a) Candidates are expected to discuss the diminishing marginal return to vehicles. For the econometric rationale, because of the economic rationale, we need to account for the non-linear relation between revenue and vehicles, hence omitting the squared term will cause omitted variable bias. We expect $\alpha_2 > 0; \alpha_3 < 0$.

(b) Candidates are expected to take the first derivative with respect of vehicles. We will obtain the marginal effect for a 100-worker firm is:

$$\alpha_2 + 2\alpha_3 driver_i.$$

The marginal effect of labour on output is $\alpha_2 + 2\alpha_3 driver$. For this to be insignificant (0) we would need $\alpha_2 = \alpha_3 = 0$. Candidates are then expected to discuss an F-test using either the R-squared or SSR. Candidates are expected to write down the restricted equation as well.

5. An economist considers the Linear Probability Model (LPM) and a Logit model to examine the determinants of admission to Ivy League universities.

(a) **(4 marks)** Give three reasons why a Logit model is preferred to the LPM. Give one potential advantage of using the LPM instead of the Logit model.

(b) **(4 marks)** A colleague suggests that it may be useful to include in the Logit estimation an interaction term between an applicant's household income ($Income$) and the highest education level of the applicant's father ($FatherEduc$). Discuss this idea.

Approaching the question

(a) Three reasons for Logit: bounding the predicted values to (0,1); no inherent heteroskedasticity; and account for non-linear relationship. One reason for LPM: Ease of interpretation and computation.

(b) The suggestion is to investigate the potential differential effect of father's education on admission by different levels of household incomes. However, it is not needed for a Logit as the marginal effect there already captures this. Best students will note that the inclusion will actually complicate the interpretation so we should not do so.

© University of London 2025

SECTION B

Answer three questions from this section. Each question carries 20 marks.

6. An economist investigates how part-time work affects students' performance in an online programme. Using the data from surveying 894 students participating in the programme, she runs the following regression, using the average of overall grades from all taken courses, $Grade_i$, (marks range from 0 - 100) as the outcome variable. The regressors include the self-reported number of self-study hours per week in logs, $\log selfstudy$; the number of hours students engaged in part-time work per week, $work_i$; and the number of concurrent courses the student took ($othercourses$).

$$\widehat{Grade_i} = \underset{(7.14)}{38.3} + \underset{(.23)}{0.51}\log(selfstudy_i) - \underset{(4.56)}{2.07}work_i - \underset{(3.25)}{4.01}othercourses_i, \qquad (6.1)$$
$$n = 894, \qquad R^2 = .355.$$

The robust standard errors are reported in parentheses.

(a) **(5 marks)** What is the interpretation of the coefficient on $othercourses$? Are the sign and size of the coefficient as you would expect? Do you think the effect is qualitatively large or small? Explain using the available information in (6.1).

(b) **(5 marks)** Can we interpret the results in (a) causally? Explain using appropriate technical details. Propose a suitable test for your answer.

(c) **(5 marks)** A colleague believes the regression (6.1) may suffer from heteroskedasticity. Explain what she means intuitively. Discuss a statistical test for this claim. What are the consequences to the regression (6.1) if she is correct?

(d) **(5 marks)** The colleague also suggests that students may provide incorrect counts of the number of hours they work part-time. Intuitively discuss the consequences on your answer in (a) if the colleague is correct. Any evidence from (6.1) that can support the colleague's claim?

Approaching the question

(a) Candidates are expected to provide economic interpretations: one extra course taken is associated with on average 4.01 marks lowered. This seems to be as expected. But a t-test shows that this is not significant, so maybe extra work is not a concern. To evaluate the magnitude, with 1 hour of self-study and not taking any other course nor work, the student will earn 38.3, still a fail. So if they take more courses, probably this can be detrimental. Candidates can be creative here with the intuition.

(b) Candidates are expected to discuss the OVB and perhaps measurement errors (from other variables, as in question d). One OVB is potentially students's academic ability. If they are better academically, they take more courses and do better, so we may overestimate the effect of additional courses. Best candidates will use OVB formula to show the result here. Top candidates discuss the potential of measurement errors.

(c) Candidates are expected discuss clearly the technical definition of heteroskedasticity (the error term variance depends on some or all of the regressors).

Candidates are expected to explain the test for heteroskedasticity: either the Breusch-Pagan version or the White version of the test.

- **For the Breusch-Pagan test:** we should
  - Save the OLS residuals as $\hat{u}_i$. Compute the squared residuals $\hat{u}_i^2$.
  - Regress $\hat{u}_i^2$ on all the regressors, with an intercept, and then compute the F test of joint significance of the regressors.

$$\hat{u}_i^2 = \delta_0 + \delta_1 \log selfstudyi + \delta_2 work_i + \delta_3 othercourses + v_i.$$

  - The null hypothesis is: $H_0 : \delta_1 = \delta_2 = \delta_3 = 0$. Students are expected to discuss the R-squared version or the SSR version of the F-statistic in this case. $k = 3, n - k - 1$.
  - We do not need to assume MLR.6 in this version of the test, under the null hypothesis, the F-stat will follow the asymptotic F distribution.
- **For the White test,** we should
  - Save the OLS residuals as $\hat{u}_i$. Compute the squared residuals $\hat{u}_i^2$.
  - Regress $\hat{u}_i^2$on all the regressors, with an intercept, and the squares and interactions of all the regressors:

$$\begin{aligned}\hat{u}_i^2 &= \delta_0 + \delta_1 \log selfstudy_i + \delta_2 worki + \delta_3 othercourses + \delta_4 \log selfstudy_i^2 + ... \\ &\quad + \delta_7 \log selfstudyi \times work_i + ... + v_i.\end{aligned}$$

    The null hypothesis is: $H_0 : \delta_1 = \delta_2 = \delta_3 = .. = \delta_9 = 0$. We can simply obtain the F-test of the joint significance of the regressors. Similarly, students are expected to discuss the R-squared version or the SSR version of the F-statistic. $k = 9, n - k - 1$.
  - We do not need to assume MLR.6 in this version of the test, under the null hypothesis, the F-stat will follow the asymptotic F distribution.
- If we reject the null hypothesis, there is evidence of heteroskedasticity.The normal standard errors are invalid. But we already use robust standard errors so there is no concern.

(d) Candidates are expected to discuss the errors-in-variables situation, at least, if the part-time hours numbers are randomly misreported: the errors may come from students misremembering their working hours in a random manner. If that's the case, we will have attenuation bias, which bias the estimate of work towards zero. The estimate for work is near zero (insignificant) so probably this concern is valid. However, we do not know the direction of bias towards other estimators. So the answers in a) won't be valid. Best students may observe that the estimator in (a) is already insignificant so probably the attenuation bias won't affect much there as well.

7. An economist uses OLS to estimate the relationship between the monthly inflation rate inf_t and the monthly consumer sentiment sen_t (the percentage of surveyed respondents who are positive about the economy) and obtains the following results:

$$\begin{aligned} \widehat{inf_t} &= \underset{(0.004)}{0.035} + \underset{(.021)}{0.12}\, sen_t - \underset{(0.216)}{0.807} sen_{t-1}, \\ & T = 223, \qquad R^2 = 0.125 \end{aligned} \tag{7.1}$$

where the HAC standard errors reported in the parentheses.

(a) **(5 marks)** What are the estimates of the short-run and long-run effect of sen_t on inf_t? Interpret the economic meaning of these estimates.

(b) **(5 marks)** The economist believes consumer sentiment growth is a leading indicator of inflation. Conduct a statistical test that a unit increase in last month's consumer sentiment results in a one unit reduction in the current month inflation. Be specific about the null hypothesis and the assumptions for the test to be valid.

Obtain e_t which is the OLS residuals from the regression (7.1). The economist runs following OLS regression of e_t on e_{t-1}:

$$\begin{aligned} e_t &= \underset{(2.21)}{0.53} + \underset{(.18)}{0.41} e_{t-1} + \underset{(.118)}{0.541} sen_t - \underset{(0.126)}{1.07}\, sen_{t-1}, \\ & T = 223, \qquad R^2 = 0.32. \end{aligned} \tag{7.2}$$

(c) **(5 marks)** What is the purpose of the regression in (7.2). (*Hint:* It is done as part of a testing procedure). Clearly indicate the null and alternative hypothesis, the test statistic, rejection rule, and the *underlying assumptions*. Be clear about the outcome of the test and how the outcome affects your answers in (a) and (b).

(d) **(5 marks)** The economist wants to test whether the long-run effect of sen_t on inf_t is statistically significant. Discuss how you can reparameterise (7.1) to make sure your regression results can provide you with appropriate information for the test. Be clear about the new regression that you would run, and how you would proceed with the test.

Approaching the question

(a) The short term effect is 0.12, that is, one percentage point increase in the current consumer sentiment leads to (or is associated with) 0.12 percentage point increase in the inflation rate.
The long-term effect is 0.12-0.807 = -0.687, so one percentage point increase in the equilibrium the current consumer sentiment leads to (or is associated with) 0.687 percentage point decrease in the long-term inflation rate. Candidates are expected to interpret the results in percentage points. The two results are opposite but economically make sense.

(b) Candidates are expected to perform a test of the coefficient for sen_{t-1} is -1, that is $\beta_{sen_{t-1}} = -1$. The TS1-TS5 assumptions (particularly no autocorrelation so the standard errors are valid) are to be listed for a full mark. Assumptions TS1'-TS5' are also fine, but the test must be asymptotic. The test statistic is:

$$t = \frac{\widehat{\beta}_{sen_{t-1}} + 1}{se(\hat{\beta})} = \frac{-0.807 + 1}{0.216} = \frac{0.193}{0.216}.$$

Under the assumptions (TS1-TS5, or TS1'-TS5'), we compare the t-stat against the critical value (at any conventional level), two sided, and we cannot reject the null of sen_{t-1} is a leading indicator.

(c) This is a test for autocorrelation in the error term of the main equation of (7.1) (Breusch-Godfrey, not Durbin-Watson, which we can skip in this course). The key assumption here is the error terms ε_t follow an AR(1) structure: $\varepsilon_t = \rho\varepsilon_{t-1} + u_t$ where u_t is a white noise.

The null hypothesis is the coefficient of e_{t-1} in (7.2) is zero. Now, if candidates assume strict exogeneity in (b), which is absolutely fine, we do not need the inclusion of the regressors. But we do in this case, so the weak exogeneity is assumed and we should perform an asymptotic t-test using t-stat $\hat{\rho}/se(\hat{\rho}) = 0.41/0.18 = 2.28$, under the asymptotic distribution $Normal(0, 1)$.

In this case, we reject the null hypothesis of no autocorrelation, in favor of the alternative that we must be dealing with autocorrelation in the error term here. Fortunately we already use the HAC standard errors so the answers in (a) and (b) remain valid.

(d) Essentially candidates are expected to test whether the coefficients for sen_t and sen_{t-1}, say β_1 and β_2 satisfy: $\theta = \beta_1 + \beta_2 = 0$. We can rewrite $\beta_1 = \theta - \beta_2$. Plug that into (7.1) and we can run the following regression:

$$\begin{aligned} inf_t &= \beta_0 + (\theta - \beta_2)sen_t + \beta_2 sen_{t-1} + \varepsilon_t \\ &= \beta_0 + \theta sen_t + \beta_2(sen_t + sen_{t-1}) + \varepsilon_t \end{aligned}$$

That is, we run a regression of inf_t on sen_t and $sen_t + sen_{t-1}$ and then perform a t-test on the significance of θ. Note that however, because of the presence of autocorrelation in ε_t, we must use the HAC standard errors and also the asymptotic version of the t-test for the significance test (essentially asymptotic t-test). Candidates must discuss these points for full marks.

8. Suppose $X_1, X_2, \ldots, X_n$ are i.i.d. random variables drawn from a distribution with the following probability density function with parameter θ:

$$f_X(x;\theta) = 2\theta x e^{-\theta x^2}, \quad x > 0,\ \theta > 0, \tag{8.1}$$

The following properties of the distribution are given:

$$\mathbb{E}[X^2] = \frac{1}{\theta}, \tag{8.2}$$

and

$$\mathbb{E}[X] = \frac{\sqrt{\pi}}{2\sqrt{\theta}}, \text{ where } \pi = 3.14159... \tag{8.3}$$

(a) **(8 marks)** Derive the Maximum Likelihood Estimator (MLE) for θ, $\hat{\theta}_n$. Be clear of the steps that you take.

(b) **(7 marks)** Is the MLE $\hat{\theta}_n$ asymptotically unbiased? Is it consistent? Prove your claim. In your answer, clearly explain the key difference between unbiasedness and consistency.

Let us first consider a sequence of estimators given by:

$$\widehat{\theta}_c = c\left(\frac{1}{\overline{X}}\right)^2, \tag{8.4}$$

where c is a positive constant and $\overline{X} = n^{-1}\sum_{i=1}^{n} X_i$. For example, $\widehat{\theta}_1 = \left(\frac{1}{\overline{X}}\right)^2$ and $\widehat{\theta}_{0.5} = 0.5\left(\frac{1}{\overline{X}}\right)^2$.

(c) **(5 marks)** Is there any value of the constant c so that the estimator $\widehat{\theta}_c$ is an unbiased estimator for θ? Any value of c so $\widehat{\theta}_c$ is consistent? Prove your claim.

Approaching the question

(a) Derivation of the MLE

i. **Likelihood Function $L(\theta)$:**
The joint likelihood function is:

$$L(\theta) = \prod_{i=1}^{n} f_X(X_i;\theta) = \prod_{i=1}^{n} 2\theta X_i e^{-\theta X_i^2}.$$

Simplify:

$$L(\theta) = (2\theta)^n \prod_{i=1}^{n} X_i \cdot e^{-\theta \sum_{i=1}^{n} X_i^2}.$$

© University of London 2025

ii. **Log-Likelihood Function $\ell(\theta)$:**
Taking the natural logarithm:

$$\ell(\theta) = \ln L(\theta) = n\ln(2\theta) + \sum_{i=1}^{n} \ln(X_i) - \theta \sum_{i=1}^{n} X_i^2.$$

Simplify further:

$$\ell(\theta) = n\ln 2 + n\ln\theta + \sum_{i=1}^{n} \ln(X_i) - \theta \sum_{i=1}^{n} X_i^2.$$

iii. **Derivative of $\ell(\theta)$:**
Differentiate $\ell(\theta)$ with respect to θ:

$$\frac{\partial \ell(\theta)}{\partial \theta} = \frac{n}{\theta} - \sum_{i=1}^{n} X_i^2.$$

Set the derivative to zero:

$$\frac{n}{\theta} = \sum_{i=1}^{n} X_i^2.$$

Solve for θ:

$$\hat{\theta}_n = \frac{n}{\sum_{i=1}^{n} X_i^2}.$$

Therefore, the MLE is:

$$\hat{\theta}_n = \frac{1}{\bar{X^2}}, \quad \text{where } \bar{X^2} = \frac{1}{n}\sum_{i=1}^{n} X_i^2.$$

(b) It is not unbiased. Indeed, the expected value of the MLE is:

$$\mathbb{E}[\hat{\theta}_n] = \mathbb{E}\left[\frac{1}{\bar{X^2}}\right].$$

And we need to show that:

$$\mathbb{E}[\hat{\theta}_n] = \mathbb{E}\left[\frac{1}{\bar{X^2}}\right] = \theta = \frac{1}{\mathbb{E}\left[X^2\right]} \text{ from (8.2).}$$

This rarely can be the case because $E(X^2) \neq E(\bar{X^2})$ and in general $E(1/Y) \neq 1/E(Y)$ due to the Jensen inequality. So in general it is not unbiased.
It is however consistent because:

$$\text{plim}[\hat{\theta}_n] = \text{plim}\left[\frac{1}{\bar{X^2}}\right] = \frac{1}{\text{plim}\left[\bar{X^2}\right]} \text{ from plim properties.}$$

By the Weak Law of Large Number we can see that $\text{plim}\left[\bar{X^2}\right] \to E(X^2) = \frac{1}{\theta}$, from also (8.2). Therefore we can show that $\text{plim}[\hat{\theta}_n]$.
Candidates then need to clearly define and contrast unbiasedness and consistency for full marks: unbiasedness (repeated samples, so that the expected value of the estimate from the estimator would be the parameter) and consistency (increasing sample size, the estimate would approach in probability the parameter) concepts.

© University of London 2025

(c) This can be a trick question. Again we need to show:

$$\mathbb{E}[\hat{\theta}_c] = c\mathbb{E}\left[\left(\frac{1}{\bar{X}}\right)^2\right] = \theta = \frac{\pi}{4E(X)^2}$$

There won't be a single value of c so this holds for any sample.

For consistency, it's simpler:

$$\text{plim}[\hat{\theta}_c] = c\,\text{plim}\left[\frac{1}{\bar{X}^2}\right] = \frac{c}{\left[(\text{plim}\,\bar{X})^2\right]} = \frac{c}{E^2(X)} \text{ from plim properties.}$$

From (8.3) we will need $c = \frac{4}{\pi}$.

Candidates need to explain each step for full marks.

9. We are interested in the relationship between the quarterly price level CPI_t and the quarterly employment level Emp_t in the following ADL(1,1) model:

$$CPI_t = \beta_0 + \beta_1 CPI_{t-1} + \beta_2 Emp_t + \beta_3 Emp_{t-1} + u_t, \quad t = 1, ..., T, \tag{9.1}$$

where Emp_t and CPI_t are both $I(1)$ variables, and $|\beta_1| < 1$. Assume also that the error term has a zero mean and is independent of CPI_{t-1}, Emp_t and Emp_{t-1}.

(a) **(5 marks)** Distinguish the concepts of spurious relationship and co-integration and explain why it is so important that we have to test for cointegration.

(b) **(5 marks)** Discuss how you can test whether the above specification is spurious or cointegrating. Clearly indicate the test equation, test statistic, the the rejection rule. In your answer, briefly explain the concepts of spurious and cointegrating relationship.

(c) **(5 marks)** A colleague ascertains that there exists an Error Correction Model (ECM) that describes the short-run and long-run dynamics between CPI_t and Emp_t. Derive the Error Correction Model and interpret the various components of the ECM. Discuss how does this claim relate to your answers to (a) and (b). *Hints:* Subtract CPI_{t-1} from both sides of (9.1) to obtain the following version:

$$\Delta CPI_t = \delta_0 + \delta_1 \Delta Emp_t + \gamma s_{t-1} + \epsilon_t, \tag{9.2}$$

with s_{t-1} is the error correction term.

(d) **(5 marks)** Discuss the Engle-Granger two-step procedure to estimate the ECM obtained in (c).

Approaching the question

(a) Candidates are expected to discuss the two concepts, and then talk about the ADF test of the residual for co-integration (in general it's called the Engle-Granger test).

- Spurious regression refers to the scenario where <u>two unrelated series</u> give the impression that they are related, purely because of the fact that both are trending. Classic example is where CPI_t and Emp_t are both unrelated $I(1)$ series (random walk or random walk with drift). Running a regression of CPI_t on Emp_t (spurious) has a large R^2 and a seemingly significant slope parameter. The problem is that here the error term is $I(1)$ as well and statistical inference is invalid. The regression is meaningless!
- [Another example of a spurious regression is where two completely unrelated trend stationary series may give the impression that they are related purely because of the omission of the "time" variable (OVB). Once account is taken of the deterministic trend, the regression will be able to detect that there is no relation between X and Y. By omitting the time trend in the regression, the parameter on X will capture their joint trending behaviour).]

- If, on the other hand, we have cointegration than we do have a long run relationship between our I(1) variables Emp_t and CPI_t. We have a linear combination of I(1) variables that becomes stationary I(0). For example, if Emp_t and CPI_t are I(1) process, but $CPI_t - \beta Emp_t$ is stationary. Then we can say X and Y are cointegrated, meaning there is a long term relationship between these two. In that case, there will be an ECM representation (b).

To test for co-integration, we can test for a unit root in the residual.

- We run the above equation to obtain the residuals:

$$\hat{u}_t = CPI_t - \hat{\beta}_0 - \hat{\beta}_1 CPI_{t-1} - \hat{\beta}_2 Emp_t - \hat{\beta}_3 Emp_{t-1}$$

  Next we use a ADF specification, given by:

$$\Delta\hat{u}_t = \rho\hat{u}_{t-1} + \delta\Delta\hat{u}_{t-1} + e_t$$

  where by having added sufficient lagged differences we ensure that e_t does not exhibit autocorrelation. We now test $H_0 : \rho = 0$ (spurious, nonstationary) against $H_1 : \rho < 0$ (cointegrating, stationary). The test statistic is $\frac{\hat{\rho}}{se(\hat{\rho})}$, and we reject at the 5% level of significance if this value is smaller than the critical value of the Engle-Granger Test (so not exactly equal to those in the DF or ADF test)..

(b) If there is an ECM model for X and Y, then X and Y must be cointegrated and the test conducted in (b) should have rejected the null. Candidates are expected to use the hints to derive the ECM, which can be written as:

$$\Delta CPI_t = (\beta_1 - 1)\left(CPI_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta_1}Emp_{t-1}\right) + \beta_2\Delta Emp_t + u_t,$$

where

- $CPI_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta 1}Emp_{t-1}$ is the error correction term, measuring the disequilibrium at time $t-1$.
- $(\beta_1 - 1)$ indicates how fast the two process are coming back to the steady state
- β_2 is the direct effect a change in Emp has on CPI

(c) The error correction model can be estimated by two steps:

- Firstly, we estimate the cointegating relationship between X and Y by OLS to obtain (need the residuals!)

$$\widehat{diseq}_{t-1} = CPI_{t-1} - \hat{\gamma}_1 - \hat{\gamma}Emp_{t-1}$$

- Next, using $\widehat{diseq}_{t-1}$ we run OLS on the error correction equation

$$\Delta CPI_t = (\beta_1 - 1)\widehat{diseq}_{t-1} + \beta_2\Delta Emp_t + u_t.$$

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	**2·55**	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	**2·60**	0·99534	**2·75**	0·99702	**2·90**	0·99813	**3·05**	0·99886	**3·20**	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	**2·65**	0·99598	**2·80**	0·99744	**2·95**	0·99841	**3·10**	0·99903	**3·25**	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918	**3·30**	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	**3·263**	0·9994	**3·731**	0·99990	**3·916**	0·99995
3·105	0·9991	**3·320**	0·9995	**3·759**	0·99991	**3·976**	0·99996
3·138	0·9992	**3·389**	0·9996	**3·791**	0·99992	**4·055**	0·99997
3·174	0·9993	**3·480**	0·9997	**3·826**	0·99993	**4·173**	0·99998
3·215	0·9994	**3·615**	0·9998	**3·867**	0·99994	**4·417**	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	**5·0**	1·6449	**3·0**	1·8808	**2·0**	2·0537	**1·0**	2·3263	**0·10**	3·0902
45	0·1257	**4·8**	1·6646	**2·9**	1·8957	**1·9**	2·0749	**0·9**	2·3656	**0·09**	3·1214
40	0·2533	**4·6**	1·6849	**2·8**	1·9110	**1·8**	2·0969	**0·8**	2·4089	**0·08**	3·1559
35	0·3853	**4·4**	1·7060	**2·7**	1·9268	**1·7**	2·1201	**0·7**	2·4573	**0·07**	3·1947
30	0·5244	**4·2**	1·7279	**2·6**	1·9431	**1·6**	2·1444	**0·6**	2·5121	**0·06**	3·2389
25	0·6745	**4·0**	1·7507	**2·5**	1·9600	**1·5**	2·1701	**0·5**	2·5758	**0·05**	3·2905
20	0·8416	**3·8**	1·7744	**2·4**	1·9774	**1·4**	2·1973	**0·4**	2·6521	**0·01**	3·7190
15	1·0364	**3·6**	1·7991	**2·3**	1·9954	**1·3**	2·2262	**0·3**	2·7478	**0·005**	3·8906
10	1·2816	**3·4**	1·8250	**2·2**	2·0141	**1·2**	2·2571	**0·2**	2·8782	**0·001**	4·2649
5	1·6449	**3·2**	1·8522	**2·1**	2·0335	**1·1**	2·2904	**0·1**	3·0902	**0·0005**	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables

© University of London 2025

UNIVERSITY OF LONDON

EC2020 Elements of Econometrics

Candidates should answer **EIGHT** of the following **NINE** questions: all **FIVE** from Section A (8 marks each) and **THREE** from Section B (20 marks each). Candidates are strongly advised to divide their time accordingly. If more than EIGHT questions are answered, only the first EIGHT questions attempted will be counted.

Please find the questions on the following page.

© University of London 2025

© University of London 2025

SECTION A

Answer all questions from this section. Each question carries 8 marks.

1. Consider the following simplified production function for a sample of firms i:

$$output_i = \alpha_0 + \alpha_1 capital_i + \alpha_2 labour_i + \alpha_3 labour_i \times capital_i + u_i, \qquad i = 1, ..., n \quad (1.1)$$

where $output_i$ denotes the total products of firm i, and $capital_i$ and $labour_i$ denote the amount of capital (machines) and workers employed by the firm. Assume that we obtain a random sample of n firms and the error terms u_i are i.i.d with zero mean and variance δ^2. We also assume that u_i and the regressors are independent.

(a) **(4 marks)** Discuss the economic and econometric rationales for including $labour_i \times capital_i$ in this model. What is the expected sign for α_3?

(c) **(4 marks)** Discuss the problem of excluding $labour_i \times capital_i$ from the model to an OLS estimation of (1.1).

Approaching the question

(a) Candidates are expected to discuss use of the interaction term to see the differential effect of either labour (or capital) on output by the different levels of capital (labour). This is to capture the non-linearity of the effect. Best answer would derive the interpretation.

(b) There will be an OVB of excluding this important variable. Candidates are expected to use the OVB formula to reason about the sign of the biases here.

2. Consider the following bivariate regression model

$$y_i = \alpha + \beta z_i + u_i, \quad (2.1)$$

for $i = 1, ..., n$. The regressor z_i is non-negative, *stochastic*, and has a non-zero finite variance, i.e., $\sum_{i=1}^{n}(z_i - \bar{z})^2 \neq 0$. The error term u_i is i.i.d. with mean zero and constant variance of σ^2, and independent of z_i. Given a random sample of observations $\{(y_i, z_i)\}_{i=1}^{n}$ from the population of interest, consider the following estimator:

$$\tilde{\beta} = \frac{\sum_{i=1}^{n} z_i^{2025} y_i}{\sum_{i=1}^{n} z_i^{2026}}. \quad (2.2)$$

(a) **(3 marks)** Derive the condition under which $\tilde{\beta}$ is a consistent estimator of β?

(b) **(5 marks)** Let us assume that the condition in (a) is satisfied so that $\tilde{\beta}$ is consistent. Will this condition also ensure that $\tilde{\beta}$ is unbiased?

Approaching the question

© University of London 2025

(a) For consistency, we need to show that $\operatorname{plim}\tilde{\beta} = \beta$. We first simplify $\tilde{\beta}$ as follows:

$$\begin{aligned}
\tilde{\beta} &= \frac{\sum_{i=1}^{n} z_i^{2025} y_i}{\sum_{i=1}^{n} z_i^{2026}} \\
&= \frac{\sum_{i=1}^{n} z_i^{2025}(\alpha + \beta z_i + u_i)}{\sum_{i=1}^{n} z_i^{2026}} \quad \text{replacing } y_i \text{ from the model} \\
&= \beta + \alpha\left(\frac{\sum_{i=1}^{n} z_i^{2025}}{\sum_{i=1}^{n} z_i^{2026}}\right) + \frac{\sum_{i=1}^{n} z_i^{2025} u_i}{\sum_{i=1}^{n} z_i^{2026}}
\end{aligned}$$

Taking the plim, we get:

$$\begin{aligned}
\text{plim}(\tilde{\beta}) &= \beta + \operatorname{plim}\alpha\left(\frac{\sum_{i=1}^{n} z_i^{2025}}{\sum_{i=1}^{n} z_i^{2026}}\right) + \text{plim}\left(\frac{\sum_{i=1}^{n} z_i^{2025} u_i}{\sum_{i=1}^{n} z_i^{2026}}\right) \\
&= \beta + \alpha\left(\frac{\operatorname{plim} 1/n \sum_{i=1}^{n} z_i^{2025}}{\operatorname{plim} 1/n \sum_{i=1}^{n} z_i^{2026}}\right) + \frac{\text{plim}\left(\frac{1}{n}\sum_{i=1}^{n} z_i^{2025} u_i\right)}{\text{plim}\left(\frac{1}{n}\sum_{i=1}^{n} z_i^{2026}\right)} \quad \text{properties of plim} \\
&= \beta + \alpha\frac{E(x^{2025})}{E(x^{2026})} + \frac{E\left(z_i^{2025} u_i\right)}{E\left(z_i^{2026}\right)} \quad \text{from the Law of Large Numbers} \\
&= \beta + \alpha\frac{E(x^{2025})}{E(x^{2026})} \quad \text{as } E(xu_i) = 0
\end{aligned}$$

At this point, for the estimator to be consistent, we would need $\alpha = 0$ because $E(z^{2025}) = 0$ is only when all of the zs are zero, contradicting the assumption. So the condition is $\alpha = 0$.

(b) When the estimator is consistent, we must have $\alpha = 0$. From (a), unbiasedness, we need to show that $E(\tilde{\beta}) = \beta$. The expectation of $\tilde{\beta}$ is given by:

$$\begin{aligned}
E(\tilde{\beta}) &= E\left(\frac{\sum_{i=1}^{n} z_i^{2025} y_i}{\sum_{i=1}^{n} z_i^{2026}}\right) = E\left(\frac{\sum_{i=1}^{n} z_i^{2025}(\alpha + \beta z_i + u_i)}{\sum_{i=1}^{n} z_i^{2026}}\right) \\
&= E\left[\frac{\alpha\left(\sum_{i=1}^{n} z_i^{2025}\right) + \beta\left(\sum_{i=1}^{n} z_i^{2026}\right) + \sum_{i=1}^{n} z_i^{2025} u_i}{\sum_{i=1}^{n} z_i^{2026}}\right] \\
&= \beta + \alpha\left(\frac{\sum_{i=1}^{n} z_i^{2025}}{\sum_{i=1}^{n} z_i^{2026}}\right) + \frac{\sum_{i=1}^{n} z_i^{2025} E(u_i)}{\sum_{i=1}^{n} z_i^{2026}} \\
&= \beta + \alpha\left(\frac{\sum_{i=1}^{n} z_i^{2025}}{\sum_{i=1}^{n} z_i^{2026}}\right) \quad \text{because } E(z_i u_i) = 0.
\end{aligned}$$

Thus, in general $\tilde{\beta}$ is a biased estimator unless the population intercept $\alpha = 0$. Best candidates would realise that when $\sum_{i=1}^{n} z_i^{2025}$ cannot be zero because of z_i being non-negative and sample variability.

Therefore, when $\tilde{\beta}$ is an consistent estimator, it is also unbiased.

3. Consider the following demand curve for public transport in London in year t

$$\ln(Q_t) = \beta_0 + \beta_1 \ln(P_t) + \beta_2 \ln(Y_t) + u_t \tag{3.1}$$

where Q_t and P_t are the quantity (number of rides) and the price of public transport (per ride) in London in year t, and Y_t is the average income in London in year t.

(a) **(2 marks)** What is the interpretation for β_2? Briefly explain.

(b) **(2 marks)** You are told that price and quantity in the public transport market are simultaneously determined in a market equilibrium. Briefly discuss how this piece of information poses problems to the OLS estimator for β_1.

(c) **(4 marks)** You plan to use the average income in the previous year Y_{t-1} to solve the problems in (b). Briefly discuss if this idea is sound.

Approaching the question

(a) 1% increase in income is associated with β_1 percent change in the quantity. It's the income elasticity of demand for public transport.

(b) There will be simultaneity bias, that is the error term is correlated with P_t. OLS estimator will be biased and inconsistent.

(c) We may want to use the variable as an IV. (1 mark), but it won't be really valid. Relevance can be weakly satisfied. But it won't be exogenous (probably related to demand for public transport, or not random at all). Candidates can be creative here, as long as discussing the three conditions (that this may be actually relevant.)

4. Consider a linear regression model of time series variables:

$$y_t = \beta_1 z_t + \beta_2 z_{t-1} + u_t, \qquad t = 1, ..., T \tag{4.1}$$

where the error terms u_t are independent of z_t and z_{t-1}, y_t and z_t are covariance stationary and weakly dependent processes.

(a) **(3 marks)** Explain the concepts of weak dependence. Why is it being assumed in the context of time series analysis?

(b) **(5 marks)** Discuss a suitable test for autocorrelation in the error term. Be clear about the null hypothesis and the assumptions needed. What should we do if we reject the test's null hypothesis?

Approaching the question

(a) A process is weakly dependent if the correlation between two random temporal values of the process will disappear as the time distance between them increases: $Corr(z_t, z_{t+h}) \to 0$ as $h \to \infty$. This assumption is important because together with stationarity it replaces the assumption of independence of the sample distribution so that the LLN and the CLT hold: asymptotic independence of the process. That is, it allows for correlation within the time series, but the correlation will disappear if the time distance gets large.

(b) Candidates are expected to perform an autocorrelation test. Assume that u_t can be expressed as follows:

$$u_t = \rho u_{t-1} + v_t, \quad \text{with } |\rho| < 1,$$

and v_t is a white noise generated randomly from a normal distribution with mean zero and constant variance. Since we only assume weak exogeneity here, it's safe to include all the regressors in the autocorrelation test. To test for autocorrelation, the null hypothesis is $H_o : \rho = 0$ against $H_A : \rho \neq 0$. The testing equation is obtained from using the OLS residuals $\hat{u}_t$ of (4.1): $\hat{u}_t = y_t - (\hat{\beta}_1 z_t + \hat{\beta}_2 z_{t-1})$. The testing equation is:

$$\hat{u}_t = \gamma_0 + \gamma_1 z_t + \gamma_2 z_{t-1} + \rho \hat{\varepsilon}_{t-1} + u_t.$$

- We must use the asymptotic t-statistics $\frac{\hat{\rho}}{s.e(\hat{\rho})}$ for a large sample.
- At the 5% level of significance, we should reject if $|\frac{\hat{\rho}}{s.e(\hat{\rho})}| > 1.96$, which follows a $N(0,1)$ distribution).
- Students can instead use the LM test statistics $(T-1)R^2 \sim \chi^2_1$ (only mentioned but not discuss in the Subject Guide).
- If we reject the null hypothesis, we should be concerned about the presence of autocorrelation. We can use the HAC robust standard error (not robust alone) to fix the standard errors.

5. An economist considers a Probit model to explain the mortgage application approval decision.

(a) **(4 marks)** Discuss intuitively how a Probit model is estimated. There is no need to provide technical details nor formulas and the discussion should be brief (fewer than 150 words).

(b) **(4 marks)** A colleague suggests it may be useful to use age and age squared to capture the potential U-shaped relationship of age and mortgage approval. Do you agree? Explain.

Approaching the question

(a) A Probit model estimates the relationship between a binary outcome and explanatory variables by linking the probability of the outcome to a latent variable that follows a normal distribution. The model uses the normal cumulative distribution function (CDF) to ensure probabilities stay between 0 and 1. Estimation adjusts coefficients to make predicted probabilities align with observed outcomes, typically using maximum likelihood. This approach captures non-linear effects, where changes in predictors impact probabilities most around 0.5.

© University of London 2025

(b) It is tempting but not necessary. The Probit model already captures the non-linear relationship. Best answers can use a simple derivation of the Probit to show this. Top students will note that doing so will complicate the interpretation further so no, it should not be included.

© University of London 2025

SECTION B

Answer three questions from this section. Each question carries 20 marks.

6. Central Banks are paying attention to economic indicators in their fight against inflation. An economist is interested in the following ADL(1,1) relationship between the monthly inflation rate inf_t and the monthly consumer sentiment sen_t (the percentage of surveyed respondents who are positive about the economy).

$$infl_t = \beta_0 + \beta_1 infl_{t-1} + \beta_2 sen_t + \beta_3 sen_{t-1} + u_t, \quad |\beta_1| < 1 \tag{6.1}$$

where sen_t and $infl_t$ are covariance stationary variables. The error terms u_t are an i.i.d. $(0, \sigma_u^2)$ and independent of $infl_{t-1}, infl_{t-2}, infl_{t-3}$..... We also assume that u_t is independent of sen_t for all t.

(a) **(7 marks)** What does it mean by "covariance stationary"? How is it different from "difference stationary" and "trend stationary"? Briefly explain the concepts and why it is important to distinguish them.

(b) **(5 marks)** Discuss the statistical properties your OLS estimators for the ADL(1,1) parameters. In particular: (i) are the estimators unbiased and consistent?, and (ii) should we use robust standard errors? Support your answers with appropriate technical derivations.

(c) **(5 marks)** Show that when we exclude the relevant variable $infl_{t-1}$ from (6.1), the new model will suffer from autocorrelation. Explain your answer.
Hint: You are expected to reformulate the model as:

$$infl_t = \beta_0 + \beta_2 sen_t + \beta_3 sen_{t-1} + v_t, \tag{6.2}$$

where the new error term is $v_t = \beta_1 infl_{t-1} + u_t$.

(d) **(3 marks)** What are the consequences of excluding $infl_{t-1}$ to an OLS estimation of (6.2). Propose a solution to the resulting problem(s).

Approaching the question

(a) Covariance stationarity is when a series has the first two moments do not change over time $(E(x_t), Var(x_t))$ and the covariance between two values x_t and x_{t+h} depends only on the distance in time (h) but not in time. We need this assumption, together with weak dependence, to replace the i.i.d assumption and to use LLT and CLN.
Difference stationarity is completely different, which means the series is not stationary and can only be made stationary by taking the first difference. Trend stationarity means the series can be made stationary by taking away the trend.
It is important to distinguish the two because the former is highly persistent while the latter is weakly dependent. For the former, shocks or policy changes will have last-

ing/permanent effect, while for the latter the effect is transitory. Therefore it is important to distinguish the two to avoid estimating meaningless and spurious relationship due to highly persistent behaviour.

(b) This question is a classic discussion on the properties of estimation with lagged dependent variables.

- Because of the presence of $infl_{t-1}$ in the above model, we can not establish unbiasedness because the strict exogeneity assumption is violated (u_t and $infl_t$ are correlated).
- The assumptions for consistency are satisfied because for this we only need to ensure that u_t is uncorrelated with the regressors $infl_{t-1}, sen_t$ and sen_{t-1}. We are given that the emp variables are independent of the errors so they are uncorrelated. As long as there is no autocorrelation, we will also not have any correlation between $infl_{t-1}$ and u_t.
- We don't need to use robust standard errors as there is no heteroskedasticity or autocorrelation.

(c) Candidates must annotate the derivations clearly for the full mark. The key is to show $Cov(v_{t+1}, v_t) \neq 0$.

- Using our definition of v_t and v_{t+1}, we note

$$\begin{aligned} Cov(v_{t+1}, v_t) &= Cov\left(\beta_1 infl_t + u_{t+1}, \beta_1 infl_{t-1} + u_t\right) \\ &= \beta_1^2 Cov\left(infl_t, infl_{t-1}\right) + \beta_1 Cov\left(infl_t, u_t\right) \\ &\quad + \beta_1 Cov\left(u_{t+1}, infl_{t-1}\right) + Cov\left(u_{t+1}, u_t\right) \end{aligned}$$

- Because u_t is white noise (i.i.d errors unrelated to anything in the past) $Cov\left(u_{t+1}, u_t\right) = 0$ and $Cov\left(u_{t+1}, infl_{t-1}\right) = 0$. e
- To show that $Cov(v_{t+1}, v_t) \neq 0$ it is sufficient to point out that $Cov\left(infl_t, u_t\right) \neq 0$. This follows quickly when noting that as our model is given by

$$infl_t = \beta_0 + \beta_1 infl_{t-1} + \beta_2 sen_t + \beta_3 sen_{t-1} + u_t, \quad |\beta_1| < 1$$

  which shows a direct relation between $infl_t$ and u_t, and hence $Cov(infl_t, u_t) \neq 0$. [Clearly $Cov(infl_t, infl_{t-1}) \neq 0$ as $\beta_1 \neq 0$.]

(d) Now there are two problems with the exclusion: autocorrelation, which can be solved by using HAC, but more importantly, OVB due to the exclusion of an important variable. The most potential solution is to use an IV (for example further lags of $infl_t$ for the sen_t variables, or some other external variables (supply-side shocks)). This is a difficult part and reserved for the top candidates.

7. An economist investigates how part-time work affects students' performance in an online programme. Using the data from surveying 894 students participating in the programme, she runs the following regression, using the average of overall grades from all taken courses, $Grade_i$, (marks range from 0 - 100) as the outcome variable. The regressors include the self-reported number of self-study hours per week in logs, $\log selfstudy$; the number of hours students engaged in part-time work per week, $work_i$; and the number of concurrent courses the student took ($othercourses$).

$$\widehat{Grade_i} = \underset{(7.14)}{38.3} + \underset{(.23)}{0.51}\log(selfstudy_i) - \underset{(4.56)}{2.07}work_i - \underset{(3.25)}{4.01}othercourses_i, \qquad (7.1)$$
$$n = 894, \qquad R^2 = .355.$$

The robust standard errors are reported in parentheses.

(a) **(5 marks)** What is the interpretation of the coefficient on $work_i$? Are the sign and size of the coefficient as you would expect? Do you think the effect is qualitatively large or small? Explain using the available information in (7.1).

(b) **(5 marks)** A colleague believes that the grade distribution is more varied for those who took many additional courses than those who only took a few. Discuss a statistical test for this claim. What are the consequences to the regression (7.1) if she is correct?

(c) **(5 marks)** Another colleague says that since the coefficient of $othercourses_i$ and $work$ are insignificant, including the variable is harmful and the team should remove the variable from the regression. Critically discuss the suggestion and propose a suitable test to see if the colleague is right.

(d) **(5 marks)** The colleague also suggests that students tend to misreport the self-study hours to the survey. Intuitively discuss the consequences on your answer in (a) if the colleague is correct.

Approaching the question

(a) Candidates are expected to provide economic interpretations: one extra hour of part-time work is associated with on average 2.07 marks lowered. This seems to be as expected. But a t-test shows that this is not significant, so maybe part-time work is not a concern. To evaluate the magnitude, with 1 hour of self-study and not taking any other course nor work, the student will earn 38.3, still a fail. So if they work more, probably this can be detrimental. Candidates can be creative here with the intuition.

(b) Candidates are expected discuss clearly the technical definition of heteroskedasticity (the error term variance depends on some or all of the regressors), and also the intuition: the GPA would be more or less variable for students who do more self-study, or part-time work than those who do less self-study/part-time work.
Candidates are expected to explain the test for heteroskedasticity: either the Breusch-Pagan version or the White version of the test.

© University of London 2025

- **For the Breusch-Pagan test:** we should
  - Save the OLS residuals as $\hat{u}_i$. Compute the squared residuals $\hat{u}_i^2$.
  - Regress $\hat{u}_i^2$ on all the regressors, with an intercept, and then compute the F test of joint significance of the regressors.

$$\hat{u}_i^2 = \delta_0 + \delta_1 \log selfstudyi + \delta_2 work_i + \delta_3 othercourses + v_i.$$

  - The null hypothesis is: $H_0 : \delta_1 = \delta_2 = \delta_3 = 0$. Students are expected to discuss the R-squared version or the SSR version of the F-statistic in this case. $k = 3, n - k - 1$.
  - We do not need to assume MLR.6 in this version of the test, under the null hypothesis, the F-stat will follow the asymptotic F distribution.
- **For the White test,** we should
  - Save the OLS residuals as $\hat{u}_i$. Compute the squared residuals $\hat{u}_i^2$.
  - Regress $\hat{u}_i^2$on all the regressors, with an intercept, and the squares and interactions of all the regressors:

$$\begin{aligned}\hat{u}_i^2 &= \delta_0 + \delta_1 \log selfstudy_i + \delta_2 worki + \delta_3 othercourses + \delta_4 \log selfstudy_i^2 + ... \\ &\quad + \delta_7 \log selfstudyi \times work_i + ... + v_i.\end{aligned}$$

    The null hypothesis is: $H_0 : \delta_1 = \delta_2 = \delta_3 = .. = \delta_9 = 0$. We can simply obtain the F-test of the joint significance of the regressors. Similarly, students are expected to discuss the R-squared version or the SSR version of the F-statistic. $k = 9, n - k - 1$.
  - We do not need to assume MLR.6 in this version of the test, under the null hypothesis, the F-stat will follow the asymptotic F distribution.
- If we reject the null hypothesis, there is evidence of heteroskedasticity.The normal standard errors are invalid. But we already use robust standard errors so there is no concern.

(c) The colleague is right that the coefficients are insignificant (candidates are expected to briefly mention the t-stat or the large standard error here). However, including an irrelevant variable is not particularly harmful, it also increases the standard errors of the other estimates, which in this case are still statistically significant at 5% level. So the effect of such inclusion is minimal. Intuitively, excluding the variable is potentially more harmful as having other courses may actually be an important factor in a student performance. So we should keep that variable in. Same for working a part-time job. One way to test if we should drop these variables is to conduct an F-test of joint significance of these variables. Candidates can briefly discuss the restricted and unrestricted versions of the testing equation.

© University of London 2025

(d) Candidates are expected to discuss the errors-in-variables situation, at least, if the self-reported numbers are randomly misreported: the errors may come from students misremembering their self-study in a random manner. If that's the case, we will have attenuation bias, which bias the estimate of self-study towards zero. However, we do not know the direction of bias towards other estimators. So the answers in a) won't be valid. Best students may observe that the estimator in (a) is already insignificant so probably the attenuation bias won't affect much there as well.

8. We are interested in the relationship between the quarterly price level CPI_t and the quarterly employment level Emp_t in the following model:

$$CPI_t = \alpha_0 + \alpha_1 Emp_t + u_t, \quad t = 1, ..., T, \tag{8.1}$$

where Emp_t and CPI_t are both $I(1)$ variables. Assume that the error terms exhibit autocorrelation and follow an AR(1) structure as:

$$u_t = \rho u_{t-1} + v_t, |\rho| < 1, \tag{8.2}$$

where v_t is an i.i.d. innovation with zero mean and constant variance, and is independent of u_{t-1}.

(a) **(5 marks)** Explain what it means to say CPI_t is an $I(1)$ variable. Discuss how you would test for it. In your answer, specify the test equation, test statistic, and assumptions.

(b) **(5 marks)** A colleague suggests to test for co-integration before proceeding with analysing the data. Discuss the distinction between co-integration and spurious relationship, and why it is important to test for the co-integration.

(c) **(5 marks)** Discuss a test for evidence of a co-integrating relationship between CPI and Emp. *Hints:* we can use (8.2) to get rid of the autocorrelation in (8.1).

(d) **(5 marks)** A colleague ascertains that there exists an Error Correction Model (ECM) that describes the short-run and long-run dynamics between CPI_t and Emp_t. Derive the Error Correction Model and interpret the various components of the ECM. Discuss how this claim relates to your answer in (b). *Hints:* Your aim is to obtain the following regression specification:

$$\Delta CPI_t = \delta_0 + \delta_1 \Delta Emp_t + \gamma s_{t-1} + \epsilon_t, \tag{8.3}$$

with s_{t-1} is the error correction term.

Approaching the question

(a) When we can make a process stationary by differencing a process once, we say that we have a process that is integrated of order 1, or I(1). CPI is a classic example of an I(1) process. To perform a test, we need a DF or ADF test in this case. For example, assume CPI follows an AR(1) structure, which is reasonable. The testing equation of the D-F test requires the estimation of:

$$\Delta CPI_t = \pi_0 + \pi_1 CPI_{t-1} + e_t,$$

We will test whether $\pi_1 = 0$ against $H_1 : \pi_1 < 0$, using the tau-statistic: $\hat{\pi_1}/se(\hat{\pi_1})$. At the 5% significance level, we need to reject if the tau-statistic is larger than the DF critical value (not the t), which is about -2.86.

(b) The key to the rest of the answers is to realise that we can lag one period of (8.1) to obtain CPI_{t-1}, u_{t-1} and then substitute (8.2) into (8.1) to obtain the following general form:

$$CPI_t = \beta_0 + \beta_1 CPI_{t-1} + \beta_2 Emp_t + \beta_3 Emp_{t-1} + u_t, \quad t = 1, ..., T, \tag{8.4}$$

Candidates are then expected to discuss the two concepts, and then talk about the ADF test for co-integration (named the Engle-Granger test).

- Spurious regression refers to the scenario where <u>two unrelated series</u> give the impression that they are related, purely because of the fact that both are trending. Classic example is where CPI_t and Emp_t are both unrelated $I(1)$ series (random walk or random walk with drift). Running a regression of CPI_t on Emp_t (spurious) has a large R^2 and a seemingly significant slope parameter. The problem is that here the error term is $I(1)$ as well and statistical inference is invalid. The regression is meaningless!
- [Another example of a spurious regression is where two completely unrelated trend stationary series may give the impression that they are related purely because of the omission of the "time" variable (OVB). Once account is taken of the deterministic trend, the regression will be able to detect that there is no relation between X and Y. By omitting the time trend in the regression, the parameter on X will capture their joint trending behaviour).]
- If, on the other hand, we have cointegration than we do have a long run relationship between our I(1) variables Emp_t and CPI_t. We have a linear combination of I(1) variables that becomes stationary I(0). For example, if Emp_t and CPI_t are I(1) process, but $CPI_t - \beta Emp_t$ is stationary. Then we can say X and Y are cointegrated, meaning there is a long term relationship between these two. In that case, there will be an ECM representation (c).

(c) To test for co-integration, we can test for a unit root in the residual. Candidates are not required to discuss the following details, only briefly.

- We run the above equation to obtain the residuals:

$$\hat{u}_t = CPI_t - \hat{\beta}_0 - \hat{\beta}_1 CPI_{t-1} - \hat{\beta}_2 Emp_t - \hat{\beta}_3 Emp_{t-1}$$

Next we use a ADF specification, given by:

$$\Delta \hat{u}_t = \rho \hat{u}_{t-1} + \delta \Delta \hat{u}_{t-1} + e_t$$

where by having added sufficient lagged differences we ensure that e_t does not exhibit autocorrelation. We now test $H_0 : \rho = 0$ (spurious, nonstationary) against $H_1 : \rho < 0$ (cointegrating, stationary). The test statistic is $\frac{\hat{\rho}}{se(\hat{\rho})}$, and we reject at the 5% level of significance if this value is smaller than -2.86 (the actual Engle-Granger critical value is slight different from this DF number).

© University of London 2025

(d) If there is an ECM model for X and Y, then X and Y must be cointegrated and the test conducted in (b) should have rejected the null. Candidates are expected to use the hints to derive the ECM, which can be written as:

$$\Delta CPI_t = (\beta_1 - 1)\left(CPI_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta_1}Emp_{t-1}\right) + \beta_2 \Delta Emp_t + v_t,$$

where

- $CPI_{t-1} - \frac{\beta_0}{1-\beta_1} - \frac{\beta_2+\beta_3}{1-\beta 1}Emp_{t-1}$ is the error correction term, measuring the disequilibrium at time $t-1$.
- $(\beta_1 - 1)$ indicates how fast the two process are coming back to the steady state
- β_2 is the direct effect a change in Emp has on CPI

9. Consider a random sample $X_1, X_2, \ldots, X_n$ drawn from a distribution having the following probability density function with parameter θ:

$$f_X(x;\theta) = 3\theta^2 x e^{-\theta x^3}, \quad x > 0, \quad \theta > 0. \tag{9.1}$$

The following property of the distribution is given:

$$\mathbb{E}[X] = \theta. \tag{9.2}$$

(a) **(8 marks)** Derive the Maximum Likelihood Estimator (MLE) for θ, $\hat{\theta}_n$. Be clear of the steps that you take.

(b) **(6 marks)** Is the MLE $\hat{\theta}_n$ asymptotically unbiased? Is it consistent? Prove your claim. In your answer, clearly explain the key difference between unbiasedness and consistency.

(c) **(6 marks)** Is the estimator $\widehat{\theta}_c = (n+1)^{-1} \sum_{i=1}^{n} X_i$ unbiased? Is it consistent? Support your answers with appropriate derivations.

Approaching the question

(a) The likelihood function for a random sample $X_1, X_2, \ldots, X_n$ is the product of the individual probability density functions:

$$L(\theta) = \prod_{i=1}^{n} f_X(X_i;\theta) = \prod_{i=1}^{n} 3\theta^2 X_i e^{-\theta X_i^3}.$$

Simplifying the likelihood function:

$$L(\theta) = \left(\prod_{i=1}^{n} 3\theta^2 X_i\right)\left(\prod_{i=1}^{n} e^{-\theta X_i^3}\right).$$

This becomes:

$$L(\theta) = (3\theta^2)^n \left(\prod_{i=1}^{n} X_i\right) e^{-\theta \sum_{i=1}^{n} X_i^3}.$$

Thus, the likelihood function is:

$$L(\theta) = (3\theta^2)^n \left(\prod_{i=1}^{n} X_i\right) e^{-\theta \sum_{i=1}^{n} X_i^3}.$$

To find the MLE, we first take the log-likelihood:

$$\ell(\theta) = \log L(\theta) = n\log(3\theta^2) + \sum_{i=1}^{n} \log X_i - \theta \sum_{i=1}^{n} X_i^3.$$

Now, differentiate the log-likelihood with respect to θ and set it equal to zero:

$$\frac{\partial \ell(\theta)}{\partial \theta} = \frac{2n}{\theta} - \sum_{i=1}^{n} X_i^3 = 0.$$

Solving for θ:

$$\hat{\theta} = \frac{2n}{\sum_{i=1}^{n} X_i^3}.$$

Thus, the MLE for θ is:

$$\hat{\theta} = \frac{2n}{\sum_{i=1}^{n} X_i^3}.$$

(b) The unbiasedness is rather simple. The expected value of the estimator is simply:

$$E[\hat{\theta}] = E\left[\frac{2n}{\sum_{i=1}^{n} X_i^3}\right].$$

Clearly this is different from θ. So it is biased.

The plim is rather disappointing too. It's inconsistent as well.

Candidates then need to clearly define and contrast unbiasedness and consistency for full marks: unbiasedness (repeated samples, so that the expected value of the estimate from the estimator would be the parameter) and consistency (increasing sample size, the estimate would approach in probability the parameter) concepts.

(c) We want to look at property of $\hat{\theta}_{c^*} = \frac{n}{n+1}\overline{X} = (n+1)^{-1}\sum\limits_{i=1}^{n} X_i$.

The estimator is biased: $E\left(\hat{\theta}_{c^*}\right) = (n+1)^{-1}\sum\limits_{i=1}^{n} EX_i = \frac{n}{n+1}\theta \neq \theta$.

Candidates can use sufficient conditions to prove consistency. That is, $\lim_{n\to\infty} E\left(\hat{\theta}_{c^*}\right) = \theta$ and $\lim_{n\to\infty} Var\left(\hat{\theta}_{c^*}\right) = \lim_{n\to\infty}\left((\frac{n}{n+1})^2\theta^2/n\right) = \lim_{n\to\infty}\left((\frac{n}{(n+1)^2}\theta^2\right) = 0$, or alternatively show $\operatorname{plim}\hat{\theta}_{c^*} = \operatorname{plim}\frac{n}{n+1} \times \operatorname{plim}\overline{X} = 1.\theta = \theta$ where LLN ensures that $\operatorname{plim}\bar{X} = E(X_i) = \theta$.

END OF PAPER

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

The function tabulated is $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$. $\Phi(x)$ is the probability that a random variable, normally distributed with zero mean and unit variance, will be less than or equal to x. When $x < 0$ use $\Phi(x) = 1 - \Phi(-x)$, as the normal distribution with zero mean and unit variance is symmetric about zero.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0·00	0·5000	0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·97725
·01	·5040	·41	·6591	·81	·7910	·21	·8869	·61	·9463	·01	·97778
·02	·5080	·42	·6628	·82	·7939	·22	·8888	·62	·9474	·02	·97831
·03	·5120	·43	·6664	·83	·7967	·23	·8907	·63	·9484	·03	·97882
·04	·5160	·44	·6700	·84	·7995	·24	·8925	·64	·9495	·04	·97932
0·05	0·5199	0·45	0·6736	0·85	0·8023	1·25	0·8944	1·65	0·9505	2·05	0·97982
·06	·5239	·46	·6772	·86	·8051	·26	·8962	·66	·9515	·06	·98030
·07	·5279	·47	·6808	·87	·8078	·27	·8980	·67	·9525	·07	·98077
·08	·5319	·48	·6844	·88	·8106	·28	·8997	·68	·9535	·08	·98124
·09	·5359	·49	·6879	·89	·8133	·29	·9015	·69	·9545	·09	·98169
0·10	0·5398	0·50	0·6915	0·90	0·8159	1·30	0·9032	1·70	0·9554	2·10	0·98214
·11	·5438	·51	·6950	·91	·8186	·31	·9049	·71	·9564	·11	·98257
·12	·5478	·52	·6985	·92	·8212	·32	·9066	·72	·9573	·12	·98300
·13	·5517	·53	·7019	·93	·8238	·33	·9082	·73	·9582	·13	·98341
·14	·5557	·54	·7054	·94	·8264	·34	·9099	·74	·9591	·14	·98382
0·15	0·5596	0·55	0·7088	0·95	0·8289	1·35	0·9115	1·75	0·9599	2·15	0·98422
·16	·5636	·56	·7123	·96	·8315	·36	·9131	·76	·9608	·16	·98461
·17	·5675	·57	·7157	·97	·8340	·37	·9147	·77	·9616	·17	·98500
·18	·5714	·58	·7190	·98	·8365	·38	·9162	·78	·9625	·18	·98537
·19	·5753	·59	·7224	·99	·8389	·39	·9177	·79	·9633	·19	·98574
0·20	0·5793	0·60	0·7257	1·00	0·8413	1·40	0·9192	1·80	0·9641	2·20	0·98610
·21	·5832	·61	·7291	·01	·8438	·41	·9207	·81	·9649	·21	·98645
·22	·5871	·62	·7324	·02	·8461	·42	·9222	·82	·9656	·22	·98679
·23	·5910	·63	·7357	·03	·8485	·43	·9236	·83	·9664	·23	·98713
·24	·5948	·64	·7389	·04	·8508	·44	·9251	·84	·9671	·24	·98745
0·25	0·5987	0·65	0·7422	1·05	0·8531	1·45	0·9265	1·85	0·9678	2·25	0·98778
·26	·6026	·66	·7454	·06	·8554	·46	·9279	·86	·9686	·26	·98809
·27	·6064	·67	·7486	·07	·8577	·47	·9292	·87	·9693	·27	·98840
·28	·6103	·68	·7517	·08	·8599	·48	·9306	·88	·9699	·28	·98870
·29	·6141	·69	·7549	·09	·8621	·49	·9319	·89	·9706	·29	·98899
0·30	0·6179	0·70	0·7580	1·10	0·8643	1·50	0·9332	1·90	0·9713	2·30	0·98928
·31	·6217	·71	·7611	·11	·8665	·51	·9345	·91	·9719	·31	·98956
·32	·6255	·72	·7642	·12	·8686	·52	·9357	·92	·9726	·32	·98983
·33	·6293	·73	·7673	·13	·8708	·53	·9370	·93	·9732	·33	·99010
·34	·6331	·74	·7704	·14	·8729	·54	·9382	·94	·9738	·34	·99036
0·35	0·6368	0·75	0·7734	1·15	0·8749	1·55	0·9394	1·95	0·9744	2·35	0·99061
·36	·6406	·76	·7764	·16	·8770	·56	·9406	·96	·9750	·36	·99086
·37	·6443	·77	·7794	·17	·8790	·57	·9418	·97	·9756	·37	·99111
·38	·6480	·78	·7823	·18	·8810	·58	·9429	·98	·9761	·38	·99134
·39	·6517	·79	·7852	·19	·8830	·59	·9441	·99	·9767	·39	·99158
0·40	0·6554	0·80	0·7881	1·20	0·8849	1·60	0·9452	2·00	0·9772	2·40	0·99180

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 4. THE NORMAL DISTRIBUTION FUNCTION

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
2·40	0·99180	**2·55**	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918
·41	·99202	·56	·99477	·71	·99664	·86	·99788	·01	·99869	·16	·99921
·42	·99224	·57	·99492	·72	·99674	·87	·99795	·02	·99874	·17	·99924
·43	·99245	·58	·99506	·73	·99683	·88	·99801	·03	·99878	·18	·99926
·44	·99266	·59	·99520	·74	·99693	·89	·99807	·04	·99882	·19	·99929
2·45	0·99286	**2·60**	0·99534	**2·75**	0·99702	**2·90**	0·99813	**3·05**	0·99886	**3·20**	0·99931
·46	·99305	·61	·99547	·76	·99711	·91	·99819	·06	·99889	·21	·99934
·47	·99324	·62	·99560	·77	·99720	·92	·99825	·07	·99893	·22	·99936
·48	·99343	·63	·99573	·78	·99728	·93	·99831	·08	·99896	·23	·99938
·49	·99361	·64	·99585	·79	·99736	·94	·99836	·09	·99900	·24	·99940
2·50	0·99379	**2·65**	0·99598	**2·80**	0·99744	**2·95**	0·99841	**3·10**	0·99903	**3·25**	0·99942
·51	·99396	·66	·99609	·81	·99752	·96	·99846	·11	·99906	·26	·99944
·52	·99413	·67	·99621	·82	·99760	·97	·99851	·12	·99910	·27	·99946
·53	·99430	·68	·99632	·83	·99767	·98	·99856	·13	·99913	·28	·99948
·54	·99446	·69	·99643	·84	·99774	·99	·99861	·14	·99916	·29	·99950
2·55	0·99461	**2·70**	0·99653	**2·85**	0·99781	**3·00**	0·99865	**3·15**	0·99918	**3·30**	0·99952

The critical table below gives on the left the range of values of x for which $\Phi(x)$ takes the value on the right, correct to the last figure given; in critical cases, take the upper of the two values of $\Phi(x)$ indicated.

3·075	0·9990	**3·263**	0·9994	**3·731**	0·99990	**3·916**	0·99995
3·105	0·9991	**3·320**	0·9995	**3·759**	0·99991	**3·976**	0·99996
3·138	0·9992	**3·389**	0·9996	**3·791**	0·99992	**4·055**	0·99997
3·174	0·9993	**3·480**	0·9997	**3·826**	0·99993	**4·173**	0·99998
3·215	0·9994	**3·615**	0·9998	**3·867**	0·99994	**4·417**	0·99999
			0·9999		0·99995		1·00000

When $x > 3{\cdot}3$ the formula $1-\Phi(x) \doteqdot \dfrac{e^{-\frac{1}{2}x^2}}{x\sqrt{2\pi}}\left[1-\dfrac{1}{x^2}+\dfrac{3}{x^4}-\dfrac{15}{x^6}+\dfrac{105}{x^8}\right]$ is very accurate, with relative error less than $945/x^{10}$.

TABLE 5. PERCENTAGE POINTS OF THE NORMAL DISTRIBUTION

This table gives percentage points $x(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{2\pi}}\int_{x(P)}^{\infty} e^{-\frac{1}{2}t^2}\,dt.$$

If X is a variable, normally distributed with zero mean and unit variance, $P/100$ is the probability that $X \geqslant x(P)$. The lower P per cent points are given by symmetry as $-x(P)$, and the probability that $|X| \geqslant x(P)$ is $2P/100$.

P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$	P	$x(P)$
50	0·0000	**5·0**	1·6449	**3·0**	1·8808	**2·0**	2·0537	**1·0**	2·3263	**0·10**	3·0902
45	0·1257	**4·8**	1·6646	**2·9**	1·8957	**1·9**	2·0749	**0·9**	2·3656	**0·09**	3·1214
40	0·2533	**4·6**	1·6849	**2·8**	1·9110	**1·8**	2·0969	**0·8**	2·4089	**0·08**	3·1559
35	0·3853	**4·4**	1·7060	**2·7**	1·9268	**1·7**	2·1201	**0·7**	2·4573	**0·07**	3·1947
30	0·5244	**4·2**	1·7279	**2·6**	1·9431	**1·6**	2·1444	**0·6**	2·5121	**0·06**	3·2389
25	0·6745	**4·0**	1·7507	**2·5**	1·9600	**1·5**	2·1701	**0·5**	2·5758	**0·05**	3·2905
20	0·8416	**3·8**	1·7744	**2·4**	1·9774	**1·4**	2·1973	**0·4**	2·6521	**0·01**	3·7190
15	1·0364	**3·6**	1·7991	**2·3**	1·9954	**1·3**	2·2262	**0·3**	2·7478	**0·005**	3·8906
10	1·2816	**3·4**	1·8250	**2·2**	2·0141	**1·2**	2·2571	**0·2**	2·8782	**0·001**	4·2649
5	1·6449	**3·2**	1·8522	**2·1**	2·0335	**1·1**	2·2904	**0·1**	3·0902	**0·0005**	4·4172

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu - 1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	99·95	99·9	99·5	99	97·5	95	90	80	70	60
$\nu = 1$	$0{\cdot}0^{6}3927$	$0{\cdot}0^{5}1571$	$0{\cdot}0^{4}3927$	$0{\cdot}0^{3}1571$	$0{\cdot}0^{3}9821$	0·003932	0·01579	0·06418	0·1485	0·2750
2	0·001000	0·002001	0·01003	0·02010	0·05064	0·1026	0·2107	0·4463	0·7133	1·022
3	0·01528	0·02430	0·07172	0·1148	0·2158	0·3518	0·5844	1·005	1·424	1·869
4	0·06392	0·09080	0·2070	0·2971	0·4844	0·7107	1·064	1·649	2·195	2·753
5	0·1581	0·2102	0·4117	0·5543	0·8312	1·145	1·610	2·343	3·000	3·655
6	0·2994	0·3811	0·6757	0·8721	1·237	1·635	2·204	3·070	3·828	4·570
7	0·4849	0·5985	0·9893	1·239	1·690	2·167	2·833	3·822	4·671	5·493
8	0·7104	0·8571	1·344	1·646	2·180	2·733	3·490	4·594	5·527	6·423
9	0·9717	1·152	1·735	2·088	2·700	3·325	4·168	5·380	6·393	7·357
10	1·265	1·479	2·156	2·558	3·247	3·940	4·865	6·179	7·267	8·295
11	1·587	1·834	2·603	3·053	3·816	4·575	5·578	6·989	8·148	9·237
12	1·934	2·214	3·074	3·571	4·404	5·226	6·304	7·807	9·034	10·18
13	2·305	2·617	3·565	4·107	5·009	5·892	7·042	8·634	9·926	11·13
14	2·697	3·041	4·075	4·660	5·629	6·571	7·790	9·467	10·82	12·08
15	3·108	3·483	4·601	5·229	6·262	7·261	8·547	10·31	11·72	13·03
16	3·536	3·942	5·142	5·812	6·908	7·962	9·312	11·15	12·62	13·98
17	3·980	4·416	5·697	6·408	7·564	8·672	10·09	12·00	13·53	14·94
18	4·439	4·905	6·265	7·015	8·231	9·390	10·86	12·86	14·44	15·89
19	4·912	5·407	6·844	7·633	8·907	10·12	11·65	13·72	15·35	16·85
20	5·398	5·921	7·434	8·260	9·591	10·85	12·44	14·58	16·27	17·81
21	5·896	6·447	8·034	8·897	10·28	11·59	13·24	15·44	17·18	18·77
22	6·404	6·983	8·643	9·542	10·98	12·34	14·04	16·31	18·10	19·73
23	6·924	7·529	9·260	10·20	11·69	13·09	14·85	17·19	19·02	20·69
24	7·453	8·085	9·886	10·86	12·40	13·85	15·66	18·06	19·94	21·65
25	7·991	8·649	10·52	11·52	13·12	14·61	16·47	18·94	20·87	22·62
26	8·538	9·222	11·16	12·20	13·84	15·38	17·29	19·82	21·79	23·58
27	9·093	9·803	11·81	12·88	14·57	16·15	18·11	20·70	22·72	24·54
28	9·656	10·39	12·46	13·56	15·31	16·93	18·94	21·59	23·65	25·51
29	10·23	10·99	13·12	14·26	16·05	17·71	19·77	22·48	24·58	26·48
30	10·80	11·59	13·79	14·95	16·79	18·49	20·60	23·36	25·51	27·44
32	11·98	12·81	15·13	16·36	18·29	20·07	22·27	25·15	27·37	29·38
34	13·18	14·06	16·50	17·79	19·81	21·66	23·95	26·94	29·24	31·31
36	14·40	15·32	17·89	19·23	21·34	23·27	25·64	28·73	31·12	33·25
38	15·64	16·61	19·29	20·69	22·88	24·88	27·34	30·54	32·99	35·19
40	16·91	17·92	20·71	22·16	24·43	26·51	29·05	32·34	34·87	37·13
50	23·46	24·67	27·99	29·71	32·36	34·76	37·69	41·45	44·31	46·86
60	30·34	31·74	35·53	37·48	40·48	43·19	46·46	50·64	53·81	56·62
70	37·47	39·04	43·28	45·44	48·76	51·74	55·33	59·90	63·35	66·40
80	44·79	46·52	51·17	53·54	57·15	60·39	64·28	69·21	72·92	76·19
90	52·28	54·16	59·20	61·75	65·65	69·13	73·29	78·56	82·51	85·99
100	59·90	61·92	67·33	70·06	74·22	77·93	82·36	87·95	92·13	95·81

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 8. PERCENTAGE POINTS OF THE χ^2-DISTRIBUTION

This table gives percentage points $\chi^2_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{2^{\nu/2}\,\Gamma(\frac{\nu}{2})} \int_{\chi^2_\nu(P)}^{\infty} x^{\frac{1}{2}\nu-1}\, e^{-\frac{1}{2}x}\, dx.$$

If X is a variable distributed as χ^2 with ν degrees of freedom, $P/100$ is the probability that $X \geqslant \chi^2_\nu(P)$.

For $\nu > 100$, $\sqrt{2X}$ is approximately normally distributed with mean $\sqrt{2\nu-1}$ and unit variance.

(The above shape applies for $\nu \geqslant 3$ only. When $\nu < 3$ the mode is at the origin.)

P	50	40	30	20	10	5	2·5	1	0·5	0·1	0·05
$\nu = 1$	0·4549	0·7083	1·074	1·642	2·706	3·841	5·024	6·635	7·879	10·83	12·12
2	1·386	1·833	2·408	3·219	4·605	5·991	7·378	9·210	10·60	13·82	15·20
3	2·366	2·946	3·665	4·642	6·251	7·815	9·348	11·34	12·84	16·27	17·73
4	3·357	4·045	4·878	5·989	7·779	9·488	11·14	13·28	14·86	18·47	20·00
5	4·351	5·132	6·064	7·289	9·236	11·07	12·83	15·09	16·75	20·52	22·11
6	5·348	6·211	7·231	8·558	10·64	12·59	14·45	16·81	18·55	22·46	24·10
7	6·346	7·283	8·383	9·803	12·02	14·07	16·01	18·48	20·28	24·32	26·02
8	7·344	8·351	9·524	11·03	13·36	15·51	17·53	20·09	21·95	26·12	27·87
9	8·343	9·414	10·66	12·24	14·68	16·92	19·02	21·67	23·59	27·88	29·67
10	9·342	10·47	11·78	13·44	15·99	18·31	20·48	23·21	25·19	29·59	31·42
11	10·34	11·53	12·90	14·63	17·28	19·68	21·92	24·72	26·76	31·26	33·14
12	11·34	12·58	14·01	15·81	18·55	21·03	23·34	26·22	28·30	32·91	34·82
13	12·34	13·64	15·12	16·98	19·81	22·36	24·74	27·69	29·82	34·53	36·48
14	13·34	14·69	16·22	18·15	21·06	23·68	26·12	29·14	31·32	36·12	38·11
15	14·34	15·73	17·32	19·31	22·31	25·00	27·49	30·58	32·80	37·70	39·72
16	15·34	16·78	18·42	20·47	23·54	26·30	28·85	32·00	34·27	39·25	41·31
17	16·34	17·82	19·51	21·61	24·77	27·59	30·19	33·41	35·72	40·79	42·88
18	17·34	18·87	20·60	22·76	25·99	28·87	31·53	34·81	37·16	42·31	44·43
19	18·34	19·91	21·69	23·90	27·20	30·14	32·85	36·19	38·58	43·82	45·97
20	19·34	20·95	22·77	25·04	28·41	31·41	34·17	37·57	40·00	45·31	47·50
21	20·34	21·99	23·86	26·17	29·62	32·67	35·48	38·93	41·40	46·80	49·01
22	21·34	23·03	24·94	27·30	30·81	33·92	36·78	40·29	42·80	48·27	50·51
23	22·34	24·07	26·02	28·43	32·01	35·17	38·08	41·64	44·18	49·73	52·00
24	23·34	25·11	27·10	29·55	33·20	36·42	39·36	42·98	45·56	51·18	53·48
25	24·34	26·14	28·17	30·68	34·38	37·65	40·65	44·31	46·93	52·62	54·95
26	25·34	27·18	29·25	31·79	35·56	38·89	41·92	45·64	48·29	54·05	56·41
27	26·34	28·21	30·32	32·91	36·74	40·11	43·19	46·96	49·64	55·48	57·86
28	27·34	29·25	31·39	34·03	37·92	41·34	44·46	48·28	50·99	56·89	59·30
29	28·34	30·28	32·46	35·14	39·09	42·56	45·72	49·59	52·34	58·30	60·73
30	29·34	31·32	33·53	36·25	40·26	43·77	46·98	50·89	53·67	59·70	62·16
32	31·34	33·38	35·66	38·47	42·58	46·19	49·48	53·49	56·33	62·49	65·00
34	33·34	35·44	37·80	40·68	44·90	48·60	51·97	56·06	58·96	65·25	67·80
36	35·34	37·50	39·92	42·88	47·21	51·00	54·44	58·62	61·58	67·99	70·59
38	37·34	39·56	42·05	45·08	49·51	53·38	56·90	61·16	64·18	70·70	73·35
40	39·34	41·62	44·16	47·27	51·81	55·76	59·34	63·69	66·77	73·40	76·09
50	49·33	51·89	54·72	58·16	63·17	67·50	71·42	76·15	79·49	86·66	89·56
60	59·33	62·13	65·23	68·97	74·40	79·08	83·30	88·38	91·95	99·61	102·7
70	69·33	72·36	75·69	79·71	85·53	90·53	95·02	100·4	104·2	112·3	115·6
80	79·33	82·57	86·12	90·41	96·58	101·9	106·6	112·3	116·3	124·8	128·3
90	89·33	92·76	96·52	101·1	107·6	113·1	118·1	124·1	128·3	137·2	140·8
100	99·33	102·9	106·9	111·7	118·5	124·3	129·6	135·8	140·2	149·4	153·2

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 10. PERCENTAGE POINTS OF THE t-DISTRIBUTION

This table gives percentage points $t_\nu(P)$ defined by the equation

$$\frac{P}{100} = \frac{1}{\sqrt{\nu\pi}} \frac{\Gamma(\frac{1}{2}\nu+\frac{1}{2})}{\Gamma(\frac{1}{2}\nu)} \int_{t_\nu(P)}^{\infty} \frac{dt}{(1+t^2/\nu)^{\frac{1}{2}(\nu+1)}}.$$

Let X_1 and X_2 be independent random variables having a normal distribution with zero mean and unit variance and a χ^2-distribution with ν degrees of freedom respectively; then $t = X_1/\sqrt{X_2/\nu}$ has Student's t-distribution with ν degrees of freedom, and the probability that $t \geqslant t_\nu(P)$ is $P/100$. The lower percentage points are given by symmetry as $-t_\nu(P)$, and the probability that $|t| \geqslant t_\nu(P)$ is $2P/100$.

The limiting distribution of t as ν tends to infinity is the normal distribution with zero mean and unit variance. When ν is large interpolation in ν should be harmonic.

P	40	30	25	20	15	10	5	2·5	1	0·5	0·1	0·05
ν = 1	0·3249	0·7265	1·0000	1·3764	1·963	3·078	6·314	12·71	31·82	63·66	318·3	636·6
2	0·2887	0·6172	0·8165	1·0607	1·386	1·886	2·920	4·303	6·965	9·925	22·33	31·60
3	0·2767	0·5844	0·7649	0·9785	1·250	1·638	2·353	3·182	4·541	5·841	10·21	12·92
4	0·2707	0·5686	0·7407	0·9410	1·190	1·533	2·132	2·776	3·747	4·604	7·173	8·610
5	0·2672	0·5594	0·7267	0·9195	1·156	1·476	2·015	2·571	3·365	4·032	5·893	6·869
6	0·2648	0·5534	0·7176	0·9057	1·134	1·440	1·943	2·447	3·143	3·707	5·208	5·959
7	0·2632	0·5491	0·7111	0·8960	1·119	1·415	1·895	2·365	2·998	3·499	4·785	5·408
8	0·2619	0·5459	0·7064	0·8889	1·108	1·397	1·860	2·306	2·896	3·355	4·501	5·041
9	0·2610	0·5435	0·7027	0·8834	1·100	1·383	1·833	2·262	2·821	3·250	4·297	4·781
10	0·2602	0·5415	0·6998	0·8791	1·093	1·372	1·812	2·228	2·764	3·169	4·144	4·587
11	0·2596	0·5399	0·6974	0·8755	1·088	1·363	1·796	2·201	2·718	3·106	4·025	4·437
12	0·2590	0·5386	0·6955	0·8726	1·083	1·356	1·782	2·179	2·681	3·055	3·930	4·318
13	0·2586	0·5375	0·6938	0·8702	1·079	1·350	1·771	2·160	2·650	3·012	3·852	4·221
14	0·2582	0·5366	0·6924	0·8681	1·076	1·345	1·761	2·145	2·624	2·977	3·787	4·140
15	0·2579	0·5357	0·6912	0·8662	1·074	1·341	1·753	2·131	2·602	2·947	3·733	4·073
16	0·2576	0·5350	0·6901	0·8647	1·071	1·337	1·746	2·120	2·583	2·921	3·686	4·015
17	0·2573	0·5344	0·6892	0·8633	1·069	1·333	1·740	2·110	2·567	2·898	3·646	3·965
18	0·2571	0·5338	0·6884	0·8620	1·067	1·330	1·734	2·101	2·552	2·878	3·610	3·922
19	0·2569	0·5333	0·6876	0·8610	1·066	1·328	1·729	2·093	2·539	2·861	3·579	3·883
20	0·2567	0·5329	0·6870	0·8600	1·064	1·325	1·725	2·086	2·528	2·845	3·552	3·850
21	0·2566	0·5325	0·6864	0·8591	1·063	1·323	1·721	2·080	2·518	2·831	3·527	3·819
22	0·2564	0·5321	0·6858	0·8583	1·061	1·321	1·717	2·074	2·508	2·819	3·505	3·792
23	0·2563	0·5317	0·6853	0·8575	1·060	1·319	1·714	2·069	2·500	2·807	3·485	3·768
24	0·2562	0·5314	0·6848	0·8569	1·059	1·318	1·711	2·064	2·492	2·797	3·467	3·745
25	0·2561	0·5312	0·6844	0·8562	1·058	1·316	1·708	2·060	2·485	2·787	3·450	3·725
26	0·2560	0·5309	0·6840	0·8557	1·058	1·315	1·706	2·056	2·479	2·779	3·435	3·707
27	0·2559	0·5306	0·6837	0·8551	1·057	1·314	1·703	2·052	2·473	2·771	3·421	3·690
28	0·2558	0·5304	0·6834	0·8546	1·056	1·313	1·701	2·048	2·467	2·763	3·408	3·674
29	0·2557	0·5302	0·6830	0·8542	1·055	1·311	1·699	2·045	2·462	2·756	3·396	3·659
30	0·2556	0·5300	0·6828	0·8538	1·055	1·310	1·697	2·042	2·457	2·750	3·385	3·646
32	0·2555	0·5297	0·6822	0·8530	1·054	1·309	1·694	2·037	2·449	2·738	3·365	3·622
34	0·2553	0·5294	0·6818	0·8523	1·052	1·307	1·691	2·032	2·441	2·728	3·348	3·601
36	0·2552	0·5291	0·6814	0·8517	1·052	1·306	1·688	2·028	2·434	2·719	3·333	3·582
38	0·2551	0·5288	0·6810	0·8512	1·051	1·304	1·686	2·024	2·429	2·712	3·319	3·566
40	0·2550	0·5286	0·6807	0·8507	1·050	1·303	1·684	2·021	2·423	2·704	3·307	3·551
50	0·2547	0·5278	0·6794	0·8489	1·047	1·299	1·676	2·009	2·403	2·678	3·261	3·496
60	0·2545	0·5272	0·6786	0·8477	1·045	1·296	1·671	2·000	2·390	2·660	3·232	3·460
120	0·2539	0·5258	0·6765	0·8446	1·041	1·289	1·658	1·980	2·358	2·617	3·160	3·373
∞	0·2533	0·5244	0·6745	0·8416	1·036	1·282	1·645	1·960	2·326	2·576	3·090	3·291

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*a*). 10 PER CENT POINTS OF THE *F*-DISTRIBUTION

The function tabulated is $F(P) = F(P|\nu_1, \nu_2)$ defined by the equation

$$\frac{P}{100} = \frac{\Gamma(\frac{1}{2}\nu_1+\frac{1}{2}\nu_2)}{\Gamma(\frac{1}{2}\nu_1)\,\Gamma(\frac{1}{2}\nu_2)}\,\nu_1^{\frac{1}{2}\nu_1}\,\nu_2^{\frac{1}{2}\nu_2}\int_{F(P)}^{\infty}\frac{F^{\frac{1}{2}\nu_1-1}}{(\nu_2+\nu_1 F)^{\frac{1}{2}(\nu_1+\nu_2)}}\,dF,$$

for $P = 10, 5, 2{\cdot}5, 1, 0{\cdot}5$ and $0{\cdot}1$. The lower percentage points, that is the values $F'(P) = F'(P|\nu_1, \nu_2)$ such that the probability that $F \leqslant F'(P)$ is equal to $P/100$, may be found by the formula

$$F'(P|\nu_1, \nu_2) = 1/F(P|\nu_2, \nu_1).$$

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	39·86	49·50	53·59	55·83	57·24	58·20	58·91	59·44	60·19	60·71	62·00	63·33
2	8·526	9·000	9·162	9·243	9·293	9·326	9·349	9·367	9·392	9·408	9·450	9·491
3	5·538	5·462	5·391	5·343	5·309	5·285	5·266	5·252	5·230	5·216	5·176	5·134
4	4·545	4·325	4·191	4·107	4·051	4·010	3·979	3·955	3·920	3·896	3·831	3·761
5	4·060	3·780	3·619	3·520	3·453	3·405	3·368	3·339	3·297	3·268	3·191	3·105
6	3·776	3·463	3·289	3·181	3·108	3·055	3·014	2·983	2·937	2·905	2·818	2·722
7	3·589	3·257	3·074	2·961	2·883	2·827	2·785	2·752	2·703	2·668	2·575	2·471
8	3·458	3·113	2·924	2·806	2·726	2·668	2·624	2·589	2·538	2·502	2·404	2·293
9	3·360	3·006	2·813	2·693	2·611	2·551	2·505	2·469	2·416	2·379	2·277	2·159
10	3·285	2·924	2·728	2·605	2·522	2·461	2·414	2·377	2·323	2·284	2·178	2·055
11	3·225	2·860	2·660	2·536	2·451	2·389	2·342	2·304	2·248	2·209	2·100	1·972
12	3·177	2·807	2·606	2·480	2·394	2·331	2·283	2·245	2·188	2·147	2·036	1·904
13	3·136	2·763	2·560	2·434	2·347	2·283	2·234	2·195	2·138	2·097	1·983	1·846
14	3·102	2·726	2·522	2·395	2·307	2·243	2·193	2·154	2·095	2·054	1·938	1·797
15	3·073	2·695	2·490	2·361	2·273	2·208	2·158	2·119	2·059	2·017	1·899	1·755
16	3·048	2·668	2·462	2·333	2·244	2·178	2·128	2·088	2·028	1·985	1·866	1·718
17	3·026	2·645	2·437	2·308	2·218	2·152	2·102	2·061	2·001	1·958	1·836	1·686
18	3·007	2·624	2·416	2·286	2·196	2·130	2·079	2·038	1·977	1·933	1·810	1·657
19	2·990	2·606	2·397	2·266	2·176	2·109	2·058	2·017	1·956	1·912	1·787	1·631
20	2·975	2·589	2·380	2·249	2·158	2·091	2·040	1·999	1·937	1·892	1·767	1·607
21	2·961	2·575	2·365	2·233	2·142	2·075	2·023	1·982	1·920	1·875	1·748	1·586
22	2·949	2·561	2·351	2·219	2·128	2·060	2·008	1·967	1·904	1·859	1·731	1·567
23	2·937	2·549	2·339	2·207	2·115	2·047	1·995	1·953	1·890	1·845	1·716	1·549
24	2·927	2·538	2·327	2·195	2·103	2·035	1·983	1·941	1·877	1·832	1·702	1·533
25	2·918	2·528	2·317	2·184	2·092	2·024	1·971	1·929	1·866	1·820	1·689	1·518
26	2·909	2·519	2·307	2·174	2·082	2·014	1·961	1·919	1·855	1·809	1·677	1·504
27	2·901	2·511	2·299	2·165	2·073	2·005	1·952	1·909	1·845	1·799	1·666	1·491
28	2·894	2·503	2·291	2·157	2·064	1·996	1·943	1·900	1·836	1·790	1·656	1·478
29	2·887	2·495	2·283	2·149	2·057	1·988	1·935	1·892	1·827	1·781	1·647	1·467
30	2·881	2·489	2·276	2·142	2·049	1·980	1·927	1·884	1·819	1·773	1·638	1·456
32	2·869	2·477	2·263	2·129	2·036	1·967	1·913	1·870	1·805	1·758	1·622	1·437
34	2·859	2·466	2·252	2·118	2·024	1·955	1·901	1·858	1·793	1·745	1·608	1·419
36	2·850	2·456	2·243	2·108	2·014	1·945	1·891	1·847	1·781	1·734	1·595	1·404
38	2·842	2·448	2·234	2·099	2·005	1·935	1·881	1·838	1·772	1·724	1·584	1·390
40	2·835	2·440	2·226	2·091	1·997	1·927	1·873	1·829	1·763	1·715	1·574	1·377
60	2·791	2·393	2·177	2·041	1·946	1·875	1·819	1·775	1·707	1·657	1·511	1·291
120	2·748	2·347	2·130	1·992	1·896	1·824	1·767	1·722	1·652	1·601	1·447	1·193
∞	2·706	2·303	2·084	1·945	1·847	1·774	1·717	1·670	1·599	1·546	1·383	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*b*). 5 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 and ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	161·4	199·5	215·7	224·6	230·2	234·0	236·8	238·9	241·9	243·9	249·1	254·3
2	18·51	19·00	19·16	19·25	19·30	19·33	19·35	19·37	19·40	19·41	19·45	19·50
3	10·13	9·552	9·277	9·117	9·013	8·941	8·887	8·845	8·786	8·745	8·639	8·526
4	7·709	6·944	6·591	6·388	6·256	6·163	6·094	6·041	5·964	5·912	5·774	5·628
5	6·608	5·786	5·409	5·192	5·050	4·950	4·876	4·818	4·735	4·678	4·527	4·365
6	5·987	5·143	4·757	4·534	4·387	4·284	4·207	4·147	4·060	4·000	3·841	3·669
7	5·591	4·737	4·347	4·120	3·972	3·866	3·787	3·726	3·637	3·575	3·410	3·230
8	5·318	4·459	4·066	3·838	3·687	3·581	3·500	3·438	3·347	3·284	3·115	2·928
9	5·117	4·256	3·863	3·633	3·482	3·374	3·293	3·230	3·137	3·073	2·900	2·707
10	4·965	4·103	3·708	3·478	3·326	3·217	3·135	3·072	2·978	2·913	2·737	2·538
11	4·844	3·982	3·587	3·357	3·204	3·095	3·012	2·948	2·854	2·788	2·609	2·404
12	4·747	3·885	3·490	3·259	3·106	2·996	2·913	2·849	2·753	2·687	2·505	2·296
13	4·667	3·806	3·411	3·179	3·025	2·915	2·832	2·767	2·671	2·604	2·420	2·206
14	4·600	3·739	3·344	3·112	2·958	2·848	2·764	2·699	2·602	2·534	2·349	2·131
15	4·543	3·682	3·287	3·056	2·901	2·790	2·707	2·641	2·544	2·475	2·288	2·066
16	4·494	3·634	3·239	3·007	2·852	2·741	2·657	2·591	2·494	2·425	2·235	2·010
17	4·451	3·592	3·197	2·965	2·810	2·699	2·614	2·548	2·450	2·381	2·190	1·960
18	4·414	3·555	3·160	2·928	2·773	2·661	2·577	2·510	2·412	2·342	2·150	1·917
19	4·381	3·522	3·127	2·895	2·740	2·628	2·544	2·477	2·378	2·308	2·114	1·878
20	4·351	3·493	3·098	2·866	2·711	2·599	2·514	2·447	2·348	2·278	2·082	1·843
21	4·325	3·467	3·072	2·840	2·685	2·573	2·488	2·420	2·321	2·250	2·054	1·812
22	4·301	3·443	3·049	2·817	2·661	2·549	2·464	2·397	2·297	2·226	2·028	1·783
23	4·279	3·422	3·028	2·796	2·640	2·528	2·442	2·375	2·275	2·204	2·005	1·757
24	4·260	3·403	3·009	2·776	2·621	2·508	2·423	2·355	2·255	2·183	1·984	1·733
25	4·242	3·385	2·991	2·759	2·603	2·490	2·405	2·337	2·236	2·165	1·964	1·711
26	4·225	3·369	2·975	2·743	2·587	2·474	2·388	2·321	2·220	2·148	1·946	1·691
27	4·210	3·354	2·960	2·728	2·572	2·459	2·373	2·305	2·204	2·132	1·930	1·672
28	4·196	3·340	2·947	2·714	2·558	2·445	2·359	2·291	2·190	2·118	1·915	1·654
29	4·183	3·328	2·934	2·701	2·545	2·432	2·346	2·278	2·177	2·104	1·901	1·638
30	4·171	3·316	2·922	2·690	2·534	2·421	2·334	2·266	2·165	2·092	1·887	1·622
32	4·149	3·295	2·901	2·668	2·512	2·399	2·313	2·244	2·142	2·070	1·864	1·594
34	4·130	3·276	2·883	2·650	2·494	2·380	2·294	2·225	2·123	2·050	1·843	1·569
36	4·113	3·259	2·866	2·634	2·477	2·364	2·277	2·209	2·106	2·033	1·824	1·547
38	4·098	3·245	2·852	2·619	2·463	2·349	2·262	2·194	2·091	2·017	1·808	1·527
40	4·085	3·232	2·839	2·606	2·449	2·336	2·249	2·180	2·077	2·003	1·793	1·509
60	4·001	3·150	2·758	2·525	2·368	2·254	2·167	2·097	1·993	1·917	1·700	1·389
120	3·920	3·072	2·680	2·447	2·290	2·175	2·087	2·016	1·910	1·834	1·608	1·254
∞	3·841	2·996	2·605	2·372	2·214	2·099	2·010	1·938	1·831	1·752	1·517	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

TABLE 12(*d*). 1 PER CENT POINTS OF THE *F*-DISTRIBUTION

If $F = \frac{X_1}{\nu_1} \Big/ \frac{X_2}{\nu_2}$, where X_1 and X_2 are independent random variables distributed as χ^2 with ν_1 and ν_2 degrees of freedom respectively, then the probabilities that $F \geqslant F(P)$ and that $F \leqslant F'(P)$ are both equal to $P/100$. Linear interpolation in ν_1 or ν_2 will generally be sufficiently accurate except when either $\nu_1 > 12$ or $\nu_2 > 40$, when harmonic interpolation should be used.

(This shape applies only when $\nu_1 \geqslant 3$. When $\nu_1 < 3$ the mode is at the origin.)

$\nu_1 =$	1	2	3	4	5	6	7	8	10	12	24	∞
$\nu_2 =$ 1	4052	4999	5403	5625	5764	5859	5928	5981	6056	6106	6235	6366
2	98·50	99·00	99·17	99·25	99·30	99·33	99·36	99·37	99·40	99·42	99·46	99·50
3	34·12	30·82	29·46	28·71	28·24	27·91	27·67	27·49	27·23	27·05	26·60	26·13
4	21·20	18·00	16·69	15·98	15·52	15·21	14·98	14·80	14·55	14·37	13·93	13·46
5	16·26	13·27	12·06	11·39	10·97	10·67	10·46	10·29	10·05	9·888	9·466	9·020
6	13·75	10·92	9·780	9·148	8·746	8·466	8·260	8·102	7·874	7·718	7·313	6·880
7	12·25	9·547	8·451	7·847	7·460	7·191	6·993	6·840	6·620	6·469	6·074	5·650
8	11·26	8·649	7·591	7·006	6·632	6·371	6·178	6·029	5·814	5·667	5·279	4·859
9	10·56	8·022	6·992	6·422	6·057	5·802	5·613	5·467	5·257	5·111	4·729	4·311
10	10·04	7·559	6·552	5·994	5·636	5·386	5·200	5·057	4·849	4·706	4·327	3·909
11	9·646	7·206	6·217	5·668	5·316	5·069	4·886	4·744	4·539	4·397	4·021	3·602
12	9·330	6·927	5·953	5·412	5·064	4·821	4·640	4·499	4·296	4·155	3·780	3·361
13	9·074	6·701	5·739	5·205	4·862	4·620	4·441	4·302	4·100	3·960	3·587	3·165
14	8·862	6·515	5·564	5·035	4·695	4·456	4·278	4·140	3·939	3·800	3·427	3·004
15	8·683	6·359	5·417	4·893	4·556	4·318	4·142	4·004	3·805	3·666	3·294	2·868
16	8·531	6·226	5·292	4·773	4·437	4·202	4·026	3·890	3·691	3·553	3·181	2·753
17	8·400	6·112	5·185	4·669	4·336	4·102	3·927	3·791	3·593	3·455	3·084	2·653
18	8·285	6·013	5·092	4·579	4·248	4·015	3·841	3·705	3·508	3·371	2·999	2·566
19	8·185	5·926	5·010	4·500	4·171	3·939	3·765	3·631	3·434	3·297	2·925	2·489
20	8·096	5·849	4·938	4·431	4·103	3·871	3·699	3·564	3·368	3·231	2·859	2·421
21	8·017	5·780	4·874	4·369	4·042	3·812	3·640	3·506	3·310	3·173	2·801	2·360
22	7·945	5·719	4·817	4·313	3·988	3·758	3·587	3·453	3·258	3·121	2·749	2·305
23	7·881	5·664	4·765	4·264	3·939	3·710	3·539	3·406	3·211	3·074	2·702	2·256
24	7·823	5·614	4·718	4·218	3·895	3·667	3·496	3·363	3·168	3·032	2·659	2·211
25	7·770	5·568	4·675	4·177	3·855	3·627	3·457	3·324	3·129	2·993	2·620	2·169
26	7·721	5·526	4·637	4·140	3·818	3·591	3·421	3·288	3·094	2·958	2·585	2·131
27	7·677	5·488	4·601	4·106	3·785	3·558	3·388	3·256	3·062	2·926	2·552	2·097
28	7·636	5·453	4·568	4·074	3·754	3·528	3·358	3·226	3·032	2·896	2·522	2·064
29	7·598	5·420	4·538	4·045	3·725	3·499	3·330	3·198	3·005	2·868	2·495	2·034
30	7·562	5·390	4·510	4·018	3·699	3·473	3·304	3·173	2·979	2·843	2·469	2·006
32	7·499	5·336	4·459	3·969	3·652	3·427	3·258	3·127	2·934	2·798	2·423	1·956
34	7·444	5·289	4·416	3·927	3·611	3·386	3·218	3·087	2·894	2·758	2·383	1·911
36	7·396	5·248	4·377	3·890	3·574	3·351	3·183	3·052	2·859	2·723	2·347	1·872
38	7·353	5·211	4·343	3·858	3·542	3·319	3·152	3·021	2·828	2·692	2·316	1·837
40	7·314	5·179	4·313	3·828	3·514	3·291	3·124	2·993	2·801	2·665	2·288	1·805
60	7·077	4·977	4·126	3·649	3·339	3·119	2·953	2·823	2·632	2·496	2·115	1·601
120	6·851	4·787	3·949	3·480	3·174	2·956	2·792	2·663	2·472	2·336	1·950	1·381
∞	6·635	4·605	3·782	3·319	3·017	2·802	2·639	2·511	2·321	2·185	1·791	1·000

Dennis V. Lindley, William F. Scott, New Cambridge Statistical Tables, (1995) © Cambridge University Press, reproduced with permission.

Durbin–Watson test statistic d : 1% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	0.81	1.07	0.70	1.25	0.59	1.46	0.49	1.70	0.39	1.96
16	0.84	1.09	0.74	1.25	0.63	1.44	0.53	1.66	0.44	1.90
17	0.87	1.10	0.77	1.25	0.67	1.43	0.57	1.63	0.48	1.85
18	0.90	1.12	0.80	1.26	0.71	1.42	0.61	1.60	0.52	1.80
19	0.93	1.13	0.83	1.26	0.74	1.41	0.65	1.58	0.56	1.77
20	0.95	1.15	0.86	1.27	0.77	1.41	0.68	1.57	0.60	1.74
21	0.97	1.16	0.89	1.27	0.80	1.41	0.72	1.55	0.63	1.71
22	1.00	1.17	0.91	1.28	0.83	1.40	0.75	1.54	0.66	1.69
23	1.02	1.19	0.94	1.29	0.86	1.40	0.77	1.53	0.70	1.67
24	1.04	1.20	0.96	1.30	0.88	1.41	0.80	1.53	0.72	1.66
25	1.05	1.21	0.98	1.30	0.90	1.41	0.83	1.52	0.75	1.65
26	1.07	1.22	1.00	1.31	0.93	1.41	0.85	1.52	0.78	1.64
27	1.09	1.23	1.02	1.32	0.95	1.41	0.88	1.51	0.81	1.63
28	1.10	1.24	1.04	1.32	0.97	1.41	0.90	1.51	0.83	1.62
29	1.12	1.25	1.05	1.33	0.99	1.42	0.92	1.51	0.85	1.61
30	1.13	1.26	1.07	1.34	1.01	1.42	0.94	1.51	0.88	1.61
31	1.15	1.27	1.08	1.34	1.02	1.42	0.96	1.51	0.90	1.60
32	1.16	1.28	1.10	1.35	1.04	1.43	0.98	1.51	0.92	1.60
33	1.17	1.29	1.11	1.36	1.05	1.43	1.00	1.51	0.94	1.59
34	1.18	1.30	1.13	1.36	1.07	1.43	1.01	1.51	0.95	1.59
35	1.19	1.31	1.14	1.37	1.08	1.44	1.03	1.51	0.97	1.59
36	1.21	1.32	1.15	1.38	1.10	1.44	1.04	1.51	0.99	1.59
37	1.22	1.32	1.16	1.38	1.11	1.45	1.06	1.51	1.00	1.59
38	1.23	1.33	1.18	1.39	1.12	1.45	1.07	1.52	1.02	1.58
39	1.24	1.34	1.19	1.39	1.14	1.45	1.09	1.52	1.03	1.58
40	1.25	1.34	1.20	1.40	1.15	1.46	1.10	1.52	1.05	1.58
45	1.29	1.38	1.24	1.42	1.20	1.48	1.16	1.53	1.11	1.58
50	1.32	1.40	1.28	1.45	1.24	1.49	1.20	1.54	1.16	1.59
55	1.36	1.43	1.32	1.47	1.28	1.51	1.25	1.55	1.21	1.59
60	1.38	1.45	1.35	1.48	1.32	1.52	1.28	1.56	1.25	1.60
65	1.41	1.47	1.38	1.50	1.35	1.53	1.31	1.57	1.28	1.61
70	1.43	1.49	1.40	1.52	1.37	1.55	1.34	1.58	1.31	1.61
75	1.45	1.50	1.42	1.53	1.39	1.56	1.37	1.59	1.34	1.62
80	1.47	1.52	1.44	1.54	1.42	1.57	1.39	1.60	1.36	1.62
85	1.48	1.53	1.46	1.55	1.43	1.58	1.41	1.60	1.39	1.63
90	1.50	1.54	1.47	1.56	1.45	1.59	1.43	1.61	1.41	1.64
95	1.51	1.55	1.49	1.57	1.47	1.60	1.45	1.62	1.42	1.64
100	1.52	1.56	1.50	1.58	1.48	1.60	1.46	1.63	1.44	1.65

n = number of observations

k' = number of explanatory variables

Durbin–Watson test statistic d : 5% significance points of d_L and d_U.

n	$k'=1$		$k'=2$		$k'=3$		$k'=4$		$k'=5$	
	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U	d_L	d_U
15	1.08	1.36	0.95	1.54	0.82	1.75	0.69	1.97	0.56	2.21
16	1.10	1.37	0.98	1.54	0.86	1.73	0.74	1.93	0.62	2.15
17	1.13	1.38	1.02	1.54	0.90	1.71	0.78	1.90	0.67	2.10
18	1.16	1.39	1.05	1.53	0.93	1.69	0.82	1.87	0.71	2.06
19	1.18	1.40	1.08	1.53	0.97	1.68	0.86	1.85	0.75	2.02
20	1.20	1.41	1.10	1.54	1.00	1.68	0.90	1.83	0.79	1.99
21	1.22	1.42	1.13	1.54	1.03	1.67	0.93	1.81	0.83	1.96
22	1.24	1.43	1.15	1.54	1.05	1.66	0.96	1.80	0.86	1.94
23	1.26	1.44	1.17	1.54	1.08	1.66	0.99	1.79	0.90	1.92
24	1.27	1.45	1.19	1.55	1.10	1.66	1.01	1.78	0.93	1.90
25	1.29	1.45	1.21	1.55	1.12	1.66	1.04	1.77	0.95	1.89
26	1.30	1.46	1.22	1.55	1.14	1.65	1.06	1.76	0.98	1.88
27	1.32	1.47	1.24	1.56	1.16	1.65	1.08	1.76	1.01	1.86
28	1.33	1.48	1.26	1.56	1.18	1.65	1.10	1.75	1.03	1.85
29	1.34	1.48	1.27	1.56	1.20	1.65	1.12	1.74	1.05	1.84
30	1.35	1.49	1.28	1.57	1.21	1.65	1.14	1.74	1.07	1.83
31	1.36	1.50	1.30	1.57	1.23	1.65	1.16	1.74	1.09	1.83
32	1.37	1.50	1.31	1.57	1.24	1.65	1.18	1.73	1.11	1.82
33	1.38	1.51	1.32	1.58	1.26	1.65	1.19	1.73	1.13	1.81
34	1.39	1.51	1.33	1.58	1.27	1.65	1.21	1.73	1.15	1.81
35	1.40	1.52	1.34	1.58	1.28	1.65	1.22	1.73	1.16	1.80
36	1.41	1.52	1.35	1.59	1.29	1.65	1.24	1.73	1.18	1.80
37	1.42	1.53	1.36	1.59	1.31	1.66	1.25	1.72	1.19	1.80
38	1.43	1.54	1.37	1.59	1.32	1.66	1.26	1.72	1.21	1.79
39	1.43	1.54	1.38	1.60	1.33	1.66	1.27	1.72	1.22	1.79
40	1.44	1.54	1.39	1.60	1.34	1.66	1.29	1.72	1.23	1.79
45	1.48	1.57	1.43	1.62	1.38	1.67	1.34	1.72	1.29	1.78
50	1.50	1.59	1.46	1.63	1.42	1.67	1.38	1.72	1.34	1.77
55	1.53	1.60	1.49	1.64	1.45	1.68	1.41	1.72	1.38	1.77
60	1.55	1.62	1.51	1.65	1.48	1.69	1.44	1.73	1.41	1.77
65	1.57	1.63	1.54	1.66	1.50	1.70	1.47	1.73	1.44	1.77
70	1.58	1.64	1.55	1.67	1.52	1.70	1.49	1.74	1.46	1.77
75	1.60	1.65	1.57	1.68	1.54	1.71	1.51	1.74	1.49	1.77
80	1.61	1.66	1.59	1.69	1.56	1.72	1.53	1.74	1.51	1.77
85	1.62	1.67	1.60	1.70	1.57	1.72	1.55	1.75	1.52	1.77
90	1.63	1.68	1.61	1.70	1.59	1.73	1.57	1.75	1.54	1.78
95	1.64	1.69	1.62	1.71	1.60	1.73	1.58	1.75	1.56	1.78
100	1.65	1.69	1.63	1.72	1.61	1.74	1.59	1.76	1.57	1.78

n = number of observations

k' = number of explanatory variables