European Transport \ Trasporti Europei n. 42 (2009): 1-25

A comparison of methods for representing random taste heterogeneity in discrete choice models

Mogens Fosgerau [1*], Stephane Hess [2]**

[1] Technical University of Denmark, Copenhagen
Center for Transport Studies, Sweden
[2] Institute for Transport Planning and Systems, ETH Zürich

Abstract

This paper reports the findings of a systematic study using Monte Carlo experiments and a real dataset aimed at comparing the performance of various ways of specifying random taste heterogeneity in a discrete choice model. Specifically, the analysis compares the performance of two recent advanced approaches against a background of four commonly used continuous distribution functions. The first of these two approaches improves on the flexibility of a base distribution by adding in a series approximation using Legendre polynomials. The second approach uses a discrete mixture of multiple continuous distributions. Both approaches allow the researcher to increase the number of parameters as desired. The paper provides a range of evidence on the ability of the various approaches to recover various distributions from data. The two advanced approaches are comparable in terms of the likelihoods achieved, but each has its own advantages and disadvantages.

Keywords: Random taste heterogeneity; Mixed logit; Method of sieves; Mixtures of distributions.

1. Introduction

The widespread use of models such as the Mixed Multinomial Logit (MMNL) model (cf. Revelt and Train, 1998; Train, 1998; McFadden and Train, 2000; Hensher and Greene, 2003; Train 2003) has made the issue of choosing a mixing distribution very important. In these models we must specify a mixing distribution, i.e. a distribution of random parameters, that may be interpreted as representing random taste heterogeneity. The trouble is that we never observe these random parameters and that we mostly have little a priori information about the shape of their distribution except possibly a sign constraint. On the other hand, the choice of a specific distribution may seriously bias results if that distribution is not suitable for the data (cf. Hess *et al.*, 2005; Fosgerau,

* Corresponding author: Mogens Fosgerau (mf@dtf.dk)
** Corresponding author: Stephane Hess (stephane.hess@ivt.baug.ethz.ch)