UNIVERSITY OF LONDON | INTERNATIONAL PROGRAMMES

Course information 2016–17
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.

Prerequisites

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)* **and** *MT105A Mathematics 1 (half course)* **or** *MT1174 Calculus*

Co-requisites

Students can only take *EC2020 Econometrics* at the same time as or after *ST104B Statistics 2* and *MT105B Mathematics 2*, not before.

The exception to this is students who are taking *MT1174 Calculus*. They must still take *ST104b Statistics 2* at the same time or before *EC2020 Elements of econometrics*, but will not take *MT105b Mathematics 2*

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. .

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.

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

Syllabus

This is a description of the material to be examined. 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 $_2$ 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.

Students should consult the appropriate *EMFSS Programme Regulations*, which are reviewed on an annual basis. The *Regulations* provide information on the availability of a course, where it can be placed on your programme's structure, and details of co-requisites and prerequisites.