

# ECON0108 Econometrics

## Contact Details

<b>Lecturer:</b>	Prof. Lars Nesheim (L.NESHEIM@UCL.AC.UK)
<b>Office:</b>	Drayton House 308
<b>Office hours:</b>	By appointment
<b>Website:</b>	<a href="http://www.cemmap.ac.uk/cemmap/person/id/58">http://www.cemmap.ac.uk/cemmap/person/id/58</a>
<b>Lecture time and location:</b>	Lecture 1: Friday January 24: IOE, 20 Bedford Way, 808 Lectures 2-10: Monday 12:00 – 14:00, 222 Euston Road, LG01

## Required Reading

The course content is based on material in the lecture notes, required readings and the problem sets. The required readings include Chapters from the two textbooks listed below as well as selected articles.

1. [Pagan and Ullah \(1999\)](#).
2. [Li and Racine \(2007\)](#)
3. Selected academic papers.

## Coursework and Exams

There will be 8-10 problem sets. These will be available the week before the tutorial and will be due the morning before the tutorial. The first problem set is due Monday and the first tutorial is Monday.

The final examination will be an online examination based on material covered in lectures, assigned readings, and tutorial class assignments.

## Course Outline

1. Lecture 1: Nonparametric density estimation
  - [Pagan and Ullah \(1999\)](#), Chapter 2.
  - [Li and Racine \(2007\)](#), Chapter 1.
2. Lecture 2: Nonparametric mean regression
  - [Pagan and Ullah \(1999\)](#), Chapter 3.
  - [Li and Racine \(2007\)](#), Chapter 2.
3. Lecture 3: Semiparametric methods
  - [Pagan and Ullah \(1999\)](#), Chapters 4-5.
  - [Li and Racine \(2007\)](#), Chapters 6-8.
  - [Andrews \(1994\)](#)
  - [Robinson \(1988\)](#)
4. Lecture 5: Penalised estimation and model selection
  - [Zou \(2006\)](#)
  - [Horowitz and Nesheim \(2021\)](#)
  - [Belloni et al. \(2014a\)](#) and [Belloni et al. \(2014b\)](#).
5. Lecture 6: Production function estimation
  - [Akerberg et al. \(2015\)](#)

- Navarro et al. (2012)
6. Lecture 7: CCP and dynamic games
- Hotz and Miller (1993)
  - Arcidiacono and Miller (2011, 2020)
  - Aguirregabiria and Mira (2007)
7. Lecture 8-9: Random coefficient models
- Berry et al. (1995)
  - Train (2009)
  - Hu (2017); Schennach (2016)
  - Hoderlein et al. (2017)
8. Lecture 10: Dynamic Factor models
- Stock and Watson (2002)
  - Bai (2009)
  - Bańbura and Modugno (2014)

## References

- Daniel A Akerberg, Kevin Caves, and Garth Frazer. Identification properties of recent production function estimators. *Econometrica*, 83(6):2411–2451, 2015.
- Victor Aguirregabiria and Pedro Mira. Sequential estimation of dynamic discrete games. *Econometrica*, 75(1):1–53, 2007.
- Donald WK Andrews. Asymptotics for semiparametric econometric models via stochastic equicontinuity. *Econometrica: Journal of the Econometric Society*, pages 43–72, 1994.
- Peter Arcidiacono and Robert A Miller. Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, 79(6):1823–1867, 2011.
- Peter Arcidiacono and Robert A Miller. Identifying dynamic discrete choice models off short panels. *Journal of Econometrics*, 215(2):473–485, 2020.
- Jushan Bai. Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279, 2009.
- Marta Bańbura and Michele Modugno. Maximum likelihood estimation of factor models on datasets with arbitrary pattern of missing data. *Journal of Applied Econometrics*, 29(1):133–160, 2014.
- Alexandre Belloni, Victor Chernozhukov, and Christian Hansen. High-dimensional methods and inference on structural and treatment effects. *The Journal of Economic Perspectives*, 28(2):29–50, 2014a.
- Alexandre Belloni, Victor Chernozhukov, and Christian Hansen. Inference on treatment effects after selection among high-dimensional controls†. *Review of Economic Studies*, 81(2), 2014b.
- Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890, 1995.
- Stefan Hoderlein, Lars Nesheim, and Anna Simoni. Semiparametric estimation of random coefficients in structural economic models. *Econometric Theory*, 33(6):1265–1305, 2017.
- Joel L Horowitz and Lars Nesheim. Using penalized likelihood to select parameters in a random coefficients multinomial logit model. *Journal of Econometrics*, 222(1):44–55, 2021.
- V Joseph Hotz and Robert A Miller. Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60(3):497–529, 1993.

- Yingyao Hu. The econometrics of unobservables—latent variable and measurement error models and their applications in empirical industrial organization and labor economics, 2017.
- Qi Li and Jeffrey Scott Racine. *Nonparametric econometrics: theory and practice*. Princeton University Press, 2007.
- Salvador Navarro, David Rivers, Amit Gandhi, et al. On the identification of production functions: How heterogeneous is productivity? In *2012 Meeting Papers*, number 105. Society for Economic Dynamics, 2012.
- Adrian Pagan and Aman Ullah. *Nonparametric econometrics*. Cambridge university press, 1999.
- Peter M Robinson. Root-n-consistent semiparametric regression. *Econometrica: Journal of the Econometric Society*, pages 931–954, 1988.
- Susanne M Schennach. Recent advances in the measurement error literature. *Annual Review of Economics*, 8:341–377, 2016.
- James H Stock and Mark W Watson. Forecasting using principal components from a large number of predictors. *Journal of the American statistical association*, 97(460):1167–1179, 2002.
- Kenneth E Train. *Discrete choice methods with simulation*. Cambridge university press, 2009.
- Hui Zou. The adaptive lasso and its oracle properties. *Journal of the American statistical association*, 101(476):1418–1429, 2006.

## ECON0108: Topics covered in lectures 1-5, 2024-5

In this 5-lecture sequence, econometric models are motivated as sets of restrictions defining admissible data generating processes. We see data and would like to know: which particular process delivered that data; with large amounts of data could all but one admissible process be ruled out as the data generator; could it be that a model is misspecified in the sense that no process it admits can generate the data we see? We set up a framework in which such questions can be answered.

We study: econometric model construction, the identifying power of models, analogue estimation built on identifying correspondences, strongly and weakly restrictive models built on parametric, semi- and non-parametric restrictions, overidentification and the falsifiability of restrictions, and partial identification. The focus is almost entirely on structural econometrics.

Economic understanding, argument and theory deliver some of the restrictions employed in structural econometric models. Other restrictions have a less tangible basis. Less restrictive models are more credible, so we study which restrictions are essential if information about interesting economic magnitudes is to be got from economic data. Central in this endeavour is the concept of the identifying power of an econometric model.

In these five lectures we make a careful study of **identification** first considering the concepts involved and then methods for determining whether a model and a probability distribution of observable variables identify a particular structural feature. Since parametric restrictions are rarely credible, we study the identifying power of semi- and non-parametric models, first considering models with additive unobservable variables and then non-additive models.

Weakly restrictive models may admit the possibility that more than one structure generates the probability distribution of the random variables whose realizations we see. In this case there may be partial identification of structural features and functionals of those features. We study methods for characterizing identified sets of economic magnitudes.

Identifying correspondences link features of models to features of the probability distributions that structures admitted by models generate. Replacing features of distributions by estimates in these identifying correspondences leads to **analogue estimators** of structural features. We introduce and motivate OLS, GLS, IV and other estimators in this way. In doing so we provide a review of basic econometric estimators. When there is partial identification, it is necessary to consider estimators of and inference on identified sets.

The way in which sources of stochastic variation feature in models and the properties of stochastic variation are important considerations. Human tastes and experience are diverse so it is interesting to consider how multidimensional latent variables can be included in econometric models. This requires consideration of nonlinear models. We consider the nature of stochastic variation in econometric models.

Economic processes deliver many endogenous outcomes and there is interest in understanding the structural equations, correspondences, and inequalities whose simultaneous satisfaction produces values of those outcomes. Econometric models allowing endogenous explanatory variables are a central topic in the course and we study a variety of identifying models.

We show how in models involving a scalar non-additive latent variable, **quantile regression functions** can carry information about structural features.

A leading case in which non-additive error models naturally arise is when **discrete responses** are observed. We study parametric and semiparametric models for discrete outcomes and

consider their identifying power when there is endogeneity. In these cases, most of the point identifying models used in practice are either complete models or employ conditional independence restrictions under which endogeneity is removed once there is conditioning on certain observed variables. We will consider how these restrictions can be relaxed using partially identifying instrumental variable models.

We study a general class of **incomplete models** in which multiple values of unobservable variables can deliver a particular value of endogenous variables and provide examples of such models in the study of auctions, empirical IO and labour markets. We give a result on the identifying power of these, generally partially identifying, models.

Extensive notes and references will be distributed via the course Moodle page where you can find links to some key articles, exercises, solutions, slides and videos.

The topics we will study, and the planned order of attack are as follows.

**Lecture 1.** Structural econometrics. The purpose and history of identification analysis. Identification – concepts, definitions. Identification in parametric and semiparametric models. Econometric estimators as analogue estimators, OLS 2SLS, ML, GMM.

**Lecture 2.** Complete models. Linear simultaneous equations models. Reduced form. The classical analysis of identification, rank and order conditions. Triangular models, linear, nonlinear and nonparametric, additive errors. Control functions.

**Lecture 3.** Quantiles and quantile regression. Triangular models with nonadditive errors. Conditional independence restrictions. Treatment effect models. Incomplete models. Instrumental variable models with additive unobservables.

**Lecture 4.** Incomplete models. Instrumental variable (IV) models. IV methods in additive error models, flexible parametric specifications and ill-posedness. Nonparametric IV models with non-additive errors. Application to modelling returns to schooling.

**Lecture 5.** Generalized IV (GIV) models. GIV models for discrete outcomes. Partial identification. Application to modelling female labour force participation.

Andrew Chesher, September 25th, 2024

# Syllabus for ECON0108, Part 2

George Kapetanios  
(Notes by Richard J. Smith)

November-December 2024

- Week 6: Review of Probability and Statistics and Linear Regression (Part I)
- Week 7: Linear Regression (Part II) and Univariate Time Series Analysis
- Week 8: Univariate and Multivariate Time Series Analysis
- Week 9: Maximum Likelihood and Instrumental Variables
- Week 10: Generalised Method of Moments and Review of Part 2 of ECON0108