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

Date
2008

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. 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.

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. Brief historical overview of discrete choice models and associated estimators
2. Extension to the vanilla logit model
3. *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)
4. *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.
5. An introduction to programming in Ox and PcGive, in conjunction with an overview of economic modelling using the Ox-Metrics suite of programs
6. Identification of Discrete Choice Models
7. Revealed versus Stated Preference Models
8. Hands-on experience in the use of Ox and PcGive to estimate binary, multinomial and ordinal discrete choice models
9. An introduction to *DCM* - a new object-oriented package for estimating a wide range of discrete choice models

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.