Course information 2020-21
EC2020 Elements of econometrics

General information

COURSE LEVEL: 5

CREDIT: 30

NOTIONAL STUDY TIME: 300 hours

Summary

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.

Conditions

Prerequisite: If taken as part of a BSc degree, the following course(s) 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 **OR** MT1186 Mathematical methods

Co-requisites: Students may only take *EC2020 Elements of econometrics* at the same time as, or after, the following courses, not before:

- ST104B Statistics 2 (half course)
- **AND** MT105B Mathematics 2 (half course)†

† If you have passed MT1174 Calculus **OR** MT1186 Mathematical methods, you do not need to 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.

Please consult the current EMFSS Programme Regulations for further information on the availability of a course, where it can be placed on your programme's structure, and other important details.

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 co-integration.

Essential reading

For full details, please refer to the reading list.

Dougherty, C. *Introduction to Econometrics.* (Oxford: Oxford University Press, 2016) fifth edition [ISBN 9780199676828].

Assessment

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

Syllabus

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 I error and the significance of a test. 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. Important results relating to OLS regressions. 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. P-values of a test. *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**.** Interpretation of the multiple regression coefficients. Properties of the multiple regression coefficients. Population variance of the regression coefficients. Decomposition of their standard errors. T test and confidence intervals. Multicollinearity. *F* tests in a multiple regression model. Adjusted R^2. 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. Reset test. Nonlinear regression.

Dummy variables: dummy variables. Dummy classification with more than two categories. The effects of changing the reference category. The dummy variable trap. 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). 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 (White) standard errors.

Stochastic regressors and measurement errors: S=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). Three requirements of an instrument. 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. Multiple instruments. Application of the Durbin–Wu–Hausman test.

Simultaneous equations estimation: definitions of endogenous variables, exogenous variables, structural form equations and reduced form equations. Inconsistency of OLS on structural form equations. Use of instrumental variables. Exact identification, under identification, and over identification. 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 and Probit analysis. Maximum likelihood estimation of the population mean and variance of a random variable. Maximum likelihood estimation of regression coefficients. Likelihood ratio tests and asymptotic t tests.

Models using time series data: assumptions for regressions with time series data. Static demand functions fitted using aggregate time series data.

Please consult the current EMFSS Programme Regulations for further information on the availability of a course, where it can be placed on your programme's structure, and other important details.

Models using time series data (continued): lagged variables and naive attempts to model dynamics. Long run and short run effects. 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.

Autocorrelation: definition of autocorrelation. Consequences of autocorrelation. Breusch–Godfrey lagrange multiplier, and Durbin–Watson d tests for autocorrelation. AR(1) nonlinear regression. Potential advantages and disadvantages of such estimation, in comparison with OLS. Cochrane–Orcutt iterative process. Use of autocorrelation and heteroscedasticity consistent (Newey-West) standard errors. 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. Difference-stationarity and Trend-stationarity. Integrated of order 1. Spurious regressions. Granger–Newbold experiments with random walks. Unit root tests. Akaike Information Criterion and Schwarz's Bayes Information Criterion. Co-integration. Error correction models.

Please consult the current EMFSS Programme Regulations for further information on the availability of a course, where it can be placed on your programme's structure, and other important details.