× ×

On The Specification, Identification, and Estimation of Discrete Choice Models:

Classical and Bayesian Perspectives

× ×

Course Tutor

Dr. Melvyn Weeks
Faculty of Economics, University of Cambridge

Location

ISEG/Universidade Tecnica de Lisboa

Lisbon, Portugal

Date

September 2005

The Course

For many years applied workers working in the area of discrete choice have been fully versed in the trade-off between the tractable, yet restrictive, logit model, and the flexible, yet difficult to operationalise, probit model. Thus, despite multinomial probit offering the potential of improving upon the restrictions on choice behaviour imposed by the logit model - namely restrictive substitutions patterns, enforced iid over individuals faced with multiple choices over time, and a neglect of random preference heterogeneity (RPH) over a population - the logit model and other Generalised Extreme Value (GEV) variants have, in general, dominated much of the literature. Following a long lineage of earlier work (from Thurstone (1927), Hausman and Wise (1978), to Manski and Lerman (1981)), independent work by McFadden (1989) and Pakes and Pollard (1989), has led to the emergence of a class of simulation-based estimators. This development, in conjunction with the continued growth in computer power, has meant that the well known potential of probit models now offers a viable alternative to logit models. However, it might be said that the mixed logit model - in essence a weighted average of logit choice probabilities evaluated over a distribution of parameter values - offers a more viable extension of vanilla logit. As with the probit model, mixed logit is capable of circumventing the fundamental limitations of vanilla logit. However, once the RPH is integrated out, a standard logit choice probability remains, with attendant tractability.

This course provides an overview of these developments in the discrete choice literature, and by introducing a new software package written in Ox, provides hands-on instruction on the issues involved in the specification, identification and estimation of this ever expanding class of models. In this regard the course will also cover material to be found in Eklof and Weeks (2004), and Eklof and Weeks (2003).

Objectives of The Course

The objectives of this course are to provide participants with both an introduction to the theory and application of binary and multiple response models, and an overview of recent developments in discrete choice modelling. The distinction between single and multiple index models is also made. This introduction progresses from the standard two choice logit and probit model, to models which in various ways facilitate the representation of a wide range of choice behaviour. The class of models include nested logit, mixed (random coefficient) logit, the multinomial probit model, and ordered (random coefficient) probit.

As a natural complement to the material on both the mixed logit and multinomial probit model, we also provide an introduction to the use of simulation in both a classical and Bayesian setting. Simulation is examined in the context of both the numerical integration of choice probabilities, and the use of simulation in constructing posterior distributions of parameters. This material will include both a brief overview of the econometric theory but also a hands-on introduction to the use of simulation methods in applied work. We emphasise that although throughout the course an exposition of the underlying econometric issues will be provided, the emphasis will be upon applied issues. To this end,

participants are also introduced to a new user-friendly package, *DCM* (Discrete Choice Models) which may be used to estimate a wide range of discrete choice models.

Course Outline

The course material will include

1. An introduction to programming in Ox and PcGive, in conjunction with an overview of economic modelling using the Ox-Metrics suite of programs
2. Brief historical overview of discrete choice models and associated estimators
3. Extension to the vanilla logit model
4. *Classical Simulation* An introduction to simulation-based inference in the context of discrete choice modelling. We begin with a very simple accept-reject (AR) procedure in the guise of the well known crude frequency simulator (CFS). Smoothed versions of this algorithm along with the Geweke, Hajivassiliou, Keane (GHK) simulator, the simulator of choice for many practitioners estimating multinomial probit model, are also introduced. We will refer to a number of texts including Van Dijk, Monfort, and Brown (1995) and Mariano, Weeks, and Schuermann (2000)
5. *Bayesian Simulation* We examine the curse of dimensionality in the context of Bayesian discrete choice models, and evaluate the use of data augmentation. The work of Albert and Chib (1993), and Chib and Greenberg (1996) is examined.
6. Identification of Discrete Choice Models
7. Revealed versus Stated Preference Models
8. Towards Flexible Substitution Patterns
9. Hands-on experience in the use of Ox and PcGive to estimate binary, multinomial and ordinal discrete choice models
10. An introduction to *DCM* - a new object-oriented package for estimating a wide range of discrete choice models

Applications

Throughout the course we make use of a wide range of economic applications. We select from:

- Stated Preference Models for Mobile Communications
- Discrete Choice Models of Labour Supply
- Discrete Choice Models with applications to Financial Markets
- Panel Data Model of Ordered Response: The Measurement and Determinants of Institutional Change
- Discrete Choice Models of Transport Behaviour
- Count and Ordered Response Models of Fertility Behaviour

Course Texts

Throughout the course we will make use of many published articles. However, two texts which will be consulted frequently are:

Mariano, R., Schuermann, T., and M. Weeks (2000) Simulation-Based Inference in Econometrics. Cambridge University Press.

Train, K. (2003) *Discrete Choice Methods with Simulation.* Cambridge University Press

Louviere, J., D. Hensher, and J. Swait, (2000). *Stated Choice Methods: Analysis and Applications.*Cambridge University Press

References

Albert, J., and S. Chib (1993): "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of the American Statistical Association*, 88, 669–679.

Chib, S., and E. Greenberg (1996): "Bayesian Analysis of Multivariate Probit Models," Research Paper, John M. Olin School of Business Washington University.

Eklof, M., and M. Weeks (2003): "On Specification, Identification and Estimation in a Class of Discrete Choice Models," Cambridge Working Papers in Economics, No. 0352.

——— (2004): "Estimation of Discrete Choice Models Using DCM for Ox," mimeo, University of Cambridge.

Hausman, J., and D. Wise (1978): "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogeneous Rpeferences," *Econometrica*, pp. 403–429.

Manski, C., and S. Lerman (1981): "On the Use of Simulated Frequencies to Approximate Choice Probabilities," in *Structural Analysis of Discrete Data with Econometric Applications*, ed. by C. Manski, and D. McFadden, pp. 305–319. MIT Press, Cambridge, Massachusetts.

Mariano, B., M. Weeks, and T. Schuermann (eds.) (2000): *Simulation Based Inference: Theory and Applications*. Cambridge University Press.

McFadden, D. (1989): "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 57, 995–1026.

Pakes, A., and D. Pollard (1989): "Simulation and Asymptotics of Optimization Estimators," *Econometrica*, 54, 755–785.

Thurstone, L. (1927): "A Law of Comparative Judgement," *Psychological Review*, 34, 273–286.

Van Dijk, H. K., A. Monfort, and B. W. Brown (eds.) (1995): *Econometric Inference Using Simulation Techniques*Chichester, West Sussex, England. John Wiley and Sons.

Agenda

Session 1. Overview

Session 2. Theory I: Simple Binary and Multinomial Choice Models

CourseNotes: SLIDiscOverview

- OLS when conditional moments are probabilities
- Binary probit and binary logit

Session 3. Theory II: Bivariate and Multinomial Choice Models

CourseNotes: SLIBvp

- The Bivariate Probit Model
- Multinomial Choice Models

Session 4. Theory III: The Random Utility Model
CourseNotes: SLIRUM

- The Random Utility Model
- Random versus Stated Preference
- Models in Characteristic Space
- Properties of Discrete Choice Models
- The Tractability of Logit
- The Limitations of Logit

Session 5. Theory IV: The Mixed Logit Model
CourseNotes: SLIMixedLogit

- Circumventing the IIA Assumption
- The Mixed Logit Model
- Panel Data Model

- Contrasting Error versus Random Components

Session 6. Theory V: Identification Conditions for Discrete Choice Models

CourseNotes: SLIIdentif

- Missing Data and the Problem of Identification
- Identification of parameters in the Multinomial Probit Model
- Identification of parameters in the Mixed Logit Model

Session 6. Theory VI: Towards Flexible Substitution Patterns

CourseNotes: SLIFlexiblesub

- Independence of Irrelevant Alternatives and Proportionate Substitution
- Non-proportional substitution and the Mixed Logit Model
- Error Components, Random Coefficients, and discrete choice elasticities

Session 7. Theory VII: Simulation-Based Inference In Classical And Bayesian settings

CourseNotes: SLIBayesianEstimation

- The Monte-Carlo principle
- A simple Crude frequency Simulator
- Advanced Simulators: GHK (*G*eweke, *H*ajivassiliou and *K*eane)
- Bias in Simulating Log-likelihoods
- Simulating from Posterior Distributions
- An Introduction to Gibbs and Metropolis Hastings Algorithms

Session 8. Introduction to Ox and the OxMetrics Family

CourseNotes: SLIOxMetrics

- Why Use Ox?

- The OxMetrics environment including OxEdit
- An overview of programming in Ox including the Object Oriented approach
- Estimation and Inference Using Ox and PcGive
- An Overview of the software package `DCM` (Discrete Choice Models)

Session 9. Practical I
CourseNotes: SLIOxMetrics

- An Overview of the Ox programming language
  OxFiles :Multiple files referenced in notes
- Input and Output: An introduction to the construction of datasets as input into estimations routines. This will include an overview of the `Database` class in Ox.
- Ox as an object-oriented matrix programming language

  Files: Refer to on-line Ox Documentation

  OxFiles: EconGrowthData.ox, BayesModAverage.ox, BAYESModAV0707.ox

  Data File: Bayesdata.xls
- A simple OLS estimator in matrix form: A first Ox program
  OxFiles :oxtut2e
- Calling libraries within Ox
- Estimating simple binary and multinomial choice models in Pc_Give
  Data File: Finney.in7
  OxFiles : probit1,probit2,probit3, probit4

Session 10. Practical II: An Introduction to `DCM` for *Ox*.

CourseNotes: SLIDCMIntro

- Key Features of `DCM`:
  i) Data Handling
  Moving between *multiple* and *single* row data structures
  (see *MixedLogit.ox*)

  ii) Main User Commands
- Specification and Estimation of a Conditional Logit Model Using `DCM`

  OxFiles: ConditLog.ox

- The Mixed Logit Model and Testing for Random Preference Heterogeneity
  OxFiles: MixedLogit.ox and ConditLog.ox

A Discrete Choice Model of Labour Supply
Files: Mnp.ox, married4.xls, Mnp.tex

- Specification and Estimation of a Multinomial Probit Model
- Stated Preference Models for Mobile Communications

Session 11. Practical III: DCM in OxPack for GiveWin

CourseNotes: DCMmanual, SLIDCMIntro

- Running DCM in GiveWin