UNIVERSITY OF LONDON | INTERNATIONAL PROGRAMMES

Course information 2015–16
EC2020 Elements of econometrics

Econometrics is the application of statistical methods to the quantification and critical assessment of hypothetical economic relationships using data. This course gives students an opportunity to develop an understanding of econometrics to a standard that will equip them to understand and evaluate most applied analysis of cross-sectional data and to be able to undertake such analysis themselves.

Prerequisite
If taken as part of a BSc degree, courses which must be passed before this course may be attempted:
EC1002 Introduction to economics
and
ST104A Statistics 1 (half course) **or**
ST104B Statistics 2 (half course)
and
MT1005A Mathematics 1 (half course) **or**
MT1005B Mathematics 2 (half course) **or**
MT1174 Calculus

Aims and objectives
The aims of this course are:
- To develop an understanding of the use of regression analysis and related techniques for quantifying economic relationships and testing economic theories.
- To equip students to read and evaluate empirical papers in professional journals.
- To provide students with practical experience of using mainstream regression programmes to fit economic models.

Essential reading
For full details please refer to the reading list.
Dougherty, C. *Introduction to Econometrics.* (Oxford: Oxford University Press)

Assessment
This course is assessed by a three hour unseen written examination.

Learning outcomes
At the end of the course and having completed the essential reading and activities students 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.

Syllabus

This is a description of the material to be examined, as published in the *Programme handbook.* On registration, students will receive a detailed subject guide which provides a framework for covering the topics in the syllabus and directions to the essential reading

This syllabus is intended to provide an explicit list of all the mathematical formulae and proofs that you are expected to know for the Elements of Econometrics examination. You are warned that the examination is intended to be an opportunity for you to display your understanding of the material, rather than of your ability to reproduce standard items.

Review: Random variables and sampling theory: Probability distribution of a random variable. Expected value of a random variable. Expected value of a function of a random variable. Population variance of a discrete random variable and alternative expression for it. Expected value rules. Independence of two random variables. Population covariance, covariance and variance rules, and correlation. Sampling and estimators. Unbiasedness. Efficiency. Loss functions and mean square error. Estimators of variance, covariance and correlation. The normal distribution. Hypothesis testing. Type II error and the power of a test. *t* tests. Confidence intervals. One-sided tests. Convergence in probability and plim rules. Consistency. Convergence in distribution (asymptotic limiting distributions) and the role of central limit theorems.

Simple regression analysis: Simple regression model. Derivation of linear regression coefficients. Interpretation of a regression equation. Goodness of fit.

Properties of the regression coefficients: Types of data and regression model. Assumptions for Model A. Regression coefficients as random variables. Unbiasedness of the regression coefficients. Precision of the regression coefficients. Gauss–Markov theorem. *t* test of a hypothesis relating to a regression coefficient. Type I error and Type II error. Confidence intervals. One-sided tests. *F* test of goodness of fit.

Multiple regression analysis: Multiple regression with two explanatory variables. Graphical representation of a relationship in a multiple regression model. Properties of the multiple regression coefficients. Population variance of the regression coefficients. Decomposition of their standard errors. Multicollinearity. *F* tests in a multiple regression model. Hedonic pricing models. Prediction.

Transformation of variables: Linearity and nonlinearity. Elasticities and double-logarithmic models. Semilogarithmic models. The disturbance term in nonlinear models. Box–Cox transformation. Models with quadratic and interactive variables. Nonlinear regression.

Dummy variables: Dummy variables. Dummy classification with more than two categories. The effects of changing the reference category. Multiple sets of dummy variables. Slope dummy variables. Chow test. Relationship between Chow test and dummy group test.

Specification of regression variables: Omitted variable bias. Consequences of the inclusion of an irrelevant variable. Proxy variables. *F* test of a linear restriction. Reparameterization of a regression model (see the *Further Material* hand-out). *t* test of a restriction. Tests of multiple restrictions. Tests of zero restrictions.

Heteroscedasticity: Meaning of heteroscedasticity. Consequences of heteroscedasticity. Goldfeld–Quandt and White tests for heteroscedasticity. Elimination of heteroscedasticity using weighted or logarithmic regressions. Use of heteroscedasticity-consistent standard errors.

Stochastic regressors and measurement errors: Stochastic regressors. Assumptions for models with stochastic regressors. Finite sample and asymptotic properties of the regression coefficients in models with stochastic regressors. Measurement error and its consequences. Friedman's Permanent Income Hypothesis. Instrumental variables (IV). Asymptotic properties of IV estimators, including the asymptotic limiting distribution of $\sqrt{n}\left(b_2^{\mathrm{IV}} - \beta_2\right)$ where b_2^{IV} is the IV estimator of β_2 in a simple regression model. Use of simulation to investigate the finite-sample properties of estimators when only asymptotic properties can be determined analytically. Application of the Durbin–Wu–Hausman test

Simultaneous equations estimation: Definitions of endogenous variables, exogenous variables, structural equations and reduced form. Inconsistency of OLS. Use of instrumental variables. Exact identification, underidentification, and overidentification. Two-stage least squares (TSLS). Order condition for identification. Application of the Durbin–Wu–Hausman test.

Binary choice models and maximum likelihood estimation: Linear probability model. Logit model. Probit model. Maximum likelihood estimation of the population mean and variance of a random variable. Maximum likelihood estimation of regression coefficients. Likelihood ratio tests.

Models using time series data: Static demand functions fitted using aggregate time series data. Lagged variables and naive attempts to model dynamics. Autoregressive distributed lag (ADL) models with applications in the form of the partial adjustment and adaptive expectations models. Error correction models. Asymptotic properties of OLS estimators of ADL models, including asymptotic limiting distributions. Use of simulation to investigate the finite sample properties of parameter estimators for the ADL(1,0) model. Use of predetermined variables as instruments in simultaneous equations models using time series data. (Section 11.7 of the text, *Alternative dynamic representations* ..., is not in the syllabus)

Autocorrelation: Assumptions for regressions with time series data. Assumption of the independence of the disturbance term and the regressors. Definition of autocorrelation. Consequences of autocorrelation. Breusch–Godfrey lagrange multiplier, Durbin–Watson d, and Durbin h tests for autocorrelation. AR(1) nonlinear regression. Potential advantages and disadvantages of such estimation, in comparison with OLS. Cochrane–Orcutt iterative process. Autocorrelation with a lagged dependent variable. Common factor test and implications for model selection. Apparent autocorrelation caused by variable or functional misspecification. General-to-specific versus specific-to-general model specification.

Introduction to nonstationary processes: Stationary and nonstationary processes. Granger–Newbold experiments with random walks. Unit root tests. Akaike Information Criterion and Schwarz's Bayes Information Criterion. Cointegration. Error correction models.

Introduction to panel data models: Definition of panel data set (longitudinal data set). Pooled OLS model. Definition of, and consequences of, unobserved heterogeneity. Within-groups fixed effects model. First differences fixed-effects model. Least squares dummy variable model. Calculation of degrees of freedom in fixed effects models.

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

The definition of consistency is required, and the sufficient condition of consistency should also be given. The probability limit or sufficient condition of consistency should be used to show the consistency of the OLS estimator of the slope. The solution is as follows:

Definition:

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

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

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

The sufficient condition for consistency comprises:

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.

We now examine the consistency of the estimator of the slope parameter:

Let the model be:

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

The OLS estimator of β_2 is:

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

and:

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

σ_{Xu} is the population covariance between X and u, which by assumption 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

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

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

The sufficient condition of consistency should be used to show the consistency of the given estimator. The solution is as follows:

We have:

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

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. We have:

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

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

As the estimator is unbiased and also as 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) The log-likelihood function should be derived, and it should be differentiated with respect to μ and equated to zero to obtain the ML estimator. The solution is as follows:

The likelihood function is:

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

Hence, taking logs, we obtain the log-likelihood function:

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

To maximise $\ln L$, differentiate with respect to μ, and set the partial derivative equal to zero. We have:

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

which gives:

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

(b) The properties of ML estimators should be discussed as follows:

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

In order to get an estimable equation we need to eliminate the expected values from the equation. First, multiply through by $(1-\gamma)$ and lag to get a new equation, then subtract the new equation from the original equation to get the result. The solution is as follows:

We have:

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

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

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

Now subtract this from the original equation:

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

or:

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

The 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 are estimated using 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 defined as:

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

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:

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

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

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

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

This can be tested by a t test. If we do not reject H_0, then we can apply OLS to $Y_t = \beta_0 + \beta_1 X_t + u_t$ and get the estimated parameters. If H_0 is rejected, then our estimated equations for the two different periods are given by (iii) and (iv).

(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 defined as:

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

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:

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

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

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

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

This can be tested by a t test. If we do not reject H_0, then we can apply OLS to $Y_t = \beta_0 + \beta_1 X_t + u_t$ and get the estimated parameters. If H_0 is rejected, then our estimated equations for the 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 defined as:

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

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:

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

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

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

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

This can be tested by an F test. If we do not reject H_0, then we can apply OLS to $Y_t = \beta_0 + \beta_1 X_t + u_t$ 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:

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

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 $\text{E}(u_t) = 0$.

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

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

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

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

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) The interpretation of α_1, α_2 and $\alpha_1 + \alpha_2$ should be given. The solution is as follows:

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 increase in the labour input, holding capital constant, will increase output by 1.498 per cent, on average.

Similarly, α_2 is the (partial) elasticity of output with respect to the capital input, holding the labour input constant. Hence, a 1 per cent increase in the capital input, holding labour constant, will increase output by 0.489 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) The significance of both slope parameters should be tested, and the assumptions should be explicitly provided. The solution is as follows:

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 values for the two-sided t distribution with 12 degrees of freedom are ± 2.179. Hence, reject the null hypothesis in both cases. The test requires the following assumptions:

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

(c) It should be mentioned that not enough information is given to test this restriction. The solution is as follows:

The estimated coefficients sum to 1.987, which gives the value of the returns 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. The solution is as follows:

The F test is given by:

$$F = \frac{(R_U^2 - R_R^2)/q}{(1 - R_U^2)/(n - k - 1)} = \frac{(0.89 - 0.77)/1}{(1 - 0.89)/12} = 13.09.$$

The 5 per cent critical value for $F_{1,\,12}$ is 4.75. Therefore, we reject the null hypothesis of constant returns to scale, and so 1.987 is significantly different from 1.

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

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

Re-arranging gives:

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

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**

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

where $\boldsymbol{x_t}$ is an exogenous variable.

i. **Examine the identification of both equations.**

(4 marks)

ii. **Obtain the ordinary least squares estimator of $\boldsymbol{\alpha}$, and examine its consistency.**

(7 marks)

iii. **Derive the two-stage least squares of $\boldsymbol{\alpha}$ 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, the order condition of identification should be used. The solution is as follows:

Order Condition of Identification (Necessary condition of identification):

$$R \geq G - 1$$

where:

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

If:

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

(b) i. The concept of the order of identification given in part (a) should be used. The solution is as follows:
In the first equation, $R = 0$ and $G - 1 = 1$, hence $R < G - 1$. Therefore, the equation is under identified.
In the second equation, $R = 1$ and $G - 1 = 1$, hence $R = G - 1$. Therefore, the equation is exactly identified.

ii. It is necessary to derive the OLS estimator. Also, the probability limit should be used to examine consistency. The solution is as follows:
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:

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

and:

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

This implies $\widehat{\alpha}_{OLS}$ is an inconsistent estimator of α.

iii. It is necessary to derive the 2SLS estimator. Again, the probability limit should be used to examine consistency. The solution is as follows:
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, we have:

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

since z_t is correlated with y_{2t}, but uncorrelated with u_{1t}. The covariances given are the population covariances.

(d) A brief discussion of the ILS estimator is required. The solution is as follows:

- 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 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 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

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

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)

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 the linear probability model the estimated probability of an event occurring may be greater than one or less than zero. The solution is as follows:

$t = 1.02/0.51 = 2$ which is significantly different from 0. As income increases by £1000, $MORT$ increases by 1.02 units, 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. The solution is as follows:

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 values, 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 maximum likelihood under the alternative hypothesis, and $\ln L_0$ is estimated by maximum likelihood under the null hypothesis.

(c) It is necessary to derive the variance of the disturbance term. The solution is as follows:

Let the model be:

$$Y_i = \beta_0 + \beta_1 X_i + u_i; \quad i = 1, 2, \ldots, n \qquad \text{(i)}$$

where:

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

As Y_i takes only two values, 1 or 0, 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.

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

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

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

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

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**

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

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

i. **Is y_t stationary? Explain.**

(5 marks)

ii. **Calculate the autocorrelation function of y_t.**

(5 marks)

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) The concept of spurious regression should be explained with a simple model. The solution is as follows:

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 on X_t, that is:

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

then even if there is no relationship between Y_t and X_t, 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) A clear concept of cointegration is required, and the assumptions should be stated explicitly. The solution is as follows:

We test:

$$\mathrm{H}_0 : \text{No cointegration} \quad \text{vs.} \quad \mathrm{H}_1 : \text{Cointegration.}$$

The cointegration test statistic is $-0.2161/0.0845 = -2.557$.

The 5 per cent critical value given in the MacKinnon table is -3.3377. Therefore, we cannot reject the null hypothesis of no cointegration.

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

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

i. We have:

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

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

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

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

(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)

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 necessary to derive the variance and covariance of u_t. For finding these it has to be shown that $E(u_t) = 0$. We have:

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

which 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:

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

ii. We have:

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

Hence:

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

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

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, whose assumptions should be stated clearly. The solution is as follows:

Durbin–Watson (DW) Statistic

The DW statistic is defined as:

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

The Durbin–Watson statistic can be applied only if:

- the disturbance term follows an AR(1) process
- the model has an intercept term
- 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:

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

(d) As ρ is unknown, the Cochrane–Orcutt method of estimation, or Prais–Winstein method of estimation, should be used. The solution is as follows:

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 $\text{E}(\varepsilon_t) = 0$, $\text{E}(\varepsilon_t^2) = \sigma_\varepsilon^2$ and $\text{E}(\varepsilon_s \varepsilon_t) = 0$ 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:

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

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 the 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 steps (b) to (d). Keep on doing so until the estimate of ρ converges, which will be the final estimate of ρ. Denote this as $\widehat{\rho}_F$.
- Replace ρ in (iii) by the final estimate of ρ, i.e. by $\widehat{\rho}_F$. Apply OLS to (iii) to obtain the final estimates of β_0 and β_1.