WORKING PAPER

ITLS-WP-15-16

Detecting dominancy and accounting for scale differences when using stated choice data to estimate logit models

By
Michiel C.J. Bliemer, John M. Rose[2] and Casper G. Chorus[3]

[2] Institute for Choice, University of South Australia
[3] Faculty of Technology, Policy and Management, Delft University of Technology

August 2015

ISSN 1832-570X

INSTITUTE of TRANSPORT and LOGISTICS STUDIES
The Australian Key Centre in
Transport and Logistics Management

The University of Sydney
Established under the Australian Research Council's Key Centre Program.

NUMBER: Working Paper ITLS-WP-15-16

TITLE: Detecting dominancy and accounting for scale differences when using stated choice data to estimate logit models

ABSTRACT: Stated choice surveys have been used for several decades to estimate preferences of agents using choice models, and are widely applied in the transportation domain. Typically orthogonal or efficient experimental designs underlie such surveys. These experimental designs may suffer from choice tasks containing a dominant alternative, which we show is problematic because it affects scale and therefore may bias parameter estimates. We propose a new measure based on minimum regret to calculate dominancy and automatically detect such choice tasks in an experimental design. This measure is then used to define a new experimental design type that ensures trade-offs within the design. Finally, we propose a new regret-scaled multinomial logit model that takes the level of dominancy within a choice task into account. Results using simulated and empirical data show that the presence of dominant alternatives can bias model estimates, but that making scale a function of a smooth approximation of normalised minimum regret can properly account for scale differences without the need to remove choice tasks with dominant alternatives from the dataset.

KEY WORDS: ***random utility, dominant alternatives, stated choice experiments, discrete choice, regret***

AUTHORS: **Bliemer, Rose and Chorus**

CONTACT: INSTITUTE OF TRANSPORT AND LOGISTICS STUDIES (H73)

The Australian Key Centre in Transport and Logistics Management

The University of Sydney NSW 2006 Australia

Telephone: +612 9114 1824

E-mail: business.itlsinfo@sydney.edu.au

Internet: http://sydney.edu.au/business/itls

DATE: August 2015

1. Introduction

Discrete choice models based on utility theory are widely used to analyse behaviour and preferences of agents (e.g., travellers) in order to estimate willingness-to-pay measures (e.g., value of travel time savings) and to predict market shares (e.g., mode shares). Stated choice surveys are often used to collect data for estimating the coefficients that describe behaviour. For example, in the transportation field, there are many such surveys for investigating behaviour in mode choice (e.g., De Luca and Di Pace, 2015; Bekhor and Shiftan, 2010) and route choice (e.g., Chorus, 2014; Hensher, 2006), but also parking choice (Axhausen and Polak, 1991) and vehicle type and fuel choice (Hess et al., 2012a). For an extensive review of stated preference studies in the transportation literature, see Bliemer and Rose (2011). In this paper we look at the impact of having dominant alternatives in a stated choice survey. We show that this can lead to biased parameter estimates, and we provide different ways of dealing with this problem.

1.1 Stages in stated choice studies

Studying travel behaviour using stated choice surveys typically follows four stages as shown in Figure 1, namely (i) survey design, (ii) data collection, (iii) data cleaning, and (iv) model estimation. In the first stage the analyst typically conducts a literature review, performs qualitative research in the form of focus groups, in-depth interviews and/or pilot studies to determine what alternatives, attributes, attribute levels and how many choice tasks are appropriate for the study. Once these survey dimensions have been determined, the analyst generates an experimental design that provides the attribute level combinations across alternatives in each choice task shown to the respondents, see e.g. Bliemer and Rose (2009) or Rose and Bliemer (2009). This experimental design is then translated into a pen and paper, web based, computer-assisted personal interviewing (CAPI) or other survey instrument. In the second stage, a sample of respondents is invited to complete the survey. Each respondent faces one or more choice tasks and the analyst captures their choices. In the third stage, the analyst combines the choices together with the attribute levels from the experimental design into a single dataset. The analyst may decide to remove problematic data by eliminating specific choice tasks and choice observations or in extreme cases even all observations obtained from a respondent in order to create a ‘clean’ dataset. Finally, in the last stage the analyst selects one or multiple model types in order to estimate behavioural coefficients.

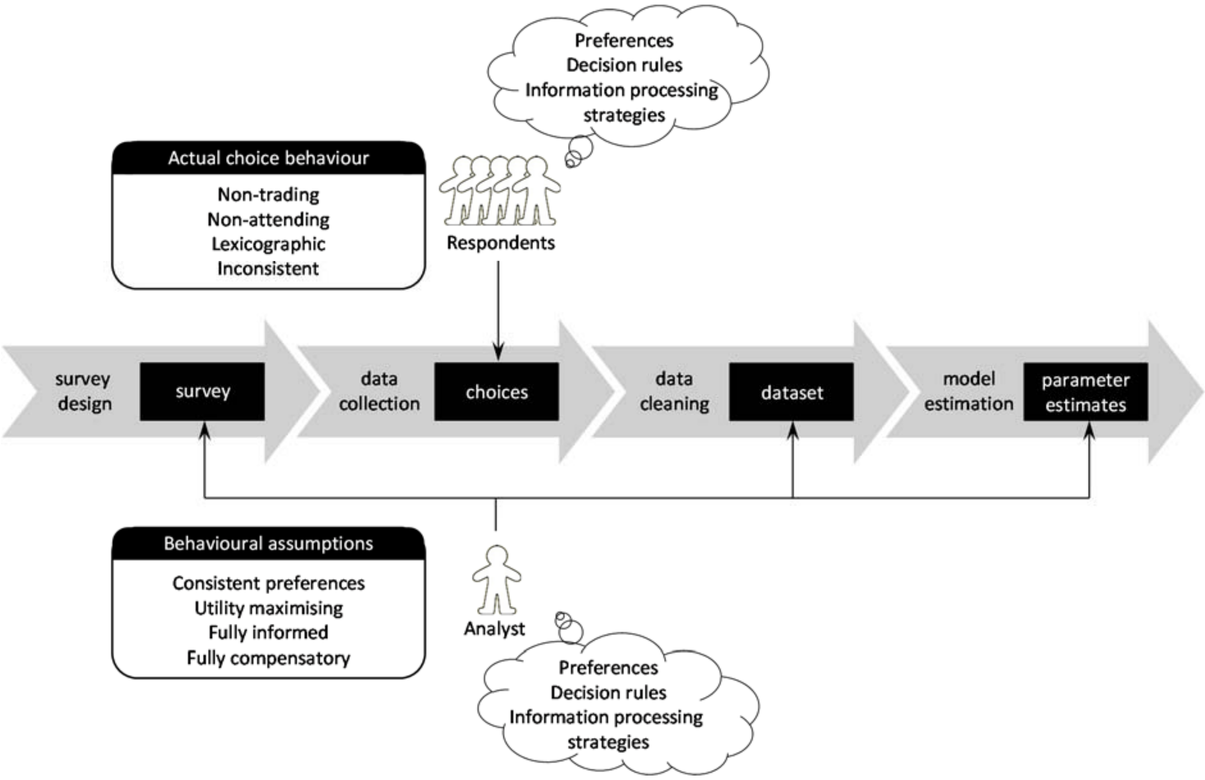


Figure 1 - Travel behaviour analysis and possible discrepancy between respondents and analyst

1.2 Discrepancies between actual behaviour and model assumptions

While this process is relatively straightforward, there is a likely discrepancy between the preferences, decision rules, and information processing strategies that the respondents make use of and the ones considered by the analyst in the choice model. In other words, the model may not be able to correctly describe the actual behaviour, such that actual preferences may be inconsistent with the parameter estimates. By adopting a (deterministic or random) utility theoretic framework, the analyst typically assumes that (i) respondents have consistent preferences that do not change during the survey, (ii) respondents are maximising utility, and (iii) respondents process all information, and (iv) each respondent make decisions based on a fully compensatory decision process in which trade-offs are made between all attributes that are relevant to him or her across all alternatives.[1]

In practice, respondents often do not behave in such a fully rational way, for example due to learning or fatigue effects, complexity of the choice task at hand, or by minimising the effort of decision making by adopting simplifying rules that do not involve some form of trading behaviour or do not maximise utility. Several types of behavioural rules are often mentioned, such has non-trading behaviour, lexicographic behaviour, attribute non-attendance, and inconsistent behaviour. Such behaviour is not compatible with the analyst’s assumption of utility maximisation and fully compensatory behaviour, and hence may lead to biased parameter estimates in the choice model. Many researchers are usually aware of such discrepancies but typically ignore it, whilst others remove ‘problematic’ choices in the

[1] An attribute is considered relevant for a certain respondent if it is of importance in the decision making and hence makes a contribution to his or her utility. Irrelevant attributes are assumed to be excluded from the model (for that respondent). Each respondent can have a different set of relevant attributes.

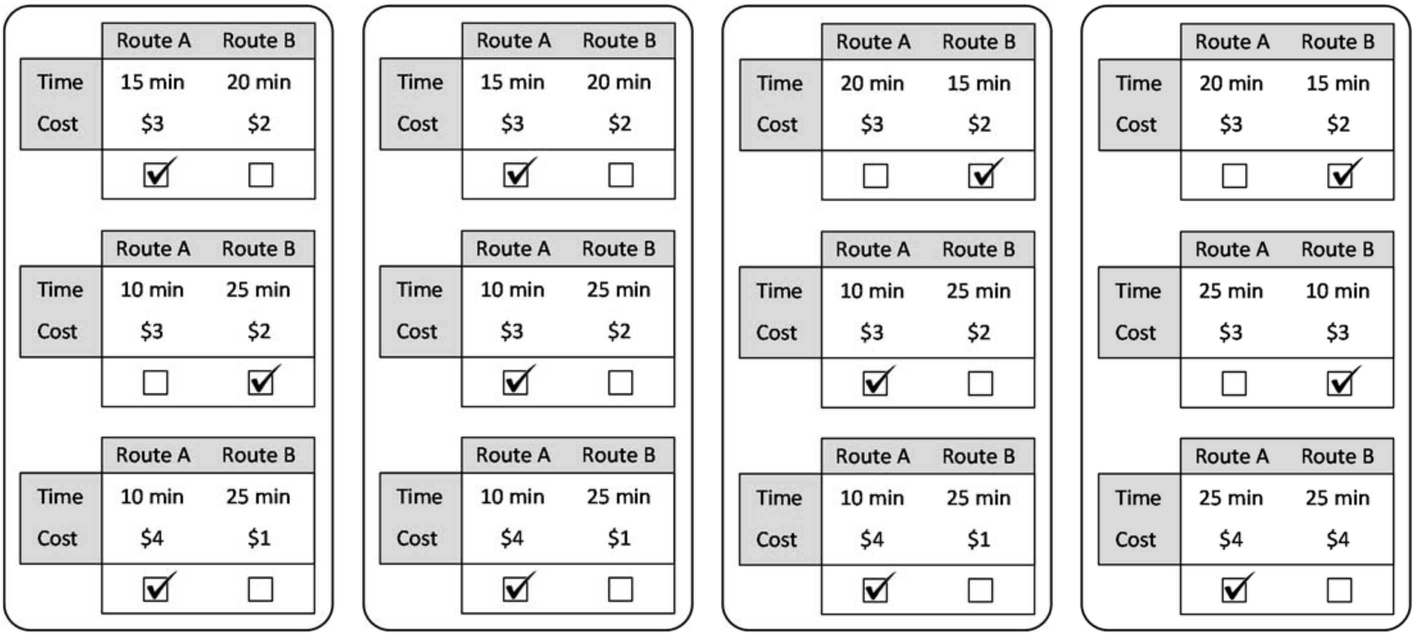
data cleaning stage or adopt more sophisticated choice models that explicitly account for alternative decision rules and information processing strategies.

Inconsistent behaviour may occur due to learning or fatigue effects, such that choices in the beginning of the survey seem to portray different preferences than in latter choice tasks, or due to choice task complexity (Sælesminde, 2001, 2002). Inconsistencies can be detected through different measures, see e.g. Sælesminde (2001) and Rezaei and Patterson (2015). Inconsistent choices are often present in data obtained through stated choice surveys and removing inconsistent respondents has a significant impact on the parameter estimates (Hess et al., 2010; Rezaei and Patterson, 2015; Rose et al., 2013; Sælesminde, 2001, 2002). Scaling approaches can be adopted to account for fatigue effects and/or choice complexity in model estimation (Bradley and Daly, 1994; DeShazo and Fermo, 2002; Swait and Adamowicz, 2001). Danthurebandara et al. (2011) proposes to take choice complexity already into account when generating the experimental design for the survey.

Respondent may also adopt strategies in order to minimise the effort and concentrate only on a subset of alternatives and/or attributes that are considered most important. This may lead to non-attendance of alternatives or attribute non-attendance, which may bias parameter estimates (Hess et al., 2013; Rose et al, 2013). Several approaches have been approached in the literature to account for attribute non-attendance in particular (Hensher et al., 2005; Hess and Rose, 2007; Collins, 2012; Collins et al., 2013).

Further, respondents may adopt different decision rules, e.g. non-trading or lexicographic rules, elimination by aspects, or hierarchical non-compensatory rules (Hess et al., 2010; Hess et al., 2012b; Rose et al., 2013; Sælesminde, 2006; Scott, 2002; Van de Kaa, 2006).

Examples of such behaviour in stated choice studies is shown in Figure 2. Assume a simple choice model in which marginal utilities of time and cost are negative. Choices in Figure 2(a) illustrate inconsistent behaviour (changing attractiveness of alternatives leads to a switch in choice towards the alternatives that becomes less attractive), choices in Figure 2(b) illustrate non-trading behaviour (always choosing the first alternative), and choices in Figure 2(c) illustrate non-attendance of attributes (always choosing the alternative with the lowest travel time).



(a) (b) (c) (d)

Figure 2 – Deviations from model assumptions due to (a) inconsistent preferences, (b) non-attendance of alternatives or non-trading, (c) non-attendance of attributes or lexicographic rules, (d) dominancy of alternatives

In this study we focus on another type of discrepancy between observed behaviour and modelled behaviour as a result of the presence of dominant alternatives in the survey. Dominant alternatives may trigger actual behaviour that is inconsistent with the analysts' model of behaviour, non-trading behaviour in particular. An alternative in a choice task is called dominant if it is better than (or equal to) another alternative in a pairwise comparison of the attribute level. Under the same assumptions as before, in Figure 2(d) it is expected that all respondents choose Route B, since this alternative strictly dominates Route A on both attributes. In the second choice task, Route A only strictly dominates Route B on one attribute (time), but again there is no trade-off possible. In the third choice task, neither route dominates, but still there is no trade-off being made by respondents, and as such does not provide any information to the analyst.

1.3 Paper contributions and outline

Whilst the presence of dominant alternatives in the dataset can lead to significantly biased parameter estimates in model estimation, there is surprisingly little research on the impacts and how best to resolve the problem. The most common way of dealing with this problem is simply removing choice tasks with dominant alternatives during the survey design stage. In many cases this is a manual exercise in which the analyst reviews the choice tasks and eliminates the ones with dominant alternatives. Such choice tasks may easily be overlooked at the design stage and end up in the survey used during data collection. Then the analyst has the choice to remove them from the dataset during the data cleaning stage, or keep them in and account for them in the choice model.

The contributions of our study are as follows. First, we define a dominancy measure based on regret that expresses whether a choice task contains a dominant alternative. Secondly, we propose a new design methodology that automatically detects problematic choice tasks by embedding our dominancy measure and generates an experimental design without problematic choice tasks. Thirdly, we study the impact of the presence of dominant alternatives in the dataset on parameter estimates in a simple multinomial logit context and show that mainly scale is affected. Finally, we propose a novel discrete choice model that aims to correct for the presence of dominant alternatives by automatically adjusting the scale for each choice task based on a smooth approximation of our newly proposed dominancy measure.

The paper is structured as follows. In Section 2 we provide a brief literature review on dominant alternatives in stated choice studies and show that such alternatives easily occur in experimental designs. In Section 3 we describe a measure that can be used to detect dominant alternatives in an experimental design or dataset. Further, we describe a new efficient experimental design methodology that we term D^*-efficiency, which ensures that the design does not contain dominant alternatives. Section 4 proposes a novel regret-scaled multinomial logit model aiming to correct for scale differences due to problematic choice tasks. Section 5 describes four experimental designs with varying numbers of problematic choice tasks in a simple route choice case study, and use these in order to simulate choice observations as well as to collect empirical data in a real-world survey. Section 6 describes the simulation results, while Section 7 discusses outcomes from the empirical dataset. Section 8 concludes with a discussion, recommendations, and limitations of this study.

2. Dominancy in stated choice studies

2.1 Literature review

First of all, it is important to note that dominancy is only defined for unlabelled experiments, i.e. surveys in which all alternatives are generic with the same attributes and in which the model does not contain an alternative specific constant. According to the literature review in Bliemer and Rose (2011), most studies conducted in the transportation literature used unlabelled experiments; however there are many exceptions that use labelled experiments, in particular in mode choice studies.

Analysts often include dominant alternatives on purpose in order to determine whether a respondent pays attention to or understands the survey. For example, in the DATIV study in Denmark in 2004 (Burge and Rohr, 2004), nine choice tasks were generated for an unlabelled experiment with two alternatives and two attributes (travel time and travel cost) and choice task six contained on purpose a dominant alternative. If the respondent failed to choose the dominant alternative, all choices from this respondent were removed from the dataset. Also in a value of time study in the Netherlands such a dominant alternative was imposed in one of the choice tasks (Van de Kaa, 2006). Bradley and Daly (1994) collected data using a design in which the first choice task contained a dominant alternative as a lead-in into the survey, which also allows the interviewer to check whether the respondent has understood the choice task. In their model estimations, they estimate a separate scale parameter for each choice task. The scale parameter of the first choice task with the dominant alternative is much larger compared to the other scale parameters. Foster and Mourato (2002) use dominant alternatives to test for consistency of responses. Also Johnson and Mathews (2001) and many others include a dominant alternative in the survey to test for consistency. It could be argued that including a dominant alternative in a survey could be problematic and may actually lead to inconsistency in subsequent choice tasks, since the respondent may no longer take the survey seriously. Therefore, putting a choice task with a dominant alternative at the end of the survey is possibly better than at the beginning of the survey.

Hensher et al. (1988) states that dominant alternatives often occur when generating experimental designs. For example, the random, orthogonal, and efficient designs generated in Walker et al. (2015) include several dominant alternatives. Crabbe and Vandebroek (2012) propose adjusting prior information in order to significantly reduce the likelihood of generating dominant alternatives in Bayesian D-efficient designs (although they cannot be avoided completely). Altering prior information may reduce the occurrence of dominant alternatives, but is not desirable since the analyst is artificially changing assumptions on preferences. Huber and Zwerina (1996) propose a utility balancing approach that limits (but not necessarily prevents) the number of dominant alternatives. However, one has to be careful since a perfectly utility balanced design may lead to efficiency losses.

Dominance is not easily modelled by choice models based on random utility theory (Huber et al., 1982). In most cases, the analyst will remove choice tasks with dominant alternatives from the dataset before model estimation, since no information is obtained from choices of dominant alternatives (Hensher et al., 1988). As discussed above, experimental designs may include dominant alternatives (by accident or on purpose). Therefore, there is often the need to check for them in the survey design stage as well as the data cleaning stage.

In practice it is often not too difficult to manually detect and remove problematic choice tasks in a survey during the survey design stage. However, as choice tasks become more complex, it may become more difficult for the analyst to detect them, so experimental designs may need to be computer generated with appropriate dominancy constraints on choice tasks in place. In practice, most analysts post-process the experimental designs and remove choice tasks with dominant alternatives. This leads in almost all cases to a loss of orthogonality and attribute level balance of the design. It is therefore desirable to make the dominancy check an integral part of the experimental design methodology.

Next we illustrate that experimental designs for surveys with unlabelled alternatives are very likely to contain dominant alternatives without proper attention.

2.2 Likelihood of dominant alternatives

Consider a stated choice survey with *M* unlabelled alternatives described by *A* attributes. Furthermore, suppose that each attribute has *L* levels. A full factorial L^{MA} experimental design contains all possible choice tasks described by combinations of attribute levels, although some of these choice tasks will essentially be the same by simply re-arranging the alternatives (because they are unlabelled).

Making assumptions on the respondents' preferences, the analyst can determine dominant alternatives. The fewer attributes are present in the survey, the higher the likelihood that dominance will occur. Also, having more alternatives and fewer attribute levels increases the chance of dominant alternatives. Table 1 illustrates the fraction of choice tasks without a dominant alternative in a full factorial L^{MA} design. For example, 89.3 per cent of a $3^{3\times 3}$ full factorial design has a dominant alternative, leaving only 350 unique choice tasks[2] without a dominant alternative. Suppose that we would like to create a design consisting of six choice tasks. In total $\binom{350}{6} = 350!/\left(6!(350-6)!\right) \approx 2.445 \cdot 10^{12}$ unique designs without dominant alternatives can be created. While this is a very large number of possible designs to choose from, since the probability of picking choice task without a dominant alternative is $1-89.3=10.7$ per cent, which means that the probability of randomly generating a design without any dominant alternatives is negligible (around 0.00015 per cent). For certain design dimensions it is not

[2] In total there are 2,100 choice tasks without a dominant alternative, but most of them are permutations of 350 unique choice tasks by re-ordering alternatives. The number of possible permutations for three alternatives is $3! = 6,$ such that the number of unique combinations is $2100/6 = 350.$

even possible to find choice tasks that do not contain a dominant alternative (e.g., a $2^{3\times 2}$ design, $2^{4\times 2}$ design, $2^{4\times 3}$ design, and $3^{4\times 2}$ design). The non-dominated choice tasks corresponding to the shaded design dimensions in Table 1 are shown in Table 2, in which we use two or three route alternatives with two (travel time and toll cost) or three attributes (travel time, fuel cost, toll cost), and two or three levels (10, 15, and 20 minutes travel time and $1, $2, or $3 costs, where the middle level is omitted in case of two levels only). Clearly, requiring non-dominance is often a very strict constraint on the experimental design.

Table 1 – Dominancy in choice tasks in an L^{MA} ***design***

L	*M*	*A*	Dominant (%)	Unique non-dominant tasks
2	2	2	87.5	1
2	2	3	71.9	9
2	2	4	57.0	55
2	3	2	100	0
2	3	3	97.7	2
2	3	4	90.6	64
2	4	2	100.0	0
2	4	3	100.0	0
2	4	4	99.1	25
3	2	2	77.8	9
3	2	3	55.6	162
3	2	4	38.3	2,025
3	3	2	99.2	1
3	3	3	89.3	350
3	3	4	72.9	24,025
3	4	2	100.0	0
3	4	3	98.6	310
3	4	4	91.1	159,300
4	2	2	71.9	36
4	2	3	47.3	1080
4	2	4	30.1	22,896
4	3	2	97.7	16
4	3	3	82.2	7,760
4	3	4	61.8	1,069,056
4	4	2	99.9	1
4	4	3	95.9	28,355
4	4	4	*	*

* Not calculated since $4^{4\times 4} \approx 4,295,000,000$ choice tasks

Table 2 – Choice tasks without dominant alternatives in several L^{MA} designs

	Route A			Route B			Route C		
Task	Time	Fuel	Toll	Time	Fuel	Toll	Time	Fuel	Toll
Unique non-dominant choice tasks in a $2^{2\times2}$ full factorial design									
1	10	–	3	20	–	1	–	–	–
Unique non-dominant choice tasks in a $2^{2\times3}$ full factorial design									
1	10	1	2	10	3	1	–	–	–
2	10	1	2	20	1	1	–	–	–
3	10	1	2	20	3	1	–	–	–
4	10	3	3	20	1	1	–	–	–
5	10	3	3	20	1	3	–	–	–
6	10	3	3	20	1	1	–	–	–
7	10	3	3	20	1	3	–	–	–
8	10	3	3	20	3	1	–	–	–
9	20	1	3	20	3	1	–	–	–
Unique non-dominant choice tasks in a $2^{3\times3}$ full factorial design									
1	10	3	3	20	1	3	20	3	1
2	10	1	3	10	3	1	20	1	1
Unique non-dominant choice tasks in a $3^{2\times2}$ full factorial design									
1	15	–	3	20	–	2	–	–	–
2	10	–	3	20	–	2	–	–	–
3	15	–	3	20	–	1	–	–	–
4	15	–	2	20	–	1	–	–	–
5	10	–	3	20	–	1	–	–	–
6	10	–	2	20	–	1	–	–	–
7	10	–	3	15	–	2	–	–	–
8	10	–	3	15	–	1	–	–	–
9	10	–	2	15	–	1	–	–	–
Unique non-dominant choice tasks in a $3^{3\times2}$ full factorial design									
1	10	–	3	15	–	2	20	–	1

3. Detecting dominant alternatives in experimental designs

In this section a simple measure is proposed that together with assumptions on preferences can be used to assess whether a choice task contains a dominant alternative, and how to generate experimental designs without dominancy.

3.1 Dominancy measure

Consider a choice model based on utility theory with systematic utilities V_{nsj} for each respondent $n \in \{1,\ldots,N\}$, alternative $j \in \{1,\ldots,J\}$ and each choice task $s \in \{1,\ldots,S\}$. Assume that each alternative has attributes indexed by $k \in \{1,\ldots,K\}$. We further assume that the systematic utilities are given by a linear or nonlinear utility function $V_{nsj}(\mathbf{x}_{nsj} \mid \boldsymbol{\beta}_n)$ in which for each respondent *n* and each choice task *s*, alternative *j* is represented by a set of attribute levels given by a $K \times 1$ vector $\mathbf{x}_{nsj} = [x_{nsjk}]_{k=1,\ldots,K}$ called a profile. Preferences of respondent *n* are given by a $K \times 1$ vector of partworths, $\boldsymbol{\beta}_n = [\beta_{nk}]_{k=1,\ldots,K}$. In case of a linear in the parameters and linear in the attributes (LPLA) utility function we can determine the utilities as $V_{nsj}(\mathbf{x}_{nsj} \mid \boldsymbol{\beta}_n) = \boldsymbol{\beta}'_n \mathbf{x}_{nsj}$, where the prime indicates the transpose. For each respondent *n*, choice task *s* is defined by the $1 \times JK$ vector consisting of profiles, $\mathbf{x}_{ns} = [x'_{ns1},\ldots,x'_{nsK}]$. As shown in Table 2, an experimental design for respondent *n* is the collection of choice tasks given by $S \times JK$ matrix $\mathbf{x}_n = [\mathbf{x}_{n1},\ldots,\mathbf{x}_{nS}]'$, where each row represents a choice task. Note that in many cases design $\mathbf{x}_n$ will be the same for all respondents, but in some cases these levels may vary. For example in a pivot design the attribute levels are based on respondent specific reference levels (Rose et al., 2008).

We define dominance of an alternative as follows. An alternative *j* is said to dominate alternative *i* for respondent *n* in choice task *s* if for each attribute *k* the utility of alternative *j* is larger than (or equal to) the utility that would be obtained if the level of that attribute in alternative *j* would be replaced by its level in alternative *i*, ceteris paribus (keeping all other attribute levels in alternative *j* the same). This is a fairly general definition of dominance that can also be applied to nonlinear utility functions.

We can formulate this definition in terms of profiles. Define $\Delta_{ns,j\leftarrow i,k}$ as the difference in utility between alternative *j* and alternative *i* in choice task *s* for respondent *n* by only comparing differences in attribute *k*. Consider alternatives *j* and *i* with profiles $\mathbf{x}_{nsj}$ and $\mathbf{x}_{nsi}$, respectively. Then $\Delta_{ns,j\leftarrow i,k}$ is defined as

$$\Delta_{ns,j\leftarrow i,k} = V_{nsj}(\mathbf{x}_{nsj} \mid \boldsymbol{\beta}_n) - V_{nsj}(\mathbf{x}_{nsj} + (x_{nsik} - x_{nsjk})\mathbf{1}_k \mid \boldsymbol{\beta}_n), \tag{1}$$

where $\mathbf{1}_k$ is a $K \times 1$ vector with zeros, except in row *k* where it takes on value 1. Profile $\mathbf{x}_{nsj} + (x_{nsik} - x_{nsjk})\mathbf{1}_k$ is identical to profile $\mathbf{x}_{nsj}$ except that the level for attribute *k* is replaced by the level of that attribute in profile $\mathbf{x}_{nsi}$.

An alternative *j* with profile $\mathbf{x}_{nsj}$ is said to dominate an alternative *i* with profile $\mathbf{x}_{nsi}$ if and only if

$$\Delta_{ns,j\leftarrow i,k} \geq 0, \quad \text{for all } k. \tag{2}$$

Alternative *j* would strictly dominate alternative *i* if the inequality sign in (1) would be strict for at least one attribute *k*. If the profiles of *i* and *j* are identical, then $\Delta_{ns,j\leftarrow i,k} = 0$ for all *k* and inequalities (2) hold by definition. In this case, neither alternative strictly dominates the other.

Utility functions can be linear and nonlinear in the attributes. Consider an LPLA utility function. Then $\Delta_{ns,j\leftarrow i,k}$ simplifies to

$$\begin{aligned}\Delta_{ns,j\leftarrow i,k} &= \boldsymbol{\beta}_n'\mathbf{x}_{nsj} - \boldsymbol{\beta}_n'(\mathbf{x}_{nsj} + (x_{nsik} - x_{nsjk})\mathbf{1}_k) \\ &= \beta_{nk}(x_{nsjk} - x_{nsik}).\end{aligned} \tag{3}$$

In order to determine whether $\Delta_{ns,j\leftarrow i,k} \geq 0$ the analyst only needs to look at the difference in the levels of attribute *k* between alternatives *i* and *j*, and only needs to know the sign of β_{nk}. The exact value of the partworth is not important. If the analyst thinks that an attribute has a negative impact on utility (e.g., travel time, toll cost), then the analyst can simply use $\beta_{nk} = -1$, while for attributes with a positive impact on utility (e.g., in-flight entertainment, on-board wifi) one can use $\beta_{nk} = +1$. This is useful for removing dominant alternatives at the survey design stage.

While utility functions are in most cases linear in the parameters, sometimes they are nonlinear in the attributes, for example $V_{nsj} = \beta_{n1}x_{nsj1} + \beta_{n2}\log(x_{nsj2}) + \beta_{n3}x_{nsj1}x_{nsj2}$, which includes an interaction effect, resulting in $\Delta_{ns,j\leftarrow i,1} = (\beta_{n1} + \beta_{n3}x_{nsj2})(x_{nsj1} - x_{nsi1})$ and $\Delta_{ns,j\leftarrow i,2} = \beta_{n2}\log(x_{nsj2}/x_{nsi2}) + \beta_{n3}x_{nsj1}(x_{nsj2} - x_{nsi2})$. If the partworths are all positive (without knowing the exact value), then $\Delta_{ns,j\leftarrow i,k} \geq 0$ if $x_{nsjk} \geq x_{nsik}$ for both attributes. Similarly, if all partworths are negative, then $\Delta_{ns,j\leftarrow i,k} \geq 0$ if $x_{nsjk} \leq x_{nsik}$. If β_{n1} and β_{n2} are both positive (negative) and β_{n3} is negative (positive), then determining the sign of $\Delta_{ns,j\leftarrow i,k}$ relies on more exact knowledge regarding the values of these partworths.

An alternative *j* is said to be dominant in choice task *s* for respondent *n* if and only if

$$\Delta_{ns,j\leftarrow i,k} \geq 0, \quad \text{for all } k\text{, for all } i \neq j. \tag{4}$$

Alternative *j* is strictly dominant if the inequality in (2) strictly holds for at least one attribute *k* and one other alternative *i*. The conditions in (1) can be combined into the following measure:

$$R_{nsj} = \sum_{i\neq j}\sum_{k=1}^{K}\max\left\{0, -\Delta_{ns,j\leftarrow i,k}\right\} = 0. \tag{5}$$

Value R_{nsj} can be seen as the regret that respondent *n* attaches to selecting alternative *j* over all other alternatives in choice task *s*. More specifically, we use conceptualization of regret as proposed in the context of the random regret minimization (RRM) model (Chorus et al., 2008; Chorus, 2010).[3] If this regret is zero, then alternative *j* is better than (or equally good as) alternative *i* in a pairwise comparison between all attributes. Therefore, in order for the respondent to make trade-offs, R_{nsj} need to be strictly

[3] This conceptualization of regret differs from the one proposed in classical regret based models such as Regret Theory (Loomes and Sugden, 1982); these conventional theories postulate that regret is a function of the relative *utilities of alternatives* and can only exist in the context of uncertainty. In contrast, RRM postulates that regret is a function of the relative *values of attributes* and that it arises – also in the absence of uncertainty – when the decision maker has to put up with a relatively poor performance on one or more attributes to arrive at a relatively strong performance on other attributes.

positive, in which a respondent always feels some regret choosing an alternative. If $R_{nsj} = 0$ for all alternatives *j*, then all their profiles are identical.

Define also for each respondent *n* and each choice task *s* the minimum regret per choice task, R_{ns}, and the minimum regret per design, R_n,

$$R_{ns} = \min_j \{R_{nsj}\}, \tag{6}$$

$$R_n = \min_s \{R_{ns}\}. \tag{7}$$

If $R_{ns} = 0$, then choice task *s* contains a strictly dominant alternative or it contains identical alternatives. In both cases no trade-offs between attributes need to be made, hence we typically would like to avoid such choice tasks in the dataset (since they provide no information). Hence, in this section, we make no distinction between the two cases. However, as will be discussed in Section 4, the two cases have a very different impact on scale in estimation. If $R_n = 0$, then the experimental design contains at least one dominant alternative.

While a deterministic model postulates that rational respondents always select a dominant alternative, this may not be the case in a stochastic model with random utilities. In this case, each respondent is assumed to maximise random utility given by $U_{njs} = V_{njs} + \varepsilon_{nsj}$, where ε_{nsj} is a random unobserved error term following a certain probability distribution. Even in a random utility framework it may be difficult to explain why a respondent would choose a dominated alternative in an unlabelled choice experiment. Assuming that the analyst has a correct understanding of the partworths of the respondent, it may be that the error is confounded with one of the attributes, i.e. the respondent may relate a small travel time to a trip on a motorway (a characteristic not included as an attribute in the model and therefore assumed to be in the error term), and may not like driving on motorways. Then even if the respondent has a negative marginal utility for travel time, he or she may still choose the alternative with a higher travel time. Hence, the error will never be exactly equal to zero, but will likely be close to zero. Understanding that dominance is related to the error term, which in turn is related to the scale parameter in a logit model, is the starting point for scaling each choice task with respect to regret in Section 4.

3.2 Efficient experimental designs without dominant alternatives

None of the existing experimental design techniques rule out the existence of dominant alternatives in one or more choice tasks. We therefore propose a constrained experimental design method that automatically checks for dominancy (i.e., strictly dominant alternatives and identical alternatives that are not strictly dominated) within the design.

Assume that X_n denotes the set of all possible experimental designs for respondent *n* that satisfy the analysts design dimensions and possibly attribute level balance and orthogonality. A D-optimal design is a matrix with attribute levels $\mathbf{x}_n \in X_n$ that minimises the determinant of the asymptotic variance-covariance (AVC) matrix under the assumption of a vector of prior partworths $\boldsymbol{\beta}_n$ (Huber and Zwerina, 1996). Such a design maximises the (Fisher) information obtained from the choice tasks. Since in most cases one cannot guarantee to have found the optimal design (as this would require evaluating all possible designs), these designs are often referred to as D-efficient instead of D-optimal. Besides minimising the determinant, one can also minimise the trace of the AVC matrix (resulting in an A-efficient design), or minimise the maximum sample size required for statistically significant parameter estimates (resulting in an S-efficient design, see Rose and Bliemer, 2013).

We define a D*-optimal (efficient) design as a design that maximises the determinant of the Fisher information matrix under the restrictions that (i) the design contains no dominant alternatives as defined in (4) and (5), and (ii) does not contain choice task replications. Permutations of profiles in a choice task results in an identical choice task. This is not necessarily a problem, and sometimes such choice tasks are included in the design on purpose in order to assess consistent choice behaviour respondents (see Section 1.2). However, most analysts would prefer to include unique attribute level combinations in each choice task and avoid any replications.

To formulate mathematically, let $I(\mathbf{x}_n \mid \boldsymbol{\beta}_n)$ denote the Fisher information matrix that depends on the experimental design and prior partworths $\boldsymbol{\beta}_n$. These prior values are best guesses from the literature or a pilot study. Then the D*-optimal design for respondent *n* is the matrix $\mathbf{x}_n \in X_n$ that solves the following nonlinear programming problem:

$$\begin{aligned} &\max_{\mathbf{x}_n \in X_n} \left| I(\mathbf{x}_n \mid \boldsymbol{\beta}_n) \right| \\ \text{subject to: } &R_n(\mathbf{x}_n \mid \boldsymbol{\beta}_n) > 0, \\ &\mathbf{x}_n \text{ does not contain choice task replications,} \end{aligned} \tag{8}$$

where $|\cdot|$ denotes the matrix determinant. Note that such a design cannot be generated if $\boldsymbol{\beta}_n = \mathbf{0},$ i.e. if the analyst has no information regarding the partworths, not even the sign, since in that case by definition $R_n(\mathbf{x}_n \mid 0) = 0.$ So in order to generate a design without dominant alternatives, there needs to be at least some trade-offs between attributes. In case the analyst only knows the signs, one can set values close to zero for the priors, i.e. $\beta_{nk} = -0.001$ or $\beta_{nk} = 0.001.$ This enables computation of the minimum regret and these small deviations from zero will only have little effect on the Fisher information matrix.

Similarly, an A*-optimal, S*-optimal or other efficient designs can be defined, where the asterisk indicates that the design is dominancy constrained. Also other more advanced designs such as Bayesian D*-optimal designs can be defined by a direct extension of Bayesian D-optimal designs (Sándor and Wedel, 2001). Such Bayesian efficient designs are more robust against misspecification of prior partworths.

Traditional column based algorithms, i.e. relabelling and swapping techniques described in Huber and Zwerina (1996), modify columns in matrix $\mathbf{x}_n$ and will generally struggle generating designs without dominant alternatives. Since the dominancy constraint is on the entire choice task, a row based algorithm that modifies a row in matrix $\mathbf{x}_n$ will therefore be more useful. Federov (1972) proposed a row based algorithm for generating optimal designs, which was modified by Cook and Nachtsheim (1980). This modified Federov algorithm can be used to first construct a candidature set that consist of all (or a select of) choice tasks that do not contain dominant alternatives. For example, when generating a fractional factorial $3^{3\times 3}$ design, we first determine the 350 unique choice tasks without dominancy and without replications (see Table 1). Then we randomly select *S* choice tasks from this set to form a design, and keep replacing rows in the design with rows in the candidature set until the best design has been found. Note that the number of designs that can be created by selecting *S* tasks out of 350 is typically very large (see Section 2.2). Therefore the algorithm is usually terminated once the Fisher information no longer improves for a certain number of iterations.

We implemented a column based as well as a row based algorithm in Ngene version 1.1 (ChoiceMetrics, 2012), which that take the constraint $R_n > 0$ into account and avoids replications of the same choice task. We use these algorithms to generate the D*-optimal designs in this paper. While a row based algorithm can easily avoid dominant alternatives, it is more difficult to generate attribute level balanced designs. If attribute level balance is required, our algorithm selects new choice tasks from the candidature set such that attribute level balance is satisfied or only marginally violated.

4. Regret-scaled multinomial logit model

4.1 Choice task based scaling

Suppose that the analyst decides not to remove choice tasks with dominant alternatives from the dataset and wishes to estimate a simple discrete choice model assuming a rational decision maker *n* selecting alternative *j* that maximise the random utility $U_{nsj} = V_{nsj} + \varepsilon_{nsj}$, where ε_{nsj} is a random unobserved (by the analyst) component of the utility.

In order to account for differences in scale due to the presence of dominant alternatives, one can estimate a nested logit model with choice task specific scale parameters, such as in Bradley and Daly (1994). Scale parameters for choice tasks with a strictly dominant alternative will be large in contrast to other choice tasks. On the other hand, a choice task with identical alternatives will also not allow any trade-offs, but is expected to have a very small scale parameter, since the choice will be mostly based on the unobserved component (i.e., the user chooses more or less randomly since all alternatives are the same).

Bradley and Daly estimated 14 scale parameters on top of four regular coefficients in the utility function. One of the scale parameters (corresponding to a base choice task) needs to be set to one, and all other scales are relative to this base. Clearly, such a choice task specific scaling significantly increases the number of parameters to be estimated. In order to avoid having to estimate a separate parameter per choice task, we adopt a parametric approach in which we make scale a function of our dominancy measure introduced in Section 3.1.

Under the assumption that ε_{nsj} are independently and identically extreme value type I distributed with variance $\frac{1}{6}\pi^2\lambda_{ns}^{-2}$, we obtain an extension of the well-known multinomial logit (MNL) model (McFadden, 1974) in which this variance of the error term in choice task *s* is inversely related to the scale in the choice task, λ_{ns}. The probability of respondent *n* selecting alternative *j* in choice task *s* is then given by

$$P_{nsj} = \frac{\exp(\lambda_{ns} V_{nsj})}{\sum_i \exp(\lambda_{ns} V_{nsi})}. \tag{9}$$

In a (homoscedastic) MNL logit model, $\lambda_{ns} = 1$ for all *s*. In case of a strictly dominant alternative in choice task *s*, the variance of error ε_{nsj} typically diminishes, which corresponds to an increase in the scale parameter. Therefore, we will relate scale parameter λ_{ns} to minimum regret R_{ns} in order to make it heteroscedastic.

Choice tasks with dominant alternatives trigger actual behaviour which is at odds with the typical compensatory logit model. Not only do such choice tasks not allow for trade-offs, the actual behaviour is much less random in these choice tasks compared to other choice tasks. We provide two reasons for lower error variances (larger scale) in choice tasks with dominant alternatives. First, most respondents will find it easy to choose from such a task, leading to homogeneous patterns in the sample (compared to other choice tasks), which implies much less error/variation in actual behaviour. Secondly, although tastes for attributes vary across the population, pretty much every decision maker is expected to have the same direction of taste (e.g., negative valuation of costs). As such, in a choice task with a dominant alternative, taste heterogeneity does not influence actual behaviour.

There are two concerns in using minimum regret R_{ns} as a descriptor for λ_{ns}. First of all, R_{ns} is bounded from below by zero, but the upper bound depends on the attribute level ranges. For interpretability reasons we prefer an upper bound that does not rely on the levels, similar to the entropy upper bound of (independent of attribute level range) used in the model of Swait and Adamowicz (2001) for scaling choice tasks according to complexity. Secondly, R_{ns} as defined in Equation (6) is not 'smooth', since it involves minimum and maximum operators. This typically leads to numerical problems in model estimation, and it also does not discriminate between a choice task with a strictly dominant alternative (with a very high scale parameter) and a choice task with identical alternatives (with a very low scale parameter). We address these two concerns in the next subsections.

4.2 Normalised minimum regret

In order to address the first issue, we simply normalise the minimum regret by the average regret. Hence, our normalised minimum regret M_{ns} becomes

$$M_{ns} = \frac{R_{ns}}{J^{-1}\sum_j R_{njs}}. \tag{10}$$

Note that $R_{njs} \geq 0$ for all alternatives *j* such that $M_{ns} \geq 0.$ Suppose that choice task *s* contains a strictly dominant alternative for respondent *n*. This means that $R_{ns} = 0$ and there exists a dominated alternative *j* for which $R_{njs} > 0.$ As a result, a choice task with a strictly dominant alternatives yields $M_{ns} = 0.$ Now suppose that choice task *s* does not contain any dominant alternatives for respondent *n*, such that $R_{njs} > 0$ for all alternatives *j*. Since R_{ns} is the minimum over these values, R_{ns} can never be greater than $\frac{1}{J}\sum_j R_{njs}$, hence $M_{ns} \leq 1.$ The upper bound of $M_{ns} = 1$ is reached when all alternatives have the same positive regret R_{nsj}, making each alternative equally attractive. In summary, it holds that $M_{ns} \in [0,1].$

In the extreme case where the profiles of all alternatives are identical, i.e. $R_{njs} = 0$ for all alternatives *j*, the normalised minimum regret in (10) is undefined (zero divided by zero). Clearly such choice tasks should be prevented at all times, but as we will show in Section 4.3, the smooth approximation of the normalised minimum regret is defined in this extreme case and will not lead to numerical problems.

It is interesting to note that normalised minimum regret M_{ns} has similarities to entropy E_{ns}, which is defined by Shannon (1948) as (our notation):

$$E_{ns} = -\sum_{j=1}^{J} P_{nsj} \log P_{nsj}. \tag{11}$$

This entropy value is bounded by $E_{ns} \in [0, -\frac{J}{2}\log(\frac{1}{2})]$. Entropy is used as a proxy for choice task complexity in Swait and Adamowicz (2001). Typically a low (high) normalised minimum regret also means a low (high) entropy, and vice versa. For example, if a choice task has a strictly dominant alternative (i.e., a relatively easy choice) such that one alternative is chosen with a probability equal to 1, then $M_{ns} = 0$ and $E_{ns} = 0$. On the other hand, if all alternatives are different on every attribute but regrets are identical (i.e., a relatively difficult choice), then all probabilities are identical and both M_{ns} and E_{ns} are maximised.

An important difference is that entropy depends on choice probabilities, which makes it dependent on the model assumptions. Normalised minimum regret only depends on the utility function and not on a specific type of discrete choice model. To illustrate the difference, consider again our simple route choice example in which Route A is described by a travel time of 10 minutes and a travel cost of \$1, while Route B has a travel time of 11 minutes and the same travel cost of \$1. Assume an LPLA utility function and negative partworths for time and cost. If one uses an MNL model, then the probabilities will be almost identical, yielding a high value for E_{ns}. In contrast, M_{ns} does not rely on a specific choice model and will be equal to zero since Route A is a strictly dominant alternative. Hence, entropy is not a good measure for dominancy in all cases, but it will likely give similar results in many cases.

4.3 Smooth approximation of minimum regret

In order to resolve the second issue, we replace the maximum operator with the 'soft maximum' operator in order to approximate the non-smooth minimum regret R_{ns} function by a smooth function $\tilde{R}_{ns}$ (we denote all smooth approximations with a tilde). The soft maximum for a series of values $a_1, \ldots, a_Z$ is defined as follows (see e.g., Cook, 2011):

$$\max_z \{a_z\} \approx \frac{1}{\xi} \log\left(\sum_{z=1}^{Z} \exp(\xi a_z) \right), \tag{12}$$

where $\xi > 0$ defines the 'hardness'. The approximation becomes exact if $\xi \to \infty$. Hence, the smooth approximation for the regret of alternative *j* defined in Equation (5) is given by

$$\tilde{R}_{nsj} = \frac{1}{\xi} \sum_{i \neq j} \sum_{k=1}^{K} \log\Big(1 + \exp\big(-\xi \Delta_{ns,\, j \leftarrow i,k}\big)\Big). \tag{13}$$

The smooth approximation for R_{ns} in Equations (6) and (7) can be calculated in the same way by taking the 'soft minimum'. Since $\min_z\{a_z\} = -\max\{-a_z\}$, we can use Equation (12) again to calculate the following smooth approximations:

$$\tilde{R}_{ns} = -\frac{1}{\xi}\log\sum_{j=1}^{J}\exp\left(-\xi\tilde{R}_{nsj}\right), \tag{14}$$

$$\tilde{R}_{n} = -\frac{1}{\xi}\log\sum_{s=1}^{S}\exp\left(-\xi\tilde{R}_{ns}\right). \tag{15}$$

It is interesting to note that Equation (13) is identical to the formulation of regret for an alternative as formulated in Chorus (2010) in the case of linear utility functions and using a moderate hardness of the soft maximum of $\xi = 1,$ resulting in

$$\tilde{R}_{nsj} = \sum_{i\neq j}\sum_{k=1}^{K}\log\Big(1+\exp\big(\beta_{nk}(x_{nsik}-x_{nsjk})\big)\Big). \tag{16}$$

Furthermore, regret for a choice task as stated in Equation (14) is identical to the random regret logsum derived by Chorus (2012) in the case of linear utility functions and $\xi = 1.$ Our generalisation with respect to hardness ξ and any nonlinear utility functions (including interactions between attributes) can also be applied in a random regret choice modelling context. Cranenburgh et al. (2015) provide an alternative derivation and interpretation of hardness ξ in the regret formulation.

The smooth approximation of normalised minimum regret M_{ns}, denoted by $\tilde{M}_{ns},$ can be calculated for each respondent *n* and for each choice task *s* using Equation (10) by replacing R_{ns} with $\tilde{R}_{ns}$ and R_{nsj} with $\tilde{R}_{nsj}$. If choice task *s* for respondent *n* has a strictly dominant alternative *j*, then $\tilde{M}_{ns} \to 0$ approaches zero for a sufficiently large ξ. In case all alternatives have an identical positive regret R', then $\tilde{M}_{ns} = 1-\log(J)(\xi R')^{-1},$ which approaches one for sufficiently large ξ. Hence, for finite ξ it holds that $\tilde{M}_{ns} \in (0,1).$ Finally, consider the case in which all alternatives are represented by identical profiles, i.e. $R_{nsj} = 0$ for all alternatives *j*. While M_{ns} in Equation (10) is undefined in this case, it can be shown that in case of identical profiles $\tilde{M}_{ns} = 1-\left(K(J-1)\log(2)\right)^{-1}\log(J),$ which equals 0.5 in when $J = K = 2.$

4.4 Scaling using smooth approximations of normalised minimum regret

Now that we have normalised minimum regret and also derived a smooth approximation, we can relate scale λ_{ns} to $\tilde{M}_{ns}$ in such a way that scale decreases with increasing normalised minimum regret. Two obvious choices would be an exponential or a power function. We propose the following power function:

$$\lambda_{ns} = \tilde{M}_{ns}^{-\gamma}, \tag{17}$$

where γ is a coefficient that needs to be estimated. If $\gamma = 0,$ then the probabilities in Equation (9) are consistent with the MNL model. Since that λ_{ns} and $\tilde{M}_{ns}$ are typically related, it is expected that $\gamma > 0.$ We also tested other functional forms, such as $\lambda_{ns} = \exp(-\gamma\tilde{M}_{ns}),$ but a power function seems to work best, especially since $\lambda_{ns} \to \infty$ if $\tilde{M}_{ns} \to 0$. We call our choice model in Equation (9) with scale determined as in Equation (17) a regret-scaled multinomial logit (RS-MNL) model.

5. Simulated and empirical datasets

In order to demonstrate how dominancy can be excluded from surveys and how it can be taken into account in estimation, we created four experimental designs for a simple route choice study. Then we used these designs to simulate choices and also to create an online survey to collect actual choice data from respondents.

5.1 Simple route choice case study

In order to demonstrate the impact of dominancy, we consider a simple route choice case study in which there are two unlabelled alternatives (Routes 1 and 2) with a generic LPLA utility function considering two attributes, namely travel time and travel cost:

$$V_{js} = \beta_T T_{js} + \beta_C C_{js}, \tag{18}$$

where β_T and β_C are the partworths for time and cost, respectively, such that the value of travel time savings (VTTS) is given by β_T / β_C. In the case study we assume that we show the same fixed design to all respondents (i.e., time and cost are not respondent specific) and that the population is homogeneous (i.e., partworths are not respondent specific), such that we can omit subindex *n*. We consider four different levels for each attribute, namely $T_{js} \in \{10, 15, 20, 25\}$ (minutes) and $C_{js} \in \{1, 2, 3, 4\}$ (Australian dollars). We generate designs that have eight choice tasks each and are all attribute level balanced. Attribute level balance ensures that the design covers the range of levels for each attribute equally, which is often seen as a desirable property.

According to Table 1, there are 36 unique choice tasks without dominant alternatives, such that there exist $\binom{36}{8} = 30,260,340$ unique designs without a dominant alternative. However, a much smaller number of designs will be attribute level balanced.

In order to assess dominancy, we need to know the signs of the partworths. We assume that partworths β_T and β_C are both negative. Further, in order to generate efficient experimental designs, we assume the following prior values (best guesses) for these partworths: $\beta_T = -0.2$ and $\beta_C = -1.2$, such that the VTTS is \$10 per hour.

5.2 Four experimental designs

The first design we generate is an orthogonal design (see e.g., Louviere et al., 2000), see Table 1. The next design is also orthogonal, but we also aimed to optimise the efficiency, leading to an orthogonal D-optimal design (see Rose and Bliemer, 2009). Based on the prior partworths, this design has a D-error of 0.076, which is much lower than the D-error of the orthogonal design (0.304) and hence is more efficient by capturing more (Fisher) information per choice task, leading to smaller standard errors. The third design is a D-optimal design (see e.g., Huber and Zwerina, 1996) that is no longer orthogonal, but has been optimised for efficiency and has the lowest possible D-error (under the requirement of attribute level balance), namely 0.057. Note that choice tasks 2 and 3 are essentially the same, as well as 1 and 6, 4 and 5, and 7 and 8, therefore this design contains only four unique combinations of attribute levels and four replications. We further generated a D^*-optimal design without dominant alternatives and choice task replications, which has a D-error of 0.064.

Inspecting the experimental designs, we notice that many choice tasks contain dominant alternatives (shaded in grey in Table 3). All choice tasks in the orthogonal design contain a strictly dominant alternative. Such dominant alternatives also occur in the last five choice tasks of the orthogonal D-optimal design, and in the last two choice tasks of the D-optimal design. Hence, mainstream design generation procedures do not rule out that such choice tasks exist in the dataset, and will in most cases have to be removed manually.

In Figure 3 we have visually represented the choice tasks in the experimental design, with travel time and cost on the horizontal and vertical axis, respectively. Each profile in the design is represented with a black dot and each choice task is represented by a line between two dots. All possible choice tasks without a dominant alternative are shown in Figure 3(a), i.e. all lines need to have a negative slope (running from north-west to south-east or vice versa). Dashed (red) lines indicate a choice task with a dominant alternative, while solid (blue) lines indicate a choice task without a dominant alternative.

Table 3 – Experimental designs

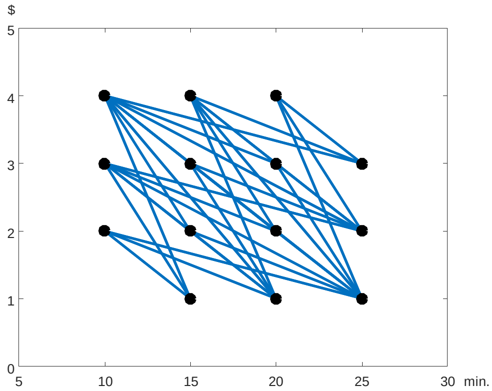
		Orthogonal design			*Orthogonal D-optimal design*			*D-optimal design*			*D^*-optimal design*		
		D-error = 0.304			D-error = 0.076			D-error = 0.057			D-error = 0.064		
s	j	T_{sj}	C_{sj}	P_{sj}	T_{sj}	C_{sj}	P_{sj}	T_{sj}	C_{sj}	P_{sj}	T_{sj}	C_{sj}	P_{sj}
1	1	15	2	0.77	20	2	0.60	10	4	0.35	10	4	0.17
1	2	15	3	0.23	10	4	0.40	25	1	0.65	20	1	0.83
2	1	25	1	0.97	10	4	0.60	10	3	0.86	25	2	0.35
2	2	25	4	0.03	25	1	0.40	25	2	0.14	10	4	0.65
3	1	20	4	0.03	25	1	0.35	25	2	0.14	15	4	0.20
3	2	20	1	0.97	15	3	0.65	10	3	0.86	20	2	0.80
4	1	10	2	0.90	10	1	0.96	20	1	0.93	20	1	0.80
4	2	15	3	0.10	20	2	0.04	15	4	0.07	15	3	0.20
5	1	10	3	0.23	20	3	0.10	15	4	0.07	15	3	0.69
5	2	10	2	0.77	15	2	0.90	20	1	0.93	25	2	0.31
6	1	20	1	0.73	15	2	0.10	25	1	0.65	25	2	0.14
6	2	25	1	0.27	10	1	0.90	10	4	0.35	10	3	0.86
7	1	15	2	0.97	25	4	0.10	20	3	0.10	10	3	0.65
7	2	20	4	0.03	20	3	0.90	15	2	0.90	25	1	0.35
8	1	25	4	0.01	15	3	0.96	15	2	0.90	20	1	0.93
8	2	10	3	0.99	25	4	0.04	20	3	0.10	15	4	0.07

Our four designs are clearly attribute level balanced, since each attribute level appears exactly twice. All choice tasks in the orthogonal design in Figure 3(b) contain a strictly dominant alternative (for each line represented by a profile closest to the origin). The orthogonal D-optimal design in Figure 3(c) contains three solid (blue) line segments and five dashed (red) lines. The D-optimal design in Figure

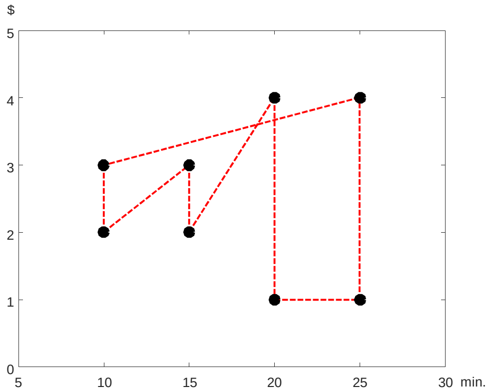
3(d) shows only four lines, since each choice task is replicated twice. The D*-optimal design in Figure 3(e) shows eight solid (blue) lines, such that there are no replications nor dominant alternatives. For completeness we also generated a D^*-optimal design without requiring attribute level balance, shown in Figure 3(f), which has a D-error of 0.053. It is clear that without the requirement of attribute level balance, profiles are pushed towards the edges since this increases trade-offs and thereby efficiency.

Even though the probabilities in the MNL model would suggest that the probability of choosing Route 1 is 0.77 for the first choice task in the orthogonal design (see Table 3), a rational decision maker would under these assumptions always choose Route 1. Hence we would expect that the observed probabilities will be (close to) 1.00 and 0.00 for routes 1 and 2, respectively. This discrepancy is due to the difference between the assumptions in the (compensatory) MNL model and the actual (non-compensatory) behaviour. Such a discrepancy between the modelled and actual choice probabilities could be diminished by increasing scale λ_1 in our RS-MNL model.

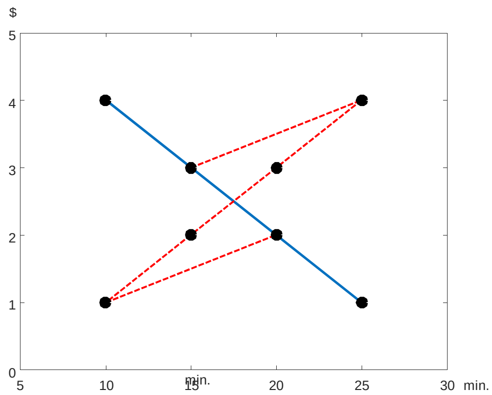
We will use these four experimental designs to generate a simulated as well as an empirical dataset.



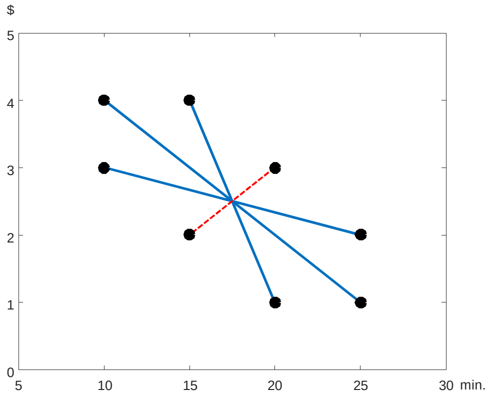
(a) Choice tasks without dominant alternatives



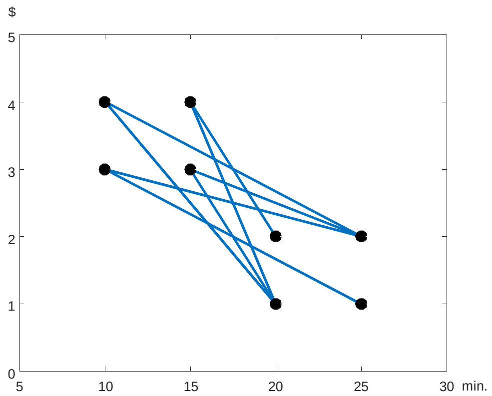
(b) Orthogonal design



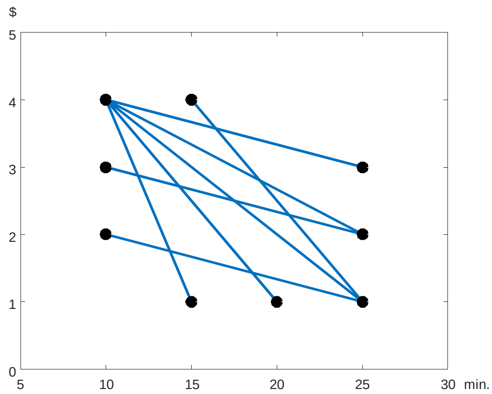
(c) Orthogonal D-optimal design



(d) D-optimal design



(e) D*-optimal design



(f) Unbalanced D*-optimal design

Figure 3 – Visualisation of choice tasks

5.3 Simulated choices

In this section we generate a dataset by simulating choices consistent with an MNL model, except when there is a dominant alternative. In such a choice task, there are no trade-offs and we assume that the actual behaviour with be non-compensatory in which all respondents choose the dominant alternative. This simulation setup is therefore similar to Rose et al. (2013). They also simulate datasets to determine the impacts of wrong model assumptions, but did not look at the case of dominant alternatives.

Let y_{nsj} denote a choice indicator that equals one if respondent *n* chooses alternative *j* in choice task *s*, and zero otherwise. Assuming an MNL model and that the true partworths are $\beta_T = -0.2$ and $\beta_C = -1.2,$ we simulate these observations by randomly drawing ε_{nsj} from an extreme value type I distribution with variance $\frac{1}{6}\pi^2$ independently for each alternative, choice task, and respondent. In case there is no dominant alternative in choice task *s* (i.e., $R_{ns} > 0$), then $y_{nsj} = 1$ if $V_{nsj} + \varepsilon_{nsj} \geq V_{nsi} + \varepsilon_{nsi}$ for all *i*, and zero otherwise. In case the choice task does contain a dominant alternative (i.e, $R_{ns} = 0$), then $y_{nsj} = 1$ for alternative *j* that has minimum regret $R_{nsj} = 0,$ and zero otherwise. Note that none of the experimental designs in Table 3 have identical alternatives in a single choice task, so there will be only one such dominant alternative.

We simulate 500 respondents per design, such that in total there are $500 \times 8 = 4,000$ choice observations in each of the four datasets.

5.4 Empirical choices

We also used the four experimental designs to create an internet survey. In total 360 respondents were asked to participate in the, in which each respondent had to face 16 choice task originating from two complete experimental designs, thereby obtaining in total 5,760 choice observations (1,440 per experimental design).

In total six different combinations of experimental designs can be made (1-2, 1-3, 1-4, 2-3, 2-4, 3-4), and the order can be reversed, such that each respondent saw one of twelve different versions of the survey. We emphasized in the survey that the choice tasks were computer generated in order to prepare the respondent for possible ‘silly’ choice tasks because of dominant alternatives.

The observed choice probabilities are listed in Table 4, in which the grey cells indicate a choice task with a strictly dominant alternative. It is interesting to see that the choice probabilities only reach 1.000/0.000 in one case (namely the second choice task in the orthogonal design in which both routes have the same travel time, but one route has a cost of $1 while the second route has a cost of $4). In all other cases, at least one respondent did not choose the dominant alternative. Taking a closer look at the data, there are 40 respondents that chose one dominated alternative, five respondents that chose two dominated alternatives, two respondents that chose three dominated alternatives, three respondents that chose four dominated alternatives, and one respondent that chose seven dominated alternatives (out of 16). We will refer to these choice observations as spurious choices. Hence, out of 5,760 choice observations there are 75 spurious choices (1.3%). There were no respondents that consistently chose routes with longer travel times and higher costs, so we can conclude that all respondents perceive time and cost as a disutility in general. The 40 respondents may have made a mistake due to fatigue, especially since the ‘mistake’ occurred mostly near the end of the survey. The 11 respondents that chose the dominated alternative multiple times may not have taken the survey seriously and may have selected their preferred option in a somewhat random fashion.

6. Results from simulated dataset

6.1 Estimates for the multinomial logit model

Using the data simulated in Section 5.3, we first estimate partworths in an MNL model based on the orthogonal design, the orthogonal D-optimal design, the D-optimal design, and the D*-optimal design. We use BIOGEME (Bierlaire, 2003) for all model estimations in this paper.

Table 4 – Observed choice probabilities in empirical dataset

s	*j*	*Orthogonal design*	*Orthogonal D-optimal design*	*D-optimal design*	*D*-optimal design*
1	1	0.989	0.689	0.356	0.206
1	2	0.011	0.311	0.644	0.794
2	1	1.000	0.417	0.756	0.517
2	2	0.000	0.583	0.244	0.483
3	1	0.017	0.644	0.233	0.061
3	2	0.983	0.356	0.767	0.939
4	1	0.967	0.961	0.939	0.889
4	2	0.033	0.039	0.061	0.111
5	1	0.011	0.050	0.056	0.578
5	2	0.989	0.950	0.944	0.422
6	1	0.983	0.050	0.633	0.367
6	2	0.017	0.950	0.367	0.633
7	1	0.994	0.033	0.078	0.500
7	2	0.006	0.967	0.922	0.500
8	1	0.028	0.978	0.972	0.956
8	2	0.972	0.022	0.028	0.044

First, we consider the dataset obtained from the two orthogonal designs. Since no information is collected from choice tasks with a dominant alternative, it was not possible to estimate any partworths with the orthogonal design. Next, we estimate the coefficients based on the choice data from the orthogonal D-optimal design. This yields $\beta_T = -1.490$ and $\beta_C = -7.605$ (such that the VTTS is \$11.76 per hour), very different from the assumed 'true' values $\beta_T = -0.2$ and $\beta_C = -1.2$. However, the standard errors are very large, respectively 4.795 and 23.977, which means that the parameters are not statistically significant. If we would estimate the partworths based on the first three choice tasks (i.e., ignore choice data on the last five problematic choice tasks in the design), then $\beta_T = -0.204$ and $\beta_C = -1.173$ (VTTS is \$10.44 per hour). While these partworths are much closer to the true values, they again have very large standard errors such that these values of the partworths are not statistically significant with a sample size of 500 respondents. It is clear that the parameter estimates are inflated (due to scale) by the presence of choice tasks with dominant alternatives.

Now we consider the two efficient designs. The D-optimal design has two choice tasks with dominant alternatives. Estimates of the partworths yields $\beta_T = -0.249$ and $\beta_C = -1.452$ (which translates into a VTTS of \$10.48 per hour) with corresponding standard errors 0.007 and 0.041. While these values are quite close to the true values, they are statistically different from -0.2 and -1.2. The analyst would be better off removing the last two choice tasks from the dataset. On the reduced dataset the partworths are estimated to be $\beta_T = -0.203$ and $\beta_C = -1.214$ (VTTS is \$10.03 per hour) with corresponding standard errors of 0.008 and 0.045. These parameter estimates are statistically different –0.249 and –1.452, while they are not statistically different from the 'true' values. Hence, failure to remove choice tasks with dominant alternatives may lead to biased parameter estimates. Using data from the D*-optimal design, the estimated partworths are $\beta_T = -0.201$ and $\beta_C = -1.209$ (the VTTS is \$9.95 per hour) very close to the 'true' values, and have a high reliability (standard errors of 0.006 and 0.038, respectively). We therefore argue that the analyst is better off using a D*-optimal design instead of a D-optimal design, even though the D-optimal design seemingly has a higher efficiency (lower D-error).
6.2 Estimates for the regret-scaled multinomial logit model

Using the same simulated dataset, we now estimated our newly proposed RS-MNL model. Again, the partworths could not be estimated using the orthogonal design. However, we were successful in estimating the coefficients using the orthogonal D-optimal design without removing the five choice tasks with dominant alternatives. This resulted in $\gamma = 0.065,$ $\beta_T = -0.271,$ and $\beta_C = -1.483$ (VTTS is \$10.98 per hour). The scale parameter for choice tasks without a dominant alternative is $\lambda_s = 1.127,$ while the scale parameter in choice tasks with a dominant alternative is either 8.623 or 10.460, see Table 5. The scaled partworths are therefore $-0.271 \times 1.127 = -0.305$ and $-0.1.483 \times 1.127 = -1.671,$ which are inflated compared to -0.2 and -1.2, but much less inflated compared to the values of -1.490 and -7.605 obtained in the MNL model. It seems that our scaling approach to account for dominancy works well. The resulting probabilities listed in Table 5 are reasonably close to the probabilities in choice tasks without a dominant alternative in Table 3, and are equal to one and zero in choice tasks with a dominant alternative.

Next, we estimate the RS-MNL model using the two efficient designs. The partworth estimates from the D-optimal design are $\gamma = 0.041,$ $\beta_T = -0.173,$ and $\beta_C = -1.005.$ The value of the scale parameter λ_s for the choice tasks that do not contain a dominant alternative (i.e., the first six choice tasks in Table 3) is between 1.166 and 1.202. This means that the estimates are very similar to the 'true' values of the partworths, because $-0.173 \times 1.166 = -0.202$ and $-0.173 \times 1.202 = -0.208$, while $-1.005 \times 1.166 = -1.173$ and $-1.005 \times 1.202 = -1.208$. For the last two choice tasks that contain dominant alternatives, $\lambda_s = 5.555,$ which is large enough to ensure that the probabilities go to 0.000/1.000 and 1.000/0.000, respectively. The probabilities of the choice tasks without dominant alternatives are very similar to the probabilities predicted by the MNL model based on the 'true' values (compare Tables 3 and 5). The RS-MNL model has a loglikelihood of -1,306, which is much better than the model fit of the MNL model with a loglikelihood value of 1,454.

Table 5 – Estimated scale parameters and probability predictions

		Orthogonal D-optimal design		*D-optimal design*		*D^*-optimal design*	
s	*j*	λ_s	P_{sj}	λ_s	P_{sj}	λ_s	P_{sj}
1	1	1.127	0.58	1.166	0.38	1.000	0.17
1	2	1.127	0.42	1.166	0.62	1.000	0.83
2	1	1.127	0.58	1.191	0.87	1.000	0.35
2	2	1.127	0.42	1.191	0.13	1.000	0.65
3	1	1.127	0.39	1.191	0.13	1.000	0.20
3	2	1.127	0.61	1.191	0.87	1.000	0.80
4	1	10.46	1.00	1.202	0.93	1.000	0.80
4	2	10.46	0.00	1.202	0.07	1.000	0.20
5	1	8.623	0.00	1.202	0.07	1.000	0.69
5	2	8.623	1.00	1.202	0.93	1.000	0.31
6	1	8.623	0.00	1.166	0.62	1.000	0.14
6	2	8.623	1.00	1.166	0.38	1.000	0.86
7	1	8.623	0.00	5.555	0.00	1.000	0.65
7	2	8.623	1.00	5.555	1.00	1.000	0.35
8	1	10.46	1.00	5.555	1.00	1.000	0.93
8	2	10.46	0.00	5.555	0.00	1.000	0.07

Finally, estimating the RS-MNL model with the data obtained from the D^*-optimal design, we obtained exactly the same results as with the MNL model in Section 6.1, namely $\gamma = 0$ (such that $\lambda_s = 1$ for all choice tasks), $\beta_T = -0.201$ and $\beta_C = -1.209$. This is encouraging and means that the RS-MNL model is able to replicate the 'true' values, and can be 'safely' be used on datasets with and without dominant alternatives.

7. Results from empirical dataset

Using the choice observations from our online survey as described in Section 5.4, we estimate partworths in an MNL model as well as our RS-MNL model based on (i) a pooled dataset, (ii) a reduced pooled data set in which observations from respondents that selected one or more dominated alternatives were removed, and (iii) separate datasets for each of the four experimental designs.

7.1 Estimates on pooled dataset

Table 6 summarises the estimates for both models on a pooled dataset of all four experimental designs. The RS-MNL model has a better model fit than the MNL model measured by the log-likelihood value and the adjusted R^2. The exponent of the smooth normalised minimum regret (see Equation (17)) is positive and significant, which means that scale is not constant over all choice tasks but needs to be adjusted for choice tasks containing dominant alternatives. With $\gamma = 0.055$, scale parameter λ_s is between 1.055 and 1.078 for all choice tasks that do not contain a dominant alternative, while the scale is between 1.369 and 1.496 for choice tasks that include a dominant alternative. This means that for choice tasks without a dominant alternative, $\lambda_s\beta_T$ is between 0.190 and 0.194, while $\lambda_s\beta_C$ is between 1.132 and 1.348. For choice tasks with a dominant alternative, $\lambda_s\beta_T$ is between 0.246 and 0.269, while $\lambda_s\beta_C$ is between 1.711 and 1.870. The MNL estimates, which assume $\lambda_s = 1$, fall as expected between these values. Clearly, including dominant alternatives in the dataset impacts upon scale and inflates the MNL partworths. The VTTSs are \$8.83 and \$8.67 per hour respectively for the MNL and R-MNL model, which are not statistically different.

7.2 Estimates on reduced pooled dataset

As described in Section 5.4, there were 51 (out of 360) respondents with one or more spurious choices. We cleaned the dataset by removing all choice observations from these 51 respondents (so not only the 1.3% spurious observations), which amounts to a removal of 14.2% of all observations. Table 7 presents the estimates for the MNL and RS-MNL model.
The RS-MNL model has a much better model fit than the MNL model. The difference in model fit is much larger than in Table 6, which suggests that the presence of the spurious choices actually diminishes the problem of scale inflation.

The γ parameter in the RS-MNL model is significantly higher than the value in Table 6 since much larger scale differences are estimated over different choice tasks. For choice tasks without a dominant alternative it holds that $1.369 \leq \lambda_s \leq 1.496$, while $11.355 \leq \lambda_s \leq 26.758$ for choice tasks that contain a dominant alternative. Clearly, without the spurious choice observations, the RS-MNL model is much better able to distinguish between choice tasks with and without dominant alternatives, and the VTTSs grow further apart (and become statistically different).

Table 6 – Estimates on pooled dataset (360 respondents, 5,760 observations)

	MNL		*RS-MNL*	
	coeff.	*s.e.*	*coeff.*	*s.e.*
Time (β_T)	-0.206	0.0059	-0.159	0.0103
Cost (β_C)	-1.400	0.0335	-1.100	0.0633
Exponent (γ)	--	--	0.055	0.0130
VTTS	$8.83/h		$8.67/h	
Loglikelihood	-2115.6		-2099.4	
Adjusted R^2	0.470		0.473	

Table 7 – Estimates on reduced pooled dataset (309 respondents, 4,944 observations)

	MNL		*RS-MNL*	
	coeff.	*s.e.*	*coeff.*	*s.e.*
Time (β_T)	-0.236	0.0072	-0.061	0.0089
Cost (β_C)	-1.610	0.0416	-0.431	0.0634
Exponent (γ)	--	--	0.540	0.0734
VTTS	$8.80/h		$8.45/h	
Loglikelihood	-1609.4		-1501.0	
Adjusted R^2	0.531		0.563	

7.3 Estimates on separate datasets for each experimental design

Next, we estimated models using data from each design separately (without removing spurious choices), see Table 8. All estimates are statistically significant. Looking at the MNL model, the first thing we notice is that the VTTSs are statistically different when using data from different designs. Looking at the partworths, scale has a clear influence on the estimates, where designs with more dominant alternatives means higher scale and therefore a more deterministic choice. Using these partworths in prediction will lead to quite different choice probabilities. The RS-MNL model has a better model fit in all four designs, especially in the dataset from the orthogonal and D-optimal design.

While in our simulated dataset no estimates could be obtained using the orthogonal design (since all choice tasks have a dominant alternative), perhaps surprisingly there is no problem estimating the partworths using the empirical dataset. This is due to the fact that not always the dominant alternative was chosen by the respondents. This indicates that choice tasks with dominant alternatives can actually contain information. Simply removing them would therefore lead to loss of information.

Looking at the estimates for the orthogonal D-optimal design, γ is negative in the RS-MNL model. This means that scale is small when dominancy in a choice task is large. We attribute this counterintuitive result to the existence of spurious choices. However, γ is small such that scale differences are very small.

The estimates obtained from the D*-optimal dataset are quite similar across the MNL and the RS-MNL model. Since this dataset does not contain any dominant alternatives, there do not seem to be scaling issues in the MNL model, such that the models perform similarly and the VTTSs are identical. An exponent of $\gamma = 0.672$ means that there are some scale differences across choice tasks, where the scale parameter takes the values $3.625 \le \lambda_s \le 5.297$. In other words, the relative difference between the highest and the lowest scale is with 46% quite small, such that the results are not so different from the MNL results.

Table 7 – Estimates on separate datasets (180 respondents, 1,440 observations per design)

	MNL							
	Orthogonal design		*Orthogonal D-optimal design*		*D-optimal design*		*D^*-optimal design*	
	coeff.	*s.e.*	*coeff.*	*s.e.*	*coeff.*	*s.e.*	*coeff.*	*s.e.*
Time (β_T)	-0.385	0.066	-0.256	0.164	-0.183	0.010	-0.143	0.010
Cost (β_C)	-2.660	0.260	-1.510	0.083	-1.180	0.051	-1.170	0.066
VTTS	$8.68/h		$10.17/h		$9.31/h		$7.33/h	
LL	-168.3		-506.3		-602.4		-741.5	
Adj. R^2	0.831		0.491		0.396		0.256	

	RS-MNL							
	Orthogonal design		*Orthogonal D-optimal design*		*D-optimal design*		*D^*-optimal design*	
	coeff.	*s.e.*	*coeff.*	*s.e.*	*coeff.*	*s.e.*	*coeff.*	*s.e.*
Time (β_T)	-0.748	0.077	-0.610	0.262	-0.019	0.016	-0.032	0.017
Cost (β_C)	-4.680	0.399	-3.290	1.310	-0.128	0.093	-0.258	0.136
Exp (γ)	-0.046	0.003	-0.022	0.003	0.719	0.020	0.672	0.234
VTTS	$9.59/h		$11.12/h		$8.88/h		$7.33/h	
LL	-130.7		-505.4		-588.2		-736.4	
Adj. R^2	0.866		0.491		0.404		0.259	

8. Discussion, recommendations, and limitations

8.1 Summary and discussion

In this paper we have discussed the impacts of the existence of choice tasks with dominant alternatives in an unlabelled stated choice study. In a simple case study with simulated choices we showed that dominant alternatives could lead to biased model estimates due to the discrepancy between actual behaviour (which is non-compensatory in the case of dominant alternatives) and assumed behaviour in the model (typically compensatory behaviour). A dominant alternative triggers non-trading behaviour, which mostly affects error variance (i.e., scale).

We discussed three ways of dealing with dominant alternatives. First, the analyst could simply make sure that dominant alternatives do not exist in the stated choice data. To this end, we proposed a new D^*-optimal design method, in which we use minimum regret as a measure to detect and eliminate choice tasks that contain a dominant alternative.

Secondly, the analyst can simply clean the data such that (i) choice tasks with a dominant alternative are removed, or (ii) all choice tasks of respondents that fail to choose a dominant alternative are removed. In our empirical analysis we show that a choice task with a dominant alternative may actually contain information, in contrast to common belief. A requirement seems that the dominant alternative is not chosen by all respondents in the dataset. Removing all choice tasks with dominant alternatives may therefore result in information loss. If we would have used this strategy to clean the dataset of our orthogonal design, we would not have been able to estimate the partworths. If the analyst removes only data from certain respondents, dominant alternatives may still exist in the dataset. Such dominant alternatives typically decreases error variance and as such increases scale, leading to biased estimates in the MNL model.

Thirdly, the analyst can compensate for scale differences in the model. We proposed a regret-scaled (RS-) MNL model, in which scale increases with a decrease in normalised minimum regret. We further proposed a smooth approximation of this normalised minimum regret in order to avoid numerical problems in model estimation. Our simulation and empirical results show that our RS-MNL model improves model fit and seem to appropriately account for scale differences.

8.2 Recommendations

Based on these findings, we would strongly recommend using a (Bayesian) D^*-optimal design instead of an orthogonal or D-optimal design in stated choice studies, since this avoids dominant alternatives in the experiment design. Further, when a dataset includes dominant alternatives, we suggest not removing these choice tasks (since they contain some information), but rather adopting our RS-MNL model that automatically accounts for scale differences.

8.3 Limitations

In our study we have focussed on limitations of the MNL model. Clearly, more advanced discrete choice models exist. Therefore, we only demonstrated the impacts on the MNL model assuming homogeneous preferences. However, we argue that dominancy has an impact on all models based on (random) utility, since they all assume compensatory behaviour. The theory and methods in our paper can be applied to each individual respondent (indicated by subindex n), and as such can be applied to for example latent class models with discrete groups of heterogeneous users or to mixed logit models with continuous preference heterogeneity. Therefore, results in this paper are expected to translate to more advanced models.

Furthermore, we have only focussed on dominancy in isolation. In our empirical dataset, many other behavioural processes may have led to the actual choices, including non-trading, lexicographic, or inconsistent behaviour. We can therefore not guarantee that our observed scale differences are purely the result of the presence of dominant alternatives, but may also be the result of learning, fatigue, and other effects.

Acknowledgments

We would like to thank Andrew Collins for his useful contributions to the literature study on information processing strategies and decision rules.

References

Axhausen, K.W. and J.W. Polak (1991) Choice of parking: stated preference approach. *Transportation*, Vol. 18, pp. 59-81.

Bekhor, S., and Y. Shiftan (2010) Specification and estimation of mode choice model capturing similarity between mixed auto and transit alternatives. *Journal of Choice Modelling*, Vol. 3(2), pp. 29-49.

Bierlaire, M. (2003) BIOGEME: A free package for the estimation of discrete choice models, *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland.

Bliemer, M.C.J., and J.M. Rose (2009) Designing stated choice experiments: state-of-the-art. In R. Kitamura, T. Yoshii and T. Yamamoto (eds) *The Expanding Sphere of Travel Behaviour Research*. Emerald, UK, pp. 499-537.

Bliemer, M.C.J., and Rose, J.M. (2011) Experimental design influences on stated choice outputs: an empirical study in air travel choice. *Transportation Research Part A*, 45, pp. 63-79.

Bradley, M., and A. Daly (1994) Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. *Transportation*, Vol. 21, pp. 167-184.

Burge, P., and C. Rohr (2004) DATIV: SP design: Proposed approach for pilot study. Tetra-plan in cooperation with RAND Europe and Gallup A/S.

ChoiceMetrics (2012) *Ngene 1.1.1 User Manual & Reference Guide*, Australia.

Chorus, C.G. (2010) A new model of random regret minimization. *European Journal of Transport and Infrastructure Research*, 10(2), pp. 181-196.

Chorus, C.G. (2014) Benefit of adding an alternative to one's choice set: a regret minimization perspective. *Journal of Choice Modelling*, Vol. 13, pp. 49-59.

Chorus, C.G., T.A. Arentze, and H.J. Timmermans (2008) A random regret-minimization model of travel choice. *Transportation Research Part B*, Vol. 42(1), pp. 1-18.

Chorus, C.G. (2012) Logsums for utility-maximizers and regret-minimizers, and their relation with desirability and satisfaction. *Transportation Research Part A: Policy and Practice*, Vol. 46(7), 1003-1012.

Collins, A.T. (2012) *Attribute nonattendance in discrete choice models: measurement of bias, and a model for the inference of both nonattendance and taste heterogeneity*. PhD Thesis, Institute of Transport and Logistics Studies, University of Sydney, Australia.

Collins, A.T., J.M. Rose, and D.A. Hensher (2013) Specification issues in a generalised random parameters attribute nonattendance model. *Transportation Research Part B*, Vol. 56, pp. 234-253.

Cook, J. (2011) *Basic properties of the soft maximum*. Working document, Department of Biostatistics, University of Texas, Houston, USA.

Cook, R.D., and C.J. Nachtsheim (1980) A comparison of algorithms for constructing exact D-optimal designs. *Technometrics*, Vol. 22, pp. 315-324.

Crabbe, M., and M. Vandebroek (2012) Using appropriate prior information to eliminate choice sets with a dominant alternative from D-efficient designs. *Journal of Choice Modelling*, Vol. 5(1), pp. 22-45.

Danthurebandara, V.M., J. Yu, and M. Vandebroek (2011) Effect of choice task complexity on design efficiency in conjoint choice experiments. *Journal of Statistical Planning and Inference*, Vol. 141(7), pp. 2278-2286.

De Luca, S., and R. Di Pace (2015) Modelling users' behaviour in inter-urban carsharing program: a stated preference approach. *Transportation Research Part A*, Vol 71, pp. 59-76.

DeShazo, J., and Fermo, G. (2002) Designing choice sets for stated preference methods: the effects of complexity on choice consistency. *Journal of Environmental Economics and Management*, Vol. 44, pp. 123-143.

Federov, V.V. (1972) *Theory of optimal experiments*. Academic Press, New York.

Foster, V., and S. Mourato (2002) Testing for consistency in contingent ranking experiments. *Journal of Environmental Economics and Management*, Vol. 44, pp. 309-328.

Hensher, D.A., P.O. Barnard, and T.P Truong (1988) The role of stated preference methods in studies of travel choice. *Journal of Transport Economics and Policy*, Vol. 22, pp. 45-57.

Hensher, D.A., J.M. Rose, and W.H. Greene (2005) The implications on willingness to pay of respondents ignoring specific attributes. *Transportation*, Vol. 32(3), pp. 203-222.

Hensher, D.A. (2006) Towards a practical method to establish comparable values of values of travel times savings from stated choice experiments with differing design dimensions. *Transportation Research Part A*, Vol. 40, pp. 829-840.

Hess, S., and J.M. Rose (2007) A latent class approach to modelling heterogeneous information processing strategies in SP studies. In: Workshop on Valuation Methods in Transport Planning, Oslo, Norway.

Hess, S., J.M. Rose, and J. Polak (2010) Non-trading, lexicographic and inconsistent behaviour in stated choice data. *Transportation Research Part D*, Vol. 15, pp. 405-417.

Hess, S., M. Fowler, T. Adler, and A. Bahreinian (2012a) A joint model for vehicle type and fuel type choice: evidence from a cross-nested logit study. *Transportation*, Vol. 39, pp. 593-625.

Hess, S., A. Stathopoulos, and A. Daly (2012b) Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies. *Transportation*, Vol. 39(3), pp. 565-591.

Huber, J., J.W. Payne, and C. Puto (1982) Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis. *Journal of Consumer Research*, Vol. 9(1), pp. 90-98.

Huber, J., and Zwerina, K. (1996) The importance of utility balance and efficient choice designs. *Journal of Marketing Research*, Vol. 33, pp. 307–317.

Louviere, J.J., Hensher, D.A., and Swait, J.D. (2000) *Stated Choice Methods: Analysis and Application*, Cambridge University Press, Cambridge.

Louviere, J.J., Islam, T., Wasi, N., Street, D., and Burgess, L. (2008) Designing discrete choice experiments: do optimal designs come at a price? *Journal of Consumer Research*, Vol. 35(2), pp. 360–375.

Loomes, G., and R. Sugden (1982) Regret Theory: An alternative theory of rational choice under uncertainty. *The Economic Journal*, 92, 805-824.

McFadden, D. (1974) Conditional logit analysis of qualitative choice behaviour. In: Zarembka, P. (ed.). *Frontiers in Econometrics*, Academic Press, New York, 105-142.

Rezaei, A., and Z. Patterson (2015) Identifying inconsistent responses in stated choice surveys using a dominance-based approach. Presented at the *94th Annual Meeting of the Transportation Research Board*, Washington DC, USA.

Rose, J.M., and M.C.J. Bliemer (2009) Constructing efficient stated choice designs. *Transport Reviews*, 29(5), pp. 587–617.

Rose, J.M. and M.C.J. Bliemer (2013) Sample size requirements for stated choice experiments. *Transportation*, Vol. 40(5), pp. 1021-1041.

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

Rose, J.M., S. Hess, and A.T. Collins (2013) What if my model assumptions are wrong? The impact of non-standard behaviour on choice model estimation. *Journal of Transport Economics and Policy*, Vol. 47(2), pp. 245-263.

Sælesminde, K. (2001) Inconsistent choices in stated choice data. *Transportation*, Vol. 28, pp. 269-296.

Sælesminde, K. (2002) The impact of choice inconsistencies in stated choice studies. *Environmental and Resource Economics*, Vol. 23, pp. 403-420.

Sælesminde, K. (2006) Causes and consequences of lexicographic choices in stated choice studies. *Ecological Economics*, Vol. 59, pp. 331-340.

Scott, A. (2002) Identifying and analysing dominant preferences in discrete choice experiments: an application in health care. *Journal of Economic Psychology*, Vol. 23, pp. 383-398.

Shannon, C.E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, Vol. 27(3), pp. 379-423.

Sándor, Z., and M. Wedel (2001) Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, Vol. 38, pp. 430-444.

Swait, J., and Adamowicz, W. (2001) Choice environment, market complexity and consumer behaviour: a theoretical and empirical approach for incorporating decision complexity into models of consumer choice. *Organization Behavior and Human Decision Processes*, Vol. 86, pp. 141-167.

Van Cranenburgh, S., C.A. Guevara, and C.G. Chorus (2015) New insights on random regret minimization models. *Transportation Research Part A*, Vol. 74, pp. 91-109.

Van de Kaa, E.J. (2006) Assessment of the value of travel time from stated choice surveys: the impact of lexicographic answering. *Proceedings of the European Transport Conference*, Strasbourg, France.

Walker, J.L., Y. Wang, M. Thorhauge, and M. Ben-Akiva (2015) D-efficient of deficient? A robustness analysis of SP experimental design in a VOT estimation context. Presented at the *Annual Meeting of the Transportation Research Board*, Washington DC, USA.