

Applied Econometrics and Big Data

LSE SUMMER SCHOOL

This course will provide a solid grounding in recent developments in applied micro-econometrics, including state-of-the art methods of applied econometric analysis.

The course will combine both analytical and computer-based (data) material to enable students to gain practical experience in analysing a wide variety of econometric problems. It will also discuss how modern data science approaches can be used to answer important economic questions. Students will be reading various applied economic papers which apply the techniques being taught. Applications that will be considered include labour, development, industrial organisation and finance.

The topics include analysis of matching methods, identification of average, local average and marginal treatment effects using instrumental variables, regression discontinuity, randomised control experiments, post-estimation diagnostics, cross section and panel data with static and dynamic models, binary choice models and binary classification methods in machine learning, maximum likelihood estimation, ridge regression, lasso regression, and principal component regression.

Lectures are complemented with computing exercises using real data in R or Stata.

This course is ideal for advanced undergraduate students, graduate students, early-career academic researchers, and researchers in the public, private or non-profit sector.

Programme details

Lectures: 36 hours

Classes: 18 hours

Assessment*: Two written examinations and two computer based-exercises

Typical credit**: 3-4 credits (US) 7.5 ECTS points (EU)

Prerequisites

Participants should have a knowledge of quantitative research methods or introductory statistics, up to linear regression analysis. We expect participants to have completed an introductory economics course. In particular, the course will assume that participants have an understanding of statistical inference using t-tests and have prior experience of interpreting the results of multiple linear regression. We will review these topics briefly during the course.

Familiarity with linear algebra, calculus and statistical software R or Stata will be helpful but are not required.

Programme structure

Overview of Statistical Reasoning, and Introduction to Causal Inference (potential outcomes model, SUTVA, ATE)

Regression models: DID, FEs, IV and LATE

Standard errors: serial correlation, clustering and the bootstrap

Binary Models, Likelihood-based inference, Numerical optimisation in practice

Introduction to GMM & Practical Problems In Applied Analysis

Post-estimations diagnostics (Goodness of fit, Tests for functional form, tests for normality of errors, Leverage, influential observations and test for outliers), quantile regression and quantile treatment effects

Panel data models (dynamic)

Regression discontinuity design. Regression kink design

Matching methods

Discrete response models. Machine learning classification methods

Model selection, information criteria, Ridge and Lasso Regression

Principal Component Regression

Course outcomes

Demonstrate a solid grounding in recent developments in applied micro-econometrics, including state-of-the art methods of applied econometric analysis and their suitability to answer important economic questions.

Demonstrate facility with implementing the techniques covered in the course using statistical software on real-world datasets.

Demonstrate ability to answer economic questions of interest by using applied econometrics techniques.

Reading materials

Josh Angrist and Steve Pischke, (2009), *Mostly Harmless Econometrics*, (Princeton University Press).

Marno Verbeek, (2017), *A Guide to Modern Econometrics*, (Wiley).

James Stock and Mark Watson, (2011), *Introduction to Econometrics*, (MIT Press).

Gareth James, Daniela Witte, Trevor Hastie and Robert Tibshirani, (2017): *An Introduction to Statistical Learning: With Application in R*. (Springer). Available for free online.

