UNIVERSITY OF WAIKATO

Hamilton
New Zealand

Experimental Designs for Environmental Valuation
with Choice-Experiments: A Monte Carlo Investigation

Silvia Ferrini

Riccardo Scarpa

Department of Economics

Working Paper in Economics 8/05

December 2005

Silvia Ferrini

Economics Department
University of Siena
Piazza S. Francesco, 7
53110 Siena, ITALY

Tel: +39 0577 232645
Fax: +39 0577 232661

Email: silvia.f@iol.it
Web: http://www.econ-pol.unisi.it

Riccardo Scarpa

Economics Department
University of Waikato
Private Bag 3105
Hamilton, New Zealand

Tel: +64 (0) 7-838-4045
Fax: +64 (0) 7-838-4331

Email: rscarpa@waikato.ac.nz
Web: http://www.mngt.waikato.ac.nz

Abstract

We review the practice of experimental design in the environmental economics literature concerned with choice experiments. We then contrast this with advances in the field of experimental design and present a comparison of statistical efficiency across four different experimental designs evaluated by Monte Carlo experiments. Two different situations are envisaged. First, a correct *a priori* knowledge of the multinomial logit specification used to derive the design and then an incorrect one. The data generating process is based on estimates from data of a real choice experiment with which preference for rural landscape attributes were studied. Results indicate the D-optimal designs are promising, especially those based on Bayesian algorithms with informative prior. However, if good *a priori* information is lacking, and if there is strong uncertainty about the real data generating process - conditions which are quite common in environmental valuation - then practitioners might be better off with conventional fractional designs from linear models. Under mis-specification, a design of this type produces less biased estimates than its competitors.

Keywords

logit experimental design
efficiency
Monte Carlo
choice experiments
non-market valuation

JEL Classification

C13; C15; C25; C99; Q26

Acknowledgements

1   Introduction

This paper reports research results on the performance of various experimental designs (henceforth abbreviated in EDs) for logit models estimated on data from choice-experiments (henceforth abbreviated in CEs). The context of study is that of the literature on non-market valuation of environmental goods.

In the last decade the use of discrete CEs for the purpose of non-market valuation of environmental goods has encountered the favour of many applied environmental economists.

CEs are used when policy alternatives may be described in terms of attributes and the objective is to infer the value attached to the respective attribute levels[1]. Attributes could be relevant policy traits and include policy cost. Choice alternatives instead could be different policy options and are called profiles. A CE consist of selected subsets of all possible profiles. Typically, respondents are asked to select the best alternative from a set of alternatives (the "choice set"), and are asked to repeat this choice for several sets.

Using the set of observed discrete choices researchers can estimate separate marginal values for each attribute used in describing the policy alternatives, rather than a unique value for the entire policy scenario. The latter is seen as a limitation of contingent valuation, which unlike CEs cannot trace out the underlying willingness to pay for each attribute. Willingness to pay estimates are typically derived from random utility assumptions and their efficiency reflect the informativeness of the study. On the other hand, in this multi-attribute context the efficiency of the estimates depends crucially on the choice of experimental designs i.e. how attributes and attribute levels are combined to create synthetic alternatives (or profiles) and eventually choice sets to provide maximum information on the model parameters.

Yet, little work has been done to systematically evaluate the effect of the experimental design (ED) on the efficiency of estimates.[2] With few exceptions, in most published papers employing CE for the purpose of valuation one finds scant information on the methodology employed to derive the ED, or its statistical properties. The most common set of arguments seems to be something vaguely like:

---

[1] This motivates the proposed term of "attribute-based stated preference" method [33].

[2] Although some work on the effect of choice set creation and some proposed measure of choice complexity has been published [21, 19].

"The total number of combinations implied by the full factorial could not be employed, so a main effects orthogonal fraction of such factorial was employed. Choice sets were then formed by blocking the resulting set of profiles into n blocks."

Fractional factorial design is frequently used in marketing research with conjoint analysis which draws on general linear-in-the-parameters models, whereas CEs data are analysed by means of models highly non-linear-in-the-parameters, usually of the multinomial logit type.

When estimating preference parameters from CE data the high non-linearity of the multinomial logit (MNL) specification affects the efficiency properties of the maximum likelihood estimator. Hence, efficient EDs[3] for MNL specifications are likely to differ in most practical circumstances from those that are efficient in linear multivariate specifications. In particular, in a MNL context the efficiency properties of the ED will depend on the unknown values of the parameters, as well as the unknown model specification.

Although it may be good to raise the awareness around the issue that EDs for linear multivariate models are only "surrogates" for proper EDs suitable for the MNL context of analysis, one must consider why this is a dominant stance in the profession. One reason might be that the cost of implementing MNL-specific algorithms to derive "optimal" or "efficient"[4] EDs is too high when compared with the practical rewards it brings in the analysis. More empirical investigations of the type conducted by Carlsson and Martinsson [18] in a health economics context are necessary to evaluate the rewards of efficient designs for non-linear-in-the-parameter models. In as much as possible these investigations should be tailored to the state of practice in environmental valuation, which is quite different from that in health economics.[5] This is what we set out to achieve with this paper. In doing so we also extend the investigation to Bayesian designs which allow the researcher to account for uncertainty about the *a-priori* knowledge on

---

[3] The concept of D-optimality (and sometimes A-optimality) has dominated the design literature for choice experiments. However, when the objective is choice prediction, rather than inference, then other optimality criteria, such as G- and V-optimality, are more useful [39].

[4] Kuhfeld *et al.* [42] Blemier *et al.* [7] suggest that it is often more appropriate to discuss D-efficient designs, rather than D-optimal ones, although the prevailing terminology in the field seems to be about D-optimality.

[5] For example, health economists are basically concerned with a private good: health status, while environmental economists are concerned with public goods. A review of the studies in health economics reveals that choice sets are often offering only two alternatives to respondents, while in environmental economics it is more frequent the format including two experimentally designed alternatives plus the status-quo (zero-option).

the parameter values.

After reviewing recent advances in ED for logit models, it stands to reason that the current approach of the profession towards ED is "improvable". However, the gains affordable from such improvement need further investigation. This paper intends to contribute to the existing literature by exploring the empirical performance of a number of recently proposed approaches to construct designs for discrete choice experiments. The investigation is conducted by means of Monte Carlo experiments designed to focus on the finite sample size properties of frequently employed estimators for value derivation in environmental valuation.

In section 2 we provide a summary of the evolution of the knowledge on design construction for CE. In section 3 we quickly revise the use of design construction techniques in the environmental economics literature of CEs for the purpose of valuation. The methodology of our empirical investigation is explained in section 4, while in section 5 we present and discuss the results. We draw our conclusions in section 6.

2   What do we know about design construction for MNL?

A number of significant theoretical and empirical developments have taken place in the field of ED in recent years, and in this paper we draw heavily on these [57, 58, 62, 63, 64, 37, 14, 55, 40, 38, 15].

Before describing our contribution we briefly sketch some recent significant research developments in this area.

The notion of describing a good on the basis of its attribute was born out of the theoretical approach of Lancaster [43] and [44]. It was then readily employed in marketing by Green and Rao [26] who propose *conjoint analysis* as a tool to model consumer's preference.

ED techniques were first introduced in multi-attribute stated preference method for marketing by Louviere and Woodworth [46] and Louviere and Hensher [47], who used the conventional factorial design developed mostly for the statistical analysis of treatment effects in agricultural and biological experiments, to derive and predict choices or market shares. Through this approach they identify a set of "profiles" with well-known statistical properties for general linear models. These profiles are basically synthetic goods described on the basis of selected at-

tributes whose levels are arranged in an orthogonal fashion. When profiles are too numerous for evaluation in a single choice context they are divided into a "manageable" series of choice sets using different blocking techniques. This procedure guarantees that the attributes of the design are statistically independent (i.e., uncorrelated). Orthogonality between the design attributes represented the foremost criteria in the generation process of fractional factorial designs.

Later, some modifications to this basic approach were brought about by the necessity of making profiles to be "realistic" and "congruent" so that orthogonality was no longer seen as a necessary property [see also 55, on the effects of lack of orthogonality on ED efficiency, and how this can easily come about even when orthogonal designs are employed], and hence a good ED may be non-orthogonal in the attribute levels and require the investigation of mixed effects and selected attribute interactions (therefore in many realistic cases main-effects only may not be deemed adequate, as shown in [48]).

Non-orthogonal designs can be optimized for linear multivariate models and guarantee to maximize the amount of information obtained from a design—this is to say that they are D-optimal [6]—but why have these EDs (in which the response variable is continuous) been used in designing CEs (where the response is discrete and a highly non-linear specification is assumed to generate response probabilities)? The answer is given by the assumption that "an efficient design for linear models is also a good design for MNL for discrete choice response" [42]. Corroborating evidence of this is provided by Lazari and Anderson [45] and Kuhfeld *et al.* [42]. More recently Lusk and Norwood [48] studied the small-sample performance of commonly employed D-efficient EDs for linear-in-the-parameters models in the context of logit models for choice-modelling. By appealing to these empirical results one may conveniently ignore the necessity of deriving design for non-linear model where assumptions on the unknown parameter vector (β) is necessary.[7]

The effects of assigning the experimentally designed alternatives to individual choice-sets

---

[6] Such linearly optimal designs can be obtained by specific software such as SPSS, MINITAB Design Ease. The most comprehensive algorithms for choice design we know of are those in the free macro MktEx (pronounced "Mark Tex" and requiring base SAS, SAS/STAT, SAS/IML, an SAS/QC) [40, 41], while CBC also provides choice designs, but only guided towards balancedness.

[7] Typically, in non-linear model the information matrix (and hence the statistical efficiency of experimental design) is a function of the (unknown) vector of the true models parameter or, equivalently, the true choice probabilities.

were investigated by Bunch *et al.* [13] who—although restrictively assuming $\beta = 0$, thereby reducing again the D-optimality problem (efficiency maximization) to a linear problem [27]— did approach the issue of choice sets construction by proposing the *object-based* and *attribute-based* strategies, which we employ later for one of our designs under comparison in Section 4. Because of the $\beta = 0$ assumption such designs take the name of D_0-optimal or "utility-neutral". They satisfy the properties of *orthogonality*, *minimum overlapping*, and *balanced levels*. Such properties, along with that of *balanced utility* are described in [34] who consider these to be essential features in the derivation of efficient EDs.

Later on, Huber and Zwerina [34] broke away from the $\beta = 0$ assumption, and championed the D_p-optimality criterion, where p stands for "*a-priori*" information on β. They demonstrated how restrictive it can be to assume $\beta = 0$ in terms of efficiency loss, and demonstrated that including pre-test results into the development of efficient ED may improve efficiency up to fifty percent.

Their strategy to obtain a D_p-optimal ED is to start from a D_0-optimal design as described in [13] and expanded upon by Burgess and Street [14], and then improve its efficiency by means of heuristic algorithms. Not only is the resulting ED more efficient under the correct *a-priori* information, but it is also robust to some mis-specifications. It is worth noting that this is a local optimum because it is based on a given vector of parameter values.

In some later work [3] it is observed that there exists uncertainty about the *a-priori* information on parameter values β and hence such uncertainty should be accounted for in the ED construction. They propose a hierarchical Bayesian approach based on the estimates of β from some pilot study, used to derive a final D_b-optimal design using Bayes' principle. Such Bayesian ED approaches are described in Atkinson and Donev [4] and in Chaloner and Verdinelli [20] and they were also used by Sandor and Wedel [57] for MNL specifications by using and modifying the empirical algorithms proposed by Huber and Zwerina [34]. This design violates the property of balanced utility but it produces more efficient designs. However, all these Bayesian designs are not $globally$ optimal because they are derived from a search that improves upon an initial fractional design, rather than a search on a full factorial.

Recent work by Burgess and Street have tackled the issue of construction of more general designs, such as [62], [14], [63] and [15] but they are limited to the case of $\beta = 0$.

An approach to derive efficient EDs unconstrained by the $\beta = 0$ hypothesis is illustrated in [38], in which the approach by Zwerina *et al.* [67] is extended and a D_b-optimal ED is obtained by using a weakly-informative[8] (uniform) prior distribution of β.

A short summary of the evolution of ED research is reported in Table 1. Notice that although in recent years the theoretical research work on efficient ED construction for non-linear logit models has intensified [see also 24, 25, for more theoretical results], it still remains mostly anchored to the basic MNL model, whereas much of the cutting edge empirical research is based on mixed logit models of some kind. For logit models with continuous mixing of parameters we found only two applied study concerning ED: by Sandor and Wedel [58] and by Blemier *et al.* [8]. We found no study addressing the issue in the context of finite mixing (latent class models).

On the other hand, there are still few empirical evaluations of the different ways of deriving efficient EDs for multinomial logit models in the various fields of applications in economics, with the exception of [18] in health economics and [55] in transportation.

In particular, Carlsson and Martinsson [18] use a set of Monte Carlo experiments to investigate the empirical performance of four EDs (orthogonal, *shifted*, D_0-optimal and D_p-optimal) for pair-wise CE—the dominant form in health economics. They assume that the investigator correctly specifies the data generating process, the *a-priori* β and the estimation process. Under these conditions—contrary to the results found by Lusk and Norwood [48]—they find that the orthogonal ED produces strongly biased estimates. An apparently worrying result considering that this is the dominant approach in environmental economics. They also find that the *shifted* (also sometimes termed *cycled*) [13] ED performs better than the D_0-optimal for generic attributes, but in general the most efficient design is the D_p-optimal. However, their experimental conditions are quite restrictive, do not extend to Bayesian design construction and are tailored to replicate features that are common in health economics, but—according to our review—not so common in environmental economics.

In transportation modelling, instead, Rose *et al.* [55] emphasized how the much sought-after property of orthogonality may well be lost in the final dataset due to the cumulative effects

---

[8] We prefer the term "weakly-informative to the more common Bayesian term "uninformative" because of the reasons spelled out in [22] where it is noted that a uniform prior is not uninformative in this context.

of sample non-response. Furthermore, while the transportation literature of experiment design for choice modelling is often dominated by labelled experiments (one label per transportation mode, with relative label-specific attributes), the typical situation in environmental valuation seem to be that of generic (unlabelled) experiments.

Finally, on the issue of sequential design Kanninen [37] illustrates how one can choose numerical attributes such as price to sequentially ensure the maximization of the information matrix of binary and multinomial model from CE data. On the other hand Raghavarao and Wiley [51] show that with sequential design and computer aided interview it is possible to include interaction effects and define Pareto-optimal choice sets. Both papers are particularly interesting for future applications with computer aided interview administration of CEs. Sequential designs, however, are beyond the scope of this paper.

3 A review of the state of practice in environmental economics

The introduction of CE in environmental economics took place in the early 90's, when the state of research on ED was still at an embryonal stage. However, environmental economists concerned with discrete choice contingent valuation were already aware of the importance of ED [2, 36, 1] on efficiency of welfare estimates.

But such concern does not seem to have carried over to CE practice, were the dominant approach, as visible from Table 2, remains that based on fractional factorial for main effects with orthogonality. This is typically derived for algorithms suitable for multivariate linear models, which is—as explained earlier—only a surrogate upon which much potential improvement can be brought by more tailored designs. But under what conditions?

The prevailing scheme in environmental economics applications seems to be the following:

1. determination of choice attributes and their levels;

2. ex-ante determination of the number of alternatives in the *choice set*;

3. alternative profiles built on linear ED approaches;

4. assignment of the profiles so derived to choice set with different combinatorial devices.

Generally, attributes and levels are selected on the basis of both the objective of the study and the information from *focus group*. The number of *choice sets* each respondent is asked to evaluate ranges from 4 to 16 and the number of alternatives in each choice set from 2 to 7. The most frequent *choice set* composition (see Table 2) is that of two alternatives and the *status-quo* (2+sq), where typically the sq is added to ED alternatives, rather than being built into the overall design efficiency.

The allocation of alternatives in the single *choice set* is either randomized or follows the method in [13].

Only in few environmental economics studies [16, 52] is the criterion of maximizing the information matrix of the MNL the guiding principle for the derivation of the ED.

On the basis of these observations we can make a few considerations:

1. The observed delay with which factorial designs tend to be substituted with D-optimal designs might be due to a lack of persuasion on the efficiency gains derivable from the latter. Hence it is of interest to evaluate empirically, in a typical environmental valuation context, to how much such gains amount and how robust they are.

2. Amongst the various D-optimal designs algorithms the only ones that have been employed so far are those for MNL specifications. This is probably due to the fact that for these EDs predefined macro are available in SAS and are well documented [40]. These macros require as input the number of attributes (and their respective levels), of alternatives, of choice sets, the specification for indirect utility, and a guess of the *a-priori* parameter estimates β.

   On the other hand, for Bayesian EDs no pre-packaged software procedures seem to be available and the researcher needs to code the algorithm for each context of study, which requires a considerable effort and time commitment. It is therefore important to empirically investigate the gains in efficiency achievable with these more elaborate designs to be able to assess when it is worth employing them in the practice of environmental valuation.

3. The dominance in the environmental valuation literature of the 2+sq choice task format, which as demonstrated elsewhere in the literature [28, 29, e.g.] is prone to give rise to

status-quo bias, introduces a specific issue of interest to environmental economists. When such bias is present it is often inadequately addressed by means of a simple inclusion of an alternative-specific-constant in the MNL specification [60], and it requires either nested logit cite cases or more flexible specifications.

4. Finally, an empirical investigation should also explore which ED approach is most robust with regards to a wrong or poor *a-priori* assumption about the model values of β .

4   Methods

In our empirical investigation[9] we compare four different ways of deriving an ED for discrete CEs for the MNL specification. We report them here in order of growing complexity of derivation.

4.1   The *shifted* design

We chose to employ a shifted design rather than the most common fractional factorial orthogonal design (FFOD). We felt this has already been thoroughly assessed by Lusk and Norwood [48]. Furthermore, based on the results of [18], the shifted design seem to produce a better performance than the FFOD, and to be just as simple to derive. The shifted design was originally proposed by [13] and it is based on the implicit assumption that the *a-priori* values of $\beta_p = 0$. Given this assumption they consider designs for general linear models and propose a procedure to assign alternatives to choice sets. The work by Burgess and Street shows how to shift so as to obtain optimal designs.

The basic ED is derived from a FFOD. Alternatives so derived are allocated to choice-sets using *attribute-based* strategies. Within this category we use a variant of the shifting technique whereby the alternatives produced by the FFOD are used as seeds for each choice set. This strategy gives the possibility to use module arithmetic which "shifts" the original columns of the FFOD in such a way that all attributes take different levels from those in the original design. We refer to this ED as the "shifted" design. For example, in our case from an initial FFOD (the

---

[9] All is necessary to replicate this study (Gauss codes, experimental designs, etc.) are available from the authors.

seed) all attribute levels were shifted by one unit. Those originally at the highest level were set to the lowest.

4.2 D_p-optimal design

A design potentially more efficient than the *shifted* one is obtainable by making use of *a-priori* information on β and deriving a D_p-optimal design through the maximization of the information matrix for the design under the MNL model assumptions, which is given by:

$$\mathbf{I}(\mathbf{X}, \beta) = \left(-\frac{\partial^2 \ln L\left(\beta\right)}{\partial\beta\partial\beta'} \right) = \sum_{s=1}^{S} \mu^2 \mathbf{X}'_s \left(\mathbf{P}_s - \mathbf{p}_s\mathbf{p}'_s\right) \mathbf{X}_s, \tag{1}$$

where s denotes choice-situations, $\mathbf{X}_s = [x_{1s}, \ldots, x_{Js}]'$ denotes the choice attribute matrix, $\mathbf{p}_s = [p_{1s}, \ldots, p_{Js}]'$ denotes the vector of the choice probabilities for the jth alternative and $\mathbf{P}_s = \mathrm{diag}[p_{1s}, \ldots, p_{Js}]$ with zero off diagonal elements and $p_{js} = e^{\mu V_j}(\sum_{i=1}^{J} e^{\mu V_i})^{-1}$.[10]

A widely accepted [42, 57] scalar measure of efficiency in the context of EDs for models non-linear-in-the-parameter is the D-criterion, which is defined as:

$$D\text{-criterion} = \left\{ \det\left(\mathbf{I}\left(\beta\right)^{-1} \right) \right\}^{1/k}, \tag{2}$$

where k is the number of attributes. We employed the modified Federov algorithm proposed by [67] to find the arrangement of the levels in the various attributes in $\mathbf{X}$ such that the D-criterion is minimized when $\beta = \beta_p$. Such algorithm is available in the macro "%ChoicEff", in SAS v. 9 [see 40, for details].

4.3 D_b-optimal designs

While the D_p-optimal design does not incorporate the uncertainty which invariably surrounds the values of β, the D_b-optimal design allows the researcher to do so.

---

[10] As commonly done in these estimations the scale parameter μ was normalized to 1 for identification.

On the other hand the derivation of Bayesian designs is computationally more demanding, and perhaps explains why previous studies have neglected them. However, they are appealing because they show robustness to other design criteria for which they are not optimized [39].

For Bayesian designs the criterion to minimize is the D_b, which is the expected value of the D-criterion with respect to its assumed distribution over β or $\pi(\beta)$:

$$D_b\text{-criterion} = E_\beta \left[\left\{ \det \mathbf{I}(\beta)^{-1} \right\}^{1/k} \right] = \int_{\Re^k} \left\{ \det \mathbf{I}(\beta)^{-1} \right\}^{1/k} \pi(\beta) d\beta. \tag{3}$$

In practice this is achieved by approximating via simulation the value of D_b: one draws R sets of values β^r from the *a-priori* $\pi(\beta)$ and computes the average of the simulated D-criterion over the R draws:

$$\tilde{D}_b = \frac{1}{R} \sum_{r=1}^{R} \left\{ \det \mathbf{I}(\beta^r)^{-1} \right\}^{1/k}. \tag{4}$$

Bayesian approaches always allow one to incorporate the information from the *a-priori* distribution, and in this application we compared two D_b-optimal designs, one with a relatively poor information on the prior implemented by a uniform distribution [38], and the second with a more informative prior implemented by means of a multivariate normal centered on the parameter estimates from the pilot study, and with variance covariance matrix as estimated from the pilot [57].

4.3.1 D_b-optimal design with *weakly-informative* prior

The distributional assumption about the prior in this case is uniform $\pi(\beta) = U[-a, a]^k$ where $-a$ and a are the extreme values of the levels of the choice attributes. We refer to this design throughout the paper as D_b^k-optimal.

4.3.2 D_b-optimal design with *informative* prior

We refer to this design as D_b^s-optimal. Following [57] we assume the prior to be distributed $\pi(\beta) = N(\hat{\beta}, \hat{\Omega})$. While [57] derive the $\hat{\beta}$ and $\hat{\Omega}$ estimates on the basis of managers' expectations, we instead derive the values from data obtained from a pilot study, as these are typically available in environmental valuation studies. The pilot data were in turn obtained on the basis

of a fractional factorial orthogonal main effects design. The search for efficiency over $\mathbf{X}$ was implemented by using the RCS algorithm developed by Sándor and Wedel [57, 58].

4.3.3   Criteria for comparing designs

Some synthetic criteria are available for design comparison. These depend on the coding of choice and on the values of the β vector. We choose to report the D-criterion in equation 2 and the A-criterion:

$$A\text{-criterion} = \left\{ trace \left(\mathbf{I} \left(\beta \right)^{-1} \right) \right\}^{1/k}. \tag{5}$$

Given some choice of parameter values and of coding, the lower this values the more informative the design matrix, and hence the more efficient the design.

Finally, as a measure of balancedness and choice complexity we report a common measure of entropy for the design, computed as:

$$\mathcal{E}(\mathbf{X}, \beta) = - \sum_{s=1}^{S=18} \sum_{j=1}^{J=3} p_{js}(\mathbf{X}, \beta) \ln(p_{js}(\mathbf{X}, \beta)) \tag{6}$$

where j denotes alternatives and s denotes choice-situations in the design. The higher this value, the higher the complexity of the choice set. These values are reported in Table 3 and show that when evaluated with dummy coding (the most frequent coding in environmental economics for qualitative attributes) and at the parameter values of the MNL model in Table 4, the most efficient design (a-priori) is the D_p-optimal and the least efficient is the D_b^s-optimal, which is also the one associated with largest entropy.

4.4   Design of Monte Carlo experiment

To assess the difference between the alternative designs, we have drawn inspiration from a study about willingness to pay (WTP) for four rural landscape components for a government programme designed to improve rural landscape. The four components were mountain land (ML), stonewalls (SW), farmyard tidiness (FT) and cultural heritage features (CH) [59]. In this CE study all the attributes where potentially improved by the proposed policy with two degrees of intensity which we succinctly describe as "some action" and "a lot of action". In the original

study respondents were obviously given photographic representations of how such levels of improvement would differ from each other and the status-quo. The interested reader is referred to an extensive report available for this study [50].

Inspired by this study, our Monte Carlo experiment is designed to investigate the relative performance of four designs under the assumption of an expected MNL specification. Such expectation is the most frequent in this context of analysis.

However, after the data collection, the data may display evidence corroborating other more flexible specifications. In particular, we examine the case of a flexible error component model with alternative specific constant, which produces a correlation structure across utilities analog to the nested logit. This specification is motivated and examined in some detail in [60] and it accounts for status-quo effects in a more flexible fashion than the more commonly employed nested logit specification.

In our CE the error component approach takes the following basic utility form[11]:

$$
\begin{aligned}
U(c_1) &= \beta \mathbf{x}_{c_1} + \tilde{u}_{c_1} = \beta \mathbf{x}_{c_1} + \varepsilon_{c_1} + u_{c_1}, \\
U(c_2) &= \beta \mathbf{x}_{c_2} + \tilde{u}_{c_2} = \beta \mathbf{x}_{c_2} + \varepsilon_{c_2} + u_{c_2}, \\
U(sq) &= Asc + \beta \mathbf{x}_{sq} + u_{sq},
\end{aligned}
\tag{7}
$$

where, in our case, $\varepsilon_{c_1} = \varepsilon_{c_2} \sim N(0, \sigma^2)$ are additional error components to the conventional Gumbel-distributed u_{c_1} and u_{c_2}, thereby leading to the following error covariance structure:

$$
Cov(\tilde{u}_{c_1}, \tilde{u}_{c_2}) = \sigma^2, \ \ Var(\tilde{u}_{c_1}, \tilde{u}_{c_2}) = \sigma^2 + \pi^2/6,
\tag{8}
$$

$$
Cov(\tilde{u}_{c_j}, \tilde{u}_{sq}) = 0, \ \ Var(\tilde{u}_{c_j}, \tilde{u}_{sq}) = \pi^2/6, \ \ j = 1, 2;
\tag{9}
$$

where $\tilde{u}_{c_j} = \varepsilon_{c_j} + u_{c_j}$. Note that this is an analog of the nested logit model in the sense that it allows for correlation of utilities across alternatives in the same nest, but different correlation for those across nests. However, there is no IIA restriction, and the Asc captures any remaining systematic effect on the sq alternative. With $\sigma^2 = 0$ the MNL model is obtained.

Conditional on the presence of the error component ε_j the choice probability is logit, and

---

[11] In fact, as expanded upon by [12], [65], [32], more general forms than this may be empirically appealing.

the assumption above leads to the following expression for each marginal choice probability:

$$P(i) = \int_{\varepsilon} \pi(i|\varepsilon) f(\varepsilon|\theta) d\varepsilon \text{ and, hence, substituting in:}$$

$$P(i) = \int_{-\infty}^{+\infty} \frac{e^{\beta \mathbf{x}_i + \varepsilon_i}}{\sum_j e^{\beta \mathbf{x}_j + \varepsilon_j}} \phi(0, \sigma^2) d\varepsilon, \quad j = c_1, c_2, sq, \tag{10}$$

where $\phi(\cdot)$ is the normal density, and $\varepsilon_j = 0$ when $j = sq$. Estimation of $\hat{\beta}$ and $\hat{\sigma}^2$ is obtained by maximum simulated likelihood [65].

The effects of the alternative designs considered are assessed by Monte Carlo experiments.

The evaluation of the performance of the four designs in the case of an incorrectly assumed data generating process (DGP) gives us the chance of examining the robustness of their performance to the MNL specification assumed *a-priori*, which is the one for which standard non-linear designs are commercially available.

Short of the differences in the form of the DGP and the alternative ED, the steps of the experiment are the same. We create $r = 1, 2, 3, \cdots, R = 550$ samples of 100, 250 and 500 observations under two different DGP: the MNL and the error components model with alternative specific constant (abbreviated henceforth with KL-Asc).

1. At each replication r individual counterfactual responses y_{ir} are produced by identifying the alternative j associated with the largest utility value $U(\beta, \varepsilon, x_j)$, where the β values are the true one and are reported in table 4, while the errors ε are drawn from the adequate distributions (Gumbel for MNL; Gumbel *and* Normal for the KL-Asc).

2. The counterfactual y_{ir} produced for the whole sample are used to get maximum likelihood or maximum simulated likelihood estimates of $\widehat{\beta}_r$ of β. Then a series of indicators of estimation performance are computed. For the sake of comparisons across models—and given their relevance in non-market valuation—we focus on the marginal rates of substitutions with the money coefficient:

$$\widehat{MRS}_r = \widehat{\tau}_r = -\frac{\widehat{\beta}_r}{\widehat{\gamma}_r}. \tag{11}$$

And then we report some additional indicators.

(a) First, we report the average values of their distribution across replications:

$$\overline{MRS} = \tfrac{1}{R} \sum_{1}^{R} \tau^r, \ r = 1, \cdots, 550 \tag{12}$$

and the associated standard deviations.

(b) Secondly we report the mean squared error:

$$\overline{MSE} = \tfrac{1}{R} \sum_{1}^{R} (\widehat{\tau}_r - \tau)^2, \ r = 1, \cdots, 550 \tag{13}$$

where τ is the true value and $\widehat{\tau}_r$ is the rth estimated in the experiment. Everything else equal the design with lowest $\overline{MSE}$ is the one with the smallest empirical bias.

(c) The third measure considered is the average of the absolute relative error:

$$\overline{RAE} = \tfrac{1}{R} \sum_{r}^{R} | \ (\widehat{\tau}_r - \tau)/\tau \ |. \tag{14}$$

This gives a relative measure of the error, which can be easily mapped into percent of error of the "true" marginal *WTP* for the attribute.

(d) Finally, as a measure of efficiency we count the percent of MRS values falling within a 5% interval of the true value:

$$\Gamma_{0.05} = \tfrac{1}{R} \sum_{r}^{R} I(\widehat{\tau}_r \in \tau \pm \tau \times 0.05). \tag{15}$$

where $I(\cdot)$ is an indicator function. This gives an idea of the empirical efficiency of each design.

5 Monte Carlo Results

A large amount of information is produced by the experiments and here we focus only on the estimation of the coefficient for the attribute that showed highest implicit value in the original study[12] [see Table n. 4 and 59]. This attribute was expressed at two levels of policy action "some" (ML_some) and "a lot of" (ML_alot) and concerned the visual aspect of mountainous

---

[12] Qualitatively similar results were obtained for the other coefficients.

rural land (ML). Tables 5 and 6 display the results from the empirical distributions of the MRS and illustrate the sensitivity of these to the four different designs.

5.1   Correct specification and correct design information

Table 5 present the results for "the best of the worlds" in which the DGP, the *a-priori* distributions of parameters and the specification used in the estimation are all the "correct" ones.

Observing the values for the efficiency indicators $\Gamma_{0.05}$ and $\overline{MSE}$ one can detect how the D_b^s-optimal design is the most efficient at all sample sizes. As expected, efficiency increases with sample size. Similar conclusions can be derived from the values of $\overline{RAE}$. However, the liner $shifted$ design at small sample sizes $N = 100$ gives a similar performance, and certainly superior to that of the D_b^k-optimal design.

A graphical illustration of what happens at large sample sizes ($N = 500$) is reported in Figure 1 where we show the kernel-smoothed [9] distributions of $MRS_{ML_{alot}}$ for all four designs. Notice that while the D_b^k-optimal design is centered on the true value, it shows a stronger variability than the other designs. The D_p-optimal and the D_b^s-optimal respectively underestimate and overestimate by very little, while the *shifted* design produces significant overestimates at this sample size.

Analog conclusions can be drawn from an inspection of Figure 2, where we report the absolute relative error ($RAE_{ML_{tot}}$). Suppose a decision rule was to be incorrectly taken if the relative absolute error is larger than 20 or 30%. From the plot in Figure 2 it is apparent that the umber of cases in which this would occur is highest for the shifted design (continuous line). In conclusion, in this case—in which the DGP is coherent with the *a-priori* expectations and estimates are derived under the correct specification—the two best performing designs are those built by assuming the least uncertainty around the true parameters, that is the D_p-optimal and the D_b^s-optimal.

Given the difficulty inherent in the computation of the latter, however, one would expect the former (that can be obtained with the macro "`%Choiceff`" in SAS) to be more frequently employed, as our review has shown.

5.2 Incorrect specification, but correct design information

As a way to investigate the sensitivity of these results to the quality of *a-priori* assumptions—where for *a-priori* here we refer to the information available in the pre-design and estimation phase—we now turn our attention to the case in which the estimation makes use of a mis-specified model, but the D-efficient experimental designs are correctly informed. The Monte Carlo statistics for such a case are reported in Table 6, where for the mis-specified model we employ the flexible error component model with Asc for the SQ (KL-Asc) while the true model is a MNL. The values show that in this case too at medium (N =250) and large (N =500) sample sizes the best performance is obtained by the D_b^s-optimal design. The one with *weakly-informed* prior (D_b^k-optimal) is the second best performer, while the non Bayesian MNL design (D_p-optimal) is dominated by the one optimized for linear specifications (*shifted* design) at sample sizes smaller than 500.

The fact that the Bayesian (informed *and* weakly-informed) designs are the most robust in the context of correct DGP prediction come across best in observing the kernel plots of absolute relative error distributions in Figure 3, which again refers to the large sample size scenario.

There is therefore evidence that as long as the *a-priori* design information is "good" the Bayesian designs are robust to mis-specifications in the estimation phase; under all criteria the shifted design is preferable to the D_p-optimal at small sample sizes; and that even at large sample sizes the latter produces large errors more frequently than the shifted design (Figure 3).

5.3 Correct specification, but incorrect design information

What happens when—instead—the *a-priori* information incorporated in the D-efficient design is "poor" and the model specification is right? Of course, under this category falls a very large number of cases, but as a way of exploring this instance we repeated the experiment with the real DGP formulated as a KL-Asc and correct estimation assumptions, but with incorrect prior (MNL) for the experimental design.

The choice of a the error component model KL-Asc is motivated by the fact that it allows for a greater variance and correlation in the errors associated with the utilities of experimentally designed alternatives than in those associated with the status-quo alternative. This is an

often-encountered situation in environmental valuation, which results in nested logit models providing a better fit than conditional logit models [60]. The KL-asc provides a similar covariance structure to the nested logit model with a degenerate nest for the status-quo alternative. It is also more flexible and has an objective function globally concave in the parameter space, it is hence deemed appropriate for a Monte Carlo simulation.

For the sake of brevity we do not report the results in a tabular form,[13] but the findings are illustrated in Figure 4: in this instance the most robust design is the one not informed at all, i.e. the *shifted* design. The more information is built into the design instead, the higher the degree of bias produced, even under correct specification. Of course, it is easy to anticipate these results rationally, however, this investigation provides ground for some less obvious considerations.

First of all, it seems that the efficiency gains made available from more advanced non-linear and Bayesian-informed designs is only available in cases in which the *a-priori* design information is good *and* this outcome is robust to substantial model mis-specification.

In the absence of good quality *a-priori* design information to be built into the design, researchers are perhaps better off using more rudimentary designs, even when these are only optimized for linear models, which is exactly what the profession has been doing, perhaps inadvertently.

6   Conclusions

Data from discrete choice experiments for the purpose of environmental valuation are predominantly analyzed by means of highly non-linear specifications of the multinomial logit family. Yet, a review of the published literature in environmental valuation discloses a prevailing use of experimental designs produced for linear-in-the-parameters, rather than for non-linear-in-the-parameters models, without any built-in *a-priori* information on the parameter values. We reviewed various notions of D-efficiency in the experimental design literature focussing on design for multinomial logit assumptions, and on how these can be improved by using *a-priori* information.

Then, by means of Monte Carlo experiments—and inspired by the results and structure of

---

[13] These are available from the authors.

a real-world application—we explored the relative performance of four alternative approaches to derive experimental designs. The simplest design to derive is the (*shifted*), and it is based on a modification of a conventional fractional factorial main effect orthogonal design. The other three were specifically optimized for the highly non-linear multinomial logit model, and contained various form of a *a-priori* information on the underlying parameter values. The D_p-optimal design did not allow for uncertainty on parameter values, while the two Bayesian designs did, with more uncertainty for the D_b^k-optimal, and with the amount of information that typically becomes available from a standard pilot study—in the form of parameter estimates and their variance-covariance matrix—built into the D_b^s-optimal design.

The features of the Monte Carlo experiments (sample size, data generating processes, choice-set construction, etc.) were chosen so as to reflect the reality commonly faced by practitioners in environmental valuation as derived from a review of published studies.

The results from the experiment showed that efficiency gains are available from the use of Bayesian D-efficient designs for non-linear-in-the-parameters models. These gains are substantial for parameter estimates of important attributes ("a lot of" action in our empirical study), but much less so for parameters of less relevant attributes ("some" action).

For important attributes and with good *a-priori* information on the values of the unknown parameters gains can be available at all sample sizes, as shown in the results for the D_b^s-optimal design in Tables 5 and 6.

Even by building into the design relatively poor information (D_b^k-optimal design) on the parameter values, efficiency gains become attractive only at medium to large sample sizes ($N > 250$) but they are more significant when both:

- the *a-priori* information on the parameters provided by the pilot is of good quality;

- *and* the data generating process is consistent with the specification chosen in the estimation.

However, when these conditions fail, the best performance is obtained with the most "rudimentary" of the designs we employed (the *shifted* design), which is derived from the common fractional factorial orthogonal design dominating the state of practice. This design ignores any information on the parameters of the true DGP.

This result suggests that—in as much as *a-priori* information on parameter values has been ignored at the stage of design construction—environmental economists might well not have missed out too much in terms of efficiency gains, and even in bias, as a consequence of the lag with which they have been adopting recent advances in experimental design construction.

On the other hand, this points to an area of potentially interesting and valuable research on methods of design construction that do incorporate *a-priori* information progressively and cumulatively at different stages of the survey. This could be of particular interest as new computer-assisted technology becomes increasingly used in choice-experiment surveys and especially given the encouraging results that bid design updating produced in the field of contingent valuation [49, 54].

Constructing designs using adaptive techniques can be a valuable strategy in choice-experiment surveys [51]. For example, one can systematically incorporate the information becoming available as the sampling progresses to derive gradually more tailored designs. The type of information needed are the parameter estimates and their variance-covariance into successive designs. A similar suggestion was put forward by Kanninen [36] for the cost attribute. On the basis of our results we speculate that this updating should possibly involve more attributes, such as those that appear to become dominant, or even all of them as we did in this application. More research on the most effective strategy to gradually incorporate such information during survey administration is needed.

Another area of potential interest may be that of deriving experimental designs based on efficiency criteria that most directly recognize the ultimate purpose of attribute based valuation studies. The focus on efficient estimation of monetary values, typically a non-linear function of parameter estimates, should be explicitly addressed in the measure of efficiency. This could translate—for example—in the maximization of the information matrix for the vector of marginal value estimate, rather than that for the parameters of the indirect utility function.

While statistical efficiency remains an important goal, more research is necessary to evaluate whether this additional efficiency comes at too high a cost in terms of increased choice complexity to respondents. This issue requires field tests and can only be partially addressed by means of simulation tools.

Finally, given the importance that discriminating between behaviorally plausible and hence

likely specifications in logit models has on estimate efficiency, future research should also focus on the construction of designs able to discriminate between competing specifications. Seminal research of this kind in the context of multivariate linear models is already available [5]. Future work in this direction can allow researchers to address the issue of uncertainty about logit model specifications from the onset into the experimental designs.

References

[1] A. Alberini. Optimal designs for discrete choice contingent valuation surveys: Single-bound, double bound and bivariate models. *Journal of Environmental Economics and Management*, 28:287–306, 1995.

[2] A. Alberini and R. T. Carson. Efficient threshold values for binary discrete choice contingent valuation surveys and economic experiments. Discussion Paper No. QE 93-14, Resources for the Future, Quality of the Environment Division, Washington, DC, 1993.

[3] N. Arora and J. Huber. Improving parameters estimates and model prediction by aggregate customization in choice experiments. *Journal of Consumer Research*, 28(2):273–283, 2001.

[4] A. Atkinson and A. N. Donev. *Optimum Experimental Designs*. Claredon Press, Oxford, UK, 1992.

[5] A. C. Atkinson and V. V. Federov. The design of experiments for discriminating between two rival models. *Biometrika*, 62:57–70, 1975.

[6] R. K. Blamey, J. W. Bennett, J. J. Louviere, M. D. Morrison, and J. C Rolfe. Attribute causality in environmental choice modelling. *Environmental and Resource Economics*, 23:167–186, 2002.

[7] C.J. Blemier and J.M. Rose. Efficient designs for alternative specific choice-experiments. Working Paper ITLS-WP-05-04, 2005.

[8] M.C. Blemier, J.M. Rose, and D.A. Hensher. Costructing efficient stated choice experiments allowing for differences in error variances across subsets of alternatives. *Submitted to Transportation Research B*, 2005.

[9] A. W. Bowman and A. Azzalini. *Applied Smoothing Techniques for Data Analysis: the Kernel Approach with S-Plus Illustrations*. Oxford University Press, Oxford, U.K., 1997.

[10] P. C. Boxall and W. L. Adamowicz. Understanding heterogenous preferences in random utility models: A latent class approach. *Environmental and Resource Economics*, 23:441–446, 2002.

[11] P. C. Boxall, W. L. Adamowicz, J. Swait, M. Williams, and J. J. Louviere. A comparison of stated preference methods for environmental valuation. *Ecological Economics*, 18:243–253, 1996.

[12] D. Brownstone and K. Train. Forecasting new product penetration with flexible substitution patterns. *Journal of Econometrics*, 89:109–129, 1999.

[13] D. S. Bunch, J. J. Louviere, and D. Anderson. A comparison of experimental design strategies for multinomial logit models: The case of generic attributes. Working Paper, 11, Graduate School of Management, University of California, Davis, 1996.

[14] L. Burgess and D. Street. Optimal designs for 2^k choice experiments. *Communications in Statistics Theory and Method*, 32:2185–2206, 2003.

[15] L. Burgess and D. Street. Optimal designs for choice experiments with asymmetric attributes. *Journal of Statistical Planning and Inference*, 134:288–301, 2005.

[16] F. Carlsson, P. Frykblom, and C. Liljenstolpa. Valuing wetland attributes: an application of choice experiment. *Ecological Economics*, 47:95–103, 2003.

[17] F. Carlsson and P. Martinsson. Do hypothetical and actual marginal willingness to pay differ in choice experiments? *Journal of Environmental Economics and Management*, 41:179–192, 2001.

[18] F. Carlsson and P. Martinsson. Design techniques for stated preference methods in health economics. *Health Economics*, 12:281–294, 2003.

[19] S. Caussade, J. de D. Ortúzar anf L. I. Rizzi, and D. A. Hensher. Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B*, 39:621–640, 2005.

[20] K. Chaloner and I. Verdinelli. Bayesian experimental design: A review. *Statistical Science*, 10(3):273–304, 1995.

[21] J.R. DeShazo and G. Fermo. Designing choice sets for stated preference methods: The effects of complexity on choice consistency. *Journal of Environmental Economics and Management*, 44:123–143, 2002.

[22] D. Firth and J.P. Hinde. Parameter neutral optimal design for non-linear models. *Journal of the Royal Statistical Society, Series B (Methodological)*, 59:799–811, 1997.

[23] V. Foster and S. Mourato. Elicitation format and sensitivity to scope. *Environmental and Resource Economics*, 24:141–160, 2003.

[24] U. Grasshoff, H. Grossmann, H. Holling, and R. Schwabe. Optimal paired comparison designs for firt-order interactions. *Statistics*, 37:373–386, 2003.

[25] U. Grasshoff, H. Grossmann, H. Holling, and R. Schwabe. Optimal designs for main effects in linear paired comparison models. *Journal of Statistical Planning Inference*, 2004. In press.

[26] P. E. Green and V. R. Rao. A rejoinder to 'how many rating scales and how many categories shall we use in consumer research' - a comment. *Journal of Marketing*, 35:61–62, October 1971.

[27] H. Grossman, H. Holling, and R. Schwabe. Advances in optimum experimental design for conjoint analysis and discrete choice models. *Econometrics Models in Marketing*, 16:93–117, 2002.

[28] M. E. Haaijer. *Modeling Conjoint Choice Experiments with the Probit Model*. PhD thesis, University of Groningen, Labyrint Publications, The Netherlands, April 1999.

[29] M. E. Haaijer, W.A. Kamakura, and M. Wedel. The no-choice alternative in conjoint choice. *International Journal of Market Research*, 43(1):93–106, 2001.

[30] N. Hanley, R. E. Wright, and V. Adamowicz. Using choice experiments to value the environment. *Environmental and Resource Economics*, 11(3-4):413–428, 1998.

[31] N. Hanley, R. E. Wright, and G. Koop. Modelling recreation demand using choice experiments: Climbing in Scotland. *Environmental and Resource Economics*, 22:449–466, 2002.

[32] J. A. Herriges and D. J. Phaneuf. Inducing patterns of correlation and substitution in repeated logit models of recreation demand. *American Journal of Agricultural Economics*, 84:1076–1090, 2002.

[33] T. P. Holmes and W. L. Adamowicz. Attribute-based methods. In P. A. Champ, K. J. Boyle, and T. C. Brown, editors, *A primer on nonmarket valuation*, chapter 6, pages 171–219. Kluwer Academic Publisher, 2003.

[34] J. Huber and K. Zwerina. The importance of utility balance in efficient choice designs. *Journal of Marketing Research*, 39:214–227, 1996.

[35] J. Jin, Wang Z., and S. Ran. Comparison of contingent valuation and choice experiment in solid waste management programs in macao. *Ecological Economics*, Forthcoming, 2005.

[36] B. J. Kanninen. Optimal experimental design for double-bounded dichotomous choice contingent valuation. *Land Economics*, 69(2):138–146, 1993.

[37] B. J. Kanninen. Optimal designs for multinomial choice experiment. *Journal of Marketing Research*, 39:214–227, May 2002.

[38] R. Kessels, P. Goos, and M. Vandebroek. Comparing algorithms and criteria for designing Bayesian conjoint choice experiments. Working Paper, Department of Applied Economics, Katholieke Universiteit Leuven, Belgium, 2004.

[39] R. Kessels, P. Goos, and M. Vandebroek. A comparison of criteria to design efficient choice experiments. *Journal of Marketing Research*, Forthcoming:*, 2005.

[40] W. F. Kuhfeld. Marketing research. methods. In *SAS Experimental Design, Choice, Conjoint, and Graphical Techniques*. SAS Institute, Cary, NC, USA, http://support.sas.com/techsup/tnote/tnote_stat.html, 2005.

[41] W. F. Kuhfeld and R. D. Tobias. Large factorial designs for product engineering and marketing research applications. *Technometrics*, 47:132–141, May 2005.

[42] W. F. Kuhfeld, R. D. Tobias, and M. Garratt. Efficient experimental design with marketing research applications. *Journal of Marketing Research*, 31:545–557, November 1994.

[43] K. Lancaster. A new approach to consumer theory. *Journal of Political Economy*, 84:132–157, 1966.

[44] K. Lancaster. Operationally relevant characteristics in the theory of consumer behavior. In M. Peston and B. Corry, editors, *Essays in honor of Lord Robbins*. Weidenfels and Nicholson, London, 1972.

[45] A. G. Lazari and D. A. Anderson. Designs of discrete choice set experiments for estimating both attribute and availability cross effects. *Journal of Marketing Research*, 31:375–383, 1994.

[46] J. J. Louviere and G. Woodworth. Designs analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data. *Journal of Marketing Research*, 20:350–367, 1983.

[47] J.J. Louviere and D.A. Hensher. Using discrete choice models with experimental design data to forecast consumer demand for unique cultural event. *Journal of Consumer Research*, 10:348–361, 1983.

[48] J.L. Lusk and F.B. Norwood. Effect of experimental design on choice-based conjoint valuation estimates. *American Journal of Agricultural Economics*, 87:771–785, 2005.

[49] H. Nyquist. Optimal designs of discrete response experiments in contingent valuation studies. *Review of Economics and Statistics*, 74:559–563, 1992.

[50] T. O'Leary, A. McCormack, G. Hutchinson, D. Campbell, R. Scarpa, and B. Riordan. Landscape impact of REPS: a quantitative assessment. Report to the Department of Agriculture and Food, Dublin, Ireland, 2005.

[51] D. Raghavarao and J.B. Wiley. Design strategies for sequential choice experiments involving economic altenatives. *Journal of Statistical Planning and Inference, in press*, 2005.

[52] M. X. V. Ròdriguez and C. J. Leòn. Altruism and the economic values of environmental and social policies. *Environmental and Resource Economics*, 28:233–249, 2004.

[53] J. Rolfe, J. Bennett, and J. Louviere. Choice modelling and its potential application to tropical rainforest preservation. *Ecological Economics*, 35:289–302, 2000.

[54] K. Rollins, W. Wistowsky, and M. Jay. Wilderness canoeing in Ontario: using cumulative results to update dichotomous choice contingent valuation offer amounts. *Canadian Journal of Agricultural Economics*, 45:1–16, 1997.

[55] J. M. Rose and M. C. J. Bliemer. The design of stated choice experiments: The state of practice and future challenges. Working Paper ITS-WP-04-09, Institute of Transport Studies, University of Sydney and Monash University, 2004.

[56] K. Sælensminde. The impact of choice inconsistencies in stated choice studies. *Environmental and Resource Economics*, 23:403–420, 2002.

[57] Z. Sándor and M. Wedel. Designing conjoint choice experiments using managers' prior beliefs. *Journal of Marketing Research*, 38:430–444, 2001.

[58] Z. Sándor and M. Wedel. Profile construction in experimental choice designs for mixed logit models. *Marketing Science*, 4:455–475, 2002.

[59] R. Scarpa, D. Campbell, and G. Hutchinson. Individual benefit estimates for rural landscape improvements: the role of sequential bayesian design and response rationality in a choice experiment study. Paper presented at the 14th Annual Conference of the European Association of Environmental and Resource Economics Bremen, 2005.

[60] R. Scarpa, S. Ferrini, and K. G. Willis. Performance of error component models for status-quo effects in choice experiments. In *Applications of simulation methods in environmental and resource economics*, chapter 13, pages 247–274. Springer, 2005.

[61] R. Scarpa, E.S.K. Ruto, P. Kristjanson, M. Radeny, and A.G. Drucker. Valuing indigenous cattle breeds in kenya: an empirical comparison of stated and revealed preference value estimates. *Ecological Economics*, 45:409–426, 2003.

[62] D. J. Street, D. S. Bunch, and B. Moore. Optimal designs for 2^k paired comparison experiments. *Comm. Statist. Theory Methods*, 30:2149–2171, 2001.

[63] D. J. Street and L. Burgess. Optimal and near-optimal pairs for the estimation of effects in 2-level choice experiments. *Journal of Statistical Planning and Inference*, 118:185–199, 2004.

[64] D. J. Street and L. Burgess. Optimal stated preference choice experiment when all choice sets contain a specific option. University of Technology, Sydney, 2004.

[65] K. Train. *Discrete Choice Methods with Simulation*. Cambridge University Press, New York, 2003.

[66] P. Wattage, S. Mardle, and S. Pascoe. Evaluation of the importance of fisheries management objectives using choice-experiments. *Ecological Economics*, 55:85–95, 2005.

[67] K. Zwerina, J. Huber, and W. F. Kuhfeld. A general method for constructing efficient choice designs. Working paper, Fuqua School of Business, Duke University, Durham, 1996.

7 Tables

Authors	Criterion	Definition	a-priori parameter	Algorithm
Lazari and Anderson, 1994	D	Det (XX)	-	Unspecified
Kuhfeld *et al.*, 1994	D	Det (XX)		Modified Fedorov
Bunch *et al.*, 1996	D	Det (XX)		
Huber and Zwerina, 1996	D_p	Det$\{I_{MNL}(X,\beta)^{-1}\}$	β_0	RS
Zwerina *et al.*, 1996	D_p	Det$\{I_{MNL}(X,\beta)^{-1}\}$	β_0	Modified Fedorov
Sandor and Wedel, 2001	D_b	$E\left[\text{Det}\{I_{MNL}(X,\beta)^{-1}\}\right]$	$N(\beta|\beta_0,\Sigma_0)$	RSC
Sandor and Wedel, 2002	D_b	$E\left[\text{Det}\{I_{RPL}(X,\beta)^{-1}\}\right]$	$N(\beta|\beta_0,\Sigma_0)$	RSC
Kanninen, 2002	D_s	Det$\{I_{MNL}(X,\beta)^{-1}\}$		Sequential update
Burgess and Street, 2003	D_p	Det$\{I_{MNL}(X,\beta)^{-1}\}$	β_0	
Kuhfeld, 2004	D_p	Det$\{I_{MNL}(X,\beta)^{-1}\}$	β_P	Modified Fedorov
Kessels *et al.*, 2004	D_b	$E\left[\text{Det}\{I_{MNL}(X,\beta)^{-1}\}\right]$	$\beta U[-1,1]_k$	Modified Fedorov

Tab. 1: Approaches to experimental design for discrete choice experiments.

Authors and paper	Number of Attributes	Choice task Alternatives	Choice tasks per respondent	Experimental Design	Model Specification	Sampled respondents
Boxall *et al.*, 1996 (EE)	6 $(4^4 2^2)$	$2 + sq$	16	-	MNL	271
Hanley *et al.*, 1998 (ERE)	4 (2^3)	$2 + sq$	4	-	MNL	181
Rolfe *et al.*, 2000 (EE)	7 $(8^1 4^6)$	$2 + sq$	16	-	MNL	105
Carlsson and Martinsson, 2001 (JEEM)	3 (3^3)	2	14	*D*-optimal Zwerina *et al.*, 1996	EVHL	350
Boxall and Adamowicz, 2002 (ERE)	5 (4^5)	$5 + sq$	8	Orthogonal main effects	LC RPL	620
Blamey *et al.*, 2002 (ERE)	6 $(4^4 3^1 5^1)$	$2 + sq$ $4 + 1$	4/8	Fractional factorial	MNL NL LC	480 620
DeShazo and Fermo, 2002(JEEM)	4/9	2/7	-	Factorial orthogonal randomised	Heteroskedastic MNL	1800/2100
Sælensminde, 2002 (ERE)	3/4	2	9	Fractional factorial orthogonal	Binary Logit	2568
Hanley *et al.*, 2002 (ERE)	6 $(4^4 2^1 6^1)$	$2 + sq$	4/8	Fractional factorial	MNL NL	267
Foster and Mourato, 2003 (ERE)	5	$2 + sq$		Fractional factorial (SPEED software)	MNL RPL	290
Horne and Petäjistö, 2003 (LE)	5 $(4^4 2^1)$	$2 + sq$	4/8	Fractional factorial	MNL	1296
Scarpa *et al.*, 2003 (EE)	5 $(3^3 2^2 4^1)$	$2 + sq$	6	Fractional factorial	MNL+Heterosk. RPL	300
Carlsson *et al.*, 2003 (EE)	7 $(2^5 3^1 4^1)$	$2 + sq$	4	Fractional factorial *D*-optimal OPTEX (SAS)	MNL RPL	5800
Rodrìguez and Leòn, 2004 (ERE)	6 $(3^2 4^2 2^2)$	$2 + sq$	8	*D*-optimal design Huber and Zwerina, 1996	MNL RPL EVHL	350
Wattage *et al.*, 2005 (EE)	3 $(3^2 4^1)$	16	—	Orthogonal main effects	MNL	30
Jin *et al.*, 2005 (EE)	3 $(2^3 4^1)$	$1 + sq$	8	Main effects factorial design	MNL	260

MNL=Multinomial Logit, EVHL=Extreme Value Heteroschedastic Logit, RPL=Random Parameter, NL=Nested Logit; LC= Latent Class (JEEM)= Journal of Environmental Economics and Management, (LE)= Land Economics, (ERE)= Environmental and Resource Economics, (EE)= Ecological Economics.

Tab. 2: Selected features of *choice experiment* studies in environmental economics.

Criteria	*Shifted*	D_p-optimal	D_b^k-optimal	D_b^s-optimal
D-criterion	0.03946	0.03858	0.03901	0.05194
A-criterion	1.00399	1.02008	1.13810	1.61498
$\mathcal{E}$	14.84	14.60	14.02	15.93

Tab. 3: Design comparison criteria evaluated at β_{MNL} and with dummy coding.

	MNL		KL-Asc	
Tax	–0.037	(–4.46)	–0.049	(–4.45)
Ml_alot	0.712	(13.84)	0.683	(10.28)
Ml_some	0.369	(7.06)	0.294	(4.03)
S_alot	0.711	(14.22)	0.662	(9.15)
S_some	0.495	(8.99)	0.413	(4.92)
P_alot	0.589	(11.90)	0.540	(7.47)
P_some	0.416	(8.01)	0.358	(4.80)
A_alot	0.545	(11.00)	0.481	(7.02)
A_some	0.443	(8.58)	0.370	(5.27)
Asc			–1.420	(–6.20)
σ			1.351	(7.73)

Asymptotic z-values in brackets.

Tab. 4: Maximum likelihood estimates of MNL model and maximum simulated estimates of KL-Asc model for the landscape study.

DGP: Multinomial logit

Assumption: Multinomial logit

	Shifted design			D_p-optimal			D_b^k-optimal			D_b^s-optimal		
	N=100	N=250	N=500	N=100	N=250	N=500	N=100	N=250	N=500	N=100	N=250	N=500
$\overline{MRS}_{ML_{alot}}$	21.38	21.25	21.44	19.36	19.56	19.73	19.03	19.45	19.35	20.41	20.43	20.25
	(4.25)	(2.81)	(1.95)	(4.78)	(2.95)	(2.04)	(5.47)	(3.36)	(2.38)	(4.18)	(2.52)	(1.85)
$\overline{MRS}_{ML_{par}}$	10.56	10.10	10.36	8.70	8.88	8.64	10.50	10.02	10.03	10.89	10.35	10.08
	(4.36)	(2.86)	(1.99)	(4.53)	(2.85)	(2.06)	(5.12)	(3.21)	(2.27)	(4.37)	(2.96)	(2.07)
$\overline{MSE}_{ML_{alot}}$	22.12	11.49	8.12	22.79	8.72	4.30	29.93	11.28	5.64	18.60	7.48	4.21
$\overline{MSE}_{ML_{par}}$	19.28	8.15	4.06	22.18	9.40	6.13	26.40	10.29	5.21	19.81	8.85	4.26
$\overline{RAE}_{ML_{alot}}$	0.19	0.14	0.12	0.20	0.12	0.09	0.22	0.14	0.10	0.18	0.11	0.08
$\overline{RAE}_{ML_{par}}$	0.35	0.23	0.16	0.38	0.24	0.20	0.40	0.26	0.18	0.36	0.23	0.17
$\Gamma_{(0.05, ML_{alot})}$	16	21	21	15	27	36	13	23	29	18	30	37
$\Gamma_{(0.05, ML_{par})}$	8	14	21	8	13	16	8	14	21	8	15	18

True WTP: $\overline{MRS}_{ML_{alot}} = 19.35$ $\overline{MRS}_{ML_{par}} = 10.02$

Tab. 5: Summary statistics from Monte Carlo experiment on data from DGP MNL and estimates from MNL specification.

	Shifted design			D_p-optimal			D_b^k-optimal			D_b^s-optimal		
	N=100	N=250	N=500	N=100	N=250	N=500	N=100	N=250	N=500	N=100	N=250	N=500
$\overline{MRS}_{ML_{alot}}$	23.04	22.46	22.58	23.51	22.79	22.29	19.54	19.85	19.71	21.52	21.25	21.05
	(5.33)	(3.22)	(2.31)	(6.61)	(4.14)	(2.89)	(6.89)	(4.10)	(2.79)	(4.79)	(2.84)	(2.07)
$\overline{MRS}_{ML_{some}}$	11.95	11.19	11.43	11.41	11.01	10.34	10.39	10.07	10.41	11.39	10.92	10.82
	(5.48)	(3.31)	(2.27)	(6.05)	(3.59)	(2.58)	(6.44)	(3.89)	(2.65)	(5.61)	(3.62)	(2.52)
$\overline{MSE}_{ML_{alot}}$	41.96	20.03	15.74	60.90	28.97	16.94	47.50	17.07	7.89	27.60	11.68	7.16
$\overline{MSE}_{ML_{some}}$	33.71	12.29	7.11	38.48	13.86	7.77	41.48	15.13	7.18	33.30	13.89	7.01
$\overline{RAE}_{ML_{alot}}$	0.27	0.19	0.18	0.32	0.22	0.17	0.28	0.17	0.12	0.22	0.14	0.11
$\overline{RAE}_{ML_{some}}$	0.46	0.28	0.22	0.48	0.30	0.22	0.49	0.30	0.21	0.46	0.29	0.21
$\Gamma_{(0.05,ML_{alot})}$	11	17	12	8	15	17	11	19	26	15	22	27
$\Gamma_{(0.05,ML_{some})}$	8	10	14	8	9	13	9	12	14	8	10	12

<div align="center">

DGP: Multinomial logit

Assumption: Kernel Logit-Asc

True WTP: $\overline{MRS}_{ML_{alot}} = 19.35$ $\overline{MRS}_{ML_{some}} = 10.02$

</div>

Tab. 6: Summary statistics from Monte Carlo experiment on data from DGP MNL and estimates from KL-Asc specification.
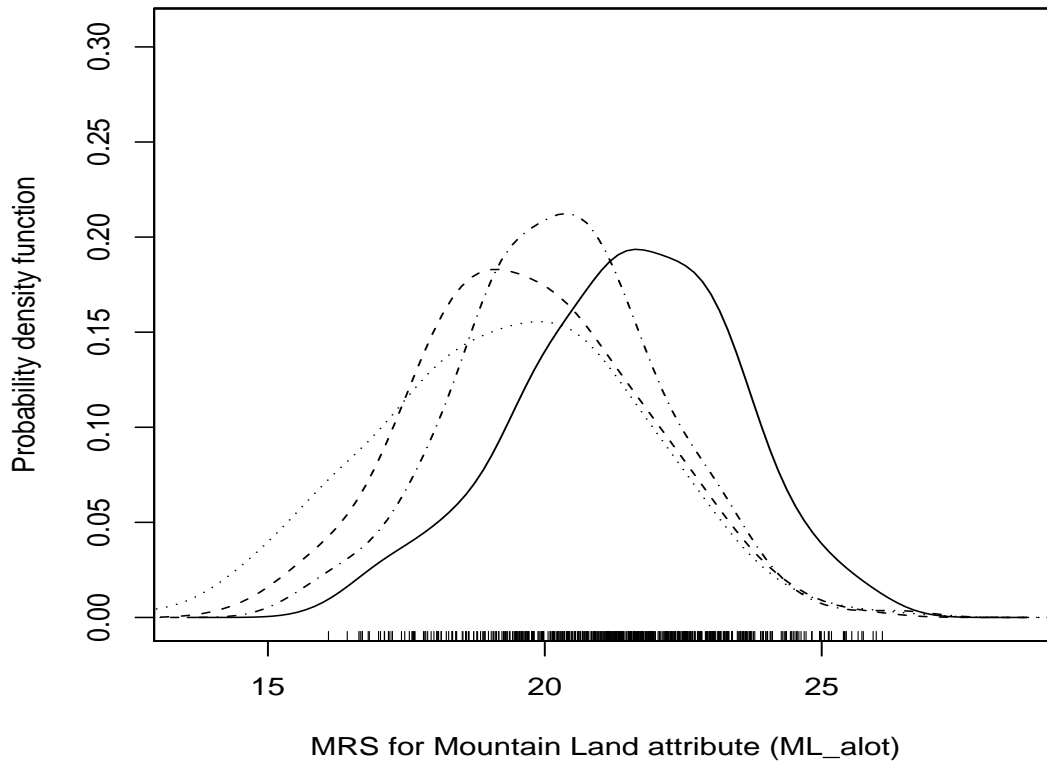
Fig. 1: DGP MNL and estimation MNL: *kernel-smoothed* distribution (optimal bandwidth) of the MRS

estimates of landscape attribute Mountain Land ML_{alot}.

Continous line: *shifted* design,

Dashed line: D_p-optimal design,

Dotted line: D_b^k-optimal design,
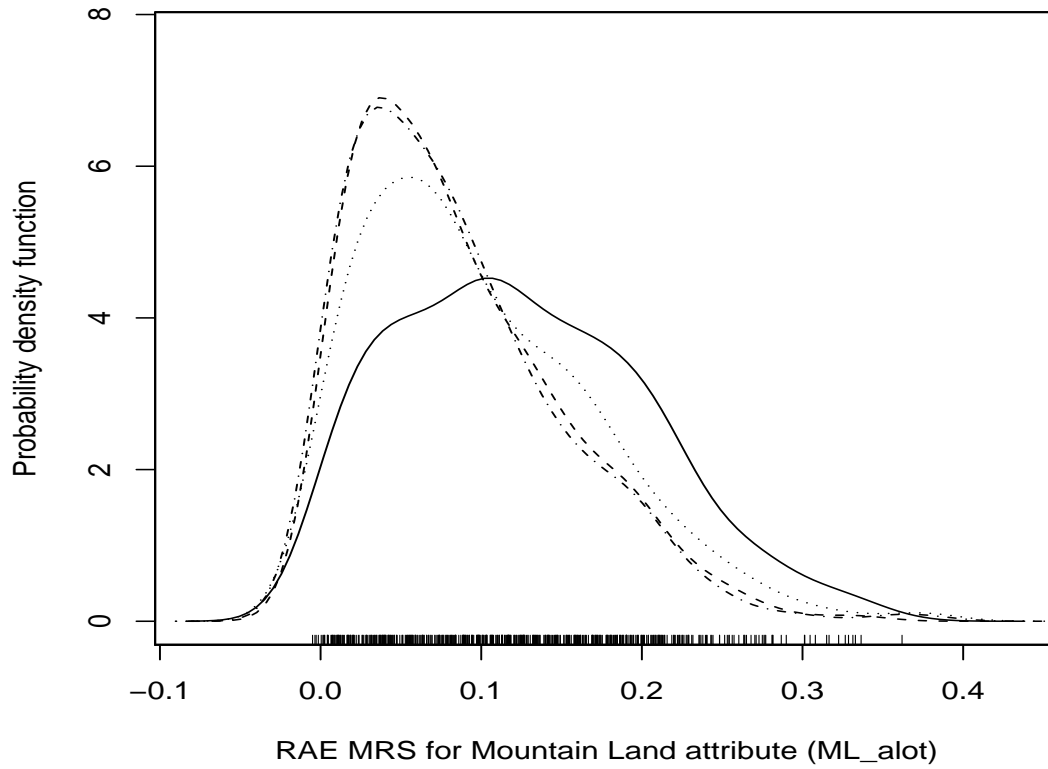
Dashed and dotted line: D_b^s-optimal design.

Fig. 2: DGP MNL and estimation MNL: *kernel-smoothed* distribution (optimal bandwidth) of the abso-

   lute relative error of landscape attribute Mountain Land ML_{alot}.

   Continous line: *shifted* design,

   Dashed line: D_p-optimal design,

   Dotted line: D_b^k-optimal design,

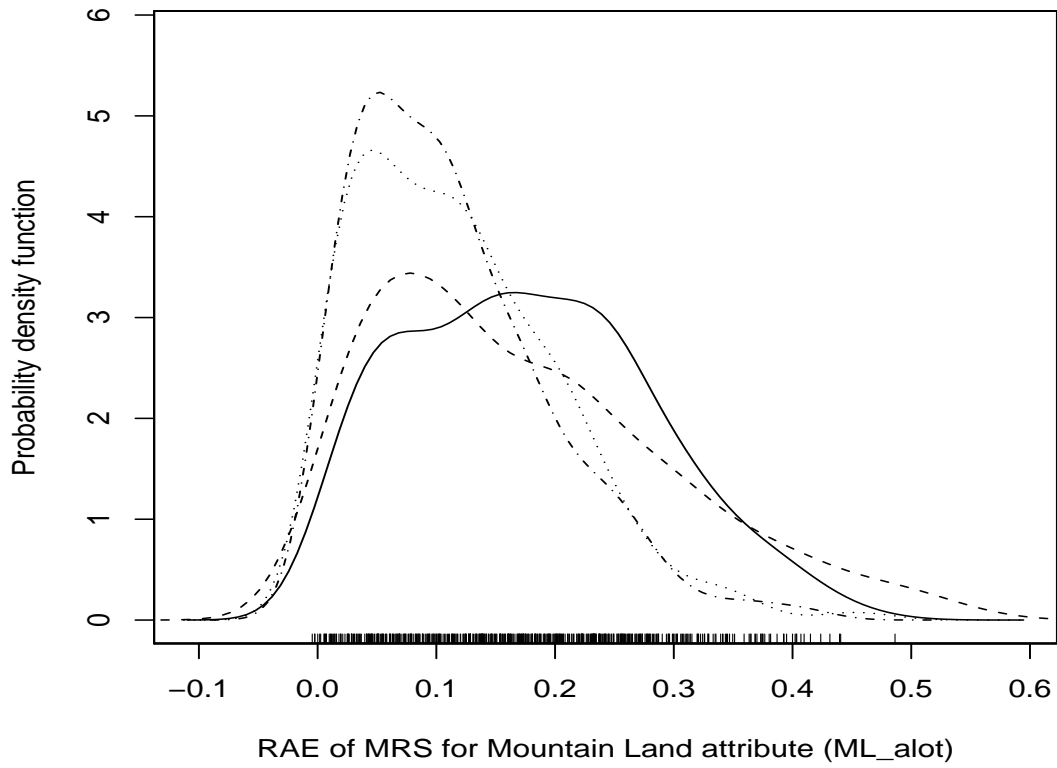   Dashed and dotted line: D_b^s-optimal design.

Fig. 3: DGP MNL and estimation KL-Asc, designed obtained under MNL assumptions: *kernel-smoothed* distribution (optimal bandwidth) of the absolute relative error of landscape attribute Mountain Land ML_{alot}.

Continous line: *shifted* design,

Dashed line: D_p-optimal design,

Dotted line: D_b^k-optimal design,
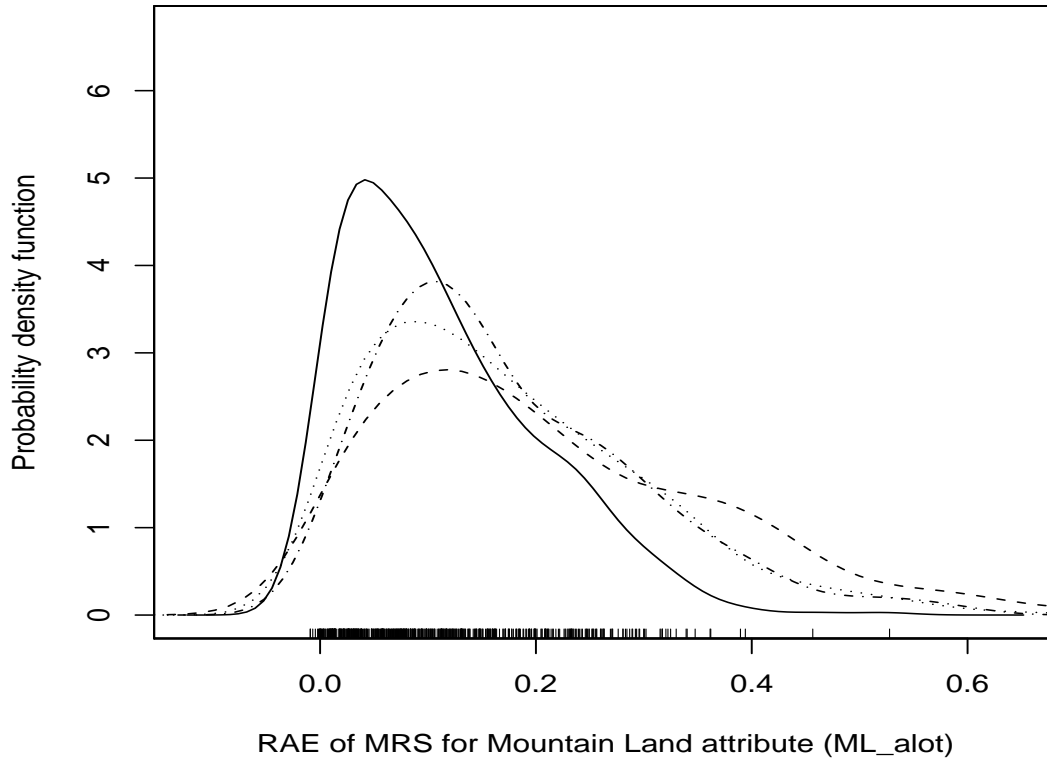
Dashed and dotted line: D_b^s-optimal design.

Fig. 4: DGP KL-Asc and estimation KL-Asc, designed obtained under MNL assumptions: *kernel-smoothed* distribution (optimal bandwidth) of the absolute relative error of landscape attribute Mountain Land ML_{alot}.

Continous line: *shifted* design,

Dashed line: D_p-optimal design,

Dotted line: D_b^k-optimal design,

Dashed and dotted line: D_b^s-optimal design.