Available online at www.sciencedirect.com
ScienceDirect
Economics Letters 96 (2007) 282–286
economics letters
www.elsevier.com/locate/econbase

A bootstrap-based minimum bias maximum simulated likelihood estimator of Mixed Logit

Athanasios G. Tsagkanos *

Department of Business Administration, University of Patras, Greece

Received 15 March 2006; received in revised form 3 December 2006; accepted 18 January 2007
Available online 18 May 2007

Abstract

We propose a new estimator for the choice probabilities of a Mixed Logit model based on Bootstrapping. We show that using Bootstrapping the resulting maximum simulated likelihood estimator exhibits minimum bias.

Keywords: Maximum simulated likelihood estimator; Minimum bias; Bootstrap

JEL classification: C15

1. Introduction

Recently, the Mixed Logit model (MXL) has emerged as the most popular discrete choice model due to our better comprehension of simulation methods in conjunction with improvements in computer speed that allowed the full power of MXL (Train, 2003).

In order to derive the MXL, let us begin with the utility function (U) of the n-th economic agent ($n=1,\ldots, N$), in a static setup, faced with J alternative choices:

$$U_{nj} = \beta_n{}' x_{nj} + \varepsilon_{nj}. \tag{1}$$

Where ε_{nj} is an iid random variable generated by an extreme value distribution, β_n denotes a vector of unobserved coefficients to be estimated with density $f\langle\beta|\theta\rangle$ and θ its relevant parameters. Finally, (x) stands for a vector of observed explanatory variables.

* University Campus - Rio, P.O. Box 1391 Patras 26500 Greece. Tel.: +30 6938113026.
E-mail address: s6980031@yahoo.com.

0165-1765/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2007.01.016

Assume that the functional form $f(.)$ has been specified then the unconditional choice probabilities are given by:

$$P_{nj}\int L_{nj}(\beta)f\langle\beta|\theta\rangle\mathrm{d}\beta = P_n(\theta) \tag{2}$$

which are functions of θ. The corresponding conditional probabilities; $L_{nj}(\beta)=\frac{e^{\beta' x_{nj}}}{\sum_{h=1}^{J}}e^{\beta' x_{nh}}$, are independent of θ given that β is fixed for all n.

The unconditional choice probabilities are approximated by simulation for any given value of θ. Initially, one draws a value of β from $f\langle\beta|\theta\rangle$ denoted as β' and then calculate the logit formula $L_{nj}(\beta^r)$. This process is repeated a number of times and results are averaged out. The resulting average is an unbiased estimator of $P_n(\theta)$ by construction and is as follows:

$$\breve{P}_{nj}=\frac{1}{R}\sum_{r=1}^{R} L_{nj}(\beta^r)=\breve{P}_n(\theta) \tag{3}$$

where R is the number of draws.

As it becomes apparent the set of parameters θ are estimated by simulation in the context of MXL. In some more detail, the estimator is expressed as

$$\hat{\theta}=\theta^*-\breve{D}^{-1}(A+B+C) \tag{4}$$

where θ^* is the true value of θ, $\breve{D}=\frac{\partial\breve{g}(\theta^*)}{\partial\theta'}$ and $\breve{g}(\theta)$ is the sample mean of the simulated values $\breve{g}_n(\theta)=\frac{\partial\ln\breve{P}_n(\theta)}{\partial\theta}$.

In addition, A is the same as arises for the traditional non-simulation based estimator, $g(\theta^*)$, B is the simulation bias and C the simulation noise. Typically, the parameters θ are estimated by the method of maximum simulated likelihood (MSL) that is asymptotically equivalent to maximum likelihood provided that R rises faster than $\sqrt{N}$. However, samples of data used in social sciences have a constant size (N) due to either increased cost of sampling or unavailability of alternative samples. In addition, estimation is conducted only once implying that R is fixed. Consequently, these constraints lead to non-zero values for B and C.

The present study focuses on the minimization of B, by proposing a bootstrap based estimator $\hat{\theta}$ exhibiting minimum bias. The extant literature has also proposed other methods of estimation such as the method of simulated moments (MSM), the method of simulated scores and hierarchical bayes (for details see Train, 2003). However, these methods although lead to unbiased estimators may suffer from other equally undesirable drawbacks. For instance, the MSM is less efficient than MSL unless the ideal instruments are used.

2. A closer look into the bias term

The simulation bias for any estimator is as follows:

$$B=E_r\breve{g}(\theta)-g(\theta)=\frac{1}{N}\sum_{n=1}^{N} E_r\breve{g}_n(\theta)-g_n(\theta). \tag{5}$$