I T L S

WORKING PAPER

ITLS-WP-08-13

Construction of experimental designs for mixed logit models allowing for correlation across choice observations

By

Michiel CJ Bliemer* and John M Rose

*ITLS/Delft Partnership Visiting Professor

June 2008

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES

The Australian Key Centre in
Transport and Logistics Management

The University of Sydney

Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITLS-WP-08-13

TITLE: **Construction of experimental designs for mixed logit models allowing for correlation across choice observations**

ABSTRACT: In each stated choice (SC) survey, there is an underlying experimental design from which the *hypothetical* choice situations are determined. These designs are constructed by the analyst, with several different ways of constructing these designs having been proposed in the past. Recently, there has been a move from so-called orthogonal designs to more efficient designs. Efficient designs optimize the design such that the data will lead to more reliable parameter estimates for the model under consideration. The main focus has been on the multinomial logit model, however this model is unable to take the dependency between choice situations into account, while in a stated choice survey usually multiple choice situations are presented to a single respondent. In this paper, we extend the literature by focusing on the panel mixed logit (ML) model with random parameters, which can take the above mentioned dependency into account. In deriving the analytical asymptotic variance-covariance matrix for the panel ML model, used to determine the efficiency of a design, we show that it is far more complex than the cross-sectional ML model (assuming independent choice observations). Case studies illustrate that it matters for which model the design is optimized, and that it seems that a panel ML model SC experiment needs less respondents than a cross-sectional ML experiment for the same level of reliability of the parameter estimates.

KEY WORDS: ***Stated choice, experimental design, D-efficiency, panel mixed logit***

AUTHORS: Michiel CJ Bliemer and John M Rose

CONTACT: Institute of Transport and Logistics Studies **(C37)**
The Australian Key Centre in Transport Management
The University of Sydney NSW 2006 Australia

Telephone: +61 9351 0071
Facsimile: +61 9351 0088
E-mail: itlsinfo@itls.usyd.edu.au
Internet: http://www.itls.usyd.edu.au

DATE: June 2008

1. Introduction

Fundamental to any stated choice (SC) study is an underlying experimental design, the generation of which is directly under the control of the research analyst. Unfortunately, not all experimental designs are equal in terms of the statistical properties that they display and as such, it is up to the analyst to select an experimental design that enhances not only the statistical properties of the design itself as well as choice data collected using that design, but also a design that compliments the statistical properties of the models on which such data are applied. Several researchers have already addressed the issue of how to construct experimental designs that display favorable statistical properties (e.g., Bliemer and Rose, 2006; Bliemer *et al*., 2007; Carlsson and Martinsson, 2002; Ferrini and Scarpa, 2007; Huber and Zwerina, 1996; Kanninen, 2002; Kessels *et al*., 2006; Sándor and Wedel, 2001, 2002, 2005; Rose and Bliemer, 2006). Central to this literature on generating experimental designs for SC data is the minimization of the standard errors obtained from data collected using the experimental design. With the minimization of the standard errors comes the maximization of the asymptotic *t*-ratios which in turn means more efficient estimators at a given sample size.

With the exceptions of Bliemer *et al*. (2007), Ferrini and Scarpa (2007), and Sándor and Wedel (2002, 2005), research into the generation of SC experiments have largely assumed that data collected using the design will be limited in analysis to the use of the simple MNL model. The MNL model remains versatile and resilient many years after its introduction by McFadden (1974) however, it does suffer from a number of limitations. The three most important limitations of the MNL model are i) that the model does not easily accommodate the presence of preference heterogeneity within choice data, ii) that it does not allow for the fact that with SC data, each decision maker typically responds to multiple choice tasks, and iii) that the MNL model imposes a constant error variance assumption across all alternatives within the model.

The incorporation of preference heterogeneity into the MNL model is possible given the inclusion of interaction terms, however incorporating these into the experimental design is a non trivial task, particularly as the interaction terms would likely be between the design attributes and the characteristics of individual respondents and not just simply between the design attributes themselves. Failure to accommodate such interaction terms into the design process will significantly impact upon the statistical properties of the design (see Rose and Bliemer, 2006). The second issue related to the presence of repeated choice observations has traditionally been handled using either bootstrapping or jackknife techniques to adjust the standard errors of the model. Nevertheless, Ortúzar *et al*. (2000) tested both methods and found using four different data sets that the standard errors varied inconsistently and against expectations and concluded that the use of these methods were unsatisfactory for correcting for the repeated choice observation problem. The final issue, that of constant error variances across alternatives may only be addressed by moving to alternative model forms, such as the nested logit (NL) model (see Carrasco and Ortúzar (2002) for a detailed discussion of the NL model and its history).

Bliemer *et al*. (2007) extended the literature on designing SC experiments by addressing the third limitation of the MNL model mentioned above by accounting for differences in the error variances of alternatives when generating experimental designs by utilizing NL models. Ferrini and Scarpa (2007) also accounted for the same issue via an error components model structure, and in doing so also addressed the second issue of repeated choice observations by allowing for a correlated error structure across choice situations. Sándor and Wedel (2002, 2005) address the first issue of accommodating preference heterogeneity by generating experimental designs expressly for mixed logit models, but did so without addressing the second issue of having repeated choice observations at the individual respondent level.

One major benefit of using the mixed logit model is that depending on how the model is specified, the model may allow for i) the incorporation of preference heterogeneity (see e.g., Ben-Akiva *et al.*, 1993; Ben-Akiva and Bolduc, 1996; Bhat, 1996; Boyd and Mellman, 1980; Cardell and Dunbar, 1980; Brownstone and Train, 1999; Hensher and Greene, 2003; McFadden and Train, 2000), ii) an accommodation of within respondent correlation across repeated choice observations (Revelt and Train, 1998) and iii) non-constant error variances across alternatives via a relaxation of the IID assumption (see e.g., Hensher and Greene, 2003, or Train, 2003). The specification utilized in Sándor and Wedel (2002, 2005) addressed only the first and third issue. In SC experiments where respondents face multiple choice situations, particularly the second issue is of importance, and has so far not been accounted for in the design of these experiments.

In this paper, we address the issue of designing SC experiments in a way that accommodates all three limitations of the MNL model. We do this by use of the panel formulation of the mixed logit model, a formulation that explicitly incorporates the possibility of correlations over the multiple choice observations made by individual respondents, whilst also allowing for preference heterogeneity and a relaxation of the IID assumption (see e.g., Revelt and Train, 1998, Train, 2003, or Ortúzar and Willumsen, 2001). The use of the panel formulation of the mixed logit model requires a different specification of the log-likelihood function of the mixed logit model to that used in Sándor and Wedel (2002, 2005). Unfortunately, the log-likelihood function for the panel specification of the mixed logit model introduces an increased degree of complexity into the generation processes of experimental designs, in that the analyst now has to deal not with summations, but rather with products. Where possible, however, we mitigate this complexity so as to make the problem of design generation as tractable as practicable. In addition to the incorporation of panel effects, we generalize the model to not only assume normally distributed random parameters, as has been done in the past within the SC experimental design literature, but also allow for other distributions such as the uniform distribution.

The remainder of the paper is organized as follows. In the next section, we outline the panel formulation of the mixed logit model which precedes discussion on the issue of generating SC experiments using this model formulation. The paper then provides a brief discussion on the differences between the panel and the cross-sectional formulations of the model before we introduce three case studies in which we compare and contrast designs generated for different discrete choice models in order to demonstrate differences in the statistical properties of each. After the case studies, we discuss the issue of misspecification of prior parameters before we move onto a general discussion and concluding remarks.

2. The panel mixed logit model

Let U_{nsj} denote the utility of alternative j perceived by respondent n in choice situation s, which consists of an observed component V_{nsj} and an unobserved component ε_{nsj},

$$U_{nsj} = V_{nsj} + \varepsilon_{nsj}. \tag{1}$$

As is common practice, the observed component is assumed to be described by a linear relationship of observed attribute levels of each alternative, x, and their corresponding weights (parameters), β,

$$V_{nsj} = \sum_{k=1}^{K} \beta_k x_{nsjk}. \tag{2}$$

In case a certain parameter β_k appears in the utility function of multiple alternatives j, it is said to be generic over these alternatives. Otherwise, the parameter is called alternative-specific. In our notation, if a certain attribute k does not appear in the utility function of a certain alternative j, then we assume that $x_{nsjk} = 0$.

When the unobserved components are assumed to be identically and independently extreme value type 1 (EV1) distributed, then the probability P_{nsj} that respondent n chooses alternative j in choice situation s is given by the *multinomial logit* model (see McFadden, 1974),

$$P_{nsj} = \frac{\exp\left(V_{nsj}\right)}{\sum_{i \in J_{ns}} \exp\left(V_{nsi}\right)}, \tag{3}$$

where J_{ns} is the set of alternatives presented to respondent n in choice situation s. Typically, the parameters β are unknown and one is interested in estimating these parameters from data, being either revealed choice data, or in our case, stated choice data. Let y_{nsj} equal one if j is the chosen alternative in choice situation s shown to respondent n, and zero otherwise. In other words, y represents the outcomes of a stated choice experiment. Then the parameters can be estimated by maximizing the likelihood function L,

$$L = \prod_{n=1}^{N} \prod_{s \in S_n} \prod_{j \in J_{ns}} \left(P_{nsj}\right)^{y_{nsj}}. \tag{4}$$

where N denotes the total number of respondents and S_n is the set of choice situations faced by respondent n.

In case of a *mixed logit* model, we assume that (some of) the parameters are random, following a certain probability distribution. In that case, the expected likelihood function is maximized in order to estimate the distributional parameters, with

$$\begin{aligned} E(L) &= E\left(\prod_{n=1}^{N} \prod_{s\in S_n} \prod_{j\in J_{ns}} \left(P_{nsj}\right)^{y_{nsj}} \right) \\ &= \prod_{n=1}^{N} E\left(\prod_{s\in S_n} \prod_{j\in J_{ns}} \left(P_{nsj}\right)^{y_{nsj}} \right), \end{aligned} \tag{5}$$

in which the second term holds since we assume that all respondents make their decisions independent of each other. If two random variables, say p and q, are independent, then $E(pq) = E(p)E(q)$. However, in a SC experiment, we cannot assume that the probabilities for a single respondent in multiple choice situations are independent, as they have the same underlying behavior. The above formulation that explicitly takes into account this dependency is called the *panel mixed logit* model. The expectation is over the random β values, which make the probabilities P_{nsj} random as well.

Instead of maximizing the likelihood, commonly the log-likelihood function is maximized, being given by

$$\log E(L) = \sum_{n=1}^{N} \log E\left(\prod_{s\in S_n} \prod_{j\in J_{ns}} \left(P_{nsj}\right)^{y_{nsj}} \right). \tag{6}$$

If the choice observations from a single respondent over a series of choice situations are assumed independent, then the likelihood function can be written as

$$E(L) = \prod_{n=1}^{N} \prod_{s\in S_n} E\left(\prod_{j\in J_{ns}} \left(P_{nsj}\right)^{y_{nsj}} \right), \tag{7}$$

such that the log-likelihood function can be simplified to

$$\log E(L) = \sum_{n=1}^{N} \sum_{s\in S_n} \sum_{j\in J_{ns}} y_{nsj} \log E\left(P_{nsj}\right). \tag{8}$$

This equation represents the log-likelihood function of well-known *cross-sectional mixed logit* model. In the remainder of this paper we will primarily focus on the panel mixed logit model, as this provides the correct modeling framework taking into account that a single respondent makes multiple choices. However, we will use the cross-sectional mixed logit model for comparison in our case studies.

Define the probability P_n^* that a certain respondent n has made a certain sequence of choices $\{j \mid y_{nsj} = 1\}_{s\in S_n}$ with respect to the set of choice situations, S_n, by

$$P_n^* = \prod_{s\in S_n}\prod_{j\in J_{ns}}\left(P_{nsj}\right)^{y_{nsj}}, \tag{9}$$

such that the (panel) log-likelihood function can be written as

$$\log E(L) = \sum_{n=1}^{N}\log E\left(P_n^*\right). \tag{10}$$

This probability P_n^* depends on the random parameters β, such that the expected probability can be written as

$$E\left(P_n^*\right) = \int_\beta P_n^*(\beta)\, f(\beta\,|\,\theta)\,d\beta, \tag{11}$$

where $f(\beta\,|\,\theta)$ is the multivariate probability density function of β, given the distributional parameters θ. By using a transformation of β such that the multivariate distribution becomes non-parametrical, we can write Eqn. (11) as

$$E\left(P_n^*\right) = \int_z P_n^*\left(\beta(z\,|\,\theta)\right)\phi(z)\,dz, \tag{12}$$

where $\beta(z\,|\,\theta)$ is a function of z with parameters θ, and where $\phi(z)$ is the multivariate non-parametrical distribution of z. It is common to use several (independent) univariate distributions[1] instead of using a single multivariate distribution, such that Eqn. (12) can be written as

$$E\left(P_n^*\right) = \int_{z_1}\cdots\int_{z_K} P_n^*\left(\beta_1(z_1\,|\,\theta_1),\ldots,\beta_K(z_K\,|\,\theta_K)\right)\phi_1(z_1)\cdots\phi_K(z_K)\,dz_1\cdots dz_K. \tag{13}$$

Having separate univariate distributions for each parameter has the benefit that different distributions can be easily mixed. For example, if $\beta_1 \sim N(\mu,\sigma)$, and $\beta_2 \sim U(a,b)$, then $E\left(P_n^*\right)$ is written as

$$E\left(P_n^*\right) = \int_{z_1}\int_{z_2} P_n^*\left(\beta_1(z_1\,|\,\mu,\sigma),\beta_2(z_2\,|\,a,b)\right)\phi_1(z_1)\phi_2(z_2)\,dz_1 dz_2, \tag{14}$$

where $\beta_1(z_1\,|\,\mu,\sigma) = \mu + \sigma z_1$ with $z_1 \sim N(0,1)$ following a standard normal distribution, and $\beta_2(z_2\,|\,a,b) = a + (b-a)z_2$ with $z_2 \sim U(0,1)$ following a standard uniform distribution. Other distributions can be used as well, such as the log-normal distribution in which the transformation $\beta(z\,|\,\mu,\sigma) = e^{\mu}e^{\sigma z}$ is used, with $z \sim N(0,1)$. Note that a fixed parameter is a special case of a random parameter, such that all equations also hold in the case that only some of the parameters are considered random. For a fixed parameter β_k we simply take $\beta_k(z_k\,|\,\mu_k) = \mu_k$, and $\phi_k(z) = 1$.

---

[1] Note that if one would not like to assume independent random variables, then one can sample directly from the multivariate distribution. In case of a multivariate normal distribution, this is possible through a Cholesky decomposition, see e.g., Greene (2002).

3. Efficient designs for the panel mixed logit model

Given the attribute levels x and the survey outcomes y, the model (distributional) parameters θ can be estimated by maximizing the log-likelihood function (Eqn. 10), see e.g., Train (2003). In a SC experiment, the attribute levels x are given by the underlying experimental design. Creating an efficient design is actually the inverse problem, in which the (distributional) parameter estimates are assumed given (as fixed or Bayesian priors), and one would like to determine optimal attribute levels x which will maximize the so-called design efficiency, which in turn is a measure for the reliability of the parameter estimates. This inverse problem of finding an efficient design is far more complex than the estimation problem. Generating efficient designs have been discussed in Huber and Zwerina (1996), Kanninen (2002), and Sándor and Wedel (2001) for multinomial logit models with all generic parameters, and in Carlsson and Martinsson (2002) and Rose and Bliemer (2005) for multinomial logit models with generic and alternative-specific parameters. Furthermore, efficient designs for the cross-sectional mixed logit model are discussed in Sándor and Wedel (2002), for the nested logit model in Bliemer *et al.* (2007), and for the error components model in Ferrini and Scarpa (2007). With the exception of the work by Ferrini and Scarpa, each of these models assume independent observations from different choice situations, which is in a SC experiment not the case when giving multiple choice situations to the same respondent. In estimation this is well-known, and panel mixed logit estimation procedures exist. However, in design generation, the dependency between choice situations has always been ignored. In this paper, we will discuss the generation of design SC experiments for panel mixed logit models, such that this dependency is taken into account for the first time.

3.1 Efficiency measures

In order to find an efficient design, the design efficiency has to be expressed in a quantitative way using a certain measure. Several measures have been proposed in the literature. All measures use the asymptotic variance-covariance (AVC) matrix of the estimates to determine the design efficiency. In simple terms, each of these measures aims to minimize the (asymptotic) standard errors, or alternatively, maximize the *t*-ratios, in case the model parameters are estimated using the design under consideration in a stated choice experiment. Rephrasing, if the design is used to construct a SC experiment and is given to a large number of respondents, then the resulting variance-covariance matrix (where the standard errors are the roots of the diagonals) when estimating the model parameters will give this AVC matrix. Hence, the AVC matrix typically depends on the survey outcomes y. However, it has been shown that for the multinomial logit model this AVC matrix can be determined independent of y (e.g., Huber and Zwerina, 1996; McFadden, 1974; Rose and Bliemer, 2005). Also, for the nested logit model (Bliemer *et al.*, 2007) and the cross-sectional mixed logit model (Sándor and Wedel, 2002, 2005), the AVC matrix can be determined independent of y. As we will show in this section, the AVC matrix for the panel mixed logit model *cannot* be determined independent of y (the same issue exists for the error components model allowing for correlated choice situations; see Ferrini and Scarpa, 2007). We will propose to create a hypothetical sample of respondents, such that the design efficiency can still be approximated. However, this comes at the cost of much longer computation times. Considering the fact that, in order to find an efficient design, thousands – if not millions – of potential experimental designs need to be evaluated, computation time is an issue.

The most widely used measure is the D-error measure (see Huber and Zwerina, 1996), which describes the *in*efficiency of a design, i.e., the lower this D-error, the more efficient the design, and is computed by taking the determinant of the AVC matrix (and scaled according to the number of parameters). Similarly, the A-error is the (scaled) trace of the AVC matrix, which is simply the sum of the diagonal elements representing (the squares of) the standard errors. Other measures can be found in e.g., Kessels *et al.* (2006). Let Ω_N denote the AVC matrix for a sample size (number of respondents) *N*. Then the D-error and A-error are defined as

$$\text{D-error} = \left(\det\left(\Omega_N\right)\right)^{1/K}, \tag{15}$$

$$\text{A-error} = \frac{\mathrm{tr}\left(\Omega_N\right)}{K}. \tag{16}$$

While in minimizing the A-error scaling of the parameters may lead to different results (standard errors of large-valued parameters are typically larger, which may overshadow the minimization of the standard errors of small-valued parameters), the D-error is not sensitive to parameter scaling by means of the determinant. In this paper we will concentrate on minimizing the D-error, such that the following problem needs to be solved:

$$\begin{aligned} &\min\left(\det\left(\Omega_N\right)\right)^{1/K} \\ &\text{s.t.} \quad x \in X. \end{aligned} \tag{17}$$

The attribute levels x are constrained to the feasible set of attribute levels X. This feasible set is determined by the experimental design dimensions chosen by the analyst and some constraints on the attribute levels. Besides the model specification (in which the number of alternatives, attributes, and parameters is chosen), the analyst has to decide on the number of choice situations for each respondent (the size of set S_n), the number of attribute levels per attribute, and the range of allowed attribute levels. In the design literature, the possible attribute levels are typically given and fixed, although approaches that pivot around respondent-specific base values exist (Rose et al., 2008). Other constraints that can be put on the attribute levels are attribute level balance (over all choice situations, each attribute level should appear an equal number of times), and orthogonality (independence between attribute levels of different attributes). For a more detailed discussion on the generation of SC experiments, see Rose and Bliemer (2008).

3.2 Deriving the AVC matrix for the panel mixed logit model

The main complexity is to determine Ω_N for a given design *x*. The remainder of this section will be devoted to the derivation of Ω_N. The AVC matrix Ω_N can be determined as the inverse of the Fisher information matrix, I_N, which in turn can be computed using the second derivatives of the log-likelihood function (10) (see Train, 2003). Mathematically,

$$\Omega_N = I_N^{-1}, \quad \text{with } I_N = -E_N\left(\frac{\partial^2 \log E(L)}{\partial\theta\partial\theta'}\right), \tag{18}$$

where $E_N(\cdot)$ is used to express the large sample population mean. Hence, the AVC matrix can be determined by calculating the Hessian matrix of the log-likelihood function for the panel mixed logit model. Let the vector of M_k parameters for the probability distribution of parameter β_k be denoted by $\theta_k = [\theta_{km}]$, assuming $m = 1,\ldots,M_k$. Using Eqn. (10), the first derivatives are

$$\frac{\partial \log E(L)}{\partial\theta_{km}} = \sum_{n=1}^{N} \frac{1}{E\left(P_n^*\right)} \frac{\partial E\left(P_n^*\right)}{\partial\theta_{km}}, \tag{19}$$

such that the second derivatives can be written as

$$\begin{aligned}\frac{\partial^2 \log E(L)}{\partial\theta_{k_1m_1}\partial\theta_{k_2m_2}} &= \sum_{n=1}^{N}\left(\frac{1}{E\left(P_n^*\right)}\frac{\partial^2 E\left(P_n^*\right)}{\partial\theta_{k_1m_1}\partial\theta_{k_2m_2}} - \frac{1}{\left(E\left(P_n^*\right)\right)^2}\frac{\partial E\left(P_n^*\right)}{\partial\theta_{k_1m_1}}\frac{\partial E\left(P_n^*\right)}{\partial\theta_{k_2m_2}}\right) \\ &= \sum_{n=1}^{N}\left(\frac{1}{E\left(P_n^*\right)}E\left(\frac{\partial^2 P_n^*}{\partial\theta_{k_1m_1}\partial\theta_{k_2m_2}}\right) - \frac{1}{\left(E\left(P_n^*\right)\right)^2}E\left(\frac{\partial P_n^*}{\partial\theta_{k_1m_1}}\right)E\left(\frac{\partial P_n^*}{\partial\theta_{k_2m_2}}\right)\right).\end{aligned} \tag{20}$$

Note that we have reversed the order of derivation and expectation, as the derivative of the expected value is the same as the expected value of the derivative since the expectation is over z, not over θ. In Eqn. (20), two terms remain to be determined,

$$\frac{\partial P_n^*}{\partial\theta_{km}} = P_n^* \frac{\partial\beta_k}{\partial\theta_{km}} \sum_{s\in S_n}\sum_{j\in J_{ns}} \frac{y_{nsj}}{P_{nsj}}\frac{\partial P_{nsj}}{\partial\beta_k}, \tag{21}$$

and

$$\begin{aligned}\frac{\partial^2 P_n^*}{\partial\theta_{k_1m_1}\partial\theta_{k_2m_2}} = \frac{1}{P_n^*}\frac{\partial P_n^*}{\partial\theta_{k_1m_1}}\frac{\partial P_n^*}{\partial\theta_{k_2m_2}} - P_n^*\frac{\partial\beta_{k_1}}{\partial\theta_{k_1m_1}}\frac{\partial\beta_{k_1}}{\partial\theta_{k_2m_2}}\sum_{s\in S_n}\sum_{j\in J_{ns}}\frac{\partial P_{nsj}}{\partial\beta_{k_2}}x_{nsjk_1} \\ + P_n^*\frac{\partial^2\beta_{k_1}}{\partial\theta_{k_1m_1}\partial\theta_{k_2m_2}}\sum_{s\in S_n}\sum_{j\in J_{ns}}\frac{y_{nsj}}{P_{nsj}}\frac{\partial P_{nsj}}{\partial\beta_{k_1}},\end{aligned} \tag{22}$$

where $\partial P_{njs}/\partial\beta_k$ is the first derivative of the multinomial logit probability,

$$\frac{\partial P_{nsj}}{\partial\beta_k} = P_{nsj}\left(x_{nsjk} - \sum_{i\in J_{ns}} P_{nsi}x_{nsik}\right). \tag{23}$$

Suppose we would like to evaluate the design efficiency of design *x*. Assuming some prior values for parameters $\theta,$ we can compute the multinomial logit probabilities P_{nsj}. Furthermore, given the probability distributions of $\beta,$ $\partial\beta/\partial\theta$ and $\partial^2\beta/\partial\theta\partial\theta'$ can be determined. Note that the latter term is zero in case of the normal and uniform distribution, such that the last term in Eqn. (22) would drop out. The only unknowns are y and therefore also P_n^*. For other model types, such as the multinomial, nested, and cross-sectional mixed logit models, the outcomes *y* could be replaced by probabilities, since $E_N(y_{nsj}) = P_{nsj}$ (*y* follows a multinomial distribution). However, $E_N\left(P_n^*\right)$ cannot be approximated that easily, as it describes a generalized multinomial distribution (Beaulieu, 1991). Rewriting it to a multinomial distribution in which an alternative is redefined as a sequence of chosen alternatives over multiple choice situations is theoretically possible; however the number of possible sequences grows exponentially,[2] such that this is not practically feasible. Instead, we will generate a hypothetical sample based on the design *x* as follows. For each respondent *n*, we draw a random parameter β_k from each given parameter distribution, then determine the observed utility V_{nsj} for each choice situation *s* based on design *x*, then draw a separate unobserved component ε_{nsj} for each alternative in each choice situation, and determine y_{nsj} by selecting the alternative with the highest utility in each choice situation. Note that the same random draw for β_k is used over all choice situations for each respondent, representing the panel formulation.

The expected values in Eqn. (19) can be approximated using simulation, similar to approximations in estimation (see Train, 2003). Instead of pseudo-random simulation taking pseudo-random draws for each parameter $\beta_k,$ closer approximations with the same number of draws can be achieved by using more intelligent quasi-random draws (e.g., using Halton or Sobol sequences) or polynomial cubature methods (e.g., Gaussian quadrature). A discussion and comparison of these simulated approximations can be found in the estimation context in e.g., Bhat (2001, 2003), Hess *et al.* (2005), Sándor and Train (2004), and in the design context in Bliemer *et al.* (2008).

3.3 Cross-sectional mixed logit as a special case

The cross-sectional mixed logit model is a special case of the mixed logit model in which all choice observations from a single respondent are treated independently. Instead of using the log-likelihood function (10) we will now use Eqn. (8), yielding a first derivative of

$$\frac{\partial \log E(L)}{\partial \theta_{km}} = \sum_{n=1}^{N} \sum_{s\in S_n} \sum_{j\in J_{ns}} \frac{1}{E\left(P_{nsj}\right)} \frac{\partial E\left(P_{nsj}\right)}{\partial \theta_{km}}, \tag{24}$$

---

[2] If for a certain respondent *n*, the number of choice situations is 10 and the number of alternatives is $J_{ns} = 2$ for all choice situations *s*, then there are already $2^{10} = 1,024$ possible choice sequences. Practical problems typically have larger numbers of alternatives and choice situations. Multinomial distributions of these large dimensions lead to prohibitive computational problems in determining $E_N(P_n^*)$.

and a second derivative of

$$\frac{\partial^2 \log E(L)}{\partial\theta_{k_1m_1}\partial\theta_{k_2m_2}} = \sum_{n=1}^{N}\sum_{s\in S_n}\sum_{j\in J_{ns}} y_{nsj}\left(\frac{1}{E\left(P_{nsj}\right)}E\left(\frac{\partial^2 P_{nsj}}{\partial\theta_{k_1m_1}\partial\theta_{k_2m_2}}\right) - \frac{1}{\left(E\left(P_{nsj}\right)\right)^2}E\left(\frac{\partial P_{nsj}}{\partial\theta_{k_1m_1}}\right)E\left(\frac{\partial P_{nsj}}{\partial\theta_{k_2m_2}}\right)\right). \quad (25)$$

When taking the large sample mean of this second equation, conform Eqn. (18), we can use the substitution $E_N\left(y_{nsj}\right) = E(P_{nsj})$, where $E_N(\cdot)$ is again the large sample mean, such that the outcomes y completely drop from the equation and collapses to the formula stated in Sándor and Wedel (2002). This makes computing the AVC matrix for the cross-sectional mixed logit model a much easier task than for the panel mixed logit model, as there is no need to generate a sample, significantly reducing the computation time.

4. Case studies

In this section we generate designs for three different case studies. In each of the three case studies, we compare and contrast three different experimental designs; the first represents an experimental design optimized assuming an MNL model, the second a design optimized for a cross-sectional mixed logit model (assuming independent choice observations) and the last design assuming the panel formulation of the mixed logit model. In the second case study, a fourth design is also generated, that being an orthogonal design[3].

In all cases, we have optimized the designs on the D-error criteria; however, for completeness we also report the efficiency of the designs using other efficiency criteria. Gaussian quadrature with 5 abscissas is being used (see e.g., Bliemer *et al.*, 2008) for simulating the probabilities in the mixed logit models, and an additional sample of 500 respondents is generated (using Halton draws for the EV1 error terms in the utility functions) for computing the panel mixed logit results. The procedures described in this report have been implemented in the Ngene[4] software, which is used to generate the efficient designs in all case studies. Reported computation times are for evaluating 100,000 designs in a randomize-and-swap algorithm, based on a notebook computer running Windows XP with a 2.0Ghz Pentium M processor and 1GB RAM.

4.1 Case study 1

The first case study follows that of Kessels *et al.* (2004) in terms of the design dimensions that we have chosen to explore. The experiment assumes each respondent observes two alternatives described by four attributes. The utility specification for the case example is given as:

3 In this paper, we have deliberately avoided discussion of orthogonal designs. Orthogonal designs are only optimal for linear models (which does not describe models within the logit class of models), or when all prior parameters are assumed to be zero for linear models (see Bliemer and Rose (2006) for a detailed discussion of this). Street and Burgess (2004) and Street et al. (2005) claim to generate optimal SC designs which are orthogonal. A careful examination of their research suggests that these designs are only optimal under the zero prior parameter assumption (indeed, they do not assume any priors) and that the designs that they generate can only be used for models assuming generic parameters.

4 Ngene is being developed by Econometric Software. Prototype version 0.9 was used.

$$V_j = \beta_1 x_{j1} + \beta_2 x_{j2} + \beta_3 x_{j3} + \beta_4 x_{j4}, \quad j = 1,2, \tag{26}$$

where each attribute takes on three levels (i.e., $x_{jk} = \{1,2,3\}$). For the case example however, we select different priors to those examined in Kessels *et al*. (2004). In the original work, Kessels *et al*. choose to use priors drawn from a uniform Bayesian distribution ranging between -1 and 1, thus suggesting that the analyst had no information as to the expected parameters, not even direction. For the present study, parameters β_3 and β_4 are treated as fixed parameters, whilst the first two parameters are specified as a random parameters drawn from normal distributions, i.e., $\beta_1 \sim N(\mu_1, \sigma_1)$ and $\beta_2 \sim N(\mu_2, \sigma_2)$. The following priors are used: $\mu_1 = 0.6$, $\sigma_1 = 0.2$, $\mu_1 = -0.9$, $\sigma_2 = 0.2$, $\beta_3 = -0.2$, and $\beta_4 = 0.8$.[5]

Three designs were generated for the first case example and are shown in Table 1. The first design was generated assuming that the model was an MNL model (computation time: 43 seconds to evaluate 100,000 designs), in which the two random parameters are assumed to be fixed and their priors equal to the means. The second and third designs were generated using the mixed logit model, where in the second design independence of choice observations is assumed (computation time: 19 minutes) whilst the third design assumes the panel formulation of the model (computation time: 67 hours).

At the base of Table 1 are the efficiency results for the three designs calculated as if the designs were used to estimate different model forms. The bold values represent the optimized D-errors. If the estimated model is of the MNL type, then the MNL design will perform best, with a D-error of 0.160 (which is smaller than the D-errors of the other two designs, 0.169 and 0.165). Similarly, the cross-sectional ML design will perform best when estimating a cross-sectional ML model (with a D-error of 0.745), and the panel ML design and will perform best for estimating the panel ML model (with a D-error of 0.248). Also shown are the associated A-errors and the minimum sample sizes required for each design to obtain significant asymptotic *t*-ratios for all parameters (see Bliemer and Rose, 2005, for a discussion of how to calculate these sample sizes).

Whenever a design is not optimized for the model to be estimated, the design will loose efficiency. Reading across the rows for example, one can observe that when the design generated for the cross-sectional ML model is used to estimate a panel ML model, the design obtains a D-error of 0.278, 12 percent larger than when the panel ML design is used (with a D-error of 0.248). Reversely, when estimating the cross-sectional ML model, it is much more efficient to use a design specifically designed for this model than using a panel ML design (a D-error of 0.745 versus a D-error of 0.933, an increase of 25 percent). Note that D-error values down the columns cannot be compared, as they are dependent on the model estimated (with possibly different numbers of parameters, for example between the MNL and ML models).

---

5 Note that any parameters could have been chosen. These were selected purely for demonstrative purposes only. Had different prior parameters been chosen instead, the results reported may have varied considerably. We discuss this in the discussion section of the paper.

Table 1: Case Study 1 designs

		MNL Model				**Cross-Sectional ML Model**				**Panel ML Model**			
		Attributes				*Attributes*				*Attributes*			
s	***j***	x_{j1}	x_{j2}	x_{j3}	x_{j4}	x_{j1}	x_{j2}	x_{j3}	x_{j4}	x_{j1}	x_{j2}	x_{j3}	x_{j4}
1	1	1	1	2	1	1	1	3	2	2	1	1	1
	2	3	3	2	3	3	3	1	1	2	3	3	3
2	1	2	2	2	1	2	2	2	1	2	3	3	3
	2	2	2	2	3	2	2	2	3	2	1	1	1
3	1	3	1	2	1	2	3	2	1	2	3	2	2
	2	1	3	2	3	2	2	2	3	2	1	2	2
4	1	1	2	1	3	1	2	3	2	1	2	2	2
	2	3	2	3	1	3	2	1	2	3	2	2	2
5	1	1	3	1	3	2	1	2	2	2	1	2	1
	2	3	2	3	1	2	3	2	2	2	3	2	3
6	1	3	3	1	2	3	3	1	3	2	2	1	3
	2	1	1	3	2	1	1	3	1	2	2	3	1
7	1	1	1	3	2	3	3	1	2	3	3	3	3
	2	3	3	1	2	1	1	3	3	1	1	1	1
8	1	2	3	1	2	3	3	3	2	3	3	1	2
	2	2	1	3	2	1	1	1	2	1	1	3	2
9	1	3	2	2	1	3	2	3	1	1	2	2	2
	2	1	1	2	3	1	3	1	3	3	2	2	3
10	1	2	1	3	1	1	3	2	3	3	1	2	1
	2	2	3	1	3	3	1	2	1	1	3	2	3
11	1	3	2	2	2	1	1	1	1	1	2	1	3
	2	1	3	2	2	3	3	3	3	3	2	3	1
12	1	2	1	1	2	3	2	2	3	1	2	3	3
	2	2	2	3	2	1	2	2	2	3	2	1	1
13	1	3	3	3	3	2	1	3	3	3	1	3	1
	2	1	1	1	1	2	1	1	1	1	3	1	3
14	1	2	3	3	3	2	1	1	1	3	3	1	1
	2	2	1	1	1	2	3	3	2	1	1	3	2
15	1	1	2	3	3	1	2	1	3	1	1	3	2
	2	3	2	1	1	3	2	3	1	3	3	1	2

Efficiency results when design is applied to different model forms

	Model Assumed for Design					
	MNL Model		**Cross-Sectional ML Model**		**Panel ML Model**	
		D-error		*D-error*		*D-error*
Model Assumed for Estimation	MNL	**0.160**	MNL	0.169	MNL	0.165
	Cross-Sectional ML	0.884	Cross-Sectional ML	**0.745**	Cross-Sectional ML	0.933
	Panel ML	0.249	Panel ML	0.278	Panel ML	**0.248**
		A-error		*A-error*		*A-error*
	MNL	0.258	MNL	0.210	MNL	0.270
	Cross-Sectional ML	8.550	Cross-Sectional ML	4.672	Cross-Sectional ML	9.774
	Panel ML	0.341	Panel ML	0.360	Panel ML	0.337
		Sample Size		*Sample Size*		*Sample Size*
	MNL	12	MNL	13	MNL	14
	Cross-Sectional ML	2781	Cross-Sectional ML	1387	Cross-Sectional ML	2932
	Panel ML	44	Panel ML	67	Panel ML	40

A further point worthy of discussion is the minimum sample sizes suggested for each of the different designs. In all cases the cross-sectional ML model – treating each choice observation as an independent pseudo individual – performs particularly badly, even when the generated design is optimized for this model form. Although not presented here, a more detailed investigation shows that particularly the standard deviations σ_1 and σ_2 need large sample sizes to be statistically significant in estimation, which explains the difference in the minimum sample sizes between estimating the MNL model and the ML models. The panel ML model accepts a much lower sample size than the cross-sectional ML model in order to estimate these standard deviations, requiring only 40 respondents at a minimum (and only a sample size of 14 is needed to estimate the means and fixed parameters to a statistically significant level).

It is interesting to see, that the sample sizes between the MNL model and the panel ML model are much more similar than the sample sizes between the two ML models. This could be explained by the fact that both the MNL as well as the panel ML model assume constant respondent behavior over the choice situations, while the cross-sectional ML model assumes varying respondent behavior over the design. In this case study, it turns out to be much better to use the MNL design than the cross-sectional ML design in order to estimate the panel ML model. This is an important finding, as currently in the state-of-the-art literature, one focuses on designs for the cross-sectional ML model (e.g., Sándor and Wedel, 2002; Kessels *et al.*, 2006), while clearly the correct model to estimate in case of stated choice experiments is the panel ML model.

4.2 Case study 2

The second case study follows that of Huber and Zwerina (1996) in terms of the number of alternatives (three), attributes (three per alternative), attributes levels (three per attribute) and choice situations (nine) chosen. In the original study, Huber and Zwerina used as attribute levels 1, 2 and 3 which were dummy coded but did not use priors, working directly with the choice probabilities. In the present study, we use a generic in the parameters utility specification and do not dummy code such that the utility specification of the model is:

$$V_j = \beta_1 x_{j1} + \beta_2 x_{j2} + \beta_3 x_{j3}, \quad j = 1,2,3, \tag{27}$$

with attribute levels $x_{j1} = \{10, 20, 30\}$ and $x_{j2} = x_{j3} = \{5,10,15\}$. As such, we break slightly from the original example. For the present study we that all parameters are randomly distributed. The first and second parameter are assumed to be uniformly distributed, $\beta_1 \sim U(a_1, b_1)$ and $\beta_2 \sim U(a_2, b_2)$, while the third parameter is assumed to be normally distributed, $\beta_3 \sim N(\mu, \sigma)$. The prior parameters for this case study are taken as $a_1 = -0.9$, $b_1 = -0.5$, $a_2 = -1.5$, $b_2 = -1.0$, $\mu = -0.8$, and $\sigma = 0.2$.

Table 2: Case Study 2 designs

		MNL Design			Cross-Sectional ML Design			Panel ML Design			Orthogonal Design		
		Attributes			*Attributes*			*Attributes*			*Attributes*		
s	***j***	x_{j1}	x_{j2}	x_{j3}	x_{j1}	x_{j2}	x_{j3}	x_{j1}	x_{j2}	x_{j3}	x_{j1}	x_{j2}	x_{j3}
	1	20	10	15	30	10	10	20	15	5	10	5	5
1	2	30	15	5	30	15	5	20	10	10	20	10	10
	3	20	10	15	20	5	15	10	5	15	30	15	15
	1	10	15	10	30	10	5	30	10	5	20	15	5
2	2	10	5	15	30	10	5	10	15	15	30	5	10
	3	20	15	5	10	15	10	20	15	10	10	10	15
	1	30	5	5	10	5	15	30	10	10	30	10	5
3	2	30	10	5	10	15	10	20	15	10	10	15	10
	3	10	5	15	30	10	5	30	10	10	20	5	15
	1	30	15	5	30	5	5	10	5	15	20	10	10
4	2	20	10	15	10	10	15	10	15	10	30	15	15
	3	10	15	15	20	10	10	30	10	5	10	5	5
	1	10	10	15	20	15	5	20	15	15	30	5	10
5	2	30	5	5	20	5	15	30	5	15	10	10	15
	3	10	15	10	10	15	15	30	10	10	20	15	5
	1	20	5	15	10	15	10	30	5	5	10	15	10
6	2	20	5	15	20	10	10	10	5	15	20	5	15
	3	30	10	5	20	5	10	20	15	5	30	10	5
	1	20	10	10	20	10	15	10	10	10	30	15	15
7	2	10	15	10	20	5	15	30	5	5	10	5	5
	3	30	5	5	30	15	5	10	5	15	20	10	10
	1	10	15	10	10	15	15	10	5	15	10	10	15
8	2	20	10	10	30	5	10	20	10	5	20	15	5
	3	20	5	10	30	10	5	10	15	5	30	5	10
	1	30	5	5	20	5	10	20	15	10	20	5	15
9	2	10	15	10	10	15	5	30	10	5	30	10	5
	3	30	10	10	10	5	15	20	5	15	10	15	10

Efficiency results when design is applied to different model forms

Model Assumed for Estimation	Model Assumed for Design							
	MNL Model		Cross-Sectional ML Model		Panel ML Model		Orthogonal Design	
		D-error		*D-error*		*D-error*		*D-error*
	MNL	**0.023**	MNL	0.042	MNL	0.037	MNL	3.604
	Cross-Sect. ML	0.414	Cross-Sect. ML	**0.104**	Cross-Sect. ML	0.114	Cross-Sect. ML	55.468
	Panel ML	0.096	Panel ML	0.082	Panel ML	**0.080**	Panel ML	0.782
		A-error		*A-error*		*A-error*		*A-error*
	MNL	0.262	MNL	0.220	MNL	0.166	MNL	34462
	Cross-Sect. ML	2.800	Cross-Sect. ML	0.970	Cross-Sect. ML	1.633	Cross-Sect. ML	$5.1\cdot10^{9}$
	Panel ML	0.433	Panel ML	0.324	Panel ML	0.298	Panel ML	7012
		Sample Size		*Sample Size*		*Sample Size*		*Sample Size*
	MNL	1	MNL	1	MNL	1	MNL	275809
	Cross-Sect. ML	24	Cross-Sect. ML	31	Cross-Sect. ML	61	Cross-Sect. ML	$1.1\cdot10^{12}$
	Panel ML	8	Panel ML	7	Panel ML	7	Panel ML	72147

As with the first case study, three designs were generated, each having nine choice situations. The first design was generated assuming an MNL functional form (taking the priors equal to the means of the random distributions), the second a cross-sectional ML form, and the third a panel ML form. The final designs and their respective efficiencies are given in Table 2. The design generation computation times were 29 seconds, 63 minutes, and 250 hours, assuming the MNL, cross-sectional ML, and panel ML model, respectively.

For the first three designs, the results for the second case study mirror those of the first. Models estimated using designs specifically generated for that model outperform designs generated for different model forms. Again, for all designs, the minimum sample size for estimating the cross-sectional ML model is much larger than the sample size for estimating the panel ML model. The MNL model and panel ML model sample sizes are more similar, although the latter one needs a higher sample size again due to the harder to estimate standard deviation σ.

Table 2 also contains a fourth design, that being an '*optimal'* orthogonal design generated using the principles outlined by Street *et al.* (2005). Following their procedure, the design is orthogonal within alternatives, not across alternatives. It is constructed such that within each choice situation, for every attribute each attribute level never appears twice across any of the alternatives. As such, the design is generated to force respondents to make trade-offs for each and every attribute in each and every choice situation of the experiment (this is the definition of design optimality put forward by Street *et al.*).

Clearly, the design process put forward by Street *et al.* (which does not consider any information on the priors) produced significantly substandard results for this case study. Indeed, the sample size requirements for all models are substantially large compared with the other designs shown in Table 2, suggesting that this design should not be used in practice. Nevertheless, the warning message here should be somewhat tempered, given that the orthogonal design chosen performed badly given the prior parameters we selected. Whilst these priors were chosen purely for descriptive reasons for the case only, it is feasible that the design may perform better given a different set of population parameter estimates.

Further, the orthogonal design was constructed using the procedures outlined in Street *et al.* (2005), which do not necessarily reflect the generation processes followed by most researchers in constructing orthogonal designs. As such, another orthogonal design may be expected to perform much better than that used here. The best orthogonal design we could find (not reported here), generated assuming an MNL model, had a D-error of 0.081, which is still significantly higher than 0.023, but clearly much better than the 'optimal' orthogonal design following Street *et al.*

All three efficient designs perform relatively under each model form, while the orthogonal design performs very poorly. The MNL design, which is the easiest to generate, could be used to estimate all models with a relatively high efficiency.

4.3 Case study 3

The final case study was constructed to demonstrate the ability of the design generation process outlined within this paper to handle both alternative specific and generic parameter estimates. The utility specification used in generating the design for the third case example is given in Eqn. (28).

$$\begin{aligned} V_1 &= \beta_0 + \beta_1 x_{11} + \beta_2 x_{12} + \beta_4 x_{13}, \\ V_2 &= \qquad \beta_1 x_{21} + \beta_2 x_{22} + \beta_3 x_{23}. \end{aligned} \tag{28}$$

For the first alternative, we include a fixed constant term. For the first and second attributes, generic parameters are assumed whilst for the third attribute of each alternative, the parameters are assumed alternative-specific. Parameters β_1, β_2, and β_3 are assumed random, following a normal distribution, $\beta_k \sim N(\mu_k, \sigma_k)$, $k = 1,2,3$. The attribute levels used are $x_{j1} = \{5,10,15\}$, $x_{j2} = x_{23} = \{0,1,2\}$ and $x_{13} = \{0,1,2,4\}$. The following prior values are used: $\beta_0 = -0.5$, $\mu_1 = -0.05$, $\sigma_1 = 0.02$, $\mu_2 = -0.9$, $\sigma_2 = 0.2$, $\mu_3 = -0.8$, $\sigma_3 = 0.2$, and $\beta_4 = -0.2$. We generate each design with 12 choice situations. The final designs are replicated in Table 3. The computation times for design generation are 28 seconds, 104 minutes, and 332 hours, for the MNL, cross-sectional ML, and the panel ML model, respectively.

Once again, the results for the third case study mirror those of the first two case studies. As would be predicted, an econometric models would be expected to be statistically more efficient when applied to data collected using experimental designs constructed specifically for that design. Also the MNL and panel ML model perform again in a quite similar fashion, while in the cross-sectional ML model very large sample sizes are required to be able to obtain reliable parameter estimates for the standard deviations. The panel ML design performs poorly when estimating the cross-sectional ML model, and vice versa, the cross-sectional ML design performs poorly when estimating the panel ML model. The MNL design looses some efficiency when estimating the ML models, but performs relatively well.

5. Misspecification of the prior parameter values

In constructing each of the designs, we have assumed that the prior parameter values correspond to the true parameter values held by the population. This represents a strong assumption that is unlikely to hold in practice, but it is necessary for creating efficient designs. To test the impact misspecification of the prior parameters has on an experimental design once generated, it is possible to fix the design and apply different sets of priors to it and in doing so recalculate the expected AVC matrix.

Let us assume that the true population parameters for the third case example were really $\beta_0 = -0.5$, $\beta_1 \sim N(-0.07, 0.03)$, $\beta_2 \sim U(-1.1, -0.8)$, $\beta_3 \sim N(-0.6, 0.15)$, and $\beta_4 = -0.3$. As such, for the first random parameter, the mean of the parameter is larger in magnitude as is the standard deviation relative to that assumed in the design generation. For the second random parameter, we assumed a normal distribution with some given priors in the design construction, but now we not just assume different priors, but actually a whole different random distribution, namely a uniform distribution. The third random parameter is assumed to have a smaller mean and standard deviation than assumed in constructing the design, while the final parameter is maintained as fixed but has an increased magnitude. Table 4 details the new efficiency measures for the three designs given in Table 3, as well as the percentage change of these values.

Table 3: Case Study 3 designs

		MNL Design			Cross-Sectional ML Design			Panel ML Design		
		Attributes			*Attributes*			*Attributes*		
s	***j***	x_{j1}	x_{j2}	x_{j3}	x_{j1}	x_{j2}	x_{j3}	x_{j1}	x_{j2}	x_{j3}
1	1	5	2	1	10	0	2	5	2	4
	2	15	0	2	10	1	2	15	0	2
2	1	15	0	3	5	0	1	15	0	4
	2	5	2	0	15	2	0	5	1	2
3	1	10	2	3	15	1	1	5	0	2
	2	10	0	1	5	1	0	15	1	0
4	1	10	2	3	15	1	4	5	1	2
	2	5	0	1	5	0	2	15	2	1
5	1	5	1	1	5	2	3	15	1	1
	2	15	1	0	15	1	2	10	0	2
6	1	15	2	2	10	2	2	5	1	3
	2	10	0	1	10	0	1	15	1	1
7	1	15	0	1	10	2	3	15	0	4
	2	5	2	0	10	0	1	5	2	0
8	1	10	1	2	10	1	1	15	1	1
	2	10	2	2	10	2	1	5	2	0
9	1	5	1	4	5	1	4	10	2	2
	2	15	2	0	15	1	0	10	0	1
10	1	15	1	4	15	0	3	10	2	1
	2	5	1	2	5	2	2	10	0	2
11	1	10	0	2	5	2	2	10	0	3
	2	10	1	2	15	0	1	5	2	1
12	1	5	0	4	15	0	4	10	2	3
	2	15	1	1	5	2	0	10	1	0

Efficiency results when design is applied to different model forms

Model Assumed for Estimation	Model Assumed for Design					
	MNL Model		Cross-Sectional ML Model		Panel ML Model	
		D-error		*D-error*		*D-error*
	MNL	**0.137**	MNL	0.149	MNL	0.146
	Cross-Sect. ML	2.026	Cross-Sect. ML	**1.007**	Cross-Sect. ML	1.567
	Panel ML	0.242	Panel ML	0.282	Panel ML	**0.222**
		A-error		*A-error*		*A-error*
	MNL	0.368	MNL	0.362	MNL	0.445
	Cross-Sect. ML	196.281	Cross-Sect. ML	9.542	Cross-Sect. ML	31.883
	Panel ML	0.700	Panel ML	0.655	Panel ML	0.622
		Sample Size		*Sample Size*		*Sample Size*
	MNL	30	MNL	35	MNL	34
	Cross-Sect. ML	252448	Cross-Sect. ML	21518	Cross-Sect. ML	104965
	Panel ML	408	Panel ML	1164	Panel ML	282

From Table 4, it can be seen that misspecification of the prior parameter values can have a significant impact upon the overall efficiency of different designs. Indeed, the D-errors increase for all designs when applied to the appropriate model forms, hence all designs are loosing efficiency due to incorrect priors. Although the D-error for the MNL design applied to the MNL model increases, the sample size actually decreases. There are two reasons for this. First, the D-error represents a form of *average* over all parameters, while the sample size requirements relate only to parameter that is most difficult to estimate (see Bliemer and Rose, 2005). Thus, it is possible that on *average*, the standard errors for each design are decreasing, but that the largest standard error within the AVC matrix actually decreases, thus requiring smaller sample sizes for all asymptotic *t*-ratios to be statistically significant. Secondly, the most difficult to estimate parameter was the standard deviation of the first normal distribution, and since this value has increased from 0.02 to 0.03 it will lead to the similar *t*-ratios with lower sample size.

Nevertheless, for design efficiency to be truly translated into estimation efficiency, the parameter priors assumed during the generation process should be as close to possible to the true, but as yet unknown, population level parameters. To ensure that this is the case, a number of strategies are available to the analyst. Firstly, a literature review or a pilot study (e.g., using an orthogonal design) conducted on a small sample may yield sufficient priors for use in constructing the design. Secondly, the analyst may consider updating the design throughout the data collection process based on priors obtained from sequential analysis of accumulated data, as suggested by Kanninen (2002). This latter strategy should not be confused with adaptive conjoint however, as no within respondent design updating occurs. Rather, more efficient designs are given to newer respondents. Thirdly, more robust efficient designs can be generated, so-called Bayesian efficient design, that are not optimized on a set of fixed priors, but rather on a set of prior distributions, reflecting the fact that priors are not known with certainty, see e.g., Bliemer and Rose (2008), Kessels *et al.* (in press), Sándor and Wedel (2001). This last strategy can be combined with the first strategy, where a pilot study is conducted and the parameter estimates together with their standard errors can be used as Bayesian priors to construct a Bayesian efficient design. Unfortunately, this adds significantly to the design generation time again, as the Bayesian efficiency needs to be simulated with draws over the random priors. In the panel ML context, due to huge computational complexity, this does not seem feasible within reasonable time constraints.

Table 4: Case Study 3 Impact of Misspecification of Prior Parameter Values

Efficiency results when design is applied to different model forms

Model Assumed for Estimation	Model Assumed for Design								
	MNL Model			**Cross-Sectional ML Model**			**Panel ML Model**		
	*D-error **								
	MNL	0.148	(+8%)	MNL	0.165	(+11%)	MNL	0.162	(+11%)
	Cross-Sect. ML	2.983	(+47%)	Cross-Sect. ML	1.723	(+71%)	Cross-Sect. ML	1.791	(+14%)
	Panel ML	0.256	(+6%)	Panel ML	0.254	(-10%)	Panel ML	0.272	(+23%)
	*A-error **								
	MNL	0.381	(+4%)	MNL	0.376	(+4%)	MNL	0.500	(+12%)
	Cross-Sect. ML	428.620	(+118%)	Cross-Sect. ML	44.721	(+369%)	Cross-Sect. ML	77.059	(+142%)
	Panel ML	0.735	(+5%)	Panel ML	0.659	(+1%)	Panel ML	0.839	(+35%)
	*Sample Size **								
	MNL	14	(-53%)	MNL	17	(-51%)	MNL	18	(-47%)
	Cross-Sect. ML	1749011	(+593%)	Cross-Sect. ML	146414	(+580%)	Cross-Sect. ML	84728	(-19%)
	Panel ML	374	(-8%)	Panel ML	279	(-76%)	Panel ML	388	(+38%)

* Values between brackets are the percentage change

6. Conclusion and discussion

The generation of SC experiments has continued to evolve over the past two decades to become an increasingly significant but complex component of SC studies. We contend that the generation of SC experiments is critical to the success of any SC study and that failure to correctly construct an appropriate design may result in erroneous findings. This paper addresses the issue of how to generate efficient SC experiments for the panel formulation of the mixed logit model, which recognizes the fact a single respondent faces multiple choice situations. We have attempted to show via the use of three case studies, the importance of generating experimental designs specifically for econometric models for which they are likely to be applied to once data has been collected. It turns out that generating efficient designs for panel mixed logit models requires a much larger computational effort than generating designs for the MNL or cross-sectional mixed logit model, due to mandatory virtual sample generation. Furthermore, designs that are efficient for the cross-sectional mixed logit model are typically not efficient for estimating the panel mixed logit model, and vice versa. In the three case studies presented, the efficient design for the MNL model performed relatively efficient for estimating panel mixed logit models, although some efficiency will be lost. Given the fact that generation of panel mixed logit designs is much more computationally intensive, a good starting point would perhaps be to generate an efficient design for the MNL model and evaluate the efficiency this design under the panel mixed logit model assumption.

A critical issue in the construction of efficient designs is what constitutes the best source for determining the priors used in generating the designs. This issue leads to several questions that analysts must address. Firstly, should a pilot study be conducted, and if so, what represents a sufficient sample size to obtain the priors? Alternatively, should the managers and other practitioner's beliefs be incorporated into the generation process and how best should such beliefs be captured? These questions remain unanswered and are in urgent need of examination.

Further, the current literature on generating experimental designs has clearly demonstrated the requirement that efficient designs be constructed such that they relate to the final model likely to be estimated as part of the study (otherwise they loose their efficiency), making the generation of efficient designs difficult, a message mirrored within this paper. Firstly, the analyst may not know the final model form until after the data has been collected. Secondly, the rapid increase in the econometric modeling available to the analyst has left the experimental design literature well and truly behind. What is urgently required is a detailed study beyond the case studies here to determine the likely consequences of mis-specifying not only the priors, but also the model form used in generating efficient SC experiments. It is imperative that orthogonal designs not be immune from such a study. Further still, the work on efficient SC experiments has largely remained theoretical. Most research tends to use only one design, whether orthogonal or efficient, and as such, there exists little evidence in practice that efficient designs will outperform other designs, despite an expectation that they should do so. As such, practical research involving data collected using multiple designs is urgently required. Indeed, the question of statistical efficiency versus behavioral efficiency remains a fertile area for future research examination.

This paper also suffers from a number of limitations. Firstly, although we have employed three different case studies, each involving designs with differing design dimensions, the paper fails to examine the impact different design dimensions play in terms of design efficiency for mixed logit designs. Does having more or less choice situations impact upon the statistical efficiency of mixed logit designs and in particular mixed logit designs constructed for panel data? Does having a wider attribute level range improve or detract from the efficiency of such designs? Similar questions can be asked as to the number of attributes and alternatives of designs? Whilst these questions have been asked about designs generated for other model types and answers attempted to be given (see e.g., Bliemer and Rose, 2005, and Rose and Bliemer, 2005), we have not attempted to address these issues here.

Secondly, in constructing our designs, we have assumed that each prior parameter is known perfectly *a priori*. For example, we have specified an exact mean and standard deviation for varying random parameters used, or exact lower and upper bounds for uniform random parameters. Sándor and Wedel (2002) however, demonstrated for the MNL that significant gains can be achieved if prior parameters are drawn using Bayesian distributions, rather than fixed values, thus incorporating a degree of uncertainty into the prior parameters used during the design generation process. For the current study, we have not done this. At issue is the fact that for random parameters, two population moments may be unknown, thus requiring the analyst to impose a Bayesian distribution for both a mean and standard deviation parameter for example. This significantly adds to the complexity of the design process, however doing so may generate designs which are more robust to misspecification of the prior parameter estimates.

Thirdly, in generating the designs within this paper, despite letting go of the principle of orthogonality, we have remained true to the wider experimental design literature in other ways. In particular, we have imposed attribute level balance within each of the designs. This means that each level of an attribute will appear an equal number of times over the course of the experiment. Whilst there may be reasons for doing this, there may also be arguments for not doing so. To understand the argument against this imposition, consider that attribute level balance acts as a constraint in terms of allocating appropriate attribute levels across a design. For example, a particular choice situation may be improved if one particular level is swapped with that of another choice situation, however the swap may detract from the second choice situation. In this way, attribute level balance may cause a number of choice situations to include dominant alternatives, despite the design being overall more efficient. This problem was identified by Kanninen (2002) who proposed allowing one attribute to be continuous, and found for the MNL model that substantial efficiency gains were possible. Similar research for more advanced models is necessary.

Finally, this paper has sought to address the issue of designing SC experiments allowing for correlation between choice responses across choice situations. The paper does not however seek to outline the algorithms, etc. for locating more efficient designs. For those interested in these algorithms, we refer the reader to other sources, in particular Bliemer and Rose (2006) and Kessels *et al.* (2006).

Acknowledgments

We would like to thank Andrew Collins for his substantial work on the implementations in the Ngene software.

References

Beaulieu, N.C. (1991) On the Generalized Multinomial Distribution, Optimal Multinomial Detectors, and Generalized Weighted Partial Decision Detectors. *IEEE Transactions on Communications*, Vol. 39(2), pp. 193-194.

Ben-Akiva, M., and D. Bolduc (1996) Multinomial Probit with a Logit Kernel and a General Parametric Specification of the Covariance Structure. *Working paper*, Department of Civil Engineering, MIT, Cambridge MA, USA.

Ben-Akiva, M., D. Bolduc, and M. Bradley (1993) Estimation of Travel Choice Models with Randomly Distributed Values of Time. *Transportation Research Record*, No. 1413, pp. 88-97.

Bhat, C. (1996) Accommodating Variations in Responsiveness to Level-of-Service Measures in Travel Model Choice Modeling. *Working paper*, Department of Civil Engineering, University of Massachusetts at Amherst MA, USA.

Bhat, C.R. (2001) Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. *Transportation Research B*, Vol. 35, pp. 677-693.

Bhat, C.R. (2003) Simulation Estimation of Mixed Discrete Choice Models using Randomized and Scrambled Halton Sequences. *Transportation Research B*, Vol. 37, pp. 837-855.

Bliemer, M.C.J., and J.M. Rose (2005) Efficiency and Sample Size Requirements for Stated Choice Studies, *ITLS working paper: itls-wp-05-08*.

Bliemer, M.C.J., and J.M. Rose (2006) Designing Stated Choice Experiments: State-of-the-art. *Proceedings of the 11th International Conference on Travel Behaviour Research*, Kyoto, Japan.

Bliemer, M.C.J., J.M. Rose and D.A. Hensher (2007) Constructing Efficient Stated Choice Experiments for Estimating Nested Logit Models. Submitted to *Transportation Research B.*

Bliemer, M.C.J., J.M. Rose, and S. Hess (in press) Approximation of Bayesian Efficiency in Experimental Choice Designs. *Journal of Choice Modelling*.

Boyd, J., and R. Mellman (1980) The Effect of Fuel Economy Standards on the U.S. Automotive Market: An Hedonic Demand Analysis. *Transportation Research A*, Vol. 14(5-6), pp. 367-378.

Brownstone, D., and K. Train (1999) Forecasting New Product Penetration with Flexible Substitution Patterns. *Journal of Econometrics*, Vol. 89(1), pp. 109-129.

Cardell, N., and F. Dunbar (1980) Measuring the Societal Impacts of Automobile Downsizing. *Transportation Research A*, Vol. 14(5-6), pp. 423-434.

Carlsson, F., and P. Martinsson (2002) Design Techniques for Stated Preference Methods in Health Economics. *Health Economics*, Vol. 12, pp. 281-294.

Carrasco, J.A., and J. de D. Ortúzar (2002) Review and Assessment of the Nested Logit Model. *Transport Reviews*, Vol. 22(2), pp. 197-218.

Ferrini, S., and R. Scarpa (2007) Designs with A-Priori Information for Nonmarket Valuation with Choice-Experiments: a Monte Carlo Study. *Journal of Environmental Economics and Management*, Vol. 53(3), pp. 342-363.

Greene, W.H. (2002) *Econometric Analysis*. Third edition, Prentice Hall, New Jersey.

Hensher, D.A., and W.H. Greene (2003) The Mixed Logit Model: The State of Practice. *Transportation*, Vol. 30(2), pp. 133-176.

Hess, S., K.E. Train, and J.W. Polak (2005) On the Use of a Modified Latin Hypercube Sampling (MLHS) Approach in the Estimation of a Mixed Logit Model for Vehicle Choice. *Transportation Research B*, Vol. 40(2), pp. 147-163.

Huber, J., and K. Zwerina (1996) The Importance of utility Balance and Efficient Choice Designs, *Journal of Marketing Research*, Vol. 33, 1996, pp. 307-317.

Kanninen, B.J. (2002) Optimal Design for Multinomial Choice Experiments. *Journal of Marketing Research*, Vol. 39(2), pp. 214-217.

Kessels, R., P. Goos, and M. Vandebroek (2004) Comparing Algorithms and Criteria for Designing Bayesian Conjoint Choice Experiments. Working paper, Department of Applied Economics, Katholieke Universiteit Leuven, Belgium.

Kessels, R., P. Goos, and M. Vandebroek (2006) A Comparison of Criteria to Design Efficient Choice Experiments. *Journal of Marketing Research*, Vol. 43, pp. 409-419.

Kessels, R., B. Jones, P. Goos, and M. Vandebroek (in press) An efficient algorithm for constructing Bayesian optimal choice designs. *Journal of Business and Economic Statistics*.

McFadden, D. (1974) Conditional Logit Analysis of Qualitative Choice Behaviour. In: P. Zarembka (ed.) *Frontiers of Econometrics*, Academic Press, New York, pp. 105-142.

McFadden, D., and K. Train (2000) Mixed MNL Models for Discrete Response. *Journal of Applied Econometrics*, Vol. 15(5), pp. 447-470.

Ortúzar, J. de D., D.A. Roncagliolo and U.C. Velarde (2000) Interactions and Independence in Stated Preference Modelling. In: J. de D. Ortúzar (ed.) *Stated Preference Modelling Techniques: PTRC Perspectives*, Vol. 4, PTRC Education and Research Services Ltd, London.

Ortúzar, J. de D., and L.G. Willumsen (2001) *Modelling Transport.* Third edition, John Wiley & Sons Inc., New York.

Revelt, D., and K. Train (1998) Mixed Logit With Repeated Choices: Households' Choices Of Appliance Efficiency Level. *Review of Economics and Statistics*, Vol. 80(4), pp 647-657.

Rose, J.M., and M.C.J. Bliemer (2005) Constructing Efficient Choice Experiments, *ITLS working paper: itls-wp-05-07*.

Rose, J.M., and M.C.J. Bliemer (2006) Designing Efficient Data for Stated Choice Experiments. Proceedings of the *11th International Conference on Travel Behaviour Research*, Kyoto, Japan.

Rose, J.M., and M.C.J. Bliemer (2008) Stated Preference Experimental Design Strategies. In: D.A. Hensher and K.J. Button (eds) *Handbook of Transport Modelling*, Elsevier, 2nd edition, pp. 151-180.

Rose, J.M., M.C.J. Bliemer, D.A. Hensher, and A.C. Collins (2008) Designing Efficient Stated Choice Experiments Involving Respondent Based Reference Alternatives. *Transportation Research B*, Vol. 42, pp. 395-406.

Sándor, Z., and M. Wedel (2001) Designing Conjoint Choice Experiments Using Managers' Prior Beliefs. *Journal of Marketing Research*, Vol. 38, pp. 430-444.

Sándor, Z., and M. Wedel (2002) Profile Construction in Experimental Choice Designs for Mixed Logit Models. *Marketing Science*, Vol. 21(4), pp. 455-475.

Sándor, Z., and M. Wedel (2005) Heterogeneous conjoint choice designs. *Journal of Marketing Research*, Vol. 42, pp. 210-218.

Sándor, Z., and K. Train (2004) Quasi-Random Simulation of Discrete Choice Models. *Transportation Research B*, Vol. 38, pp. 313-327.

Street, D.J., and L. Burgess (2004) Optimal and Near Optimal Pairs for the Estimation of Effects in 2-Level Choice Experiments. *Journal of Statistical Planning and Inference*, Vol. 118 (1–2), pp. 185–99.

Street, D.J., L. Burgess, and J.J. Louviere (2005) Quick and Easy Choice Sets: Constructing Optimal and Nearly Optimal Stated Choice Experiments. *International Journal of Research in Marketing*, Vol. 22, pp. 459-470.

Train, K. (2003) *Discrete Choice Methods with Simulation.* Cambridge University Press, Cambridge.