

Crabbe, Marjolein; Vandebroek, Martina

Article

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

Journal of Choice Modelling

Provided in Cooperation with:

Journal of Choice Modelling

Suggested Citation: Crabbe, Marjolein; Vandebroek, Martina (2012) : Using appropriate prior information to eliminate choice sets with a dominant alternative from D-efficient designs, Journal of Choice Modelling, ISSN 1755-5345, University of Leeds, Institute for Transport Studies, Leeds, Vol. 5, Iss. 1, pp. 22-45

This Version is available at:

<https://hdl.handle.net/10419/66843>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by-nc/2.0/uk/>

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

Marjolein Crabbe^{1,*} Martina Vandebroek^{2,†}

¹ Katholieke Universiteit Leuven, Faculty of Business and Economics, Naamsestraat 69, 3000 Leuven, Belgium

² Katholieke Universiteit Leuven, Faculty of Business and Economics and Leuven Statistics Research Centre, Naamsestraat 69, 3000 Leuven, Belgium

Received 5 January 2011, revised version received 25 October 2011, accepted 4 January 2012

Abstract

Discrete choice experiments are frequently applied in transportation research to explore the value of travel time savings. Most profile attributes in these studies, such as the travel time or the travel cost of a mode or road choice, have a clear rank order in their attribute levels. Therefore one option in a choice set can dominate the other alternatives in the set. This research shows, for a specific setup in the transportation field, that a proper choice of prior information, adequately incorporating the dominance of preferred attribute levels, eliminates choice sets with a dominant alternative from Bayesian \mathcal{D} -efficient designs.

Keywords: discrete choice experiments; dominant alternative; prior information; Bayesian \mathcal{D} -efficient designs; conditional logit model; nested logit model

1 Introduction

Discrete choice experiments know a long tradition in many research fields such as marketing, transportation, environmental and health economics, and have become indispensable in the analysis of choice behavior. Over the past years, not only choice modeling but also the design of choice experiments progressed substantially. Aiming to improve the quality of choice data, the focus in discrete choice design gradually shifted from orthogonal to statistically efficient designs. Efficiently constructed designs elicit choice data more effectively than traditional

*T: +32 16 3 26961, marjolein.crabbe@kuleuven.be

†T: +32 16 3 26975, Martina.Vandebroek@econ.kuleuven.be



orthogonal designs as a smaller sample size is needed to attain the same level of estimation accuracy (Bliemer and Rose 2005; Sándor and Wedel 2001). Unfortunately, despite the benefits of efficient designs, experiments revealed that they may include choice sets with dominant alternatives (Hess and Adler 2010).

The alternatives in the choice sets of a choice experiment are defined by attributes, for most of which a positive or a negative effect on utility can be argued. Increasing the attribute level either increases or decreases the attractiveness of the product and consequently the probability of choosing it. Alternatives with more preferred levels with respect to all attributes considered dominate the remaining options in the set in the sense that each rational participant in the survey will choose this dominating alternative. Choice sets with a dominant alternative do not give any additional information on the relative importance of attributes and are therefore not really informative in the analysis of preferences. Researchers have been aware of this, yet little attention has been devoted to this issue in much of the discrete choice literature. By emphasizing so on the importance of statistical efficiency, the presence of choice sets with a dominant alternative in efficient designs has been ignored.

Although one may argue that respondents might be relieved when receiving trivial choice tasks (as sometimes stated choice surveys can be quite intensive), we believe that offering too many choice sets with a dominant alternative is detrimental for the quality of the choice data and the research in general. Therefore it is relevant to find out how choice sets with a dominant alternative can be excluded from statistically efficient designs.

This research extends the discrete choice literature by highlighting and tackling the problem of dominance in efficient experimental design. We advocate the use of appropriate prior information, adequately incorporating the dominance of preferred attribute levels, to construct efficient designs and show that this avoids choice sets with a dominant alternative.

We illustrate by means of a choice experiment in which travelers are requested to judge and compare road options given travel time and travel cost (see Hess and Adler (2010)), how to verify whether prior values are appropriate and how to adjust them if needed. Considering the choice set with the least and the most attractive road alternative and their corresponding choice probabilities, it is shown that the prior values used in the paper do not reflect the dominance well and therefore the \mathcal{D} -efficient design includes choice sets with dominated alternatives. This study explains how shifting the given prior values for both attribute coefficients further away from zero pulls the choice probabilities of dominating and dominated alternatives further apart, yielding prior information that expresses the dominance of a shorter travel time with a lower travel cost more appropriately. Efficient designs based on those adjusted priors indeed show no choice sets with a dominant alternative.

Table 1: Choice set with a dominant tolled road option

	<i>Alternative I</i>	<i>Alternative II</i>
<i>Travel time</i>	32 minutes	38 minutes
<i>Toll price</i>	\$2	\$3

2 Experimental setup

Among many other fields of application, choice experiments are frequently applied in transportation studies (see for instance Bliemer et al. 2009; Rose and Bliemer 2009). The majority of these studies aims at investigating and quantifying 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 choice experiment analyzed in this research was discussed in Hess and Adler (2010) and deals with the attractiveness and profitability of a new and speedy toll road in comparison to existing toll-free but slower alternatives. 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.

The choice sets in the survey consist of road options for a specific trip defined by the travel time and, if present, 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. Bringing 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 to 40, 44, 48 or 52 minutes. Starting with a fairly simple setup, discarding toll-free alternatives, designs consisting of choice sets with only two tolled road options are dealt with in the following section. The fourth section generalizes ideas to a different and more complicated survey setup and consequently to a different model form. Here we discuss the experiment from Hess and Adler (2010), in which choice sets, besides two tolled road options, also include one toll-free option.

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 that one of both toll road options in a choice set dominates 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 alternative is given in Table 1.

The first option clearly dominates the second as it is faster without being more expensive. Obviously, such choice sets are not valuable for travel time savings estimation and should not appear in the experimental design. Therefore

we advocate the use of appropriate prior information in generating Bayesian \mathcal{D} -efficient designs to get rid of choice sets with a dominant tolled road option.

We construct efficient designs for two well known discrete choice models. For the first setup with choice sets including only two tolled road options, the conditional logit model (Train 2003) is used. To analyze the data from the second setup, in which a toll-free alternative is added to each choice set, the nested logit model (Train 2003) is applied as the conditional logit model turned out to be inappropriate for modeling these choice data.

3 Choice sets with two tolled alternatives

Before we turn to the experiment discussed in Hess and Adler (2010), in which each choice set comprises one toll-free and two tolled 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 (TT) and travel cost (TC), the utility a traveler receives from tolled alternative j ($j = 1$ or 2) in choice set s ($s = 1, \dots, S$) equals

$$U_{js} = (\beta_{TT} \times TT_{js}) + (\beta_{TC} \times TC_{js}) + \varepsilon_{js}, \quad (1)$$

with β_{TT} and β_{TC} the model parameters or 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 for the probability that a person chooses the tolled road option j in choice set s is found

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

To construct efficient choice designs, the \mathcal{D} -optimality criterion is used in this study (see for instance Atkinson et al. 2007). 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 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

$$\mathbf{I}(\mathbf{X}, \boldsymbol{\beta}) = \sum_{s=1}^S \mathbf{X}_s' (\mathbf{P}_s - \mathbf{p}_s \mathbf{p}_s') \mathbf{X}_s, \quad (3)$$

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

Moreover, taking uncertainty about prior values into account, we compute

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

$$\begin{aligned}\mathcal{D}_B - \text{error} &= E_{\boldsymbol{\beta}} \left[\det[\mathbf{I}(\mathbf{X}, \boldsymbol{\beta})]^{-1/2} \right] \\ &= \int_{\mathbb{R}^2} \det[\mathbf{I}(\mathbf{X}, \boldsymbol{\beta})]^{-1/2} \pi_0(\boldsymbol{\beta}) d\boldsymbol{\beta}.\end{aligned}\quad (4)$$

Note that in general the criterion is scaled to the power $1/p$, with p the number of parameters in the model which is two for the present study (β_{TT} and β_{TC}). In real applications, the Bayesian \mathcal{D} -error is approximated by the average \mathcal{D} -error over R draws $\boldsymbol{\beta}^r$ from the prior distribution

$$\tilde{\mathcal{D}}_B - \text{error} = \frac{1}{R} \sum_{r=1}^R \det[\mathbf{I}(\mathbf{X}, \boldsymbol{\beta}^r)]^{-1/2}. \quad (5)$$

The choice of the prior values or the prior distribution for the model parameters is important input in the search for statistically efficient designs. Prior information should represent the knowledge on the true population parameters, which, if any, is not always easy to provide. Specifically, we show that the presence of dominant alternatives in the choice sets of \mathcal{D} -efficient designs is due to the (improper) choice of prior information. As priors should reflect prior knowledge, they should also reflect the prior belief that rational respondents always choose the dominant alternative in a choice set. If one adequately incorporates the dominance of preferred attribute levels in the prior information, choice sets with a dominant alternative will not appear in \mathcal{D} -efficient designs.

For the travel time and the travel cost partworth in the model, a multivariate normal prior distribution will be assumed.

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

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 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. Note that different prior covariance structures were considered but that this had no effect on the conclusions that can be drawn from this simplified case.

The first set of prior means $(-0.2, -1)$ are the priors used in the second case

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

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

study in Hess and Adler (2010), dealing with the experiment considered in this research. This prior mean vector corresponds 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. Note that in Hess and Adlers' (2010) first case study and in Hess et al. (2008) prior values $(-0.04, -0.19)$ are chosen for a similar setup, again corresponding to a value of travel time savings of approximately \$12 per hour.

In the following we claim that the first prior specification does not adequately incorporate the dominance structure in the experiment and therefore choice sets with a dominant alternative will appear in the \mathcal{D} -efficient designs based on this prior information. Following 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 tolled road option (Table 2). 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 values $(-0.2, -1)$, the probability of choosing the dominant alternative is very high, yet approximately 0.82% of the travelers would still prefer the least attractive option over the most attractive one. As this is quite implausible, we propose moving the prior means further away from the zero vector, e.g. $(-0.4, -2)$, making the alternatives' utilities more unique and their choice probabilities more extreme. The probability of choosing the dominant alternative is now higher, approximately 100%, and more realistic. The dominance of lower attribute levels is thus more adequately incorporated in this prior information. One can expect \mathcal{D} -efficient designs constructed with respect to the first prior to include choice sets with a dominant alternative, whereas such choice sets will be excluded making use of the latter prior specification.

It seems that the prior values used in Hess and Adler (2010) are incapable of expressing the true preference structure of travelers as rational respondents should always choose the dominant alternative in a choice set. Prior values re-

flect the prior information on the true population parameters (for instance from a pilot study or previous experiments), but should additionally incorporate prior beliefs concerning e.g. the rationality of participants as this is also an important part of the prior knowledge. Even though the values $(-0.2, -1)$ are believed to correspond approximately to the true parameters, in reality one is of course not sure about the population values (why else would one do all the effort of setting up an experiment to estimate them). Therefore one should be willing to adjust prior values (and beliefs) in order to represent consumers' preferences better.

Moreover, one should account for potential scale effects when interpreting estimated partworths in choice models. As Swait and Louviere (1993) showed for instance that differences in partworth estimates from different data sources can occur simply due to differences in scale even when the true parameters are the same in both populations. Further note that although the absolute values of the priors have been changed, the belief about the value of travel time savings (what the experiment here is all about) remains the same as the willingness to pay is still assumed to be \$12 per hour. Multiplying each coefficient by the same constant does not affect the value of time, but is merely an increase in scale (Bliemer and Rose 2005; Bunch et al. 1996; Train 2003).

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 starting 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. We respectively refer to the efficient design for the prior distribution with mean values $(-0.2, -1)$ and $(-0.4, -2)$ as Design 1 and Design 2. A well considered choice of prior information has indeed a beneficial effect on the applicability of \mathcal{D} -efficient designs. Design 2 does not include choice sets with a dominant tolled road option as the latter prior specification expresses the dominance and preference structure in the experiment better. In conclusion, to what extent prior information expresses dominance in choice experiments can be verified by comparing the choice probabilities of the least and the most attractive feasible alternative. When a considerable portion of respondents prefers a dominated alternative, the prior information reflects irrational choice behavior and is therefore inappropriate. Shifting prior values further away from zero makes choice probabilities more extreme, expressing better that dominating alternatives should indeed be dominant. The multiplicative constant must be increased until almost all respondents choose the dominating option in the choice set.

Finally note that in practice often a more straightforward approach is applied. Instead of tackling the dominance problem before constructing efficient designs, most experimenters generate statistically efficient designs for specific (improper) prior values and simply remove manually the choice sets with a dominant alternative afterwards. As this may completely ruin the statistical efficiency of the

Table 3: Probabilities of choosing the most attractive and the least attractive toll road option and the fraction of choice sets with a dominant alternative in the Bayesian \mathcal{D} -efficient designs

		(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>Design 1</i>	<i>Design 2</i>
<i>Fraction of sets in the design with a dominant tolled road option</i>		4/8	0/8

Table 4: Local \mathcal{D} -errors of the Bayesian efficient designs for the conditional logit model

(β_{TT}, β_{TC})	<i>Design 1</i>	<i>Design 2</i>
$(-0.2, -1)$	0.0629	0.0754

remaining design, we do not recommend this. Moreover, the present and Hess and Adlers' (2010) study show that this practice may leave very few choice sets to submit to the respondents. Further note that constructing a larger efficient design will not necessarily yield more useful choice sets as, whatever the length of the design, the same choice sets (possibly with a dominant alternative) will remain most efficient with respect to the specific prior values and will pop up in the efficient design. Therefore, although more complex, we believe that avoiding dominance in choice experiments with appropriate prior information is the preferred strategy.

In the following, we check the robustness of the designs' efficiency in case the assumed priors are misspecified. In particular, even though we believe the new prior settings express the dominance of lower attribute levels and therefore travelers' preferences better, what is the observed efficiency loss using Design 2 over Design 1 in case $(-0.2, -1)$ are the true population parameters? Table 4 reports the local \mathcal{D} -errors, obtained by fixing prior values instead of averaging over a prior distribution as in (4), for Design 1 and Design 2 and parameters $(-0.2, -1)$. Additionally, Figure 1 plots the ratio of the local \mathcal{D} -errors for Design 2 relative to Design 1 for different values of β_{TT} and β_{TC} .

The sensitivity analysis reveals that the \mathcal{D} -error (and therefore a design's efficiency) is a function of the priors assumed and that it increases (decreases) in case priors are misspecified. From Table 4 it is clear that when the true population parameters indeed equal $(-0.2, -1)$, Design 1 (0.0629) turns out to be more

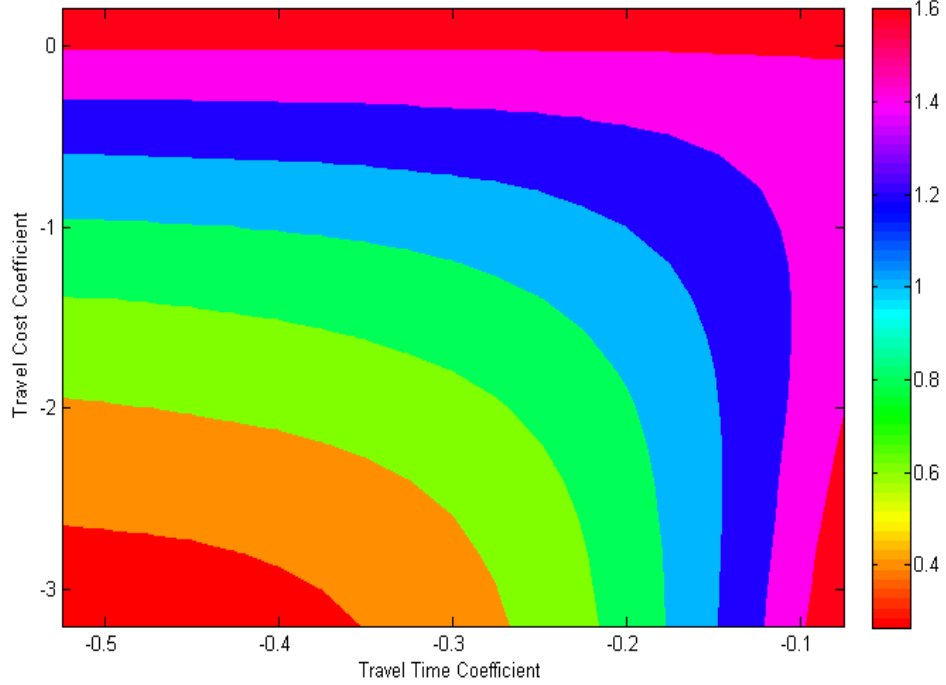


Figure 1: Ratio of local \mathcal{D} -errors for Design 2 relative to Design 1 for different values of β_{TT} and β_{TC}

efficient than Design 2 (0.0754). This makes sense as the prior distribution used to construct the Bayesian efficient Design 1 was centered around these values. Note however that the efficiency loss perceived when using Design 2 instead is only minor. Design 2's efficiency is thus quite robust with respect to the values $(-0.2, -1)$. This is in agreement with the results from Bliemer and Rose (2005), who found that as long as the ratio of the prior values of two coefficients is not changed much in comparison to the true ratio in the population, the efficiency of the \mathcal{D} -efficient design is quite stable in case of ill-assumed priors.

The stability of Design 2's efficiency also appears from the contourplot in Figure 1. As Bayesian designs were constructed, taking uncertainty about the prior values into account, Design 2 remains most efficient in a wide range around $(-0.4, -2)$. Moreover, the loss in efficiency with respect to Design 1 near $(-0.2, -1)$ is small or even missing. Therefore we advocate to adjust prior information to avoid choice sets with a dominant alternative in efficient designs, exchanging a little decrease in design efficiency with respect to generally accepted priors for a considerable increase in design applicability. Once more, it is a strong assumption to believe that one knows the true population parameters.

The travel time and the travel cost attribute each have four levels, therefore one could also obtain a model nonlinear in the attributes by coding them. By means of effects type coding efficient designs with 24 choice sets including two tolled road alternatives were constructed for the two prior settings discussed. The coding scheme has no effect on the results as the efficient design for the initial prior distribution still contains choice sets with a dominant alternative, whereas the adjusted and improved prior yields a design excluding these sets.

To generalize our findings, we also considered an alternative experimental setup based on the case study in Scarpa and Rose (2008). Their choice experiment consists of choice sets with two alternatives, defined by four attributes. More specifically, three quality attributes with four levels (0,1,2,3) each and a price attribute with five levels (0, 1, 2, 3, 4). In the example some a priori knowledge about the effect of the attributes on utility is assumed. The effect of price is obviously assumed negative and fixed to -1 , whereas the remaining attributes are assumed to have a positive effect on product utility. Moreover, the quality attribute coefficients are assumed ordered and equal to 1, 2 and 3 respectively. The priors (1, 2, 3, -1) yield a probability for choosing the least attractive alternative, i.e. lowest in quality and highest in price with attribute levels (0, 0, 0, 4), and the most attractive alternative, i.e. highest in quality and lowest in price with levels (3, 3, 3, 0), of approximately 0% and 100% according to the conditional logit model.

Assume now that the only prior knowledge is the order and relative size of the priors but not their absolute values. For example prior values (0.01, 0.02, 0.03, -0.01) could also be used to construct \mathcal{D} -efficient designs. Yet, for these values the choice probability for the least and the most attractive option respectively is 44.52% and 55.48%, not adequately expressing the dominance structure in the setup. Locally efficient designs with 20 choice sets are computed. The results are as expected, the use of the improper priors (0.01, 0.02, 0.03, -0.01) yields a design including choice sets with a dominant alternative whereas no such choice sets appear in the efficient design with respect to the values (1, 2, 3, -1). These results are in line with the previous findings.

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

In the second part of this research, the original toll road experiment from Hess and Adler (2010) is studied. Now, the choice sets in the design comprise not only two tolled roads but also one toll-free 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

Table 5: Probabilities of choosing a toll-free option and the two most extreme tolled road options for the conditional logit model

	(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%

alternatives, the first one a toll-free option and the final two toll roads.

We considered first the conditional logit model to analyze the choice data and the two prior distributions introduced in the previous section to obtain Bayesian \mathcal{D} -efficient designs. The appropriateness of the prior information is verified by two choice sets with three alternatives. The first set includes the toll-free option with the shortest (40 minutes) travel time and the least and most attractive tolled road option, the second set also includes the least and most attractive tolled road option but has the toll-free option with the longest (52 minutes) travel time. The choice probabilities of the alternatives in these two choice sets are given in Table 5 for both prior mean vectors.

One can conclude that the prior values $(-0.4, -2)$, in contrast to $(-0.2, -1)$, adequately incorporate the dominance of lower attribute levels as approximately 0% of the travelers would choose the dominated alternative in both choice sets. Yet, with prior means $(-0.4, -2)$, the Bayesian \mathcal{D} -efficient designs still include choice sets with a dominant tolled road option. Moreover, it does not help to specify the prior means even further away from zero as choice sets with a dominant alternative keep popping up in the designs.

The conditional logit model assumes that the error terms in the utility functions 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 probable that the unobserved utility of the two tolled road options, in contrast to the toll-free alternative, is correlated, the conditional logit model is no longer appropriate to analyze the choice data at hand. The use of this improper model can explain why, even with adequate 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 into meaningful subsets of similar alternatives, it is more suitable to use the nested logit choice model.

Based on the ideas of Goos et al. (2010) and Vermeulen et al. (2008), the

nested logit model is applied to the data given in this section, i.e. choices between one toll-free and two tolled road options. Two nests are considered, the first containing the toll-free road option and the second including the two tolled alternatives. 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

$$p_{F,s} = \frac{e^{\lambda_F V_{F,s}}}{e^{\lambda_T V_{T,s}} + e^{\lambda_F V_{F,s}}} \quad (6)$$

and

$$p_{T,s} = \frac{e^{\lambda_T V_{T,s}}}{e^{\lambda_T V_{T,s}} + e^{\lambda_F V_{F,s}}}, \quad (7)$$

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 captures the utility of the nest.

Because 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

$$V_{F,s} = \log \left[e^{(\beta_{TT} \times TT_{F,s}) + (\beta_{TC} \times TC_{F,s})} \right] = \beta_{TT} \times TT_{F,s}, \quad (8)$$

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,

$$p_{F,s} = \frac{e^{\beta_{TT} \times TT_{F,s}}}{e^{\lambda V_{T,s}} + e^{\beta_{TT} \times TT_{F,s}}} \quad (9)$$

and

$$p_{T,s} = \frac{e^{\lambda V_{T,s}}}{e^{\lambda V_{T,s}} + e^{\beta_{TT} \times TT_{F,s}}}. \quad (10)$$

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

$$V_{T,s} = \log \left[e^{(\beta_{TT} \times TT_{1s}) + (\beta_{TC} \times TC_{1s})} + e^{(\beta_{TT} \times TT_{2s}) + (\beta_{TC} \times TC_{2s})} \right], \quad (11)$$

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

$$p_{js} = p_{T,s} \frac{e^{(\beta_{TT} \times TT_{js}) + (\beta_{TC} \times TC_{js})}}{e^{(\beta_{TT} \times TT_{1s}) + (\beta_{TC} \times TC_{1s})} + e^{(\beta_{TT} \times TT_{2s}) + (\beta_{TC} \times TC_{2s})}}. \quad (12)$$

The information matrix for the nested logit model equals

$$\begin{aligned} \mathbf{I}(\mathbf{X}, \boldsymbol{\beta}) &= \sum_{s=1}^S \mathbf{I}_s(\mathbf{X}_s, \boldsymbol{\beta}) \\ &= \sum_{s=1}^S \begin{bmatrix} \mathbf{T}'_s \mathbf{D}_s \mathbf{T}_s + i_1 & \mathbf{T}'_s \mathbf{D}_s \mathbf{C}_s + i_2 & p_{F,s} \lambda V_{T,s}(\mathbf{T}'_s \mathbf{p}_s) + i_3 \\ \mathbf{C}'_s \mathbf{D}_s \mathbf{T}_s + i_2 & \mathbf{C}'_s \mathbf{D}_s \mathbf{C}_s & p_{F,s} \lambda V_{T,s}(\mathbf{C}'_s \mathbf{p}_s) \\ p_{F,s} \lambda V_{T,s}(\mathbf{T}'_s \mathbf{p}_s) + i_3 & p_{F,s} \lambda V_{T,s}(\mathbf{C}'_s \mathbf{p}_s) & p_{F,s} p_{T,s} V_{T,s}^2 \end{bmatrix}, \end{aligned} \quad (13)$$

with

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

and

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

The derivations for obtaining these expressions are given in Appendix B. Analogous to the preceding section, we will show that the choice of prior information affects the presence of choice sets with a dominant alternative in Bayesian \mathcal{D} -efficient designs for the nested logit model. But 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 Bayesian efficient design for the nested logit model with respect to the \mathcal{D}_s -optimality criterion is the design that minimizes the expectation of the \mathcal{D}_s -error, which is the inverse of

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

with $\mathbf{I}(\mathbf{X}, \boldsymbol{\beta})$ the information matrix in (13). Also here, Bayesian efficient designs will be obtained by an approximation of the Bayesian \mathcal{D}_s -error.

Consider again the choice sets from Table 5. The probabilities of choosing the alternatives in these choice sets according to the nested logit model are given in Table 6 and Table 7 for prior means $(-0.2, -1)$ and $(-0.4, -2)$ respectively. The dissimilarity parameter λ is consecutively set equal to 0, 0.25, 0.5, 0.75 and 1. For values $(-0.4, -2)$ the dominated road alternative (38 min - \$4) has approximately zero probability of being chosen in both sets, whereas for $(-0.2, -1)$ this probability ranges from 0.63% to 0.82%. Based on these results, we expect the Bayesian \mathcal{D}_s -efficient designs for the nested logit model obtained with prior

Table 6: Probabilities of choosing a toll-free option and the two most extreme tolled road options for the nested logit model and prior values $(-0.2, -1)$

$(\beta_{TT}, \beta_{TC}) = (-0.2, -1)$	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 7: Probabilities of choosing a toll-free option and the two most extreme tolled road options for the nested logit model and prior values $(-0.4, -2)$

$(\beta_{TT}, \beta_{TC}) = (-0.4, -2)$	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%

means $(-0.4, -2)$ to exclude choice sets with a dominant tolled road option as these express the dominance structure in the experiment better than the values $(-0.2, -1)$.

The multivariate normal distributions from the previous section are used as prior distribution to obtain Bayesian \mathcal{D}_s -efficient designs for the nested logit model. For the dissimilarity parameter λ , a uniform prior $\mathcal{U}[0, 1]$ is assumed. The designs consist of eight choice sets and are given in Appendix C. Three hundred random starting designs were updated making use of 1000 random draws from the prior distributions with the modified Fedorov algorithm. The fraction of choice sets in the designs with a dominant tolled road alternative are reported in Table 8, the designs are denoted as Design 1 and Design 2 for prior means $(-0.2, -1)$ and $(-0.4, -2)$ respectively. Again it can be concluded that a proper choice of prior information positively influences the applicability of the efficient design as choice sets with a dominant alternative are excluded.

Finally, we explore the robustness of the designs' efficiency in case of miss-

Table 8: Fraction of choice sets with a dominant tolled road option in the Bayesian \mathcal{D}_s -efficient designs

	<i>Design 1</i> (-0.2, -1)	<i>Design 2</i> (-0.4, -2)
<i>Fraction of sets in the design with a dominant tolled road option</i>	4/8	0/8

Table 9: Local \mathcal{D}_s -errors of the Bayesian efficient designs for the nested logit model

(β_{TT}, β_{TC})	<i>Design 1</i>	<i>Design 2</i>
(-0.2, -1)	0.0587	0.0725

specified priors. Table 9 gives the local \mathcal{D}_s -errors of Design 1 and Design 2 for fixed parameter values $(-0.2, -1)$. Figure 2 plots the ratios of the local \mathcal{D}_s -errors for Design 2 relative to Design 1 for different settings of the parameters β_{TT} and β_{TC} . Both in Table 9 and Figure 2 the errors are averaged over the values 0, 0.25, 0.5, 0.75 and 1 for dissimilarity parameter λ .

Similar conclusions as for the case with only two tolled road options in each choice set of the experiment can be drawn. Although Design 2, optimized for a prior distribution centered around $(-0.4, -2)$, is less efficient than Design 1 with respect to the coefficients near $(-0.2, -1)$, the efficiency loss is restricted. Furthermore, this design does not include choice sets with a dominant alternative, being more useful and informative as a survey design for analyzing travelers' choice behavior and willingness to pay for more speedy road alternatives.

5 Conclusion

The use of statistically efficient designs is well established in discrete choice analysis. Yet, additional to the choice of some optimality criterion, prior information about the partworths in the model should also be adequately chosen in order to exclude irrelevant choice sets with a dominant alternative from the efficient designs.

This research focussed mainly on a specific transportation problem in which policy makers want to know more about 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, two different experimental setups were discussed. First choice sets including only two tolled road options were considered, next choice data from sets with three alternatives, i.e. two tolled road options

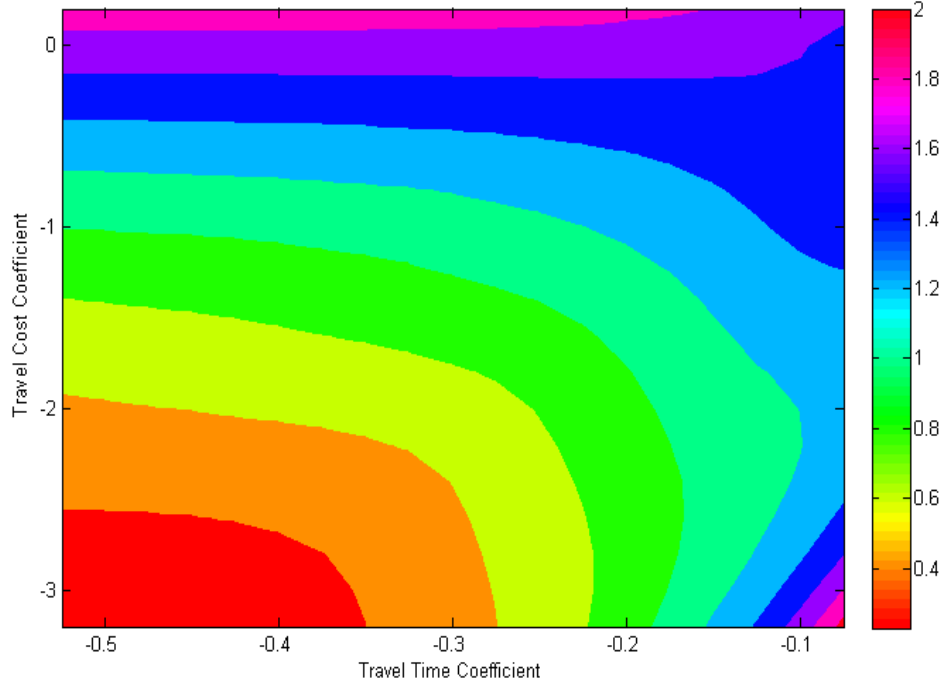


Figure 2: Ratio of local \mathcal{D}_s -errors for Design 2 relative to Design 1 for different values of β_{TT} and β_{TC}

and an additional toll-free alternative, was analyzed. The aim of the study was to find Bayesian \mathcal{D} -efficient designs not including choice sets with a dominant tolled alternative for the conditional and the nested logit model, modeling respectively choice data from the first and the second setup.

It is shown that when the dominance structure in the experiment is appropriately integrated in the prior information, the \mathcal{D} -efficient designs do not take up choice sets with a dominant alternative. Regardless the covariance structure in the prior distribution, the coding of the attribute levels and the experimental setup, the solution to dominance in efficient design appeared to be a proper choice of prior information. Moreover, the efficiency of the obtained designs was quite robust with respect to miss-specification of the prior.

Although this research applies to the conditional and the nested logit model, we are aware that nowadays more advanced models, such as the mixed logit model (Hess et al. 2005) accounting for preference heterogeneity in the population, are often used in discrete choice analysis. We expect choice sets with a dominant alternative to also pop up in efficient designs for the mixed logit model in case prior values are irrationally chosen. Further we believe that adjusting

the prior in a similar way such that it adequately incorporates the dominance of preferred attribute levels should solve the problem. Nevertheless, we applied more straightforward choice models in this research to illustrate the dominance problem because efficient and computationally fast algorithms for constructing Bayesian efficient designs for the mixed logit model are still lacking (Bliemer and Rose 2010; Yu et al. 2011).

This research advocates the use of adequate prior information to eliminate choice sets with a dominant alternative from \mathcal{D} -efficient designs, instead of computing efficient designs with respect to improper prior information and deleting useless choice sets from these designs afterwards. It is a very strong assumption that prior values correspond to the true parameters in the population, especially when they reflect irrational choice behavior. One should be willing to adjust and improve priors obtained from a pilot study or previous experiments in case they inadequately express preferences. If it is integrated in the prior information that respondents do not choose inferior alternatives, this will be reflected in the design.

Table 10: Bayesian \mathcal{D} -efficient design for prior means $(-0.2, -1)$

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 11: Bayesian \mathcal{D} -efficient design for prior means $(-0.4, -2)$

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

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

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

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

For choice set s , the information matrix is computed as

$$\mathbf{I}_s = \mathbb{E} \left[\begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TT}} \\ \frac{\partial l_s}{\partial \beta_{TC}} \\ \frac{\partial l_s}{\partial \lambda} \end{pmatrix} \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TT}} \\ \frac{\partial l_s}{\partial \beta_{TC}} \\ \frac{\partial l_s}{\partial \lambda} \end{pmatrix}' \right] = \mathbb{E} \begin{bmatrix} \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TT}} \end{pmatrix} \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TT}} \end{pmatrix}' & \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TT}} \end{pmatrix} \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TC}} \end{pmatrix}' & \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TT}} \end{pmatrix} \begin{pmatrix} \frac{\partial l_s}{\partial \lambda} \end{pmatrix}' \\ \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TC}} \end{pmatrix} \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TT}} \end{pmatrix}' & \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TC}} \end{pmatrix} \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TC}} \end{pmatrix}' & \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TC}} \end{pmatrix} \begin{pmatrix} \frac{\partial l_s}{\partial \lambda} \end{pmatrix}' \\ \begin{pmatrix} \frac{\partial l_s}{\partial \lambda} \end{pmatrix} \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TT}} \end{pmatrix}' & \begin{pmatrix} \frac{\partial l_s}{\partial \lambda} \end{pmatrix} \begin{pmatrix} \frac{\partial l_s}{\partial \beta_{TC}} \end{pmatrix}' & \begin{pmatrix} \frac{\partial l_s}{\partial \lambda} \end{pmatrix} \begin{pmatrix} \frac{\partial l_s}{\partial \lambda} \end{pmatrix}' \end{bmatrix}.$$

Because

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

and

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

we find that

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

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

We also find that

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

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

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

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

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,

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

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

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

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

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

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

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

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

Table 12: Bayesian \mathcal{D}_s -efficient design for prior means $(-0.2, -1)$

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 13: Bayesian \mathcal{D}_s -efficient design for prior means $(-0.4, -2)$

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

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

Acknowledgements

Research funded by ZKC1090 / DBOF/08/014 - DBOF project of the KULeuven

References

- Atkinson, A.C., Donev, A.N., Tobias, R.D., 2007. Optimum experimental designs, with SAS. Clarendon Press, Oxford U.K.
- 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.
- Bliemer, M.C.J., Rose, J.M., 2005. Efficiency and sample size requirements for stated choice studies. Working paper: ITLS-WP-05-08.
- Bliemer, M.C.J., Rose, J.M., 2010. Construction of experimental designs for mixed logit models allowing for correlation across choice observations. *Transportation Research Part B*, 44, 720-734.
- 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.
- Bunch, D.S., Louviere, J.J., Anderson, D.A., 1996. A comparison of experimental design strategies for multinomial logit models: the case of generic attributes. Working paper, Graduate School of Management, University of California at Davis.
- Goos, P., 2002. The optimal design of blocked and split-plot experiments. Springer, New York.
- 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.
- 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.
- 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.

- 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.
- 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.
- Kessels, R., Goos, P., Vandebroek, M., 2006a. A comparison of criteria to design efficient choice experiments. *Journal of Marketing Research*, 43, 409-419.
- 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.
- Rose, J.M., Bliemer, M.C.J., 2009. Constructing efficient stated choice experimental designs. *Transport Reviews*, 29, 587-617.
- Sándor, Z., Wedel, M., 2001. Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, 38, 430-444.
- Scarpa, R., Rose, J.M., 2008. Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why. *The Australian Journal of Agricultural and Resource Economics*, 52, 253-282.
- Swait, J., Louviere, J., 1993. The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research*, 30, 305-314.
- Train, K., 2003. *Discrete choice methods with simulation*. Cambridge University Press, New York.
- Vermeulen, B., Goos, P., Vandebroek, M., 2008. Models and optimal designs for conjoint choice experiments including a no-choice option. *International Journal of Research in Marketing*, 25, 94-103.
- Yu, J., Goos, P., Vandebroek, M., 2011. Individually adapted sequential Bayesian conjoint-choice designs in the presence of consumer heterogeneity. *International Journal of Research in Marketing*. In press.