



Using appropriate prior information to eliminate choice sets with a dominant alternative from D-efficient designs

Marjolein Crabbe and Martina Vandebroek

DEPARTMENT OF DECISION SCIENCES AND INFORMATION MANAGEMENT (KBI)

Using Appropriate Prior Information to Eliminate Choice Sets with a Dominant Alternative from \mathcal{D} -efficient Designs

Marjolein Crabbe
Martina Vandebroek

Abstract

Most attributes in transportation studies, such as the travel time and the travel cost of a travel mode or road alternative, have a clear rank order in their attribute levels. Therefore one option in a choice set of an experimental design can dominate the other alternatives in the set. This research finds Bayesian \mathcal{D} -efficient designs for a specific setup in the transportation field. It is shown that with a proper choice of prior information which adequately incorporates the dominance of lower attribute levels, no choice sets with a dominant alternative will be included in the efficient designs.

Keywords: Stated preference data · Conditional logit model · Nested logit model · Bayesian \mathcal{D} -efficient designs · Dominant alternative · Prior information

1 Introduction

Among many other fields of application, choice-based conjoint choice experiments are frequently applied and highly valued in transportation studies (Bliemer et al. 2009; Rose and Bliemer 2009). The majority of these studies aims at investigating and quantifying the travelers' valuation of travel time savings in a mode choice or road choice experiment (Hensher 2001; Hess et al. 2005; Axhausen et al. 2008). How much are travelers willing to pay to save a unit of travel time? The stated preference experiment discussed in this research deals with the attractiveness and profitability of a new and speedy toll road in comparison to existing toll-free but slower alternatives (Hess and Adler 2010). More specifically, one wants to find out if, and if so how much, travelers are willing to pay for a more expensive, yet faster road option.

In the experiment travelers are presented with a list of choice sets which comprise multiple road options for a specific trip each defined by the travel time and the travel cost, i.e. the toll price of the alternative. Both attributes can take on four levels. For the tolled options the travel time and the toll price

levels respectively equal 29, 32, 35 or 38 minutes and 1, 2, 3 or 4 dollar. To bring in a trade-off between travel time and travel cost, the travel time levels for the toll-free road options are chosen to be higher and equal 40, 44, 48 or 52 minutes.

For the toll road problem in this study a specific setup is considered for the choice sets in the experimental design, i.e. each choice set in the design consists of one toll-free option and two toll road options. To start, the more simple setup of a design including choice sets with two tolled road options only will be discussed in the following section. Toll-free alternatives are discarded in the design of this paired comparison experiment. The third section deals with the main experimental setup of choice sets with one toll-free and two toll road options.

Since both attributes of the road options, i.e. the travel time and the toll price, have a clear rank order in their levels (travelers generally prefer a shorter travel time and a lower travel cost), it is possible for one toll road option in a choice set to dominate the other tolled alternative in the set. As lower attribute levels are more attractive, a cheaper and at the same time faster option would for instance dominate a more expensive and slower option. An example of a choice set with a dominant tolled alternative is given in Table 1.

	<i>Alternative I</i> <i>(toll-free)</i>	<i>Alternative II</i> <i>(tolled)</i>	<i>Alternative III</i> <i>(tolled)</i>
<i>Travel time</i>	40 minutes	32 minutes	38 minutes
<i>Toll price</i>	\$0	\$2	\$3

Table 1: Choice set with dominant toll road option

It is clear that for the choice set in Table 1, the first tolled option (Alternative II) dominates the second tolled option (Alternative III). The first toll road option is faster without being more expensive. Such choice sets are not valuable in travel time savings estimation, as each rational participant would choose Alternative II from the two toll road options. Moreover, the presence of such non-sense choice sets in a questionnaire puts the quality and credibility of the research on the line. Respondents could loose their interest and will for cooperation, potentially resulting in reduced response quality and finally in estimation bias. Therefore it is very important to avoid choice sets with a dominant alternative in the design of the experiment.

To model the travelers' choices from the experiment, two well-known discrete choice models are considered. For the first setup, in which the choice sets comprise only two tolled road options, the conditional logit model (Train 2003) is used to model the choice data. To analyze the data from the second setup, in which, besides two toll roads, an additional toll-free alternative is included in each choice set, the nested logit model (Train 2003) is applied as the conditional logit model will prove to be inappropriate for choice sets with three such alternatives.

To attain choice data that yield accurate estimates of the model parameters, i.e. the travel time and the travel cost coefficient, \mathcal{D} -efficient designs (Atkinson et al. 2007) can be used. In this study \mathcal{D} -efficient designs will be constructed for both the conditional logit and the nested logit model. These efficient designs are obtained through the maximization of the \mathcal{D} -optimality criterion, which is the determinant of the Fisher information matrix. As the conditional logit as well as the nested logit model are nonlinear in the parameters, the \mathcal{D} -optimality criterion depends on the unknown model parameters in both cases. So, one needs to assume prior values (local \mathcal{D} -efficient designs) or a prior distribution (Bayesian \mathcal{D} -efficient designs) for the model coefficients in order to compute \mathcal{D} -efficient designs. Unfortunately, without carefully chosen prior values or distributions, the designs can include choice sets with a dominant toll road option.

This research advocates an appropriate choice of prior information to avoid the presence of choice sets with a dominant alternative in \mathcal{D} -efficient designs. By incorporating the dominance of lower attribute levels in an adequate way in the prior information, these choice sets are excluded from the efficient designs.

2 Choice sets with two tolled alternatives

Before we turn to the main setup of the experimental design, in which each choice set comprises one toll-free and two toll road options, a more straightforward setup is considered first. In this section, the toll-free alternatives are discarded and the choice sets in the design include two toll road options only. As these options are defined by two attributes, i.e. travel time and toll price, the random utility a traveler receives from tolled alternative j ($j = 1$ or 2) in choice set s ($s = 1, \dots, S$) equals

$$U_{js} = (\beta_{TT} \times TT_{js}) + (\beta_{TC} \times TC_{js}) + \varepsilon_{js},$$

with β_{TT} and β_{TC} the model partworths for respectively the travel time and the travel cost attribute.

In the conditional logit model, the error terms ε_{js} are assumed to be independently and identically extreme value distributed. Therefore, a closed form solution for the probability that a person chooses the tolled road option j in choice set s is found

$$p_{js} = \frac{e^{(\beta_{TT} \times TT_{js}) + (\beta_{TC} \times TC_{js})}}{\sum_{t=1}^2 e^{(\beta_{TT} \times TT_{ts}) + (\beta_{TC} \times TC_{ts})}}.$$

To construct efficient choice designs for the conditional logit model, the \mathcal{D} -optimality criterion is used in this study. This criterion outperforms other design criteria with respect to accuracy, predictive validity and computational effectiveness (Kessels et al. 2006a). \mathcal{D} -efficient designs are obtained by the minimization of the \mathcal{D} -error, i.e. the inverse of the determinant of the Fisher

information matrix which for the conditional logit model is given by

$$\mathbf{I}(\mathbf{X}, \beta) = N \sum_{s=1}^S \mathbf{X}'_s (\mathbf{P}_s - \mathbf{p}_s \mathbf{p}'_s) \mathbf{X}_s,$$

with $\mathbf{X}_s = [\mathbf{x}'_{1s}, \mathbf{x}'_{2s}]'$ the design matrix for choice set s and $\mathbf{x}_{ts} = [TT_{ts}, TC_{ts}]'$, $\mathbf{p}_s = [p_{1s}, p_{2s}]'$ and $\mathbf{P}_s = \text{diag}[p_{1s}, p_{2s}]$.

We will compute Bayesian \mathcal{D} -efficient designs (Sándor and Wedel 2001) by minimizing the expectation of the \mathcal{D} -error over a prior distribution $\pi_0(\beta)$ of the model parameters

$$\begin{aligned} \mathcal{D}_B - \text{error} &= E_{\beta} \left[\det[\mathbf{I}(\mathbf{X}, \beta)]^{-1/2} \right] \\ &= \int_{\mathbb{R}^2} \det[\mathbf{I}(\mathbf{X}, \beta)]^{-1/2} \pi_0(\beta) d\beta. \end{aligned}$$

Note that the \mathcal{D} -criterion is taken to the power 1/2 to obtain independence from the dimensionality of the vector of model parameters. For practical purposes, the Bayesian \mathcal{D} -error is approximated by the average \mathcal{D} -error over R draws β^r from the prior distribution

$$\tilde{\mathcal{D}}_B - \text{error} = \frac{1}{R} \sum_{r=1}^R \det[\mathbf{I}(\mathbf{X}, \beta^r)]^{-1/2}.$$

The choice of the prior distribution $\pi_0(\beta)$ for the model parameters is an important aspect in the computation of Bayesian efficient designs. Specifically, we show that the presence of dominant alternatives in the choice sets of \mathcal{D} -efficient designs is due to the (incorrect) choice of prior information. If one adequately incorporates the dominance of lower attribute levels in the prior information, choice sets with dominant alternatives will not appear in the \mathcal{D} -efficient designs. For the travel time and the travel cost partworth in the model, a multivariate normal prior distribution will be assumed.

$$\begin{pmatrix} \beta_{TT} \\ \beta_{TC} \end{pmatrix} \sim MVN \left(\begin{pmatrix} \bar{\beta}_{TT} \\ \bar{\beta}_{TC} \end{pmatrix}, \begin{pmatrix} \sigma_{TT}^2 & \sigma_{TT,TC} \\ \sigma_{TT,TC} & \sigma_{TC}^2 \end{pmatrix} \right)$$

To verify whether the prior information indeed has an effect on the number of choice sets with a dominant alternative, two different vectors of prior means are considered, i.e.

1. $(\bar{\beta}_{TT}, \bar{\beta}_{TC}) = (-0.2, -1)$,
2. $(\bar{\beta}_{TT}, \bar{\beta}_{TC}) = (-0.4, -2)$.

The two prior mean vectors correspond to a value of travel time savings of \$12 per hour. More specifically, travelers are willing to pay \$12 in order to save one hour of travel time, a realistic and frequently observed value of travel time

savings (Hess et al. 2008). The variances σ_{TT}^2 and σ_{TC}^2 are fixed to 0.001 and 0.01 respectively and the covariance $\sigma_{TT,TC}$ equals 0 in each case.

In the following we claim that only the second prior specification adequately specifies the dominance structure in the experiment and therefore choice sets with dominant alternatives will not appear in the \mathcal{D} -efficient designs obtained with this prior information. Based on the ideas of Kessels et al. (2006b), the appropriateness of the prior information is evaluated by considering the choice set with the most attractive and the least attractive toll road option (Table 2).

	<i>Travel time</i>	<i>Toll price</i>
<i>Most attractive</i>	29 min	\$1
<i>Least attractive</i>	38 min	\$4

Table 2: Choice set with the most attractive and the least attractive toll road option

As the least attractive road alternative corresponds to the longest travel time and the highest travel cost, obviously, the probability of choosing the most attractive option, i.e. the fastest and the cheapest alternative, should be approximately 100% in this choice set. The most attractive option should clearly dominate the least attractive one.

In Table 3 the probabilities for choosing the two most extreme alternatives are given for the two sets of prior means considered. For prior means $(-0.2, -1)$, a very high probability of choosing the dominant alternative is observed, yet approximately 0.82% of the travelers would still prefer the least attractive option over the most attractive one. As for the other prior mean specification, the probability for choosing the dominant alternative is higher, approximately 100%, and thus more realistic. By taking the prior means further away from the zero vector, the dominance of lower attribute levels is more adequately incorporated in the prior information. Therefore one can expect \mathcal{D} -efficient designs constructed with the first prior to include choice sets with a dominant alternative, whereas such choice sets would be excluded from the designs when making use of the latter prior specification.

For each prior a Bayesian \mathcal{D} -efficient design was computed by means of a Bayesian modification of the modified Fedorov choice algorithm (Kessels et al. 2006b). Three hundred random start designs are updated using 1000 random draws from the prior distribution. As there are only two model parameters to estimate, i.e. β_{TT} and β_{TC} , designs with eight choice sets are constructed. The designs are given in Appendix A. For each prior specification the fraction of choice sets in the \mathcal{D} -efficient design with a dominant alternative is given at the bottom of Table 3. Indeed, a well-considered choice of prior information, which accurately expresses the dominance structure in the experiment, has a beneficial effect on the suitability of the \mathcal{D} -efficient design as there are no choice sets with a dominant toll road option in the design corresponding to the latter prior specification. Note that different prior covariance structures were considered

		(β_{TT}, β_{TC})	
		$(-0.2, -1)$	$(-0.4, -2)$
<i>Most attractive</i>	29 min - \$1	99.18%	99.99%
<i>Least attractive</i>	38 min - \$4	0.82%	0.01%
<i>Fraction of sets in the design with a dominant toll road option</i>		4/8	0/8

Table 3: Probabilities of choosing the most attractive and the least attractive toll road option

and that with respect to the fraction of choice sets with a dominant alternative in the design identical results were obtained.

3 Choice sets with one toll-free and two tolled alternatives

In the second stage of this research, the main design setup for the toll road experiment is studied. Now, the choice sets in the design comprise not only two tolled roads but also one toll-free road option. The addition of a toll-free alternative in the choice sets makes it possible to investigate whether travelers are at all interested in a new toll road or whether, if possible, they would continue to use the existing road options which are slower but free of charge. Analogous to the preceding setup, we search for Bayesian \mathcal{D} -efficient designs. With respect to the \mathcal{D} -optimality criterion, we optimize designs including choice sets with three alternatives, the first a toll-free option and the final two toll roads.

To obtain the \mathcal{D} -efficient designs, we first considered the conditional logit model to analyze the choice data and used the two prior specifications introduced in the previous section. To verify the appropriateness of the prior information, two choice sets with three alternatives are considered. The first set includes the toll-free option with the shortest travel time, i.e. 40 minutes, and the two most extreme toll road options, the second set includes the toll-free option with the longest travel time, i.e. 52 minutes, and the two most extreme toll road options. The probabilities of choosing these alternatives in the choice sets are given in Table 4 for the two sets of prior means.

One can conclude that the prior mean values $(-0.4, -2)$, in contrast to the values $(-0.2, -1)$, adequately incorporate the dominance of lower attribute levels as approximately 0% of the travelers would choose the dominated alternative. Yet, when this prior information is used to compute Bayesian \mathcal{D} -efficient designs of choice sets with three alternatives for the conditional logit model, the designs include choice sets with a dominant toll road option. When specifying the prior means even further away from the zero vector, the designs still include nonsense choice sets such as sets with a dominant toll road option or sets with two equal toll road options.

	(β_{TT}, β_{TC})	
	$(-0.2, -1)$	$(-0.4, -2)$
40 min - \$0	23.00%	8.31%
29 min - \$1	76.37%	91.68%
38 min - \$4	0.63%	0.01%
52 min - \$0	2.64%	0.07%
29 min - \$1	96.57%	99.92%
38 min - \$4	0.79%	0.01%

Table 4: Probabilities of choosing a toll-free option and the two most extreme toll road options

The conditional logit model assumes that the error terms in the utilities are independently and identically extreme value distributed. This independence assumption is better known as the property of independence from irrelevant alternatives (IIA). As it seems obvious that the unobserved utility of the two toll road options in the choice sets considered here are correlated, the application of the conditional logit model to analyze the choice data at hand is no longer appropriate. The use of this improper model can explain why, even with correct prior information, the \mathcal{D} -efficient designs still include choice sets with a dominant alternative. When it makes sense to divide the alternatives in the choice sets of a design into meaningful subsets of similar alternatives, the nested logit model should be used instead.

Based on the ideas of Goos et al. (2010), the nested logit model is applied to the specific choice data considered in this section, i.e. choices between one toll-free and two toll road options. Two nests are considered, the first nest contains the toll-free road option and the second nest includes the two tolled alternatives. As such, the probabilities of choosing the toll-free nest (F), or equivalently the toll-free road option, and the toll nest (T) in choice set s respectively equal

$$p_{F,s} = \frac{e^{\lambda_F V_{F,s}}}{e^{\lambda_T V_{T,s}} + e^{\lambda_F V_{F,s}}}$$

and

$$p_{T,s} = \frac{e^{\lambda_T V_{T,s}}}{e^{\lambda_T V_{T,s}} + e^{\lambda_F V_{F,s}}},$$

with λ and V respectively the dissimilarity parameter and the intrinsic value of the nest. The dissimilarity parameter of a nest is an index for the correlation between the alternatives in the nest, more specifically the correlation between the unobserved utility of the alternatives. The smaller the parameter value, the higher the correlation. Generally, all dissimilarity parameters in the model lie between zero and one. The intrinsic value of a nest captures the utility of the nest.

As the toll-free nest includes only one alternative, its dissimilarity parameter

λ_F is set equal to one and its intrinsic value $V_{F,s}$ simplifies to

$$V_{F,s} = \log \left[e^{(\beta_{TT} \times TT_{F,s}) + (\beta_{TC} \times TC_{F,s})} \right] = \beta_{TT} \times TT_{F,s},$$

with $TT_{F,s}$ the travel time of the toll-free alternative in choice set s . Obviously $TC_{F,s}$ equals zero. To facilitate notation, the dissimilarity parameter of the nest with the two toll road options λ_T is denoted by λ . Consequently,

$$p_{F,s} = \frac{e^{\beta_{TT} \times TT_{F,s}}}{e^{\lambda V_{T,s}} + e^{\beta_{TT} \times TT_{F,s}}}$$

and

$$p_{T,s} = \frac{e^{\lambda V_{T,s}}}{e^{\lambda V_{T,s}} + e^{\beta_{TT} \times TT_{F,s}}}.$$

In these expressions the intrinsic value of the toll nest $V_{T,s}$ is

$$V_{T,s} = \log \left[e^{(\beta_{TT} \times TT_{1s}) + (\beta_{TC} \times TC_{1s})} + e^{(\beta_{TT} \times TT_{2s}) + (\beta_{TC} \times TC_{2s})} \right],$$

with TT_{1s} , TT_{2s} and TC_{1s} , TC_{2s} respectively the travel times and the toll prices for the two toll road options in choice set s . Finally, the probability that the j th tolled alternative is chosen in choice set s equals

$$p_{js} = p_{T,s} \frac{e^{(\beta_{TT} \times TT_{js}) + (\beta_{TC} \times TC_{js})}}{e^{(\beta_{TT} \times TT_{1s}) + (\beta_{TC} \times TC_{1s})} + e^{(\beta_{TT} \times TT_{2s}) + (\beta_{TC} \times TC_{2s})}}.$$

As in the preceding section, we will show that the choice of prior information affects the presence of choice sets with a dominant alternative in the Bayesian \mathcal{D} -efficient designs. The information matrix for the nested logit model equals

$$\begin{aligned} \mathbf{I}(\mathbf{X}, \beta) &= \sum_{s=1}^S \mathbf{I}_s(\mathbf{X}_s, \beta) \\ &= \sum_{s=1}^S \begin{bmatrix} \mathbf{T}'_s \mathbf{D}_s \mathbf{T}_s + i_1 & \mathbf{T}'_s \mathbf{D}_s \mathbf{C}_s + i_2 & p_{F,s} \lambda V_{T,s}(\mathbf{T}'_s \mathbf{p}_s) + i_3 \\ \mathbf{C}'_s \mathbf{D}_s \mathbf{T}_s + i_2 & \mathbf{C}'_s \mathbf{D}_s \mathbf{C}_s & p_{F,s} \lambda V_{T,s}(\mathbf{C}'_s \mathbf{p}_s) \\ p_{F,s} \lambda V_{T,s}(\mathbf{T}'_s \mathbf{p}_s) + i_3 & p_{F,s} \lambda V_{T,s}(\mathbf{C}'_s \mathbf{p}_s) & p_{F,s} p_{T,s} V_{T,s}^2 \end{bmatrix}, \end{aligned}$$

with

$$\begin{aligned} i_1 &= p_{F,s} p_{T,s} TT_{F,s}^2 - 2\lambda p_{F,s} TT_{F,s}(\mathbf{T}'_s \mathbf{p}_s) \\ i_2 &= -\lambda p_{F,s} TT_{F,s}(\mathbf{C}'_s \mathbf{p}_s) \\ i_3 &= -p_{F,s} p_{T,s} TT_{F,s} V_{T,s} \end{aligned}$$

and

$$\begin{aligned} \mathbf{D}_s &= \mathbf{P}_s + p_{T,s}^{-1}(\lambda^2 p_{F,s} - 1)\mathbf{p}_s \mathbf{p}'_s \\ \mathbf{T}_s &= [TT_{1s}, TT_{2s}]' \\ \mathbf{C}_s &= [TC_{1s}, TC_{2s}]' \\ \mathbf{X}_s &= \{TT_{F,s}, \mathbf{T}_s, \mathbf{C}_s\}. \end{aligned}$$

The derivations for obtaining these expressions are given in Appendix B. Since we are only interested in accurately estimating the model partworths β_{TT} and β_{TC} , and not the dissimilarity parameter λ , \mathcal{D}_s -efficient designs (Atkinson et al. 2007; Goos 2002) are obtained instead. The optimal design for the nested logit model with respect to the \mathcal{D}_s -optimality criterion is the design that maximizes

$$\frac{\det[\mathbf{I}(\mathbf{X}, \beta)]}{\sum_{s=1}^S p_{F,s} p_{T,s} V_{T,s}^2}.$$

Here also, Bayesian \mathcal{D}_s -efficient designs are obtained by minimizing an approximation of the Bayesian \mathcal{D}_s -error

$$\tilde{\mathcal{D}}_{sB} - \text{error} = \frac{1}{R} \sum_{r=1}^R \left[\frac{\det[\mathbf{I}(\mathbf{X}, \beta^r)]}{\sum_{s=1}^S p_{F,s} p_{T,s} V_{T,s}^2} \right]^{-1/2}.$$

To examine the effect of the prior information on the number of choice sets with a dominant alternative in the \mathcal{D}_s -efficient designs, we compare the two sets of prior means introduced in the previous section. Consider again the choice sets from Table 4. The probabilities of choosing the alternatives in these choice sets according to the nested logit model are given in Table 5 and Table 6 for the prior mean values $(-0.2, -1)$ and $(-0.4, -2)$ respectively. The dissimilarity parameter λ is respectively set equal to 0, 0.25, 0.5, 0.75 and 1. Similar as for the conditional logit model, the prior means $(-0.4, -2)$ express the dominance structure in the experiment better than the prior values $(-0.2, -1)$. Therefore we expect the Bayesian \mathcal{D}_s -efficient designs for the nested logit model obtained with this prior information to exclude choice sets with a dominant toll road option.

$(\beta_{TT}, \beta_{TC}) = (-0.2, -1)$	λ				
	0	0.25	0.5	0.75	1
40 min - \$0	0.03%	0.18%	0.99%	5.19%	23.00%
29 min - \$1	99.15%	99.00%	98.20%	94.04%	76.37%
38 min - \$4	0.82%	0.82%	0.81%	0.77%	0.63%
52 min - \$0	$\pm 0\%$	0.01%	0.09%	0.49%	2.64%
29 min - \$1	99.18%	99.17%	99.09%	98.69%	96.57%
38 min - \$4	0.82%	0.82%	0.82%	0.82%	0.79%

Table 5: Probabilities of choosing a toll-free option and the two most extreme toll road options for prior values $(-0.2, -1)$

$(\beta_{TT}, \beta_{TC}) = (-0.4, -2)$	λ				
	0	0.25	0.5	0.75	1
40 min - \$0	$\pm 0\%$	$\pm 0\%$	0.01%	0.30%	8.32%
29 min - \$1	99.99%	99.99%	99.98%	99.69%	91.67%
38 min - \$4	0.01%	0.01%	0.01%	0.01%	0.01%
52 min - \$0	$\pm 0\%$	$\pm 0\%$	$\pm 0\%$	$\pm 0\%$	0.07%
29 min - \$1	99.99%	99.99%	99.99%	99.99%	99.92%
38 min - \$4	0.01%	0.01%	0.01%	0.01%	0.01%

Table 6: Probabilities of choosing a toll-free option and the two most extreme toll road options for prior values $(-0.4, -2)$

The two multivariate normal distributions from the previous section are used as priors for the model partworths. For the dissimilarity parameter λ , a $\mathcal{U}[0, 1]$ prior is assumed first. Also, a $\mathcal{U}[0, 0.5]$ prior is assumed as we believe that the unobserved utility of the two toll road options in the choice sets is highly correlated. Bayesian \mathcal{D}_s -efficient designs with eight choice sets are computed (Appendix C). Three hundred random start designs are updated making use of 1000 random draws from the prior distributions with the modified Fedorov algorithm. The fractions of choice sets with a dominant toll road alternative in the efficient designs are given in Table 7. Again it can be concluded that a proper choice of prior information positively influences the suitability of the efficient design as nonsense choice sets are excluded from the designs.

	(β_{TT}, β_{TC})	
	$(-0.2, -1)$	$(-0.4, -2)$
$\lambda \sim \mathcal{U}[0, 0.5]$	4/8	0/8
$\lambda \sim \mathcal{U}[0, 1]$	4/8	0/8

Table 7: Fraction of choice sets with a dominant toll road option in the Bayesian \mathcal{D}_s -efficient designs

4 Conclusion

Choice-based conjoint choice experiments are an excellent tool to elicit information about consumers' product or service attribute preferences (Sawtooth 2008). Managers and policy makers frequently rely on such discrete choice data to make marketing decisions or to define a policy plan. Obviously, the response data are only reliable and useful for decision taking if they result from realistic choice situations. Therefore the quality of the data is a major concern and well-chosen designs should be used to obtain consumers' choice data. Firstly, the design must be efficient with respect to model parameter estimation. \mathcal{D} -efficient designs are constructed, as we wish to obtain accurate estimates. Moreover, the

design must be meaningful in the way that no nonsense choice sets, including a dominant alternative or including two equal alternatives, should be presented to the respondents.

This research focussed on a specific problem from the transportation field. Based on stated preference data from travelers, policy makers will decide on the profitability of a new toll road, as an alternative to existing and slower toll-free road options. To obtain information about the travelers' willingness to pay, efficient designs with two different types of choice sets were computed. First choice sets including only two toll road options, with different travel time and travel cost levels, were considered. These choice data were analyzed with the conditional logit model. Second, the choice data from sets with three alternatives, i.e. two toll road options and an additional toll-free alternative, was modeled with the nested logit model. The aim of the study was to find Bayesian \mathcal{D} -efficient designs that didn't include choice sets with a dominant alternative. It is shown that with an adequate choice of prior information, this goal can be achieved.

Assuming a multivariate normal prior distribution for the two model parts, two different prior mean vectors were compared to study the effect of the prior information. One can conclude from the design results that when the dominance structure is integrated in the prior in an appropriate way, the \mathcal{D} -efficient designs will not bring in choice sets with a dominant alternative. Choosing the prior means further away from the zero vector expresses more adequately that a road option with longer travel time and higher toll price is dominated by faster and cheaper options. This research advocates a proper choice of prior information to eliminate impractical and useless choice sets from \mathcal{D} -efficient designs, instead of computing efficient designs with improper prior information and deleting nonsense choice sets from these designs afterwards.

Appendix A. \mathcal{D} -efficient designs with choice sets including two toll road options

Choice set	Alt	Attr	
		TT	TC
1	I	29	4
	II	38	1
2	I	38	2
	II	29	2
3	I	29	2
	II	38	2
4	I	29	4
	II	38	1
5	I	29	4
	II	38	1
6	I	38	2
	II	29	2
7	I	29	4
	II	38	1
8	I	29	2
	II	38	2

Table 8: \mathcal{D} -efficient design for prior mean values $(-0.2, -1)$

Choice set	Alt	Attr	
		TT	TC
1	I	29	4
	II	38	1
2	I	29	4
	II	38	2
3	I	38	1
	II	29	2
4	I	29	4
	II	38	1
5	I	29	2
	II	38	1
6	I	38	1
	II	29	4
7	I	29	2
	II	38	1
8	I	29	3
	II	38	2

Table 9: \mathcal{D} -efficient design for prior mean values $(-0.4, -2)$

Appendix B. Information matrix for the nested logit model

Fix $y_{F,s}$ and $y_{T,s}$ to one if respectively a toll-free or a toll road option is chosen in choice set s . Similarly, y_{js} equals one if the j th tolled alternative is chosen in choice set s . As such $\sum_{j=1}^2 y_{js} = y_{T,s}$ and $y_{F,s} + \sum_{j=1}^2 y_{js} = 1$. The likelihood

function and the loglikelihood function for choice set s now respectively equal

$$\begin{aligned}
L_s &= p_{F,s}^{y_{F,s}} \prod_{j=1}^2 p_{j,s}^{y_{j,s}} \\
&= \left(\frac{e^{\beta_{TT} TT_{F,s}}}{e^{\lambda V_{T,s}} + e^{\beta_{TT} TT_{F,s}}} \right)^{y_{F,s}} \prod_{j=1}^2 \left(\frac{e^{\lambda V_{T,s}} e^{\mathbf{x}'_{j,s} \beta}}{e^{\lambda V_{T,s}} + e^{\beta_{TT} TT_{F,s}} \sum_{k=1}^2 e^{\mathbf{x}'_{k,s} \beta}} \right)^{y_{j,s}} \\
&= \frac{(e^{\beta_{TT} TT_{F,s}})^{y_{F,s}} e^{\lambda V_{T,s} \sum_{j=1}^2 y_{j,s}}}{(e^{\lambda V_{T,s}} + e^{\beta_{TT} TT_{F,s}})^{y_{F,s} + \sum_{j=1}^2 y_{j,s}} (\sum_{k=1}^2 e^{\mathbf{x}'_{k,s} \beta})^{\sum_{j=1}^2 y_{j,s}}} \prod_{j=1}^2 e^{y_{j,s} \mathbf{x}'_{j,s} \beta} \\
&= \frac{(e^{\beta_{TT} TT_{F,s}})^{y_{F,s}} e^{y_{T,s} \lambda V_{T,s}}}{(e^{\lambda V_{T,s}} + e^{\beta_{TT} TT_{F,s}})^{y_{T,s}} (\sum_{k=1}^2 e^{\mathbf{x}'_{k,s} \beta})^{y_{T,s}}} \prod_{j=1}^2 e^{y_{j,s} \mathbf{x}'_{j,s} \beta},
\end{aligned}$$

$$\begin{aligned}
l_s &= y_{F,s} \beta_{TT} TT_{F,s} + y_{T,s} \lambda V_{T,s} - \log(e^{\lambda V_{T,s}} + e^{\beta_{TT} TT_{F,s}}) - y_{T,s} \log \left(\sum_{j=1}^2 e^{\mathbf{x}'_{j,s} \beta} \right) + \sum_{j=1}^2 y_{j,s} \mathbf{x}'_{j,s} \beta \\
&= y_{T,s} (\lambda - 1) V_{T,s} - \log(e^{\lambda V_{T,s}} + e^{\beta_{TT} TT_{F,s}}) + y_{F,s} \beta_{TT} TT_{F,s} + \sum_{j=1}^2 y_{j,s} \mathbf{x}'_{j,s} \beta.
\end{aligned}$$

For choice set s , the information matrix is computed as

$$\mathbf{I}_s = \mathbf{E} \left[\begin{array}{cc} \left(\frac{\partial l_s}{\partial \beta_{TT}} \right) & \left(\frac{\partial l_s}{\partial \beta_{TC}} \right) \\ \left(\frac{\partial l_s}{\partial \beta_{TC}} \right) & \left(\frac{\partial l_s}{\partial \lambda} \right) \end{array} \right]' = \mathbf{E} \left[\begin{array}{ccc} \left(\frac{\partial l_s}{\partial \beta_{TT}} \right) \left(\frac{\partial l_s}{\partial \beta_{TT}} \right)' & \left(\frac{\partial l_s}{\partial \beta_{TT}} \right) \left(\frac{\partial l_s}{\partial \beta_{TC}} \right)' & \left(\frac{\partial l_s}{\partial \beta_{TT}} \right) \left(\frac{\partial l_s}{\partial \lambda} \right)' \\ \left(\frac{\partial l_s}{\partial \beta_{TC}} \right) \left(\frac{\partial l_s}{\partial \beta_{TT}} \right)' & \left(\frac{\partial l_s}{\partial \beta_{TC}} \right) \left(\frac{\partial l_s}{\partial \beta_{TC}} \right)' & \left(\frac{\partial l_s}{\partial \beta_{TC}} \right) \left(\frac{\partial l_s}{\partial \lambda} \right)' \\ \left(\frac{\partial l_s}{\partial \lambda} \right) \left(\frac{\partial l_s}{\partial \beta_{TT}} \right)' & \left(\frac{\partial l_s}{\partial \lambda} \right) \left(\frac{\partial l_s}{\partial \beta_{TC}} \right)' & \left(\frac{\partial l_s}{\partial \lambda} \right) \left(\frac{\partial l_s}{\partial \lambda} \right)' \end{array} \right].$$

Because

$$\frac{\partial V_{T,s}}{\partial \beta_{TT}} = \frac{\sum_{j=1}^2 e^{\mathbf{x}'_{j,s} \beta} TT_{j,s}}{\sum_{j=1}^2 e^{\mathbf{x}'_{j,s} \beta}} = p_{T,s}^{-1} \sum_{j=1}^2 p_{j,s} TT_{j,s}$$

and

$$\frac{\partial V_{T,s}}{\partial \beta_{TC}} = \frac{\sum_{j=1}^2 e^{\mathbf{x}'_{j,s} \beta} TC_{j,s}}{\sum_{j=1}^2 e^{\mathbf{x}'_{j,s} \beta}} = p_{T,s}^{-1} \sum_{j=1}^2 p_{j,s} TC_{j,s},$$

we find that

$$\frac{\partial l_s}{\partial \beta_{TT}} = y_{T,s} p_{T,s}^{-1} (\lambda - 1) \sum_{j=1}^2 p_{j,s} TT_{j,s} + \sum_{j=1}^2 y_{j,s} TT_{j,s} - \lambda \sum_{j=1}^2 p_{j,s} TT_{j,s} + (y_{F,s} - p_{F,s}) TT_{F,s},$$

$$\frac{\partial l_s}{\partial \beta_{TC}} = y_{T,s} p_{T,s}^{-1} (\lambda - 1) \sum_{j=1}^2 p_{js} TC_{js} + \sum_{j=1}^2 y_{js} TC_{js} - \lambda \sum_{j=1}^2 p_{js} TC_{js}.$$

We also find that

$$\frac{\partial l_s}{\partial \lambda} = y_{T,s} V_{T,s} - \frac{e^{\lambda V_{T,s}} V_{T,s}}{e^{\lambda V_{T,s}} + e^{\beta_{TT} TT_{F,s}}} = (y_{T,s} - p_{T,s}) V_{T,s}.$$

Since $y_{T,s}^2 = y_{T,s}$ and $y_{T,s} y_{js} = y_{js}$, we have that

$$\left(\frac{\partial l_s}{\partial \lambda} \right) \left(\frac{\partial l_s}{\partial \lambda} \right)' = (y_{T,s} - 2y_{T,s} p_{T,s} + p_{T,s}^2) V_{T,s}^2,$$

$$\begin{aligned} \left(\frac{\partial l_s}{\partial \beta_{TT}} \right) \left(\frac{\partial l_s}{\partial \lambda} \right)' &= (y_{T,s} p_{T,s}^{-1} - y_{T,s}) (\lambda - 1) V_{T,s} \sum_{j=1}^2 p_{js} TT_{js} + (1 - p_{T,s}) V_{T,s} \sum_{j=1}^2 y_{js} TT_{js} - \\ &\quad (y_{T,s} - p_{T,s}) V_{T,s} \lambda \sum_{j=1}^2 p_{js} TT_{js} + TT_{F,s} V_{T,s} (y_{T,s} - p_{T,s}) (y_{F,s} - p_{F,s}) \end{aligned}$$

and

$$\begin{aligned} \left(\frac{\partial l_s}{\partial \beta_{TC}} \right) \left(\frac{\partial l_s}{\partial \lambda} \right)' &= (y_{T,s} p_{T,s}^{-1} - y_{T,s}) (\lambda - 1) V_{T,s} \sum_{j=1}^2 p_{js} TC_{js} + (1 - p_{T,s}) V_{T,s} \sum_{j=1}^2 y_{js} TC_{js} - \\ &\quad (y_{T,s} - p_{T,s}) V_{T,s} \lambda \sum_{j=1}^2 p_{js} TC_{js}. \end{aligned}$$

Consequently,

$$\mathbb{E} \left\{ \left(\frac{\partial l_s}{\partial \lambda} \right) \left(\frac{\partial l_s}{\partial \lambda} \right)' \right\} = (p_{T,s} - p_{T,s}^2) V_{T,s}^2 = p_{F,s} p_{T,s} V_{T,s}^2,$$

$$\begin{aligned} \mathbb{E} \left\{ \left(\frac{\partial l_s}{\partial \beta_{TT}} \right) \left(\frac{\partial l_s}{\partial \lambda} \right)' \right\} &= p_{F,s} \lambda V_{T,s} \sum_{j=1}^2 p_{js} TT_{js} - p_{F,s} p_{T,s} TT_{F,s} V_{T,s} \\ &= p_{F,s} \lambda V_{T,s} (\mathbf{T}'_s \mathbf{p}_s) - p_{F,s} p_{T,s} TT_{F,s} V_{T,s} \end{aligned}$$

and

$$\begin{aligned} \mathbb{E} \left\{ \left(\frac{\partial l_s}{\partial \beta_{TC}} \right) \left(\frac{\partial l_s}{\partial \lambda} \right)' \right\} &= p_{F,s} \lambda V_{T,s} \sum_{j=1}^2 p_{js} TC_{js} \\ &= p_{F,s} \lambda V_{T,s} (\mathbf{C}'_s \mathbf{p}_s). \end{aligned}$$

Further, we find that

$$\begin{aligned}
\left(\frac{\partial l_s}{\partial \beta_{TT}}\right)\left(\frac{\partial l_s}{\partial \beta_{TT}}\right)' &= [y_{T,s}p_{T,s}^{-2}(\lambda-1)^2 - 2y_{T,s}p_{T,s}^{-1}\lambda(\lambda-1) + \lambda^2]\left(\sum_{j=1}^2 p_{js}TT_{js}\right)\left(\sum_{j=1}^2 p_{js}TT_{js}\right)' + \\
&\quad \left(\sum_{j=1}^2 y_{js}TT_{js}\right)\left(\sum_{j=1}^2 y_{js}TT_{js}\right)' + [p_{T,s}^{-1}(\lambda-1) - \lambda]\left(\sum_{j=1}^2 p_{js}TT_{js}\right)\left(\sum_{j=1}^2 y_{js}TT_{js}\right)' \\
&\quad + [p_{T,s}^{-1}(\lambda-1) - \lambda]\left(\sum_{j=1}^2 y_{js}TT_{js}\right)\left(\sum_{j=1}^2 p_{js}TT_{js}\right)' + \\
&\quad 2(y_{F,s} - p_{F,s})TT_{F,s}[y_{T,s}p_{T,s}^{-1}(\lambda-1) - \lambda]\left(\sum_{j=1}^2 p_{js}TT_{js}\right) + \\
&\quad 2(y_{F,s} - p_{F,s})TT_{F,s}\left(\sum_{j=1}^2 y_{js}TT_{js}\right) + (y_{F,s} - p_{F,s})^2TT_{F,s}^2,
\end{aligned}$$

$$\begin{aligned}
\left(\frac{\partial l_s}{\partial \beta_{TC}}\right)\left(\frac{\partial l_s}{\partial \beta_{TC}}\right)' &= [y_{T,s}p_{T,s}^{-2}(\lambda-1)^2 - 2y_{T,s}p_{T,s}^{-1}\lambda(\lambda-1) + \lambda^2]\left(\sum_{j=1}^2 p_{js}TC_{js}\right)\left(\sum_{j=1}^2 p_{js}TC_{js}\right)' + \\
&\quad \left(\sum_{j=1}^2 y_{js}TC_{js}\right)\left(\sum_{j=1}^2 y_{js}TC_{js}\right)' + [p_{T,s}^{-1}(\lambda-1) - \lambda]\left(\sum_{j=1}^2 p_{js}TC_{js}\right)\left(\sum_{j=1}^2 y_{js}TC_{js}\right)' \\
&\quad + [p_{T,s}^{-1}(\lambda-1) - \lambda]\left(\sum_{j=1}^2 y_{js}TC_{js}\right)\left(\sum_{j=1}^2 p_{js}TC_{js}\right)'
\end{aligned}$$

and

$$\begin{aligned}
\left(\frac{\partial l_s}{\partial \beta_{TC}}\right)\left(\frac{\partial l_s}{\partial \beta_{TT}}\right)' &= [y_{T,s}p_{T,s}^{-2}(\lambda-1)^2 - 2y_{T,s}p_{T,s}^{-1}\lambda(\lambda-1) + \lambda^2]\left(\sum_{j=1}^2 p_{js}TC_{js}\right)\left(\sum_{j=1}^2 p_{js}TT_{js}\right)' + \\
&\quad \left(\sum_{j=1}^2 y_{js}TC_{js}\right)\left(\sum_{j=1}^2 y_{js}TT_{js}\right)' + [p_{T,s}^{-1}(\lambda-1) - \lambda]\left(\sum_{j=1}^2 p_{js}TC_{js}\right)\left(\sum_{j=1}^2 y_{js}TT_{js}\right)' \\
&\quad + [p_{T,s}^{-1}(\lambda-1) - \lambda]\left(\sum_{j=1}^2 y_{js}TC_{js}\right)\left(\sum_{j=1}^2 p_{js}TT_{js}\right)' + \\
&\quad (y_{F,s} - p_{F,s})TT_{F,s}[y_{T,s}p_{T,s}^{-1}(\lambda-1) - \lambda]\left(\sum_{j=1}^2 p_{js}TC_{js}\right) + \\
&\quad (y_{F,s} - p_{F,s})TT_{F,s}\left(\sum_{j=1}^2 y_{js}TC_{js}\right).
\end{aligned}$$

As $E(y_{js}) = p_{js}$ and $E(y_{T,s}) = p_{T,s}$,

$$E\left\{\left(\frac{\partial l_s}{\partial \beta_{TT}}\right)\left(\frac{\partial l_s}{\partial \beta_{TT}}\right)'\right\} = \mathbf{T}'_s[\mathbf{P}_s + p_{T,s}^{-1}(\lambda^2 p_{F,s} - 1)\mathbf{p}_s\mathbf{p}'_s]\mathbf{T}_s + p_{F,s}p_{T,s}TT_{F,s}^2 - 2\lambda p_{F,s}TT_{F,s}(\mathbf{T}'_s\mathbf{p}_s),$$

$$E\left\{\left(\frac{\partial l_s}{\partial \beta_{TC}}\right)\left(\frac{\partial l_s}{\partial \beta_{TC}}\right)'\right\} = \mathbf{C}'_s[\mathbf{P}_s + p_{T,s}^{-1}(\lambda^2 p_{F,s} - 1)\mathbf{p}_s\mathbf{p}'_s]\mathbf{C}_s$$

and

$$E\left\{\left(\frac{\partial l_s}{\partial \beta_{TC}}\right)\left(\frac{\partial l_s}{\partial \beta_{TT}}\right)'\right\} = \mathbf{C}'_s[\mathbf{P}_s + p_{T,s}^{-1}(\lambda^2 p_{F,s} - 1)\mathbf{p}_s\mathbf{p}'_s]\mathbf{T}_s - \lambda p_{F,s}TT_{F,s}(\mathbf{C}'_s\mathbf{p}_s).$$

Appendix C. \mathcal{D}_s -efficient designs with choice sets including one toll-free and two toll road options

Choice set	Alt	Attr	
		TT	TC
1	I	52	0
	II	38	1
	III	29	4
2	I	52	0
	II	38	1
	III	29	4
3	I	44	0
	II	29	1
	III	38	1
4	I	44	0
	II	29	1
	III	38	1
5	I	44	0
	II	29	1
	III	38	1
6	I	40	0
	II	38	4
	III	29	4
7	I	52	0
	II	29	4
	III	38	1
8	I	52	0
	II	29	4
	III	38	1

Table 10: \mathcal{D}_s -efficient design for prior mean values $(-0.2, -1)$ and prior distribution $\mathcal{U}[0, 0.5]$ for λ

Choice set	Alt	Attr	
		TT	TC
1	I	44	0
	II	29	4
	III	38	1
2	I	44	0
	II	38	1
	III	29	4
3	I	40	0
	II	29	2
	III	38	1
4	I	40	0
	II	29	2
	III	38	1
5	I	40	0
	II	38	3
	III	29	4
6	I	44	0
	II	29	3
	III	38	1
7	I	44	0
	II	29	4
	III	38	1
8	I	44	0
	II	29	3
	III	38	2

Table 11: \mathcal{D}_s -efficient design for prior mean values $(-0.4, -2)$ and prior distribution $\mathcal{U}[0, 0.5]$ for λ

Choice set	Alt	Attr	
		TT	TC
1	I	48	0
	II	29	1
	III	38	1
2	I	52	0
	II	38	1
	III	29	4
3	I	48	0
	II	29	1
	III	38	1
4	I	52	0
	II	29	4
	III	38	1
5	I	48	0
	II	38	1
	III	29	1
6	I	40	0
	II	38	4
	III	29	4
7	I	52	0
	II	29	4
	III	38	1
8	I	52	0
	II	29	4
	III	38	1

Table 12: \mathcal{D}_s -efficient design for prior mean values $(-0.2, -1)$ and prior distribution $\mathcal{U}[0, 1]$ for λ

Choice set	Alt	Attr	
		TT	TC
1	I	52	0
	II	38	2
	III	29	4
2	I	48	0
	II	29	3
	III	38	1
3	I	40	0
	II	29	4
	III	38	1
4	I	52	0
	II	38	3
	III	29	4
5	I	48	0
	II	29	4
	III	38	1
6	I	40	0
	II	29	4
	III	38	3
7	I	52	0
	II	29	4
	III	38	3
8	I	48	0
	II	38	1
	III	29	4

Table 13: \mathcal{D}_s -efficient design for prior mean values $(-0.4, -2)$ and prior distribution $\mathcal{U}[0, 1]$ for λ

References

- [1] Atkinson, A.C., Donev, A.N., & Tobias, R.D. (2007). Optimum experimental designs, with SAS. Clarendon Press, Oxford U.K.
- [2] Axhausen, K.W., Hess, S., König, A., Abay, G., Bates, J.J., & Bierlaire, M. (2008). Income and distance elasticities of values of travel time savings: New Swiss results. *Transport Policy*, 15, 173-185.
- [3] Bliemer, M.C.J., Rose, J.M., & van Blokland, R.B. (2009). Experimental design influences on stated choice outputs. Paper to be presented at the European Transport Conference, Nordwijkerhout, The Netherlands.
- [4] Goos, P., (2002). The optimal design of blocked and split-plot experiments. Springer, New York.
- [5] Goos, P., Vermeulen, B., & Vandebroek, M. (2010). \mathcal{D} -optimal conjoint choice designs with no-choice options for a nested logit model. *Journal of Statistical Planning and Inference*, 140, 851-861.
- [6] Hensher, D.A. (2001). The sensitivity of the valuation of travel time savings to the specification of unobserved effects. *Transportation Research Part E*, 37, 129-142.
- [7] Hess, S., & Adler, T. (2010). Experimental designs for the real world. Paper presented at the 89th Annual Meeting of the Transportation Research Board, Washington, D.C.
- [8] Hess, S., Bierlaire, M., & Polak, J.W. (2005). Estimation of value of travel time savings using mixed logit models. *Transportation Research Part A*, 39, 221-236.
- [9] Hess, S., Smith, C., Falzarano, S., & Stubits, J. (2008). Measuring the effects of different experimental designs and survey administration methods using an Atlanta Managed Lanes Stated Preference survey. *Transportation Research Record*, 2049, 144-152.
- [10] Kessels, R., Goos, P., & Vandebroek, M. (2006a). A comparison of criteria to design efficient choice experiments. *Journal of Marketing Research*, 43, 409-419.
- [11] Kessels, R., Jones, B., Goos, P., & Vandebroek, M. (2006b). Recommendations on the use of Bayesian optimal designs for choice experiments. Research Report KBI 0617, Department of Decision Sciences and Information Management, Katholieke Universiteit Leuven, 18 pp.
- [12] Rose, J.M., & Bliemer, M.C.J. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews*, 29, 587-617.
- [13] Sándor, Z., & Wedel, M. (2001). Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, 38, 430-444.

- [14] Sawtooth Software Inc. (2008). The CBC Technical Paper (version 6). *Sawtooth Software Technical Paper Series* 1-27. Retrieved from <http://www.sawtoothsoftware.com/download/techpap/cbctech.pdf>.
- [15] Train, K. (2003). *Discrete choice methods with simulation*. Cambridge University Press, New York.