Transportation Research Part B 117 (2018) 1–17

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Bayesian estimator for Logit Mixtures with inter- and intra-consumer heterogeneity

Felix Becker[a,*], Mazen Danaf[b], Xiang Song[b], Bilge Atasoy[c], Moshe Ben-Akiva[d]

[a] *ETH Zurich, Department of Civil, Environmental, and Geomatic Engineering, Stefano-Franscini-Platz 5, Room HIL F 34.1, Zurich 8093, Switzerland*

[b] *Massachusetts Institute of Technology, Department of Civil and Environmental Engineering, Intelligent Transportation Systems Lab, Room 1-249, USA*

[c] *TU Delft, Department of Maritime and Transport Technology, Netherlands*

[d] *Massachusetts Institute of Technology, Department of Civil and Environmental Engineering, Room 1-181, USA*

ARTICLE INFO

Article history:
Received 30 May 2017
Revised 25 May 2018
Accepted 14 June 2018

Keywords:
Hierarchical Bayes
Mixed Logit
Logit Mixture
Inter-consumer heterogeneity
Intra-consumer heterogeneity
Panel data

ABSTRACT

Estimating discrete choice models on panel data allows for the estimation of preference heterogeneity in the sample. While the Logit Mixture model with random parameters is mostly used to account for variation across individuals, preferences may also vary across different choice situations of the same individual. Up to this point, Logit Mixtures incorporating both inter- and intra-consumer heterogeneity are estimated with the classical Maximum Simulated Likelihood (MSL) procedure. The MSL procedure becomes computationally expensive with an increasing sample size and can be burdensome in the presence of a multi-modal likelihood function. We therefore propose a Hierarchical Bayes estimator for Logit Mixtures with both levels of heterogeneity. It builds on the Allenby-Train procedure, which considers only inter-consumer heterogeneity. To test the proposed procedures, we analyze how well the true patterns of heterogeneity are recovered in a simulation environment. Results from the Monte Carlo simulation suggest that falsely ignoring intra-consumer heterogeneity despite its presence in the data leads to biased estimates and a decreased goodness of fit. The latter is confirmed by a real-world example of explaining mode choices for GPS traces. We further show that the runtime of the proposed estimator is substantially faster than for the corresponding MSL estimator.

1. Introduction

Research about taste heterogeneity has traditionally focused on variation across respondents (inter-consumer). Nonetheless, some researchers (Bhat and Castelar 2002; Bhat and Sardesai 2006; Cherchi 2009; Hess and Rose 2009) emphasize the importance of considering varying preferences among different choice situations, also called menus, for one individual (intra-consumer). Accounting for variations among menus is especially important when the data are collected over a long period of time. Hess and Rose (2009) further argue that in a survey setting, individuals' preferences may alter during the course of time in which they complete the survey. For example, respondents who are in the learning phase tend to consider only a fraction of presented attributes.

* Corresponding author.
E-mail addresses: fbecker@ethz.ch (F. Becker), mdanaf@mit.edu (M. Danaf), bensong@mit.edu (X. Song), b.atasoy@tudelft.nl (B. Atasoy), mba@mit.edu (M. Ben-Akiva).

https://doi.org/10.1016/j.trb.2018.06.007
0191-2615/

In terms of estimating Logit Mixtures with inter- and intra-consumer heterogeneity, it has so far been proposed to use Maximum Simulated Likelihood (MSL) estimators (Bhat and Castelar, 2002; Hess and Rose, 2009). From the Bayesian perspective, the Allenby-Train procedure, a Hierarchical Bayes estimator for Logit mixtures with inter-consumer heterogeneity, is available. This procedure was first mentioned by Allenby in 1997 in tutorial notes at the Advanced Research Techniques forum (as cited by Train, 2009, 300), and later generalized by Train (2001). Furthermore, Dekker et al. (2016) used a Gibbs Sampler to estimate an integrated-choice latent variable (ICLV) model with inter-consumer preference heterogeneity and intra-consumer scale heterogeneity. Scale was represented as a function of individual- and menu-specific characteristics. A maximum approximate composite marginal likelihood estimator has been proposed to estimate inter- and intra-consumer heterogeneity with a Probit kernel (Bhat 2011; Bhat and Sidharthan 2011). Patil et al. (2017) further showed that MACML can outperform the Bayesian Markov Chain Monte Carlo (MCMC) approach for the multinomial probit.

Sampling from high-dimensional posterior distributions by applying MCMC methods has numerous advantages and benefits. As an example, Huber and Train (2001) accentuate that in some cases multiple local maxima exacerbate the search for the global maximum in the MSL. Regarding the estimation of variance-covariance matrices, they argue that it is computationally expensive to calculate the derivative of every element of the upper triangular matrix when using MSL. Drawing from the conditional posterior of a full variance-covariance matrix is less problematic. Section 4.2 further shows that the runtime for double mixtures is substantially shorter for MCMC than MSL. Furthermore, Train (2009, p. 283) points out that the posterior mean and standard deviation are similar to classical estimates and standard errors, provided an uninformative prior has been applied. This result enables a classical analysis of the results of Bayesian estimation where consistency and efficiency can be achieved under more relaxed conditions as compared to MSL (Train, 2009, p. 283). Rossi and Allenby (2003) highlight another advantage of MCMC methods: both population- and individual-level parameters are produced in the estimation process. With MSL, post estimation Bayesian analysis is required to compute individual-parameters.

The development of a Hierarchical Bayes estimator to incorporate both inter- and intra-consumer heterogeneity is the main contribution of our paper.[1] The estimation procedure also provides the modeler with menu-level coefficients in addition to the already existing estimation of individual-level coefficients allowing for a valuable new application. In a system that is continuously learning from customers, menu-level coefficients of new choice situations can now be used to update existing individual-level coefficients. This idea is elaborated in Danaf et al. (2017).

The estimator is analyzed from three different perspectives. We simulate data in order to test the estimator's ability to recover the true parameters and its forecasting performance. We further compare the estimator to its MSL counterpart in terms of estimates and runtime. Lastly, we apply the estimator to transportation mode choice for GPS traces.

The remainder of the paper is broken down as follows: Section 2 describes the methodology behind the model formulation for Logit Mixtures with inter- and intra-consumer heterogeneity as well as the new estimator. Section 3 describes the framework to test the new estimator, and Section 4 presents the results. Discussion and conclusion follow in Sections 5 and 6.

2. Methodology

2.1. Model for Logit Mixtures with inter- and intra-consumer heterogeneity

The model used in this paper is assumed to have a logit kernel with a linear utility specification of choice j in menu m as shown in Eq. (1):

$$\mathrm{U_{jmn}} = \mathrm{X_{jmn}}\eta_{\mathrm{mn}} + \epsilon_{\mathrm{jmn}} \tag{1}$$

with $\mathrm{U_{jmn}}$ indicating individual n's unobserved utility of alternative j in menu m and $\mathrm{X_{jmn}}$ denoting alternative attributes. Note that each individual n is presented $\mathrm{M_n}$ menus and each menu m has $\mathrm{J_{mn}}$ alternatives. The error term $\varepsilon_{\mathrm{jmn}}$ follows the Gumbel distribution.

A model formulation for Logit Mixtures with only inter-consumer heterogeneity has three sets of parameters: the vector of sample-level parameters μ, the individual parameters ζ_n for every individual n, and the inter-consumer covariance matrix Ω_b. In order to account for intra-consumer heterogeneity, we add the menu-level parameters η_{mn} for every menu m of every individual n in the sample as well as the intra-consumer covariance matrix Ω_w.

We assume that ζ_{n} and η_{mn} are normally distributed as shown in Eqs. (2) and (3). Readers interested in varying the distributional assumptions of the parameters are referred to Train (2009, pp. 305–7).

$$\eta_{mn} \sim \mathcal{N}(\zeta_n,\ \Omega_w) \tag{2}$$

$$\zeta_n \sim \mathcal{N}(\mu, \Omega_b) \tag{3}$$

The probability not conditional on the hyperparameters η and ζ is presented in Eq. (4):

$$P(d_n|\mu,\ \Omega_b,\ \Omega_w) = \int_{\zeta_n} \prod_{m=1}^{M_n} \left[\int_{\eta_{mn}} \prod_{j=1}^{J_{mn}} P_j(\eta_{mn})^{d_{jmn}} h(\mathrm{d}\eta_{mn}|\zeta_n,\ \Omega_w) \right] f(\mathrm{d}\zeta_n|\mu, \Omega_b), \tag{4}$$

[1] The code is available on request.