Transportation Research Part B 43 (2009) 19–35

Contents lists available at ScienceDirect

Transportation Research Part B

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

Efficient stated choice experiments for estimating nested logit models

Michiel C.J. Bliemer [a,b], John M. Rose [a,*], David A. Hensher [a]

[a] Institute of Transport and Logistics Studies, Faculty of Business and Economics, The University of Sydney, NSW 2042, Australia
[b] Transport and Planning Section, Faculty of Civil Engineering and Geosciences, Delft University of Technology, The Netherlands

ARTICLE INFO

Article history:
Received 3 April 2007
Received in revised form 21 May 2008
Accepted 21 May 2008

Keywords:
Stated choice
Efficient experimental designs
Nested logit
Sample size

ABSTRACT

The allocation of combinations of attribute levels to choice situations in stated choice (SC) experiments can have a significant influence upon the resulting study outputs once data is collected. Recently, a small but growing stream of research has looked at using what have become known as efficient SC experimental designs to allocate the attribute levels to choice situations in a manner designed to produce better model outcomes. This research stream has shown that the use of efficient SC designs can lead to improvements in the reliability of parameter estimates derived from discrete choice models estimated on SC data for a given sample size. Unlike orthogonal designs, however, efficient SC experiments are generated in such a manner that their efficiency is related to the econometric model that is most likely to be estimated once the choice data is collected. To date, most of the research on efficient SC designs has assumed an MNL model format. In this paper, we generate efficient SC experiments for Nested logit models and compare and contrast these with designs specifically generated assuming an MNL model form. We find that the overall efficiency of the design is maximized only when the model assumed in generating the design is the model that is fitted during estimation.

1. Introduction

Stated choice (SC) data has proven useful in studying many transportation related problems over the past two to three decades. For example, SC data has been used to examine the demand for a cycle-way networks (e.g., Ortúzar et al., 2000), to examine the benefits derived from various calming measures on traffic (e.g., Garrod et al., 2002), to study the influences on parking choice (e.g., Shiflan and Bard-Eden, 2001; Hensher and King, 2001; Van der Waerden et al., 2002) and to establish the value of travel-time savings (VTTS) of commuters and non-commuters (e.g., Hensher, 2001a,b). Typically, SC experiments present sampled respondents with a number of different choice situations, each consisting of a universal but finite set of alternatives defined on a number of attribute dimensions. Respondents are then asked to specify their preferred alternatives given a specific hypothetical choice context. These responses may then be used by transport modelers to estimate models of choice behavior, which depending on the type of experiment conducted, may allow for the estimation of the direct or cross-elasticities (or marginal effects) of the alternatives as well as on the marginal rates of substitution respondents are willing to make in trading between two attributes (i.e., willingness to pay measures, for example, VTTS).

Unlike most data, SC data requires that the analyst designs the experiment in advance by assigning attribute levels to the attributes that define each of the alternatives which respondents are asked to consider. Traditionally, the attribute levels are allocated to each of the alternatives according to some generated experimental design, with the most common approach being to use a fractional factorial design to generate a series of single alternatives which are then allocated to choice

* Corresponding author. Tel.: +61 2 9351 0168; fax: +61 2 9351 0088.
E-mail addresses: m.c.j.bliemer@tudelft.nl, michielb@itls.usyd.edu.au (M.C.J. Bliemer), johnr@itls.usyd.edu.au (J.M. Rose), davidh@itls.usyd.edu.au (D.A. Hensher).

0191-2615/$ - see front matter
doi:10.1016/j.trb.2008.05.008

situations using randomized, cyclical, Bayesian or foldover procedures (see for example, Bunch et al., 1994; Louviere and Woodworth, 1983; Huber and Zwerina, 1996; Sándor and Wedel, 2001).

While historically, researchers have tended to rely on orthogonal experimental designs (designs in which the attribute levels between different attributes are uncorrelated, see e.g., Louviere et al., 2000) when conducting SC studies, a small but growing number of researchers have called into question this practice (e.g., Bliemer and Rose, 2006; Carlsson and Martinsson, 2003; Ferrini and Scarpa, 2007; Huber and Zwerina, 1996; Kanninen, 2002; Kessels et al., 2006; Sándor and Wedel, 2001, 2002, 2005; Rose and Bliemer, 2006). The central argument against the use of orthogonal designs is that the properties of orthogonality in SC data are not aligned with the properties of the discrete choice models typically estimated on SC data. In linear models, such as linear regression, orthogonality is important in that it avoids problems with multicollinearity in the estimated model, but more importantly, also results in the elements of the models variance–covariance matrix being minimized. It is this second point which is of primary importance. By minimizing the elements of the variance–covariance matrix of the model, the standard errors of the parameter estimates are also minimized, which in turn ensures that the *t*-ratios of the model are maximized.

Unfortunately, discrete choice models are not linear models and the variance–covariance matrices of the parameters of such models are obtained very differently to the variance–covariance matrices of linear regression models. McFadden (1974) showed that the asymptotic variance–covariance (AVC) matrix of the multinomial logit (MNL) model can be derived from the second derivatives of the log-likelihood function of the model. The same also holds for more advanced discrete choice models. Given that (i) the log-likelihood function and second derivatives of discrete choice models are dependent on the choice probabilities obtained from choice data and (ii) that only differences in the utility of the chosen and non-chosen alternatives matter, the orthogonality of a SC design says little about the expected AVC matrix of the design.

Acknowledgement of this fact has resulted in a small but growing stream of research into experimental designs generated specifically to minimize the elements of the AVC matrices for discrete choice models. Such designs are known as efficient designs (see Bliemer and Rose, 2006 for a review of such designs). To date, most research on efficient designs have assumed an MNL model form (see Ferrini and Scarpa, 2007; Sándor and Wedel, 2002, 2005 for the sole exceptions). In this paper, we examine the generation of efficient SC experimental designs to the nested logit (NL) model form, which has become a popular tool in estimating models based on SC data in the transportation area (e.g., Bhat and Castelar, 2002; Brownstone and Small, 2005; Cherchi and Ortúzar, 2002; Hess and Polak, 2006a,b; Hess et al., 2007; Polydoropoulou and Ben-Akiva, 2001; Yao and Takayuki, 2005). Given the wide scale use of the nested logit model in SC related transportation studies, understanding how better to generate the SC designs for this model is an important issue, particularly given that the AVC matrix of the nested logit model is very different to that of an MNL model.

The NL model is a significant extension to the traditional MNL model. The primary motivation to switch from the MNL model to the NL model is the restrictive MNL assumption of independent and identically distributed (IID) error terms (and the related behavioral assumption – the independence of irrelevant alternatives (IIA) assumption). A particularly important behavioral consequence of IID and IIA is that all pairs of alternatives are equally similar or dissimilar in terms of their unobserved influences (see, for example, Ben-Akiva and Lerman, 1985; Louviere et al., 2000; Hensher and Greene, 2002; Koppelman and Wen, 1998a,b; Hensher et al., 2005). This has implications for the treatment of any attributes not observed.

In practice, it is often the case that different subsets of alternatives will share similar unobserved information content, which may translate into correlation between these unobserved influences amongst pairs of alternatives (i.e., non-zero and varying covariances for pairs of alternatives). Differences in error variance and non-zero covariances represent violations of the IID and IIA assumption. By relaxing the IID (and IIA) assumption(s) of the MNL model, the NL model overcomes these problems by allowing for different treatments of the error (co)variances across subsets of the alternatives contained within the model, hence negating the problems often associated with the MNL model.

The main contributions of the paper are twofold. First, the research presented in this paper generalizes the current state-of-the-art of efficient SC designs towards the NL model, of which the MNL model is a special case. In Section 4, we show that unlike the MNL model, dependence on the choice observations is an issue in NL models, thus making the derivations more complex. To overcome this, it is necessary to rely on analytical approximations. Secondly, through the use of case studies, we demonstrate that the choice of model type (i.e., MNL or NL in this case) and also the nesting structure during the design generation process is important for the efficiency of the choice data at the time of estimation.

The remainder of this paper is organized as follows. In Section 2, we derive the NL model as necessary background before Section 3 discusses the theory on generating efficient experimental designs. In Section 4, we derive the AVC matrix for the NL model. Section 5 presents a case study in which we generate and compare SC experimental designs, and also illustrate losses in efficiency if a different model type or nesting is used for estimation than the design is generated for.

2. Nested logit model

Adopting the definitions used in Hensher and Greene (2002),[1] the elements in the NL model have a tree structure in which the top-level alternatives are referred to as branches, and the alternatives residing at the bottom of the tree structure as ele-

[1] We will restrict ourselves to NL models with two levels, being the most common. However, the theory can be extended to include more than two levels.