
This Issue Completes Volume 53
Volume 53 Issue 3 May 2007 ISSN 0095-0696
ELSEVIER
JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT
MANAGING EDITOR
Charles F. Mason
The Official Journal of the Association of Environmental and Resource Economists
Available online at
ScienceDirect
www.sciencedirect.com

Available online at www.sciencedirect.com
ScienceDirect
Journal of Environmental Economics and Management 53 (2007) 342–363
JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT
www.elsevier.com/locate/jeem

Designs with *a priori* information for nonmarket valuation with choice experiments: A Monte Carlo study

Silvia Ferrini, Riccardo Scarpa*

Department of Economics, University of Siena, Siena, Italy
Department of Economics, Waikato Management School, University of Waikato, Private Bag 3105, Hamilton 3240, New Zealand

Received 6 October 2005
Available online 6 March 2007

Abstract

Good practice in experimental design is essential for choice experiments used in nonmarket valuation. We review the practice of experimental design for choice experiments in environmental economics and we compare it with advances in experimental design. We then evaluate the statistical efficiency of four different designs by means of Monte Carlo experiments. Correct and incorrect specifications are investigated with gradually more precise information on the true parameter values. The data generating process (DGP) is based on estimates from data of a real study. Results indicate that *D*-efficient designs are promising, especially when based on Bayesian algorithms with informative prior. However, if good quality *a priori* information is lacking, and if there is strong uncertainty about the real DGP—conditions which are quite common in environmental valuation—then practitioners might be better off with shifted designs built from conventional fractional factorial designs for linear models.

Keywords: Experimental design; Choice experiments; *D*-criterion; Nonmarket valuation

1. Introduction

In the last decade the use of discrete choice experiments (CEs) for the purpose of nonmarket valuation of environmental goods has gained favor with many applied environmental economists. CEs are used when policy outcomes may be usefully 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 typically include the policy cost to the respondent.[2] A CE then consists of selected subsets of all possible ‘profiles’ obtainable by combining and varying attribute levels. Typically, respondents are asked to select the best from a set of alternatives (the ‘choice set’), and to repeat this ‘choice task’ several times over the course of the interview, each time choosing from a choice set with different alternatives.

*Corresponding author. Fax: +64 07 838 43 31.
E-mail addresses: silvia.ferrini@unisi.it (S. Ferrini), rscarpa@waikato.ac.nz (R. Scarpa).

[1]This is the basis for the proposed term of ‘attribute-based stated preference’ method [34].

[2]The notion of describing a good on the basis of its attributes was born out of the theoretical approach of Lancaster’s [46,47]. It was then readily employed in marketing by Green and Rao [27] who propose *conjoint analysis* as a tool to model consumer preference.

0095-0696/$ - see front matter
doi:10.1016/j.jeem.2006.10.007

Using the set of observed discrete choices, researchers can then estimate separate marginal values for each attribute or attribute level used in describing the policy alternatives. In essence CEs are repeated referendum contingent valuation responses where the choice situation requires the respondent to select from possibly two or more policy situations, each succinctly described in terms of attributes and their levels. Willingness to pay (WTP) estimates are typically derived assuming random utility maximization and their efficiency determines how informative the study will be.

In a multiattribute multilevel context of choice the identification and efficiency of the estimates depends crucially on the choice of experimental design (ED), i.e. how attributes and attributes levels are combined to create alternatives in the choice sets to be presented to respondents. Ideally, the ED should be statistically *efficient*, providing the maximum accuracy of the estimates for the unknown population parameters given the available sample size. At the same time the ensuing choice tasks should require a relatively low cognitive effort from respondents, so as not to impair respondents' *efficiency* [66].

Yet, little work has been done to systematically evaluate the effect of various approaches to ED on the quality of the estimates in environmental valuation.[3] 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 to the effect of

> 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. CEs data, instead, are analyzed by means of specifications that are highly nonlinear-in-the-parameters, usually of the multinomial logit (MNL) type. When estimating preference parameters from CEs data, the high nonlinearity of the MNL specification affects the efficiency properties of the maximum likelihood estimator. Hence, statistically efficient EDs for MNL specifications potentially differ in most practical circumstances from those that are efficient in linear multivariate specifications. In particular, in an MNL context the statistical efficiency of the ED will depend on the unknown values of the parameters, as well as the unknown model specification.[4]

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 nonlinear-in-the-parameter models. These investigations should be tailored to the state of practice in environmental valuation, which is somewhat 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 the parameter values.

After reviewing recent advances in ED for logit models, it becomes apparent that the profession's current approach to ED can be improved upon. However, the gains afforded by such improvement need further investigation. With this paper we contribute to the existing literature by exploring the empirical performance of a number of recently proposed approaches to construct designs for discrete CEs. The investigation is conducted by means of Monte Carlo experiments designed to focus on the finite sample size properties of frequently employed logit estimators for value derivation in environmental valuation. In particular, we are

[3]We note, though, that some results on the effect of choice set creation and some proposed measures of choice complexity have been published [21,19].

[4]The concept of statistical efficiency is defined by different criteria of optimality, such as *A*- *D*- *G*- and *V*-optimality, which are intended to improve the precision of parameter estimates (variances and covariance matrices) or the prediction variance over the design region. However, the concept of *D*-optimality or, more appropriately, *D*-efficiency [45,7]—which is based on the minimization of the variance–covariance matrix—has dominated the design literature for CEs because—despite being geared towards parameter estimation—it still performs well in prediction, and it is easier to obtain [42].

[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 often choice sets only include two alternatives, while in environmental economics the most frequent format includes two experimentally designed alternatives plus the *status quo* (zero option). The latter is included to avoid the undesirable effects associated with forced choice [22].

concerned with measuring efficiency gains and the effects of model mis-specification under various EDs over a range of commonly used sample sizes.

In Section 2 we provide a summary of the evolution of the knowledge on design construction for CEs. In Section 3 we revise briefly 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 [61,62,67–69,40,14,59,43,41,15]. ED techniques were first introduced in the practice of multiattribute stated preference methods for market research by Louviere and Woodworth [51] and Louviere and Hensher [50], 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 product 'profiles' with well-known statistical properties for general linear models. These profiles are basically synthetic goods described on the basis of selected attributes 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). For some time orthogonality between the design attributes represented the primary criterion in the generation process of fractional factorial designs.

Modifications to this basic approach were later brought about by the necessity of making profiles 'realistic' and 'congruent', with orthogonality no longer seen as a necessary property.[6] Hence a good ED may be nonorthogonal in its attribute levels and require the investigation of mixed effects and selected attribute interactions, since in many realistic settings main-effects only may not be adequate (see, e.g. [52]).

Nonorthogonal designs can be optimized for linear multivariate models so as to maximize the amount of information obtained from a design. However, it is not clear that these EDs, conceived for continuous response variables, should be used in designing CEs, where the response is discrete and a highly nonlinear specification is assumed to generate response probabilities.[7] Fortunately, it appears that '…an efficient design for linear models is also a good design for MNL for discrete choice response' [45]. Corroborating evidence of this is provided by Lazari and Anderson [48] and Kuhfeld et al. [45], and more recently by Lusk and Norwood [52] who studied the small sample performance of commonly employed *D*-efficient EDs for linear-in-the-parameters models in the context of logit models for CEs, focussing on the role of attribute interactions. By relying on these empirical results one may conveniently ignore the necessity of deriving EDs specific for nonlinear models, which would require assumptions on the unknown parameter vector.[8]

The effects of assigning the experimentally designed alternatives to individual choice sets were investigated by Bunch et al. [13] who approached the issue of choice sets construction by proposing the *object-based* and *attribute-based* strategies, which we employ later for one of our designs in Section 4. However, they restrictively assumed the parameter vector of the indirect utility function $\boldsymbol{\beta} = 0$, which implies equiprobability of alternative selection in each choice set. Under this assumption the optimization problem for the determinant of the information matrix in discrete choice models remains the same as in a linear framework [28], and obviously does not require any degree of knowledge or assumptions on the *true* population values of $\boldsymbol{\beta}$. Because of the $\boldsymbol{\beta} = 0$ 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

[6]See also [59] on the effects of lack of orthogonality on ED efficiency, and how this can easily come about even when orthogonal designs are employed.

[7]Linearly *D*-optimal designs can be obtained by specific software such as SPSS, MINITAB, Design Easy, etc. The most comprehensive algorithms for CE 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) [43,44]. CBC also provides algorithms for choice design, but only guided towards balancedness.

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

Table 1
Approaches to experiment design for discrete choice experiments

Authors	Criterion definition	Assumed *a priori* parameters	Algorithm
Lazari and Anderson [48]	D-optimal	—	Unspecified
Kuhfeld et al. [45]	D-optimal	—	Modified Fedorov
Bunch et al. [13]	D_0-optimal	—	—
Huber and Zwerina [36]	D_p-optimal	$\boldsymbol{\beta}_0$	RS
Zwerina et al. [74]	D_p-optimal	$\boldsymbol{\beta}_0$	Modified Fedorov
Sándor and Wedel [61]	D_b-optimal	$\mathcal{N}(\boldsymbol{\beta}\vert\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$	RSC
Sándor and Wedel [62]	D_b-optimal	$\mathcal{N}(\boldsymbol{\beta}\vert\boldsymbol{\beta}_0, \boldsymbol{\Sigma}_0)$	RSC
Kanninen [40]	D_s-optimal	—	Sequential update
Burgess and Street [14]	D_p-optimal	$\boldsymbol{\beta}_0$	—
Kuhfeld [43]	D_p-optimal	$\boldsymbol{\beta}_p$	Modified Fedorov
Kessels et al. [41]	D_b-optimal	$\boldsymbol{\beta}\sim U[a,b]$	Modified Fedorov
Burgess and Street [15]	D_p-optimal	$\boldsymbol{\beta}_0$	—

RS, relabeling and swapping; RSC, relabeling, swapping and cycling.

balanced utility, are described in [36] where these are considered to be essential features in the derivation of efficient EDs.

Huber and Zwerina [36] subsequently argued that in most practical research situations some kind of prior knowledge is available (e.g. from the results of a pilot survey) and broke away from the $\boldsymbol{\beta} = 0$ assumption. They championed—instead—the D_p-optimality criterion, where p stands for '*point*' (i.e. local) information on $\boldsymbol{\beta} \neq 0$. They demonstrated how restrictive it can be to assume $\boldsymbol{\beta} = 0$ as it induces efficiency loss, and verified that including pre-test results into the development of efficient ED may improve efficiency up to 50%. 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 how at the ED construction stage of a study there typically exists significant uncertainty about the *a priori* information on parameter values $\boldsymbol{\beta}$ and hence such uncertainty should be explicitly accounted for in the ED construction. Hence a sequential Bayesian approach is proposed, based on the estimates of $\boldsymbol{\beta}$ from some pilot study and deriving 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 first used in the CE literature by Sándor and Wedel [61] for MNL specifications by modifying the empirical algorithms proposed by Huber and Zwerina [36]. This approach 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 over the full factorial set.

Recent work by Burgess and Street has tackled the issue of construction of more general designs, such as [67,14,68,15,70] but they are limited to the case of $\boldsymbol{\beta} = 0$. A comparison of criteria to derive efficient EDs is illustrated in [42], in which the method by Zwerina et al. [74] is extended and a D_b-optimal ED is obtained by using a weakly informative (uniform) prior distribution of $\boldsymbol{\beta}$.[9]

A short summary of the evolution of ED research is reported in Table 1 which includes the D_p-optimal design proposed in [43] and the D_s-optimal design proposed in [40], which describes a sequential design for discrete CEs. Notice that although in recent years the theoretical research work on efficient ED construction for logit models has intensified (see also for more theoretical results) [25,26], 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 studies

[9] We prefer the term 'weakly informative' to the more common Bayesian term 'uninformative' because of the reasons spelled out in [23], where it is noted that a uniform prior is not uninformative in this context.

concerning ED: Sándor and Wedel [62] and Blemier et al. [8]. We found no study addressing the issue in the context of discrete mixing (latent class models).

There are also few empirical evaluations of the different ways of deriving efficient EDs for MNL models in the various fields of applications in economics, with the exception of [18,72] in health economics, [59] in transportation and [52] in agricultural economics. 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 of choice set in health economics research. They assume that the investigator correctly specifies the data generating process (DGP), the *a priori* $\boldsymbol{\beta}$ at the stage of decision about the ED and at the estimation stage. Under these favorable conditions they find that fractional factorial main effects orthogonal EDs are inferior to D_0-optimal and D_p-optimal designs. This is an apparently worrying result considering that this has been the dominant approach in environmental economics. They also find that the *shifted* ED (also sometimes termed *cycled*) [13] performs better than the D_0-optimal for generic attributes, but in general the most efficient design is the D_p-optimal. Viney et al. [72], instead, focus on the effect of design on cognitive effort as captured by the variance of the error component of utility. They report the following:

> …three experimental design approaches are investigated: a standard six-attribute, orthogonal main effects design; a design that combines alternatives to achieve utility balance, ensuring no alternatives are dominated; and a design that combines alternatives randomly. The different experimental designs did not impact on the underlying parameter estimates, but imposing utility balance increases the random variability of responses.

However, in both [18,72] the 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 relatively uncommon in environmental economics.

In transportation modeling, instead, Rose and Bliemer [59] emphasized how the much sought-after property of orthogonality may well be lost in the final data set due to the cumulative effects of sample nonresponse. Furthermore, while the transportation literature on ED for choice modeling is often dominated by labeled alternatives (one label per transportation mode, with relative label-specific attributes), the typical situation in environmental valuation seems to be that of generic (unlabeled) alternatives.

Finally, on the issue of sequential design, Kanninen [40] drawing on an earlier CV idea [38] illustrates how one can choose numerical attributes, such as price, to sequentially ensure the maximization of some measure (e.g. the determinant) of the information matrix of binary and multinomial models from CE data. On the other hand, Raghavarao and Wiley [55] show that with sequential design and computer-aided interviews 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. The state of practice in environmental economics

The adoption of CE in environmental economics began in the early 1990s, when research on ED for MNL models was still at an embryonic stage. However, researchers concerned with discrete choice contingent valuation were already aware of the importance of ED on the efficiency of welfare estimates [2,39,1]. This concern does not seem to have carried over to CE practice, where the dominant approach, as visible from Table 2, has been that based on fractional factorials for main effects with orthogonality. This is typically derived using algorithms developed for multivariate linear models, which are—as explained earlier—only a surrogate with much scope for improvement that can be reached via more tailored designs. But under what conditions?

The prevailing approach to ED in environmental economics applications seems to involve the following steps:

Step 1: Determine attributes and their levels.

Step 2: Determine ex ante the number of alternatives in the *choice set* and the number of choice situations for each respondent.

Step 3: Define the combination of attribute and levels describing each alternative in each choice set based on linear ED approaches.

Table 2
Selected features of *choice experiment* studies in environmental economics

Authors and paper	Number of attributes	Choice task alternatives	Choice tasks per respondent	Experimental design	Model specification	Sampled respondents
Boxall et al. [11] (EE)	6 ($4^4 2^2$)	$2+sq$	16	—	MNL	271
Hanley et al. [31] (ERE)	4 (2^3)	$2+sq$	4	—	MNL	181
Rolfe et al. [57] (EE)	7 ($8^1 4^6$)	$2+sq$	16	—	MNL	105
Carlsson and Martinsson [17] (JEEM)	3 (3^3)	2	14	*D*-optimal Zwerina et al. [74]	EVHL	350
Boxall and Adamowicz [10] (ERE)	5 (4^5)	$5+sq$	8	Orthogonal main effects	LC RPL	620
Blamey et al. [6] (ERE)	6 ($4^4 3^1 5^1$)	$2+sq$ $4+1$	4/8	 Fractional factorial	MNL NL LC	480 620
DeShazo and Fermo [21] (JEEM)	4/9	2/7	—	Factorial orthogonal randomised	Heteroskedastic MNL	1800/2100
Sælensminde [60] (ERE)	3/4	2	9	Fractional factorial orthogonal	Binary Logit	2568
Hanley et al. [32] (ERE)	6 ($4^4 2^1 6^1$)	$2+sq$	4/8	Fractional factorial	MNL NL	267
Foster and Mourato[24] (ERE)	5	$2+sq$		Fractional factorial (SPEED software)	MNL RPL	290
Horne and Petäjistö [35] (LE)	5 ($4^4 2^1$)	$2+sq$	4/8	Fractional factorial	MNL	1296
Scarpa et al. [65] (EE)	5 ($3^3 2^2 4^1$)	$2+sq$	6	Fractional factorial	MNL + Heterosk. RPL	300
Carlsson et al. [16] (EE)	7 ($2^5 3^1 4^1$)	$2+sq$	4	Fractional factorial *D*-optimal OPTEX (SAS)	MNL RPL	5800
Rodriguez and Leòn [56] (ERE)	6 ($3^2 4^2 2^2$)	$2+sq$	8	*D*-optimal design Huber and Zwerina [36]	MNL RPL EVHL	350
Wattage et al. [73] (EE)	3 ($3^2 4^1$)	16	—	Orthogonal main effects	MNL	30
Jin et al. [37] (EE)	3 ($2^3 4^1$)	$1+sq$	8	Main effects factorial design	MNL	260

MNL, Multinomial logit; EVHL, Extreme value heteroskedastic 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.

Step 4: Assign the profiles so derived to choice sets either randomly or with different combinatorial devices.

Generally, attributes and levels are selected on the basis of both the objective of the study and information gathered from *focus groups*. 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 a few environmental economics studies [16,56] is the criterion of maximizing the determinant of the information matrix of the MNL the guiding principle for the derivation of the ED.

From our review of the literature, we draw the following observations:

1. In the practice of CE design the profession continues to use orthogonal main effects designs for linear models rather than replacing them with *D*-optimal designs for logit models. This may be due to a lack of

appreciation for the efficiency gains derivable from D-optimal designs, and the robustness of these gains to mis-specification. Hence, it is of interest to empirically evaluate both the size of such expected gains and their robustness in a typical environmental valuation context.

2. Amongst the various D-optimal design algorithms the only ones that have been employed so far are those for MNL specifications with a fixed *a priori* parameter ($\boldsymbol{\beta} = \boldsymbol{\beta}_p$). This is probably due to the fact that for these EDs are well documented and predefined macros are available in SAS [43]. 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 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 environmental valuation literature is dominated by the $2+sq$ choice task format, which—as demonstrated elsewhere in the literature (e.g. [29,30])—is prone to give rise to *status quo* bias, raising 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, and it may require more flexible specifications based on mixed logit models [64]. Robustness of the efficiency gains to common model mis-specifications when the design is developed on standard MNL assumptions is therefore an important criterion in the evaluation of alternative approaches to ED. Tests for model specification are only meaningful at reasonable sample sizes. So, while the researcher can resolve the issue of specification when the sample data have become available, they must often decide on the features of the ED at an earlier stage.
4. Finally, an empirical investigation should also explore which ED approach is most robust with regards to the quality of *a priori* information and assumptions about the true values of $\boldsymbol{\beta}$.

4. Methods

In our empirical investigation 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.[10]

4.1. The shifted design

We chose to employ a shifted (or cycled) design rather than the most common fractional factorial orthogonal design (FFOD), which we felt has already been thoroughly assessed by Lusk and Norwood [52]. Furthermore, based on the results of [18], the shifted design seems to perform better than the FFOD, and to be just as simple to derive. The shifted design is based on the implicit assumption of identical selection probabilities across alternatives as implied by the assumed values of $\boldsymbol{\beta} = 0$. This approach was originally proposed by Bunch et al. [13] who consider designs for general linear models and propose a procedure to assign alternatives to choice sets. The work by Burgess and Street [14,15] shows how to shift attribute levels 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. With this method one 'shifts' the original columns of the FFOD in such a way that all attribute levels are varied. For example, in our case from an initial FFOD (the *seed*) all attribute levels were shifted (increased) by one unit, except those already at the highest level, which were assigned the lowest values in the ladder of levels. We refer to this ED as the 'shifted' design. The codified design used in the Monte Carlo experiments is reported in Table 3.

4.2. D_p-optimal design

In practice, the *a priori* information typically suggests $\boldsymbol{\beta} \neq 0$. In this case, potentially more efficient designs than the shifted one can be obtained by making use of such information on the values of $\boldsymbol{\beta}$. This can be done

[10] The necessary Gauss codes to replicate this study are available from the authors.

Table 3
Shifted and D_p-optimal experiment designs

Choice set	Alt	Shifted					D_p-optimal				
		Attributes									
		ML	SW	FT	CH	Cost	ML	SW	FT	CH	Cost
1	I	1	1	1	1	1	2	2	1	3	2
	II	2	2	2	2	2	1	1	3	1	1
2	I	1	2	2	2	1	2	2	3	3	2
	II	2	3	3	3	2	1	1	1	2	1
3	I	1	3	3	3	1	1	1	3	3	2
	II	2	1	1	1	2	3	3	2	1	1
4	I	2	1	1	1	2	1	2	2	2	1
	II	2	2	2	3	1	2	1	3	1	2
5	I	2	2	2	3	1	2	1	2	3	1
	II	3	3	3	1	2	3	2	1	1	2
6	I	2	3	3	1	1	1	3	1	2	2
	II	3	1	1	2	2	3	1	2	1	1
7	I	3	1	2	1	1	3	2	1	3	1
	II	1	2	3	2	2	2	3	3	2	2
8	I	3	2	3	2	1	1	3	3	1	1
	II	1	3	1	3	2	3	1	1	2	2
9	I	3	3	1	3	1	3	2	3	3	1
	II	1	1	2	1	2	2	3	2	1	2
10	I	1	1	3	3	2	1	1	2	1	1
	II	2	2	1	1	1	3	3	3	2	2
11	I	1	2	1	1	2	2	1	1	2	1
	II	2	3	2	2	1	3	2	2	1	2
12	I	1	3	2	2	2	1	1	3	1	2
	II	2	1	3	3	1	2	3	2	3	1
13	I	2	1	2	3	2	2	1	1	1	1
	II	3	2	3	1	1	1	2	2	2	2
14	I	2	2	3	1	2	1	1	2	3	2
	II	3	3	1	2	1	2	2	1	1	1
15	I	2	3	1	2	2	3	1	2	2	2
	II	3	1	2	3	1	1	2	3	3	1
16	I	3	1	3	2	2	1	3	1	1	2
	II	1	2	1	3	1	3	1	3	2	1
17	I	3	2	1	3	2	2	1	2	3	2
	II	1	3	2	1	1	3	3	3	2	1
18	I	3	3	2	1	2	1	1	1	3	1
	II	1	1	3	2	1	2	2	3	2	1

Effective coding in the utility function was (1) (0 1), (2) (1 0), (3) (0 0).

with a D_p-optimal design, derived by maximization the information provided in the arrangement of attributes and attribute levels.

An obvious objective function in the context of maximum likelihood estimation can be based on the information matrix for the design. Under the MNL model assumptions, this is given by

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

where s denotes choice situations, N is the number of respondents, $\mathbf{X}_s = [\mathbf{X}_{1s}, \ldots, \mathbf{X}_{Js}]'$ 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 $\mathbf{P}_s = \text{diag}[p_{1s}, \ldots, p_{Js}]$ with zero off diagonal elements and $p_{js} = \mathrm{e}^{\mu\beta\mathbf{X}_j}(\sum_{i=1}^{J}\mathrm{e}^{\mu\beta\mathbf{x}_i})^{-1}$.[11]

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

A widely accepted [45,61,42] scalar measure of efficiency in the context of EDs for models that are nonlinear in the parameter is the D-criterion (or D_p-error), which is defined as

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

where k is the number of attributes. We employed the modified Federov algorithm proposed by Zwerina et al. [74] to find the arrangement of the levels in the various attributes in $\mathbf{X}$ such that the D-criterion is minimized (and the $\det[\mathbf{I}(\boldsymbol{\beta})]$ *maximized*) when $\boldsymbol{\beta} = \boldsymbol{\beta}_p$. Such algorithms are available in the macro '`%ChoicEff`', in SAS v. 9 (see [43], for details) and the codified design used for the simulation is reported in Table 3.

4.3. D_b-optimal designs

Bayesian designs explicitly account for the uncertainty of subjective *a priori* information on the values of the population $\boldsymbol{\beta}$. 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 $\boldsymbol{\beta}$ or $\pi(\boldsymbol{\beta})$:

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

where k is the number of attributes. In practice this is achieved by approximating via simulation the value of D_b: one draws R sets of values $\boldsymbol{\beta}^r$ from the *a priori* distribution $\pi(\boldsymbol{\beta})$ and computes the average of the simulated D-criterion over the R draws:

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

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 prior incorporating relatively poor information and implemented by a uniform distribution [41]. 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 a variance–covariance matrix as estimated from the pilot [61].

While the D_p-optimal design ignores the uncertainty which invariably surrounds the values of $\boldsymbol{\beta}$, the D_b-optimal design allows the researcher to explicitly account for it. On the other hand the derivation of Bayesian designs is computationally more demanding, and perhaps explains why previous studies have neglected them.

4.3.1. D_b-optimal design with weakly informative prior

The distributional assumption about the prior in this case is that each of the k parameters is distributed i.i.d. uniform with $\pi(\beta_k) = U[a_k, b_k]$ where a_k and b_k are the extreme values bounding the intervals within which the true value of β_k is expected to be. We refer to this design throughout the paper as D_b^{w}-optimal, where the superscript w stands for *weakly informative* prior. We note that a_k and b_k can be chosen to incorporate directional prior information from economic theory on the sign of parameters. For example, the coefficient on money β_{m} (the cost of a policy alternative) is expected to have a negative impact on utility, so for β_{m} one may set $b_{\mathrm{m}} = 0$ and a_{m} equal to some negative number.

4.3.2. D_b-optimal design with informative prior

Following [61] we assume the prior to be distributed $\pi(\beta) = \mathcal{N}(\hat{\boldsymbol{\beta}}, \hat{\boldsymbol{\Omega}})$. While [61] derive the $\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\Omega}}$ on the basis of managers' expectations, we derived them from real data obtained from the pilot study preceding the field survey. Such data are commonly 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 [61,62]. In what follows we refer to this design as D_b^{i}-optimal, where the superscript i stands for *informative* prior. The two codified Bayesian designs used in the Monte Carlo study are reported in Table 4.

4.4. Design of Monte Carlo experiment

To assess the difference between the alternative designs, we have drawn inspiration from a study about 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 (CH) features [63]. In this CE study all the attributes were 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 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 [54].

Based on the estimation results obtained from these data through MNL and Kernel logit (KL) with alternative-specific constant models (Table 5), our Monte Carlo experiment is designed to investigate the

Table 4
Bayesian experiment designs

Choice set	Alt	D_b^w-optimal					D_b^i-optimal				
		Attributes									
		ML	SW	FT	CH	Cost	ML	SW	FT	CH	Cost
1	I	2	2	2	2	1	2	3	3	3	1
	II	1	1	1	1	2	3	2	1	2	2
2	I	1	2	2	1	2	2	3	1	2	1
	II	2	1	1	3	1	1	2	2	3	2
3	I	2	1	2	1	2	1	2	1	3	2
	II	1	2	1	2	1	3	3	3	1	1
4	I	2	2	2	1	1	1	3	2	2	2
	II	3	3	1	3	2	3	1	3	3	1
5	I	2	3	2	2	1	3	3	1	3	1
	II	3	1	1	1	2	2	1	3	2	2
6	I	2	2	1	1	2	1	3	2	3	2
	II	1	1	2	2	1	2	2	1	1	1
7	I	2	3	1	1	2	3	2	2	1	2
	II	1	2	3	2	1	1	3	1	2	1
8	I	2	2	1	2	1	2	3	2	1	1
	II	1	3	2	1	2	1	2	3	3	2
9	I	1	2	3	3	1	3	2	3	1	1
	II	3	3	1	2	2	2	3	1	2	2
10	I	2	1	1	2	1	1	2	3	2	1
	II	3	2	3	3	2	2	1	2	3	2
11	I	3	1	2	3	2	3	3	1	3	1
	II	1	3	1	2	1	1	2	2	1	2
12	I	1	2	1	1	2	1	3	3	2	1
	II	3	1	2	2	1	2	2	1	1	2
13	I	3	2	3	1	2	3	1	3	3	2
	II	1	1	2	3	1	1	3	2	1	1
14	I	2	3	3	3	2	2	2	3	1	2
	II	3	2	2	1	1	3	1	2	2	1
15	I	1	1	3	1	1	3	2	1	2	2
	II	3	2	2	3	2	2	3	3	3	1
16	I	1	3	2	3	1	2	1	2	3	1
	II	2	1	3	2	2	3	3	3	1	2
17	I	1	1	1	2	2	3	1	2	2	2
	II	2	3	3	1	1	1	3	1	3	1
18	I	3	3	3	2	1	2	1	3	2	2
	II	2	2	2	3	2	3	3	2	1	1

Effective coding in the utility function was (1) (0 1), (2) (1 0), (3) (0 0).

Table 5
Maximum likelihood estimates of MNL model and maximum simulated likelihood estimates of KL-Asc model for the landscape study

	MNL		KL-Asc	
Cost	−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)
SQ-Asc			−1.420	(−6.20)
σ			1.351	(7.73)

Asymptotic z-values in brackets.

relative performance of the four designs obtained with the methods described above. All designs are developed under the assumption that the true DGP has an MNL specification. Such an assumption is the most frequently made in this line of research. However, after the data collection, the response pattern may display evidence corroborating other logit 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 that in the nested logit. This specification is examined in some detail in [64] 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 (or 'kernel' logit (KL) specification) takes the following basic utility form:

$$
\begin{aligned}
U(c_1) &= \boldsymbol{\beta}\mathbf{x}_{c_1} + \tilde{u}_{c_1} = \boldsymbol{\beta}\mathbf{x}_{c_1} + \varepsilon + u_{c_1},\\
U(c_2) &= \boldsymbol{\beta}\mathbf{x}_{c_2} + \tilde{u}_{c_2} = \beta\mathbf{x}_{c_2} + \varepsilon + u_{c_2},\\
U(sq) &= Asc + \boldsymbol{\beta}\mathbf{x}_{\mathrm{sq}} + u_{\mathrm{sq}},
\end{aligned}
\tag{5}
$$

where $U(c_j)$ denotes the utility associated with choice alternative j ($j = 1, 2$) and $U(sq)$ denotes the utility associated with the *status quo* alternative. In our case, $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ are additional error components to the conventional Gumbel-distributed u_{c_1} and u_{c_2}, thereby leading to the following error covariance structure:

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

$$Cov(\tilde{u}_{c_j}, \tilde{u}_{sq}) = 0, \quad Var(\tilde{u}_{sq}) = \pi^2/6, \ \ j = 1, 2, \tag{7}$$

where $\tilde{u}_{c_j} = \varepsilon + u_{c_j}$.[12] Note that this is an analog of the nested logit model in the sense that it allows for correlation of utilities related to alternatives different from the *status quo* (i.e. in the same nest) [12,33,71]. 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 for the j alternative the choice probability is logit, and the assumption above leads to the following expression for each marginal choice probability:

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

$$P(i) = \int_{-\infty}^{+\infty} \frac{\exp(\boldsymbol{\beta}\mathbf{x}_i + 1(i)\varepsilon)}{\exp(\boldsymbol{\beta}\mathbf{x}_{sq}) + \exp(\boldsymbol{\beta}\mathbf{x}_{c1} + \varepsilon) + \exp(\boldsymbol{\beta}\mathbf{x}_{c2} + \varepsilon)} \phi(0, \sigma^2) d\varepsilon, \tag{8}$$

[12] As expanded upon by Brownstone and Stone [12], Train [71], Herriges and Phaneuf [33], more general forms than this may be empirically appealing.

where $\phi(\cdot)$ is the normal density, $1(i)$ is an indicator function that takes the value of 0 if $i = sq$, 1 otherwise, and $\boldsymbol{\beta}$ is a vector of parameters including an alternative-specific constant for the *status quo* (sq). Estimation of $\hat{\boldsymbol{\beta}}$ and $\hat{\sigma}^2$ is obtained by maximum simulated likelihood [71].

The alternative designs subject to comparison are assessed by Monte Carlo experiments. The evaluation of the performance of the four designs in the case of an incorrectly assumed 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 nonlinear 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 $r = 1, 2, 3, \ldots, R = 1000$ samples of 100, 250 and 500 observations under two different DGP: the MNL and the error components model with alternative-specific constant.

1. At each replication r individual counterfactual responses y_i^r are produced by identifying the alternative j associated with the largest utility value $U(\boldsymbol{\beta}, \varepsilon_j, u_j, x_j)$, where the $\boldsymbol{\beta}$ values are those for the DGP and are reported in Table 5, while the errors ε_j, u_j are drawn from the adequate distributions (Gumbel for MNL; Gumbel *and* Normal for the KL-Asc).
2. The counterfactual y_i^r produced for the whole sample are used to get maximum likelihood or maximum simulated likelihood estimates $\widehat{\boldsymbol{\beta}}^r$. Then a series of indicators of estimation performance are computed. Given their relevance in nonmarket valuation, we focus on estimates of marginal rates of substitution between the generic k attribute and money:

$$\widehat{MRS}_k^r = \widehat{\tau}_k^r = -\frac{\widehat{\beta}_k^r}{\widehat{\beta}_m}. \tag{9}$$

We then report some standard efficiency indicators.

(a) First, we report the empirical mean squared error:

$$MSE = \frac{1}{R}\sum_{r=1}^{R}(\widehat{\tau}^r - \tau)^2, \quad r = 1, \ldots, 1000, \tag{10}$$

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

(b) The second measure reported is the empirical bias:

$$Bias(\widehat{\tau}) = \frac{1}{R}\sum_{r=1}^{R}(\widehat{\tau}^r - \tau), \tag{11}$$

from which one can derive the variance, since $[Bias(\widehat{\tau})]^2 = MSE - Var(\widehat{\tau})$.

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

$$\overline{RAE} = \frac{1}{R}\sum_{r=1}^{R} \mid (\widehat{\tau}^r - \tau)/\tau \mid . \tag{12}$$

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, we enlarge the set of standard measures for Monte Carlo studies and as an additional measure of efficiency we report the fraction of MRS values falling within a 10% interval around the true value:

$$\Gamma_{0.05} = \frac{1}{R}\sum_{r=1}^{R} 1(\widehat{\tau}^r \in \tau \pm \tau \times 0.05), \tag{13}$$

where $1(\cdot)$ is an indicator function. This gives an easy to remember 'rule-of-thumb' measure 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 MRS for the attribute ML that showed highest implicit value in the original study (see Table 5 and [63]). Qualitatively similar results were obtained for the other attributes. All attributes in this study were expressed at two levels of policy action 'some' (ML_some) and 'a lot of' (ML_alot) and concerned the visual aspect of mountainous rural land (ML). Tables 6–8 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 6 presents the results for 'the best of both worlds'. It portrays the case in which the analyst has correctly guessed the DGP at the stage of design development, and the correct specification is employed to obtain the estimates of marginal WTP. The contribution to the empirical bias by model mis-specification is therefore nil in this case, and the observed empirical bias is due to the finite sample size. Similarly, most of the empirical MSE is made up of variance.

Table 6
Summary statistics from Monte Carlo experiment on data from DGP MNL and estimates from MNL specification

DGP: MNL, Design: MNL Assumption: MNL				
Designs	*MSE*	*Bias*	$\overline{RAE}$	$\Gamma_{(0.05)}$
ML_A lot				
Shifted $N = 100$	21.54	1.81	0.19	0.165
Shifted $N = 250$	11.55	1.88	0.14	0.213
Shifted $N = 500$	8.14	2.06	0.12	0.226
D_p-Optimal $N = 100$	22.86	−0.02	0.20	0.157
D_p-Optimal $N = 250$	9.55	0.20	0.12	0.259
D_p-Optimal $N = 500$	5.99	0.37	0.09	0.349
D_b^{w}-Optimal $N = 100$	29.17	−0.50	0.22	0.140
D_b^{w}-Optimal $N = 250$	10.28	−0.05	0.14	0.213
D_b^{w}-Optimal $N = 500$	5.25	−0.08	0.09	0.320
D_b^{i}-Optimal $N = 100$	18.28	1.05	0.17	0.182
D_b^{i}-Optimal $N = 250$	7.76	1.04	0.11	0.288
D_b^{i}-Optimal $N = 500$	4.28	0.89	0.09	0.365
ML_Some				
Shifted $N = 100$	18.86	0.48	0.34	0.086
Shifted $N = 250$	7.73	0.12	0.22	0.138
Shifted $N = 500$	4.02	0.37	0.16	0.204
D_p-Optimal$N = 100$	22.38	−1.27	0.38	0.078
D_p-Optimal $N = 250$	9.56	−1.07	0.24	0.135
D_p-Optimal $N = 500$	5.91	−1.30	0.19	0.165
D_b^{w}-Optimal $N = 100$	25.28	0.30	0.39	0.081
D_b^{w}-Optimal $N = 250$	9.92	−0.20	0.25	0.128
D_b^{w}-Optimal $N = 500$	4.96	0.17	0.18	0.195
D_b^{i}-Optimal $N = 100$	21.72	0.91	0.37	0.074
D_b^{i}-Optimal $N = 250$	8.67	0.26	0.24	0.139
D_b^{i}-Optimal $N = 500$	4.28	0.10	0.17	0.183

Table 7
Summary statistics from Monte Carlo experiment on data from DGP MNL and estimates from KL-Asc specification

DGP: MNL, Design: MNL
Assumption: Kernel logit

Designs	*MSE*	*Bias*	$\overline{RAE}$	$\Gamma_{(0.05)}$
ML_A lot				
D_b^{w}-Optimal $N = 100$	44.42	0.17	0.27	0.129
D_b^{w}-Optimal $N = 250$	16.03	0.49	0.17	0.192
D_b^{w}-Optimal $N = 500$	7.42	0.39	0.11	0.288
D_b^{i}-Optimal $N = 100$	27.87	2.32	0.22	0.154
D_b^{i}-Optimal $N = 250$	12.77	1.99	0.15	0.207
D_b^{i}-Optimal $N = 500$	7.63	1.74	0.12	0.261
ML_Some				
Shifted $N = 100$	31.89	2.04	0.44	0.090
Shifted $N = 250$	11.82	1.31	0.27	0.110
Shifted $N = 500$	7.31	1.47	0.22	0.132
D_p-Optimal $N = 100$	39.42	1.44	0.49	0.078
D_p-Optimal $N = 250$	14.95	1.13	0.31	0.097
D_p-Optimal $N = 500$	7.04	0.47	0.21	0.152
D_b^{w}-Optimal $N = 100$	38.64	0.39	0.48	0.082
D_b^{w}-Optimal $N = 250$	14.46	0.02	0.30	0.118
D_b^{w}-Optimal $N = 500$	7.03	0.37	0.21	0.142
D_b^{i}-Optimal $N = 100$	35.16	1.67	0.47	0.074
D_b^{i}-Optimal $N = 250$	13.81	0.97	0.29	0.111
D_b^{i}-Optimal $N = 500$	7.03	0.86	0.21	0.126

Based on the performance of the three D-optimal designs as described by the MSE it can be seen how D_b^{i}-optimal is the most efficient at sample sizes of $N = 250$ and 500 across both attributes. However, at small sample sizes ($N = 100$) the shifted design gives a performance superior to that of D_p- and D_b^{w}-optimal designs for both attributes, and for the attribute with the lowest impact on utility (ML_some) it even outperforms the D_b^{i}-optimal design.

A graphical illustration of what happens at large sample sizes ($N = 500$) is reported in Fig. 1 where we show the kernel-smoothed [9] distributions of $MRS_{\mathrm{ML_{alot}}}$ for all four designs. Notice that while the D_b^{w}-optimal design is centered on the true value of 19.32, the D_p-optimal and the D_b^{i}-optimal, respectively, underestimate and overestimate slightly, while the *shifted* design produces significant overestimates at this sample size.

Analog conclusions can be drawn from an inspection of Fig. 2, where we report the absolute relative error ($RAE_{\mathrm{ML_{alot}}}$). Suppose a decision rule was to be incorrectly taken if the relative absolute error is larger than 20% or 30%; from the plot in Fig. 2 it is apparent that the number 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 clear-cut superiority of designs that incorporate prior information on $\boldsymbol{\beta}$ over the shifted design only emerges at a high sample size ($N = 500$), and for the attribute with high relative weight on utility. For the attribute with low relative weight on utility at all sample sizes the MSE values for the shifted design are smaller than those for the other D-optimal designs.

5.2. Incorrect model specification and correct design information

As a way to investigate the sensitivity of these results to potential model mis-specification at the estimation phase, we now turn our attention to the case in which the estimation makes use of a mis-specified model

Table 8
Summary statistics from Monte Carlo experiment on data from DGP KL-Asc and estimates from KL-Asc specification

DGP: Kernel logit, Design: MNL
Assumption: Kernel logit

Designs	*MSE*	*Bias*	$\overline{RAE}$	$\Gamma_{(0.05)}$
ML_A lot				
Shifted $N = 100$	35.44	2.02	0.33	0.198
Shifted $N = 250$	13.75	1.73	0.21	0.293
Shifted $N = 500$	7.96	1.74	0.16	0.368
D_p-Optimal $N = 100$	59.86	2.80	0.41	0.168
D_p-Optimal $N = 250$	20.86	1.84	0.25	0.262
D_p-Optimal $N = 500$	12.19	1.68	0.20	0.308
D_b^{w}-Optimal $N = 100$	51.76	1.52	0.40	0.168
D_b^{w}-Optimal $N = 250$	18.88	1.22	0.25	0.271
D_b^{w}-Optimal $N = 500$	9.03	1.01	0.17	0.368
D_b^{i}-Optimal $N = 100$	29.07	2.16	0.31	0.215
D_b^{i}-Optimal $N = 250$	12.90	1.88	0.21	0.298
D_b^{i}-Optimal $N = 500$	8.61	1.92	0.17	0.337
ML_Some				
Shifted $N = 100$	32.28	1.12	0.74	0.087
Shifted $N = 250$	11.58	0.70	0.44	0.141
Shifted $N = 500$	6.02	0.77	0.32	0.194
D_p-Optimal$N = 100$	35.41	1.09	0.76	0.093
D_p-Optimal $N = 250$	14.50	0.90	0.50	0.137
D_p-Optimal $N = 500$	7.15	0.70	0.35	0.180
D_b^{w}-Optimal $N = 100$	42.66	0.82	0.83	0.086
D_b^{w}-Optimal $N = 250$	14.48	0.52	0.50	0.133
D_b^{w}-Optimal $N = 500$	7.23	0.46	0.36	0.162
D_b^{i}-Optimal $N = 100$	24.58	0.42	0.81	0.081
D_b^{i}-Optimal $N = 250$	9.83	0.68	0.50	0.122
D_b^{i}-Optimal $N = 500$	5.48	0.74	0.36	0.184

(KL-Asc), but the *D*-efficient experiment design is correctly informed with parameter values from the right DGP (MNL in Table 5). This would indeed be an unusual situation in real life, but it serves as a pedagogical tool to identify the effect of mis-specification. The Monte Carlo statistics for such a case are reported in Table 7, and the mis-specified model we employ is the flexible error component model with Asc for the *sq* (KL-Asc) in Eq. (5), while the DGP is an MNL.

In this case the bias will include a component due to model mis-specification, and the MSE of the effect of simulation variance. The results show a positive bias in all designs, i.e. an overestimate of marginal WTP. The values show that in this case too at medium ($N = 250$) and large ($N = 500$) sample sizes the best MSE-based performance is obtained by the D_b^{i}-optimal design, followed by the D_b^{w}-optimal one. However, there is a noticeable difference in the performance of the D_p-optimal and shifted designs between ML_alot and ML_some. While for ML_some these two designs achieve efficiency levels as good as those obtained by Bayesian designs (indeed at all sample sizes), the results are mixed at intermediate and low sample sizes for the attribute with high impact on utility ML_alot.

The fact that the Bayesian (informed *and* weakly informed) designs are the most robust in the presence of model mis-specification comes across best when observing the kernel plots of absolute relative error

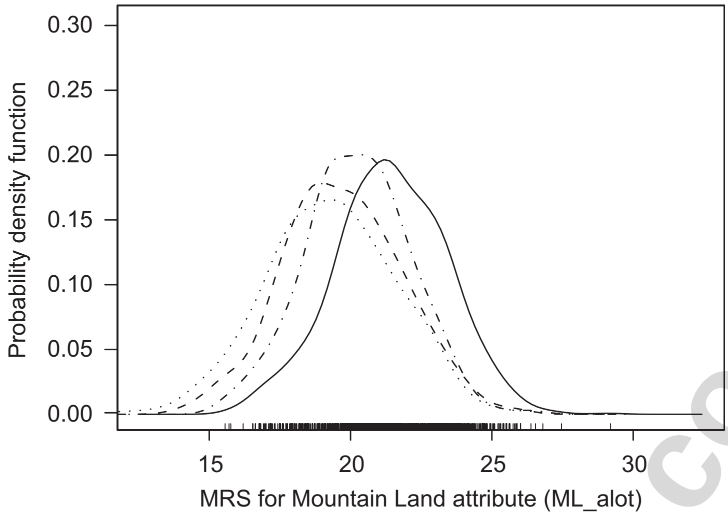

Fig. 1. DGP MNL and estimation MNL: *kernel-smoothed* distribution (optimal bandwidth) of the MRS estimates of landscape attribute mountain land ML_{alot}. MRS value of the DGP is 19.32: continuous line, *shifted* design; dashed line, D_p-optimal design; dotted line, D_b^{w}-optimal design; dashed and dotted line, D_b^{i}-optimal design.

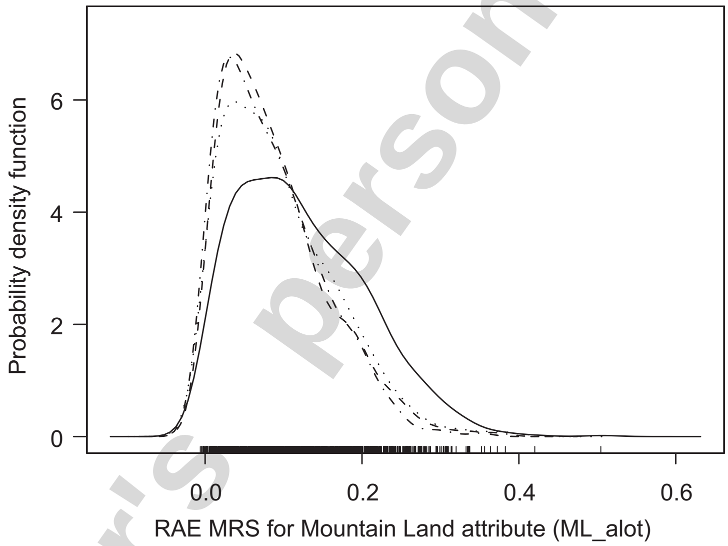

Fig. 2. DGP MNL and estimation MNL: *kernel-smoothed* distribution (optimal bandwidth) of the absolute relative error of landscape attribute mountain land ML_{alot}: continuous line, *shifted* design; dashed line, D_p-optimal design; dotted line, D_b^{w}-optimal design; dashed and dotted line, D_b^{i}-optimal design.

distributions in Fig. 3, which again refers to the large sample size scenario for the coefficient ML_alot. There is therefore evidence that

1. as long as the *a priori* design information is 'good' the Bayesian designs are robust to mis-specifications in the estimation phase when the sample is sufficiently large;
2. under all criteria the shifted design is preferable to the D_p-optimal at small and intermediate sample sizes;
3. and from Fig. 3, that at large sample sizes the D_p-optimal design produces large errors more frequently than the shifted design.

It is clear from the above that the D_p-optimal design incorporated with excessive precision the prior information on $\boldsymbol{\beta}$, thereby foregoing robustness to mis-specification. The Bayesian designs, instead, by

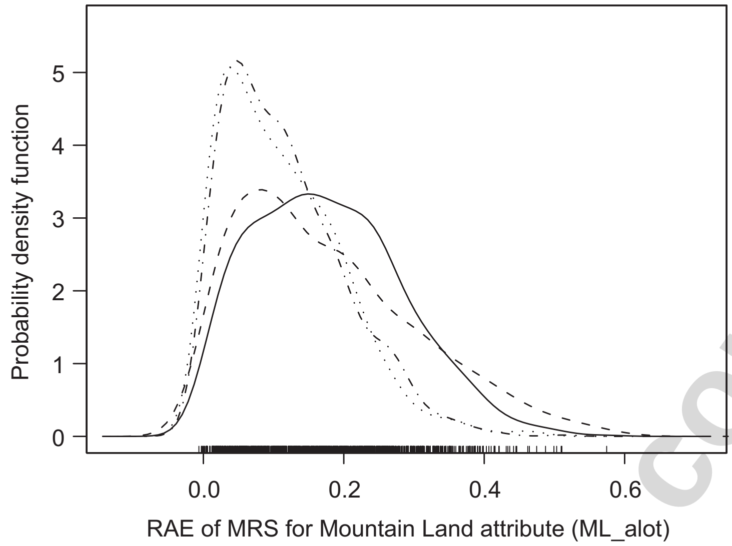

Fig. 3. DGP MNL and estimation KL-Asc, designed obtained assuming an MNL specification: *kernel-smoothed* distribution (optimal bandwidth) of the absolute relative error of landscape attribute mountain land ML_{alot}: continuous line, *shifted* design; dashed line, D_p-optimal design; dotted line, D_b^{w}-optimal design; dashed and dotted line, D_b^{i}-optimal design.

allowing some degree of uncertainty are robust to mis-specification. Such robustness, however, needs large sample sizes to produce a better performance than that obtained in the shifted design.

5.3. Correct specification, but incorrect design information

What happens—instead—when the *a priori* information incorporated in the *D*-efficient designs is 'poor' and the model specification subsequently chosen is correct (i.e. consistent with the DGP)? This case can occur in a variety of forms, and so this question can only be crudely answered by our investigation, and certainly—given its relevance—deserves a more tailored study. 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 values from Table 5) for the experiment design.

The choice of 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 (see, for example, [6,49]), which results in nested logit models providing a better fit than conditional logit models. The KL-Asc provides a similar covariance structure to the nested logit model with a degenerate nest for the *status quo* alternative [64]. It is also more flexible and has an objective function which is globally concave in the parameter space, it is hence deemed appropriate for a Monte Carlo simulation.

Table 8 reports the relevant statistics and shows the bias to be positive in all designs. The values show that under these conditions, under the $R\bar{A}E$ and $\Gamma_{0.05}$ criteria the shifted design outperforms all the *D*-efficient designs at all sample sizes and for both coefficient magnitudes. In terms of bias and mean squared error the shifted design does not perform much worse at intermediate to high sample sizes, while it dominates at low sample sizes. The distribution of RAE is illustrated in Fig. 4 for $N = 500$ and the high value coefficient ML_alot. The more incorrect information is built into the design, the higher the RAE produced, even when the specification used in estimation is the same as the DGP. While it is easy to anticipate this result rationally, our investigation provides ground for some less obvious considerations.

Importantly, results suggest that efficiency gains made available from advanced non linear and Bayesian-informed designs are only available in cases in which the *a priori* design information is good *and* that this outcome is robust to substantial model mis-specification if the sample size is large enough. In the absence of good quality *a priori* design information to build into the design, and at intermediate to small sample sizes, it would appear that researchers are better off using more rudimentary designs, even if these are only optimized for linear models. And this is exactly what the profession has been doing, perhaps inadvertently.

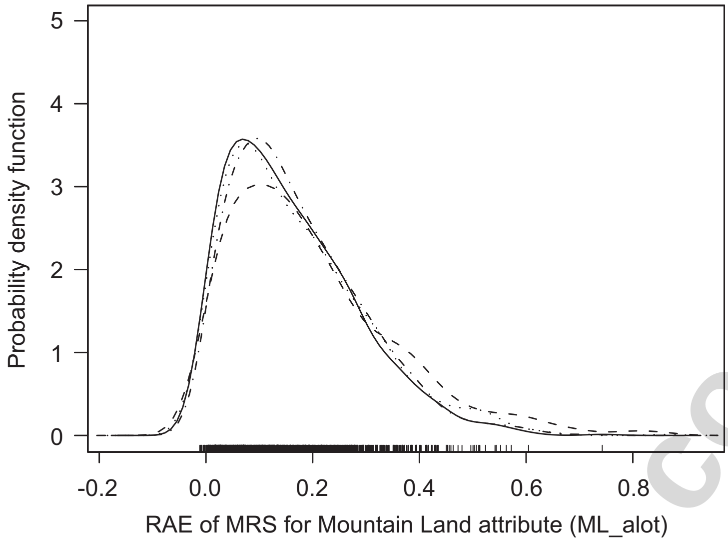

Fig. 4. DGP KL-Asc and estimation KL-Asc, designed obtained assuming an MNL specification: *kernel-smoothed* distribution (optimal bandwidth) of the absolute relative error of landscape attribute mountain land ML_{alot}: continuous line, *shifted* design; dashed line, D_p-optimal design; dotted line, D_b^{w}-optimal design; dashed and dotted line, D_b^{i}-optimal design.

6. Conclusions and further research

Although discrete choice responses from CEs are analyzed using nonlinear models, our literature review reveals that researchers in environmental economics applications have mostly used experimental designs for multivariate linear models. Experimental designs for logit models can be derived but require the analyst to formalize some guess on the true values. Bayesian methods can be used to account for the degree of uncertainty the analysts have about such guess. We review the literature in this area of research, then, by means of Monte Carlo experiments—and inspired by the results and structure of 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 which does not incorporate any *a priori* information on the true parameter values. The other three were specifically optimized for the highly nonlinear MNL model, and contained various forms of *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^{w}-optimal. Such designs can be derived on the basis 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^{i}-optimal design.

The features of the Monte Carlo experiments (sample size, DGPs, choice-set construction, etc.) were chosen so as to reflect the reality commonly faced by practitioners in environmental valuation. It is always difficult to generalize from Monte Carlo experiments, however, the results from our experiments suggest that efficiency gains are available from the use of Bayesian *D*-efficient designs for nonlinear-in-the-parameters models.

With good *a priori* information on the values of the unknown parameters gains can be available at all sample sizes. Even by building into the design relatively poor information (e.g. by means of D_b^{w}-optimal designs), efficiency gains become attractive at medium to large sample sizes ($N \geqslant 250$), but they are more significant when

- the attribute has a relatively large weight in the utility function;
- the *a priori* information on the parameters provided by the pilot is of good quality;
- *and* the DGP is consistent with the specification chosen for the derivation of the design.

However, when these conditions fail, the most 'rudimentary' of the designs we employed (the *shifted* design) does not perform much worse than *D*-efficient designs. The shifted design is derived from the common FFOD that currently dominates 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 may 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 experiment design construction.

6.1. Further research

Our research points to an area of potentially interesting and valuable contributions 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 CE surveys and especially given the encouraging results that bid design updating has produced in the field of contingent valuation [53,58].

Constructing designs using adaptive techniques can be a valuable strategy in choice-experiment surveys [55]. For example, one can systematically incorporate the information becoming available as the sampling progresses to derive gradually more tailored designs. The types of information needed is the parameter estimates and their variance–covariance into successive designs. A similar suggestion was put forward in [39] for the cost attribute. On the basis of our results we speculate that this updating should possibly involve a selection of attributes, such as those that are relevant for policy issues, 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 nonlinear function of parameter estimates, should be explicitly addressed in the measure of efficiency. This could translate—for example—as the maximization of the expected determinant 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 the efficiency of the estimates, future research should also focus on the construction of designs able to discriminate between competing specifications, perhaps by mixing measures of design information. Seminal research of this kind in the context of multivariate linear models is already available [5].

Acknowledgment

We thankfully acknowledge the provision of experiment designs, algorithms and comments from Z. Sándor, F. Carlsson and R. Kessels. We are also indebted to W. F. Kuhfeld and J. J. Louviere for various suggestions and encouragement and to K. Train for the GAUSS code for estimation of mixed logit, which we modified for our purposes. Comments by the editor, two reviewers, V. Adamowicz, M. Holmes, J. Lusk, E. Morey, F. Norwood, J. Rose, D. Street and from participants in the 2005 EAERE meeting in Bremen, in the Seminar Series of the Departments of Economics at the University of Waikato and at the Catholic University of Santiago helped us to substantially improve the paper. Special thanks must go to Z. Sándor who introduced the authors to the issue and provided very insightful suggestions. All the remaining errors are entirely our responsibility.

References

[1] A. Alberini, Optimal designs for discrete choice contingent valuation surveys: single-bound, double bound and bivariate models, J. Environ. Econ. Manage. 28 (1995) 287–306.

[2] A. Alberini, 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, J. Huber, Improving parameters estimates and model prediction by aggregate customization in choice experiments, J. Cons. Res. 28 (2) (2001) 273–283.

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

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

[6] R.K. Blamey, J.W. Bennett, J.J. Louviere, M.D. Morrison, J.C. Rolfe, Attribute causality in environmental choice modelling, Environ. Res. Econ. 23 (2002) 167–186.

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

[8] M.C. Blemier, J.M. Rose, D.A. Hensher, Constructing efficient stated choice experiments allowing for differences in error variances across subsets of alternatives, 2005. Unpublished Manuscript.

[9] A.W. Bowman, A. Azzalini, Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations, Oxford University Press, Oxford, UK, 1997.

[10] P.C. Boxall, W.L. Adamowicz, Understanding heterogenous preferences in random utility models: a latent class approach, Environ. Res. Econ. 23 (2002) 441–446.

[11] P.C. Boxall, W.L. Adamowicz, J. Swait, M. Williams, J.J. Louviere, A comparison of stated preference methods for environmental valuation, Ecol. Econ. 18 (1996) 243–253.

[12] D. Brownstone, K. Train, Forecasting new product penetration with flexible substitution patterns, J. Econs. 89 (1999) 109–129.

[13] D.S. Bunch, J.J. Louviere, 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, D. Street, Optimal designs for 2^k choice experiments, Commun. Statist. Theory Methods 32 (2003) 2185–2206.

[15] L. Burgess, D. Street, Optimal designs for choice experiments with asymmetric attributes, J. Statist. Plann. Inference 134 (2005) 288–301.

[16] F. Carlsson, P. Frykblom, C. Liljenstolpa, Valuing wetland attributes: an application of choice experiment, Ecol. Econ. 47 (2003) 95–103.

[17] F. Carlsson, P. Martinsson, Do hypothetical and actual marginal willingness to pay differ in choice experiments?, J. Environ. Econ. Manage. 41 (2001) 179–192.

[18] F. Carlsson, P. Martinsson, Design techniques for stated preference methods in health economics, Health Econ. 12 (2003) 281–294.

[19] S. Caussade, J. de D. Ortúzar, L.I. Rizzi, D.A. Hensher, Assessing the influence of design dimensions on stated choice experiment estimates, Transp. Res. B 39 (2005) 621–640.

[20] K. Chaloner, I. Verdinelli, Bayesian experimental design: a review, Statist. Sci. 10 (3) (1995) 273–304.

[21] J.R. DeShazo, G. Fermo, Designing choice sets for stated preference methods: the effects of complexity on choice consistency, J. Environ. Econ. Manage. 44 (2002) 123–143.

[22] R. Dhar, I. Simonson, The effect of forced choice on choice, J. Marketing Res. XL (2003) 146–160.

[23] D. Firth, J.P. Hinde, Parameter neutral optimal design for non-linear models, J. R. Statist. Soc. Ser. B Statist. Methodol. 59 (1997) 799–811.

[24] V. Foster, S. Mourato, Elicitation format and sensitivity to scope, Environ. Res. Econ. 24 (2003) 141–160.

[25] U. Grasshoff, H. Grossmann, H. Holling, R. Schwabe, Optimal paired comparison designs for first-order interactions, Statistics 37 (2003) 373–386.

[26] U. Grasshoff, H. Grossmann, H. Holling, R. Schwabe, Optimal designs for main effects in linear paired comparison models, J. Statist. Plann. Inference 126 (2004) 361–376.

[27] P.E. Green, V.R. Rao, A rejoinder to 'How many rating scales and how many categories shall we use in consumer research'—a comment, J. Marketing 35 (1971) 61–62.

[28] H. Grossman, H. Holling, R. Schwabe, Advances in optimum experimental design for conjoint analysis and discrete choice models, Econs. Models Marketing 16 (2002) 93–117.

[29] M.E. Haaijer, Modeling conjoint choice experiments with the probit model, Ph.D. Thesis, University of Groningen, Labyrint Publications, The Netherlands, April 1999.

[30] M.E. Haaijer, W.A. Kamakura, M. Wedel, The no-choice alternative in conjoint choice, Internat. J. Market Res. 43 (1) (2001) 93–106.

[31] N. Hanley, R.E. Wright, V. Adamowicz, Using choice experiments to value the environment, Environ. Res. Econ. 11 (3–4) (1998) 413–428.

[32] N. Hanley, R.E. Wright, G. Koop, Modelling recreation demand using choice experiments: climbing in Scotland, Environ. Res. Econ. 22 (2002) 449–466.

[33] J.A. Herriges, D. Phaneuf, Inducing patterns of correlation and substitution in repeated nested logit models of recreation demand, Am. J. Agric. Econ. 84 (4) (2002) 1076–1090.

[34] T.P. Holmes, W.L. Adamowicz, Attribute-based methods, in: P.A. Champ, K.J. Boyle, T.C. Brown (Eds.), A Primer on Nonmarket Valuation, Kluwer Academic Publisher, Dordrecht, 2003, pp. 171–219.

[35] P. Horne, L. Petäjistö, Preferences for alternative moose management regimes among Finnish landowners: a choice experiment approach, Land Econ. 79 (4) (2003) 472–482.
[36] J. Huber, K. Zwerina, The importance of utility balance in efficient choice designs, J. Marketing Res. 39 (1996) 214–227.
[37] J. Jin, Z. Wang, S. Ran, Comparison of contingent valuation and choice experiment in solid waste management programs in Macao, Ecol. Econ. 57 (2005) 430–441.
[38] B.J. Kanninen, Design of sequential experiments for CV studies, J. Environ. Econ. Manage. 25 (1993) s-1–s-11.
[39] B.J. Kanninen, Optimal experimental design for double-bounded dichotomous choice contingent valuation, Land Econ. 69 (2) (1993) 138–146.
[40] B.J. Kanninen, Optimal designs for multinomial choice experiment, J. Marketing Res. 39 (2002) 214–227.
[41] R. Kessels, P. Goos, M. Vandebroek, Comparing algorithms and criteria for designing Bayesian conjoint choice experiments, Working Paper, Department of Applied Economics, Katholieke Universiteit Leuven, Belgium, 2004.
[42] R. Kessels, P. Goos, M. Vandebroek, A comparison of criteria to design efficient choice experiments, J. Marketing Res. 43 (3) (2006).
[43] W.F. Kuhfeld, Marketing research: Methods, SAS Experimental Design, Choice, Conjoint, and Graphical Techniques, SAS Institute, Cary, NC, USA ⟨http://support.sas.com/techsup/tnote/tnote_stat.html⟩, 2005.
[44] W.F. Kuhfeld, R.D. Tobias, Large factorial designs for product engineering and marketing research applications, Technometrics 47 (2005) 132–141.
[45] W.F. Kuhfeld, R.D. Tobias, M. Garratt, Efficient experimental design with marketing research applications, J. Marketing Res. 31 (1994) 545–557.
[46] K. Lancaster, A new approach to consumer theory, J. Polit. Econ. 84 (1966) 132–157.
[47] K. Lancaster, Operationally relevant characteristics in the theory of consumer behavior, in: M. Peston, B. Corry (Eds.), Essays in honor of Lord Robbins, Weidenfels and Nicholson, London, 1972.
[48] A.G. Lazari, D.A. Anderson, Designs of discrete choice set experiments for estimating both attribute and availability cross effects, J.Marketing Res. 31 (1994) 375–383.
[49] E. Lehtonen, J. Kuuluvainen, E. Pouta, M.M. Rekola, C.-Z. Li, Non-market benefits of forest conservation in southern Finland, Environ. Sci. Policy 6 (2003) 195–204.
[50] J.J. Louviere, D.A. Hensher, Using discrete choice models with experimental design data to forecast consumer demand for unique cultural event, J. Cons. Res. 10 (1983) 348–361.
[51] J.J. Louviere, G. Woodworth, Designs analysis of simulated consumer choice or allocation experiments: an approach based on aggregate data, J. Marketing Res. 20 (1983) 350–367.
[52] J.L. Lusk, F.B. Norwood, Effect of experimental design on choice-based conjoint valuation estimates, Am. J. Agric. Econ. 87 (2005) 771–785.
[53] H. Nyquist, Optimal designs of discrete response experiments in contingent valuation studies, Rev. Econ. Statist. 74 (1992) 559–563.
[54] T. O'Leary, A. McCormack, G. Hutchinson, D. Campbell, R. Scarpa, B. Riordan, Landscape impact of REPS: a quantitative assessment, Report to the Department of Agriculture and Food, Dublin, Ireland, 2005.
[55] D. Raghavarao, J.B. Wiley, Design strategies for sequential choice experiments involving economic alternatives, J. Statist. Plann. Inference, 136 (2006) 3287–3306.
[56] M.X.V. Ròdriguez, C.J. Leòn, Altruism and the economic values of environmental and social policies, Environ. Res. Econ. 28 (2004) 233–249.
[57] J. Rolfe, J. Bennett, J. Louviere, Choice modelling and its potential application to tropical rainforest preservation, Ecol. Econ. 35 (2000) 289–302.
[58] K. Rollins, W. Wistowsky, M. Jay, Wilderness canoeing in Ontario: using cumulative results to update dichotomous choice contingent valuation offer amounts, Canad. J. Agr. Econ. 45 (1997) 1–16.
[59] J.M. Rose, 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.
[60] K. Sælensminde, The impact of choice inconsistencies in stated choice studies, Environ. Res. Econ. 23 (2002) 403–420.
[61] Z. Sándor, M. Wedel, Designing conjoint choice experiments using managers' prior beliefs, J. Marketing Res. 38 (2001) 430–444.
[62] Z. Sándor, M. Wedel, Profile construction in experimental choice designs for mixed logit models, Marketing Sci. 4 (2002) 455–475.
[63] R. Scarpa, D. Campbell, 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.
[64] R. Scarpa, S. Ferrini, 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, Springer, Berlin, 2005, pp. 247–274.
[65] R. Scarpa, E.S.K. Ruto, P. Kristjanson, M. Radeny, A.G. Drucker, Valuing indigenous cattle breeds in Kenya: an empirical comparison of stated and revealed preference value estimates, Ecol. Econ. 45 (2003) 409–426.
[66] V. Severin, Comparing statistical efficiency and respondent efficiency in choice experiments, Ph.D., School of Marketing. University of Sydney, Sydney, 2000.
[67] D.J. Street, D.S. Bunch, B. Moore, Optimal designs for 2^k paired comparison experiments, Commun. Statist. Theory Methods 30 (2001) 2149–2171.
[68] D.J. Street, L. Burgess, Optimal and near-optimal pairs for the estimation of effects in 2-level choice experiments, J. Statist. Plann. Inference 118 (2004) 185–199.
[69] D.J. Street, L. Burgess, Optimal Stated Preference Choice Experiment When all Choice Sets Contain a Specific Option, University of Technology, Sydney, 2004.

[70] D.J. Street, L. Burgess, J.J. Louviere, Quick and easy choice sets: constructing optimal and nearly optimal stated choice experiments, Int. J. Res. Marketing 22 (2005) 459–470.
[71] K. Train, Discrete Choice Methods with Simulation, Cambridge University Press, New York, 2003.
[72] R. Viney, E. Savage, J. Louviere, Empirical investigation of experimental design properties of discrete choice experiments in health care, Health Econ. 14 (2005) 349–362.
[73] P. Wattage, S. Mardle, S. Pascoe, Evaluation of the importance of fisheries management objectives using choice-experiments, Ecol. Econ. 55 (2005) 85–95.
[74] K. Zwerina, J. Huber, W.F. Kuhfeld, A general method for constructing efficient choice designs, Working paper, Fuqua School of Business, Duke University, Durham, 1996.