Bayesian Flexible Modelling of Mixed Logit Models

L. Scaccia
Dip. di Istituzioni Economiche e Finanziarie
Università di Macerata

E. Marcucci
Dip. di Istituzioni Pubbliche, Economia e Società,
Università di Roma Tre

Outline

- Random utility models and multinomial logit model
- Mixed logit models and the problem of choosing the mixing distribution
- A possible solution based on a discrete mixture of normal distributions
- An application to public transport demand data
- Conclusions

Random utility models

- Decision maker *n* chooses the alternative *i*, which maximises his own utility, among *J* possible alternatives:

$$U_{ni} > U_{nj} \quad \forall j \neq i$$

- The utility function can be written as:

$$U_{ni} = V_{ni} + \varepsilon_{ni}$$

- Thus, the probability that subject *n* chooses alternative *i* is:

$$P_n(i) = P(U_{ni} > U_{nj}) = P(\varepsilon_{nj} - \varepsilon_{ni} < V_{ni} - V_{nj}) \quad \forall j \neq i$$

Multinomial logit models (MNL)

- Assuming $\varepsilon_{ni} \sim$ i.i.d. *Extreme Value*, for all n, i

$$P_n(i) = \frac{e^{V_{ni}}}{\sum_{j=1}^{J} e^{V_{nj}}}$$

- If the utility function is linear:

$$P_n(i) = \frac{e^{\beta' x_{ni}}}{\sum_{j=1}^{J} e^{\beta' x_{nj}}}$$

Power and limitations of MNL

- ✓ Closed form formula for the choice probabilities
- ✘ Unobserved factors are uncorrelated over alternatives
- ✘ Independence from Irrelevant Alternatives (IIA)
- ✘ Proportional substitution between alternatives
- ✘ Temporal independence of unobserved factors (panel data)
- ✘ Homogeneous preference structure

Random parameter logit models

- The mixed MNL (MMNL) assumes the vector β having density $f(\beta|\theta)$,

$$P_n(i) = \int \frac{e^{\beta' x_{ni}}}{\sum_{j=1}^{J} e^{\beta' x_{nj}}} f(\beta|\theta) d\beta$$

- The latent class MNL assumes that the population is composed of C homogeneous groups

$$P_n(i) = \sum_{c=1}^{C} \left(\frac{e^{\beta_c' x_{ni}}}{\sum_{j=1}^{J} e^{\beta_c' x_{nj}}} \right) s_c$$

MMNL specification issues

- Selecting the parameters that are to be random parameters
  - Lagrange multiplier test (McFadden & Train, 2000)

- Selecting the distribution of the random parameters
  - some parameter needs to have a specific sign
  - most distributions have long tails

  - Hensher & Greene (2003) suggest empirical methods to help in the choice
  - Hess *et al.* (2005) compare different distributions using simulated and real data
  - Fosgerau and Hess (2009) proposed two approaches:
    1) improving on the flexibility of a base distribution through Legendre polynomials
    2) making use of a semi-parametric mixing distribution consisting of a discrete mixture of normal distributions

The proposed approach

- The utility of person *n* from alternative *i* in period *t* is:

$$U_{nit} = \alpha' w_{nit} + \beta'_n x_{nit} + \varepsilon_{nit}$$

- We assume:

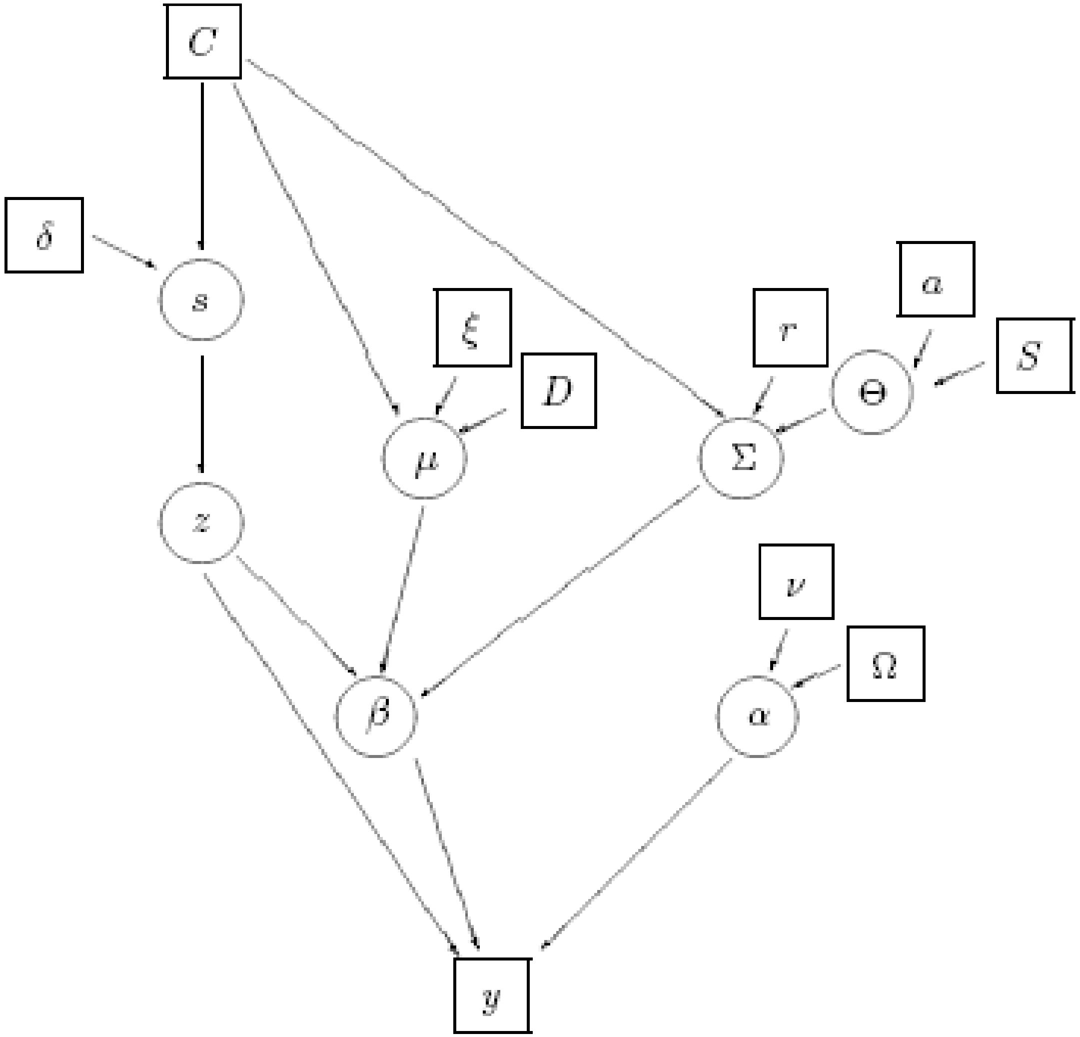
$$\beta_n \mid \mu, \Sigma \quad \sim \quad \sum_{c=1}^{C} s_c \phi\left(\cdot \mid \mu_c, \Sigma_c\right)$$

- Or, analogously, introducing latent variables $z = \{z_1, \ldots, z_N\}$:

$$\beta_n \mid z, \mu, \Sigma \quad \sim \quad \phi\left(\cdot \mid \mu_{z_n}, \Sigma_{z_n}\right)$$

Prior distributions on the parameters

- $C = 2;3$
- $(s_1, \ldots, s_C) \sim D(\delta, \ldots, \delta)$
- $z_n \sim p(z_n = c) = s_c$
- $\mu_c \sim N(\xi, D)$
- $\Sigma_c \sim IW(r, \Theta^{-1})$
- $\Theta \sim IW(a, S)$
- $\alpha \sim N(\nu, \Omega)$



Bayesian Inference

- We use MCMC to sample from the posterior joint distribution of the parameters
  - update s, z, μ, Θ and Σ through Gibbs steps
  - update α and β through Metropolis steps

- From the sample $\left(s^{(m)}, z^{(m)}, \mu^{(m)}, \Theta^{(m)}, \Sigma^{(m)}, \alpha^{(m)}, \beta^{(m)}\right)$, for $m = 1, \ldots, M$, we estimate quantities of interest, i.e.:
  - estimated individual-level taste parameters

$$\hat{\beta}_n = \frac{1}{M} \sum_{m=1}^{M} \beta_n^{(m)}$$

An application to public transport demand

- Aim of the study: analyze attributes of local public transport in Urbino (Italy) and investigate possible intervention to improve the service

- Attributes considered:
  - cost of monthly ticket (5 levels)
  - headway (5 levels)
  - first and last run (5 levels)
  - real time information displays (2 levels)
  - bus shelters (2 levels)

- Questionnaire: 15 choice exercises of which
  - 11 random
  - 2 aimed at testing the quality of the answers
  - 2 aimed at testing preference stability

- Data set: 50 respondents took part in the study, providing a data set of 750 observations

An application to public transport demand

<table>
<tr><td colspan="4">If you were to use the public transport Mercatale-Sogesta and the service would have the following characteristics, which alternative would you choose?</td></tr>
<tr><td>Prize 15.40 €</td><td>Prize 20.80 €</td><td>Prize 12.80 €</td><td rowspan="5">None of these alternatives</td></tr>
<tr><td>Headway:
30 minutes</td><td>Headway:
15 minutes</td><td>Headway:
60 minutes</td></tr>
<tr><td>First and last run
07:15 -- 01:00</td><td>First and last run
08:15 -- 24:00</td><td>First and last run
07:15 -- 01:00</td></tr>
<tr><td>No real time information displays</td><td>Real time information displays</td><td>Real time information displays</td></tr>
<tr><td>Bus shelters at Mercatale & Sogesta</td><td>Bus shelters at Mercatale & Sogesta</td><td>Bus shelters at Mercatale</td></tr>
<tr><td>A</td><td>B</td><td>C</td><td>D</td></tr>
</table>

Results

- Lagrange multiplier test to decide which parameters are to be random
  - headway
  - first and last run
- 100,000 sweeps of the MCMC algorithm, with a burn-in of 50,000 sweeps

	MMNL		MOD (2 components)		MOD (3 components)	
Parameter	Est.	Std. dev.	Est.	Std. dev.	Est.	Std. dev.
α_{cost}	-0.2671	0.0336	-0.2751	0.0346	-0.2764	0.0349
$\alpha_{displays}$	-0.0452	0.1753	-0.0190	0.1678	-0.0242	0.1828
$\alpha_{shelters}$	0.3258	0.1771	0.3167	0.1835	0.3320	0.1747
$\mu_{headway}$	-0.1039	0.0247	-0.0673	0.0806	-0.0485	0.1525
			-0.1116	0.0254	-0.0750	0.0864
					-0.1127	0.0265
$\mu_{run\ time}$	0.5322	0.0537	0.0403	0.1454	0.0180	0.1926
			0.6561	0.0798	0.0863	0.1697
					0.6840	0.0867
s	1.0000	0.0000	0.1599	0.0891	0.0503	0.0398
			0.8401	0.0891	0.1623	0.0818
					0.7873	0.0957

Table 1. *Posterior mean estimates of relevant parameters.*

Results

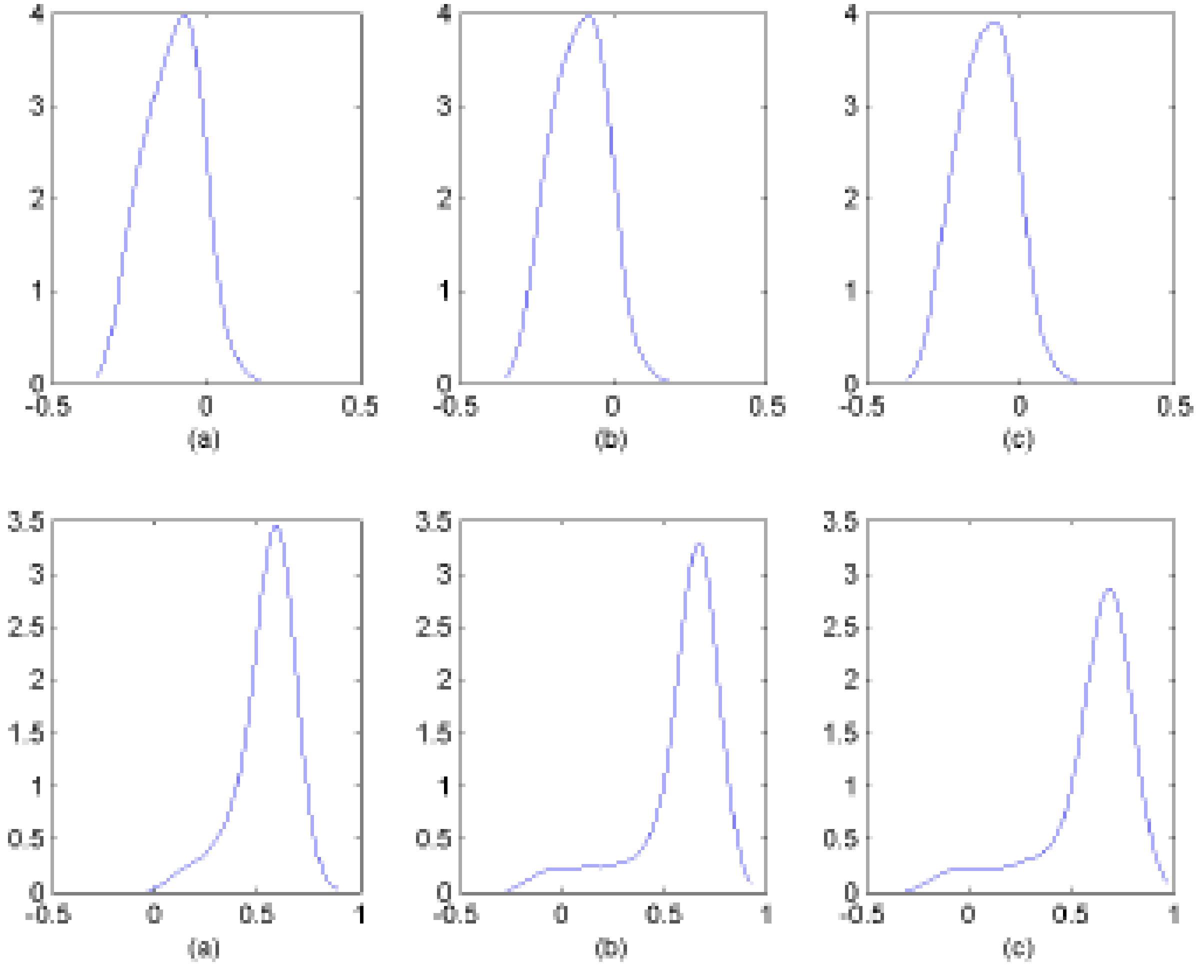


Fig. 1. Estimated marginal posterior densities for the random parameters β_{headway} (upper panel) and $\beta_{\text{run time}}$ (lower panel) under a) the MMNL model, b) the MOD model with 2 components, c) the MOD model with 3 components.

References

ALLENBY, G.M., ARORA, N. and GINTER, J.L. (1998): On the heterogeneity of demand. *Journal of Marketing Research 35 (3), 384-389.*

CIRILLO, C. and AXHAUSEN, K.W. (2006): Evidence on the distribution of values of travel-time savings from a six-week travel diary. *Transportation Research Part A: Policy and Practice 40 (5), 444-457.*

FOSGERAU, M. (2006): Investigating the distribution of the value of travel time savings. *Transportation Research Part B: Methodological 40 (8), 688-707.*

FOSGERAU, M. and HESS, S. (2009): A comparison of methods for representing random taste heterogeneity in discrete choice models. *European Transport Trasporti Europei 42, 1-25.*

HESS, S., BIERLAIRE, M. and POLAK, J.W. (2005): Estimation of value of travel-time savings using mixed logit models. *Transportation Research Part A: Policy and Practice 39 (2-3), 221-236.*

HO, R.K.W. and HU, I. (2008): Flexible modelling of random effects in linear mixed models-A Bayesian approach. *Computational Statistics and Data Analysis 52 (3), 1347-1361.*

MARCUCCI, E. and SCACCIA, L. (2005): Alcune applicazioni dei modelli a scelta discreta al settore dei trasporti. In: E. Marcucci (Ed.): *I Modelli a Scelta Discreta nel Settore dei Trasporti. Teoria, Metodologia e Applicazioni*, Carocci editore, Roma.

McFADDEN, D. (1974): Conditional logit analysis of qualitative choice behaviour. In: P.C. Zarembka (Ed.): *Frontiers in Econometrics.* Academic Press, New York, 105–142.

McFADDEN, D. and TRAIN, K. (2000): Mixed MNL models for discrete response. *Journal of Applied Econometrics 15 (5), 447-470.*

SCACCIA, L. (2010): Random parameters logit models applied to public transport demand. *Forthcoming.*

TRAIN, K. (1998): Recreation demand models with taste differences over people. *Land Economics, 74 (2), 230-239.*

TRAIN, K. (2001): A Comparison of hierarchical bayes and maximum simulated likelihood for mixed logit. *Working Paper*, Department of Economics, University of California, Berkeley.