“The importance of second-order interactions in a forest choice experiment. A partial log-likelihood analysis”

Marek Giergiczny (Poland), Pere Riera (Spain), Joan Mogas (Spain), Pierre-Alexandre Mahieu (France)

The importance of second-order interactions in a forest choice experiment. A partial log-likelihood analysis

Abstract

Although it is known that main effects models could misrepresent individuals' preferences in choice experiment applications, a review of major journal articles suggests that the current practice is to obviate the second-order interaction designs. This article presents a partial log-likelihood analysis of a forest valuation study where interactions are found to be more important than some of the main effects.

Keywords: choice experiment, two-way interactions, experimental design, partial log-likelihood analysis.
JEL Classification: H41, Q51.

Introduction

Environmental goods are often valued by means of choice modeling techniques. The most widely used variant of these techniques is probably the pair-wise choice experiment (CE). This typically involves a survey in which respondents are faced with one or several choice sets, each containing three alternatives, including the business-as-usual (BAU). Each alternative involves a bid amount to be paid, which generally equals zero for the BAU, as well as the level of each relevant non-monetary attribute of the good. The respondent's task is to state his/her preferred option (Hensher et al., 2005). This setting is consistent with the random utility maximization model (RUM) and several econometric treatments have been developed, as it will be explained in the methodology section.

The indirect utility function (IUF) of the participants can take several forms. Among other things, the practitioner has to decide whether to include interaction variables in the IUF. There are two main types of interactions: between attributes of the good, and between attributes (or alternative specific constants) and socio-demographic characteristics. The latter is often interpreted as a way of dealing with systematic taste variations. The former is usually regarded as individuals perceiving attributes as complements or substitutes. This paper focuses on this form of non-linearity in the utility function specification.

A current practice in non-market valuation is to assume that all two-way or higher order interaction terms among attributes in the IUF are equal to zero. Consequently, most authors focus on main effect designs and discard designs that allow for the estimation of interactions. A review from 2009 of seven major journals in which environmental valuation estimations are published – Ecological Economics, Environmental and Resource Economics, Journal of Choice Modelling, Journal of Environmental Economics and Management, Journal of Forest Economics, Land Economics, and Resource and Energy Economics – rendered 36 papers reporting choice modeling applications. Six articles specified that their design enabled the estimation of two-way interactions (Alvarez-Farizo et al., 2009, p. 791; Bateman and Munro, 2009, p. 125; Boyle and Ozdemir, 2009, p. 253; Chattopadhyay, 2009, p. 2837; Hoyos et al., 2009, p. 2375). Of these 6 papers, only Bateman and Munro (2009) report testing some of the two-way interactions between attributes: "we experimented with a number of specifications, including nonlinear interactions between the attributes" (p. 127).

A reason for this lack of practical assessment of second-order effects in environmental valuation may be that it demands more complex designs than for main effects only. Another possible reason is the belief that the importance of two-way interactions might be rather modest compared to the main effects. Two-way interactions have been said to explain about 6%-10% of the data variance (Louviere, 1988; Louviere et al., 2000), based on research carried by Dawes and Corrigan (1974) referring to linear models. However, discrete choice models are highly non-linear and it is unclear whether two-way interactions ought to be discarded.

Another belief is that a large sample size may be a substitute for a poor experimental design. Lusk and Norwood (2005) tested the effect of several experimental designs on valuation estimates in a Monte-Carlo simulation and found that under certain conditions, the estimated Willingness to Pay (WTP) from the main effects design is not statistically different from the true WTP even if nonlinear effects are present. However, the study raised some controversies. For instance, Carson et al. (2009) criticized the choice of IUF parameters. The interaction terms considered by Lusk and Norwood (2005) were five to ten times smaller than the main effects parameters, which resulted in a nearly linear IUF.

The significance of two-way interactions in forest valuation had previously been highlighted by Mogas et al. (2006). This paper analyzes a subset of data from the same valuation exercise reported by Mogas et al. (2006), but departs from it by using a different econometric model and by quantifying and discussing the relative explanatory power of second order effects based on a partial log-likelihood analysis (Lancsar et al., 2007). The procedure quantifies the contribution of each attribute and two-way interaction to the log-likelihood (LL) value of the model estimation, thus allowing the comparative analysis. The remainder of the paper is organized as follows. The next Section deals with some issues related to designing stated preference experiments. It is followed by a methodology section and a description of the field experiment. It concludes with the results, including some discussions, and conclusions. The overall intend of the paper is to discuss the importance of second-order interactions, and consequently the discussion and conclusions focus on the methodology rather than the policy implications of the case study.

1. Design

A number of significant advances have taken place in the field of experimental design in recent years. Orthogonality is no longer perceived as a requirement for a "good" experimental design by some researchers (Ferrins and Scarpa, 2007; Rose and Bliemer, 2008). As mentioned by Rose and Bliemer (2008), the experimental design literature was oriented to linear models. In such models, orthogonality of data is deemed important as it ensures that linear regression models do not suffer from multicollinearity. Orthogonal designs produce zero off-diagonals in the variance-covariance matrix, thus ensuring that the parameter estimates are unconfounded with one another. However, it does not fully apply to discrete choice models, which are non-linear (Train, 2003). Bliemer and Rose (2005) show that orthogonal designs are only efficient under the assumption that all parameters are equal to zero (i.e., all attributes in the design are likely to not influence the choice decisions observed) and the alternatives are unlabelled.

Efficient designs are now frequently used. This type of design supposes some prior knowledge of the parameters of the IUF. Some researchers consider the zero priors null hypothesis (e.g., Street et al., 2005; Burgess and Street, 2003, Street and Burgess, 2004; Street et al., 2001), while others assume some prior knowledge of the parameters (e.g., Huber and Swerina, 1996; Kanninen, 2002; Carlsson and MArtinsson, 2003; Bliemer and Rose, 2006; Ferrini and Scarpa, 2007; Bliemer et al., 2009). Prior knowledge on the parameters can be used to construct the so-called D-efficient designs. When doing so, the D-error, which is related to the variance covariance matrix, is minimized.

Another advance in constructing experiments is related to specific experimental designs for econometric models dealing with actual collected data – such as when some of the requested choices are missing. Sandor and Wedel (2002; 2005) addressed preference heterogeneity by generating experimental designs specifically for mixed logit models. Ferrini and Scarpa (2007) extended the literature on designing stated choice experiments by generating experimental designs assuming an error components model structure. Bliemer and Rose (2010) generated designs that are appropriate for a panel formulation of the mixed logit model.

Despite the recent advances in experimental design, we decided to construct two orthogonal designs: a main effect design and a two-way interactions design. The main reason for this was to assess the relative importance of two-way interactions compared to main effects by means of a partial log-likelihood analysis, a statistical approach that requires an orthogonal design (Lancsar et al., 2007).

Yu et al. (2006) showed that when interaction influences choices, interaction designs perform better than main effect designs in terms of prediction accuracy. They also conducted a sensitivity analysis of the prediction accuracy when wrong priors for two-way interactions were used. As expected, better prior information yields higher accuracy. They showed that constructing an interaction design and specifying an interaction model is the most advisable procedure even when completely wrong priors are defined (Yu et al., 2006). The orthogonal design for interactions used in our study can be regarded as an efficient one in which wrong priors are assumed and according to findings by Yu et al. (2006), greater accuracy should be achieved using an interaction design than a main effect design.

The valuation exercise estimates the impact of alternative afforestation programs on non-market forest values. The afforestation program was assumed to take place in Catalonia, a region in the Northeast of Spain, which has 1.3 million ha of forests, accounting for circa 40% of its total area. Pine is the predominant species, accounting for half of the population, with holm oak being the next most important specie, covering some 10% of the forest land (Ministerio de Ambiente, 1996).

The questionnaire contains three parts. The first part described some positive and negative effects of the afforestation program to the respondent. The program would result in an increase in the forest coverage from the current 40% of the Catalonian area to 50%. The planting was to take place in marginal agricultural land. The central part of the questionnaire contained the CE exercise. The final part included some debriefing questions and questions related to the socio-demographic characteristics of the respondent.

The attributes and their levels were set in collaboration with policymakers and forest researchers, and adjusted after focus groups and a pre-test of the survey. Table 1 shows the attributes and levels finally used in the questionnaire. The payment vehicle was a compulsory annual contribution from Catalan residents to a fund exclusively dedicated to the afforestation program.

Table 1. Attributes and levels used in the choice experiment exercise

Attribute	Description	Level
Picnic	Picnicking allowed in the new forests (BAU = No)	Yes No
Drive	Car driving allowed in the new forests (BAU = No)	Yes No
Mushroom	Picking mushrooms allowed in the new forests (BAU = No)	Yes No
CO_2	CO_2 sequestered annually by the new forests. Equivalent to the pollution produced annually by a city of … (BAU = 0)	300,000 people 400,000 people 500,000 people 600,000 people
Erosion	Number of years the productivity of the soil will remain (BAU = 0)	300 years 500 years 700 years 900 years
Cost	The afforestation cost per person and year (1999 values) (BAU = 0)	5 € 10 € 15 € 20 €

The questionnaire was administered to a sample of Catalan residents. Face-to-face interviews were conducted in respondents' homes. Two versions of the CE questionnaire were randomly assigned to two sub-samples of 400 individuals each. The locations of the survey were randomly drawn for each subsample. Quotas were used in order for the samples to be representative of the general population in terms of location, age and gender. The average response rate in the CE was 95%.

Two designs were constructed. The first accounted for main effects only and resulted in 16 pair-wise comparisons of afforestation scenarios, which were randomly blocked to 4 different versions, each of which had 4 choice sets of two generic (unlabeled) afforestation alternatives. BAU was also included in each choice set. In each choice set, respondents were thus asked for their preferred choice between BAU and the two alternative afforestation scenarios. The D-error of the main effect design (assuming all parameters are equal to zero) equals 0.52.

The second design allowed the estimation of all main effects and two-way attribute interactions between non-monetary attributes. It contained 64 scenarios. The experiment was blocked into 16 versions of four choice sets. Each version was displayed to an equal number of respondents to obtain an orthogonal data set (given that all choice sets are completed). The D-error of the interaction design (assuming all parameters to be zero) equals 0.15.

Ex-post statistics were produced to explore the efficiency of the designs. The minimum sample size requirements to obtain statistically significant coefficients (at 0.05 level) and WTP estimates were calculated. It was assumed that the estimated parameters were true. The details of the calculation of sample size requirements are available in Scarpa and Rose (2008) and Bliemer and Rose (2005). Efficiency measures were calculated with Ngene software and the results are reported in Tables 2 and 3. The necessary sample size for most attributes is smaller than the sample size used (1,600 observations). Given the level of efficiency of the design, reliable comparison of the WTP obtained from the main and interaction effect designs is therefore possible.

Table 2. Minimum design replication requirements by attribute for interaction design – MNL model

	Picnic	Drive	Mushroom	CO_2	Erosion	Picnic×Drive	Mushroom×Drive	Cost
B	4.41	20.30	1.75	14.86	8.39	15.53	3.79	1.65
WTP	5.17	23.35	2.72	15.51	8.29	17.07	4.98	

Notes: *The interaction design contained 64 scenarios.

Table 3. Minimum design replication requirements by attribute for main effect design – MNL model*

	Picnic	Drive	Mushroom	CO_2	Erosion	Cost
B	13.12	3.93	29.06	9.62	4.31	20.69
WTP	6.71	20.16	52.62	24.75	17.05	

Notes: *The main effect design contained 16 scenarios.

2. Method

In a CE exercise individuals are asked to identify their preferred choice i from among a given set of J alternatives. The data analysis follows a RUM (McFadden, 1974). Under RUM, it is assumed that the observed choice from individual n is the one she expected to provide her with the highest utility. Her utility function, U_{ni}, can be decomposed into a systematic part, V_{ni}, and a stochastic part,

$\varepsilon_{ni}, U_{ni} = V_{ni} + \varepsilon_{ni}$.

The probability P_{ni} that individual n chooses alternative i instead of another alternative j of the choice set is

$$P_{ni} = Pr(V_{ni} + \varepsilon_{ni} > V_{nj} + \varepsilon_{nj} \forall j \neq i).$$

If ε_{nj} is assumed to be an independently and identically distributed extreme value type 1, this probability has a closed form expression:

$$P_{ni} = \frac{e^{\beta' x_{ni}}}{\sum_j e^{\beta' x_{nj}}}, \tag{1}$$

where x is a vector of variables which may or may not include interaction terms and β a vector of parameters. Equation (1) is often referred to as a logit choice probability function.

The standard multinomial logit model (MNL) has some limitations, as listed by Train (2003):

1. It exhibits a property of independence from irrelevant alternatives.
2. MNL can represent only the systematic taste variation but not random taste variations.
3. It cannot handle situations where the unobserved part of the utility function is correlated over time.

The mixed logit model can be considered to deal with these limitations. Mixed logit probabilities can be expressed as the integrals of standard logit probabilities over a density of parameters. As explained in Train (2003), a mixed logit model is a model in which choice probabilities take the form

$$P_{ni} = \int \frac{e^{\beta'_n x_{ni}}}{\sum_j e^{\beta'_n x_{nj}}} \varphi(\beta | b, \Omega) d\beta,$$

where $\frac{e^{\beta'_n x_{ni}}}{\sum_j e^{\beta'_n x_{nj}}}$ is a standard logit formula, $\varphi(\beta|b,\Omega)$ is the density of the random coefficients with mean b and covariance Ω. For example, the logit expression in equation (1) can be treated as a special case of mixed logit with β being fixed.

The BAU option in a choice set is usually experienced and can cause respondents to consider it in a systematically different manner from the hypothetical program alternatives. As a result, there is more likely to be a correlation between the utility from the experimentally designed hypothetical alternatives than with the utility associated to the status-quo alternative. This may be captured by a specification with additional errors accounting for this difference in the correlation across utilities (Campbell et al., 2008). Correlation is a consequence of sharing this extra error component, which is absent from the BAU alternative. Empirical evidence supports the BAU bias (see, for example, Kontoleon and Yabe, 2003; Lehtonen et al., 2003; Scarpa et al., 2005) and possible explanations to account for it have been proposed in the literature (Samuelson and Zeckhauser, 1988; Haijer et al., 2001). The mixed logit with an error component structure has been used to account for the correlation in unexplained part of utility between the program alternatives. This model specification allows for different patterns of correlations between utilities implying change and those referring to BAU (Brownstone and Train, 1999).

The unobserved factors in MNL that affect respondents are assumed to be independent of the repeated choices, which may be considered unrealistic in the CE exercises where respondents usually make more than one choice. There might be some unobserved factors that are constant over the choices made by the same individual facing several choice sets, and unobserved parts of the utilities over the choices may consequently be correlated. Mixed logit models can account for dependence across repeated choices by the same respondent by specifying a panel version of the model, which overcomes the MNL limitation (3) mentioned above. Conditional on β, the probability that individual n makes a sequence of T choices is the product of logit formulas, as in

$$P_{ni} = \prod_{t=1}^{T} \left[\frac{e^{\beta_n' x_{nit}}}{\sum_j e^{\beta_n' x_{njt}}} \right],$$

where t denotes the sequence of choices made by the same respondent.

Since β_n is not known, the unconditional probability is given by the integral over all possible values of β_n, i.e.,

$$P_{ni} = \int \prod_{t=1}^{T} \left[\frac{e^{\beta_n' x_{nit}}}{\sum_j e^{\beta_n' x_{njt}}} \right] \varphi(\beta | b, \Omega) d\beta, \qquad (2)$$

with $\varphi(\beta | b, \Omega)$ being the density of a random parameter with mean b and covariance matrix Ω.

3. Results and discussion

An MNL model was estimated in the first step and a panel version of the error components random parameter logit (RPL) model was used in the second step. Since the integral in equation (2) cannot be evaluated analytically, the estimation of the probabilities has to rely on a simulation method (Train, 1999). A simulated maximum likelihood estimator with Halton draws was used. In total, 500 Halton draws were generated in each run, producing an approximation similar to 5,000 pseudo-random draws. The random parameters were assumed to be independent. In the first run, we let all attribute parameters be normally distributed, except for the cost, which was fixed. In the final models, only parameters that have standard deviations significant at the 5% level were assumed to be random. All models included error components that combined the afforestation alternatives, in addition to random parameters. The presence of the error components demonstrates that two afforestation alternatives share some characteristics that are not captured by the attributes included in the utility functions on which the data were modeled.

Two IUF forms were fitted to both datasets – interaction and main effects designs. The combination of two different designs with two IUFs and two models (MNL and RPL) yields 8 different models, for which the estimates are shown in Tables 4 (MNL) and 5 (RPL). The nonlinear effects in CO_2 and Erosion were tested against a linear specification. Since the change in the LL value due to imposing the restriction on equal slopes turned out to be statistically not significant (which holds for all specifications and models), a linear specification for CO_2 and Erosion was assumed in the final model. Similarly, restricting the alternative specific constants (ASCs) so that they were all equal did not have a significant impact on the fit of the models. The parameters of the utility function were estimated using NLOGIT 4.0 statistical software. The results based on the interaction design are discussed in Section 3.1, and they are compared with the results based on the main effect design in Section 3.2.

Table 4. MNL results

	Model 1	Model 2	Model 3	Model 4
	Interaction design Interaction IUF	Main effect design Interaction IUF	Interaction design Main effect IUF	Main effect design Main effect IUF
Picnic	0.453*** (4.398)	.173 (1.311)	0.286*** (4.051)	0.157* (1.858)
Mushroom	0.780*** (7.027)	1.016*** (6.233)	0.393*** (5.084)	0.473*** (6.191)
Drive	-0.264** (-2.047)	-0.147 (-0.855)	-0.897*** (-5.940)	-0.813*** (-11.061)
CO_2	0.088** (2.131)	0.124*** (2.967)	0.080*** (2.867)	0.180*** (4.522)
Erosion	0.108*** (2.948)	0.088** (1.960)	0.109*** (3.323)	0.119*** (3.099)
Cost	-0.269*** (-7.193)	-0.251*** (-19.609)	-0.267*** (-7.216)	-0.451*** (-10.022)
Picnic×Drive	-0.384** (-2.400)	-0.148 (-0.722)		
Mushroom×Drive	-0.808*** (-6.623)	-1.155*** (-4.170)		
ASC	0.634*** (3.938)	-0.302 (-1.405)	0.854*** (10.118)	0.858*** (5.018)
LL Full model	-1622.69	-1572.13	-1638.93	-1585.27
LL constant only	-1736.44	-1722.04	-1736.44	-1722.04
ρ^2	0.067	0.087	0.057	0.081
N	1600	1600	1600	1600

Notes: t-statistic is in brackets; ***, ** and * refer to statistically significant at 1%, 5% and 10% levels, respectively.

Table 5. RPL results

	Model 5		Model 6		Model 7		Model 8	
	Interaction design Interaction IUF		Main effect design nteraction IUF		Interaction design Main effect IUF		Main effect design Main effect IUF	
Picnic	0.495*** (4.091)	0.671*** (3.901)	0.120 (0.828)	0.525*** (2.692)	0.297*** (3.379)	0.720*** (4.472)	0.106 (1.140)	0.477*** (2.385)
Mushroom	0.875*** (6.629)	0.731*** (3.951)	1.151*** (6.430)	0.461** (2.067)	0.447*** (4.796)	0.666*** (3.469)	0.516*** 6.143)	0.392** (2.096)
Drive	-0.341** (-2.229)	0.516** (2.205)	-0.131 (-0.728)		-1.004*** (-9.654)	0.526** (2.307)	-0.864*** (-10.524)	
Erosion	0.106*** (2.812)		0.098* (1.943)		0.097*** (2.640)		0.148*** (3.459)	
CO_2	0.144*** (3.881)		0.153*** (3.485)		0.149*** (4.027)		0.223*** (5.259)	
Cost	-0.325*** (-7.437)		-0.057*** (-4.028)		-0.315*** (-7.412)		-0.515*** (-9.626)	
Picnic×Drive	-0.467** (-2.420)		-0.129 (-0.571)					
Mushroom×Drive	-0.915*** (-4.621)		-1.342*** (-4.453)					
ASC	0.961*** (4.854)		-0.567** (-2.343)		1.232*** (6.519)		1.174*** (6.147)	
Error Component	1.146*** (7.539)		1.195*** (8.004)		1.135*** (7.517)		1.198*** (8.150)	
LL Full model	-1580.56		-1536.71		-1597.12		-1557.94	
LL constant only	-1736.44		-1722.04		-1736.44		-1722.04	
ρ^2	0.100		0.127		0.091		0.116	
N	1600		1600		1600		1600	

Notes: t-statistic is in brackets; ***, ** and * refer to statistically significant at 1%, 5% and 10% levels, respectively.

3.1. Results based on the interaction design. The signs of all main effects are as it has ben expected in all the models. In addition, all the program attributes are significant factors in the choice of afforestation scenario. All parameters, including two-way interactions, are statistically significant at 0.05 level. The positive sign of the Picnic, Mushroom, CO_2, and Erosion parameters suggests that afforestation programs were more likely to be chosen when picnicking and picking mushrooms were permitted, the amount of CO_2 sequestered was high, and erosion was postponed for longer. However, afforestation programs with higher costs and driving allowed were less likely to be chosen by the participants.

The standard deviations of Mushroom, Picnic and Drive are statistically significant at 0.05 level with the mixed logit model for the interaction design data. The standard deviation of the error components are significant ($p < 0.0001$) for both IUFs.

Both interaction terms, Picnic×Drive and Picnic×Mushroom, are statistically significant and are included in the final model. The hypothesis that Picnic×Drive and Picnic×Mushroom jointly equal 0 is rejected, with $p < 0.0001$ in both MNL and RPL models.

The coefficients for Picnic×Drive and Picnic×Mushroom are both negative, indicating that the utility derived from picnicking and mushrooming depends on whether driving is permitted. The result is consistent with some of the statements made in the focus groups. A few participants complained about the presence of cars in the new forests, and about the current number of off-road vehicles in the existing forests.

3.1.1. Partial log-likelihood analysis. Based on Lancsar et al. (2007), a partial log-likelihood analysis was also undertaken. The attributes were ranked according to their statistical importance, which was deemed to be their contribution to the LL value of the model estimation. For each attribute, the change in the LL value, their relative effect, and the cumulative effect were estimated after including/removing the attribute. The relative effect was calculated as the percentage change in the LL value. The results for the MNL model with interaction design and interaction IUF are shown in Table 6. A similar pattern holds for the RPL model, which is the reason the results are not reported here.

Table 6. Partial log-likelihood analysis for the MNL model with interaction design and interaction IUF

Attribute excluded	Log-likelihood	Partial effect change in log-likelihood	Relative effect in log-likelihood	Cumulative (%)
None	-1622.69			
Cost	-1649.18	-26.49	29.20	29.20
Mushroom	-1647.62	-24.93	27.48	56.68
Mushroom×Drive	-1638.93	-16.24	17.90	74.58
Picnic	-1632.26	-9.57	10.54	85.12
Erosion	-1628.04	-5.35	5.89	91.02
CO_2	-1626.11	-3.42	3.77	94.79
Picnic×Drive	-1625.51	-2.82	3.10	97.89
Drive	-1624.60	-1.91	2.11	100.00

The payment alone accounts for almost 30% of the LL value. Contrary to the results reported in Lancsar et al. (2007), two-way interactions substantially increase the LL value, with the change in LL due to including Mushroom×Drive being larger than most of the main effects. The change due to including Picnic×Drive is of a similar magnitude to CO_2 and Drive. Jointly Mushroom×Drive and Picnic×Drive account for a 21% change in the LL value.

3.1.2. Magnitude of the interactions and their impact on WTP. In order to facilitate comparison of the relative impact of the attributes on the utility, the continuous variables CO_2, Erosion and Cost were coded using the same levels (1, 2, 3, 4), taking advantage of the equal spacing between all the continuous variables. Driving, Picnic and Mushroom were dummy coded. The magnitude of the Mushroom×Driving parameter from Models 1 and 5 is larger than the magnitude of the Mushroom, Driving and Picnic parameters alone, and it is also larger than the magnitude of the highest level of CO_2 and Erosion.

The importance of the two-way interactions in this case study can also be appreciated with the marginal rate of substitution. When two-way interactions are included in the model, the WTP for the attribute that enters the interaction term is not simply the ratio between the coefficient of this attribute and the cost coefficient, because the attribute enters the IUF multiplicatively. For example, when calculating the marginal utility of Driving, the utility function is partially differentiated with respect to Driving, including both the main effects and all the two-way attribute interactions related to driving. The WTP of Driving will thus depend on whether mushrooming and picnicking are permitted, given that the corresponding interaction variables are statistically significant. The WTP values for MNL are shown in Table 7 and those for RPL in Table 8. The calculations of estimates for attributes that enter two-way interactions are conditional on whether other attributes are present.

Table 7. WTP estimates based on MNL (euro of 2006)

	Model 1 WTP	Model 2 WTP		Model 3 WTP	Model 4 WTP
	Interaction design Interaction IUF	Main effect design Interaction IUF		Interaction design Main effect IUF	Main effect design Main effect IUF
Drive Mushroom = 0, Picnic = 0	-4.89** (-2.022)	-2.86 (-0.938)	Drive	-15.92*** (-6.219)	-8.95*** (-7.584)
Drive Mushroom = 1, Picnic = 0	-19.85*** (-5.281)	-25.38*** (-2.795)			
Drive Mushroom = 0,Picnic = 1	-12.01*** (-5.229)	-5.75* (-1.701)			
Drive Mushroom = 1, Picnic = 1	-26.97*** (-6.158)	-28.27*** (-3.083)			
Picnic Drive = 0	8.38*** (3.758)	3.39 (1.233)	Picnic	5.33*** (3.577)	1.73* (1.696)
Picnic Drive = 1	1.26 (0.615)	0.493 (0.186)			
Mushroom Drive = 0	14.4*** (5.330)	19.9*** (2.686)	Mushroom	7.33*** (4.416)	5.20*** (5.049)
Mushroom Drive = 1	-0.51 (-0.243)	-2.75 (-0.779)			
Erosion	2.01*** (3.092)	1.71** (2.250)	Erosion	2.05*** (3.130)	1.32*** (3.474)
CO_2	1.63** (2.403)	2.41*** (2.975)	CO_2	1.49*** (2.619)	1.99*** (5.853)

Note: t-statistic is in brackets.

Table 8. WTP estimates based on RPL (Euro of 2006)

	Model 5 WTP	Model 6 WTP		Model 7 WTP	Model 8 WTP
	Interaction design Interaction IUF	Main effect design Interaction IUF		Interactiondesign Main effect IUF	Main effect design Main effect IUF
Drive Mushroom = 0, Picnic = 0	-5.28** (-2.181)	-2.28 (-0.781)	Drive	-15.84*** (-6.413)	-8.33*** (7.372)
Drive Mushroom = 1, Picnic = 0	-19.45*** (-5.200)	-25.71*** (-2.860)			
Drive Mushroom = 0, Picnic = 1	-12.52*** (-4.244)	-4.54 (-1.351)			
Drive Mushroom = 1, Picnic = 1	-26.69*** (-6.280)	-27.97*** (-3.101)			
Picnic Drive = 0	7.67*** (3.618)	2.11 (0.796)	Picnic	4.68*** (3.059)	1.02 (1.063)
Picnic Drive = 1	0.44 (0.201)	-0.15 (-0.057)			
Mushroom Drive = 0	13.55*** (5.272)	20.10*** (2.757)	Mushroom	7.05*** (4.325)	4.98*** (4.98)
Mushroom Drive = 1	-0.62 (-0.278)	-3.32 (-0.941)			
Erosion	2.22*** (3.675)	1.71** (2.352)	Erosion	2.35*** (3.784)	1.42*** (3.978)
CO_2	1.64*** (2.555)	2.68*** (3.264)	CO_2	1.53*** (2.860)	2.15*** (7.50)

Note: t-statistic is in brackets.

For example, the WTP for driving in the interaction IUF Models 1 and 5 varies between a relatively low negative value when Picnic and Mushroom are excluded (-4.89 and -5.28 EUR respectively), to a substantial negative value when they are included (-26.97 EUR and -26.69 EUR respectively). This suggests that the disutility related to driving is relatively low, given that driving is permitted in areas where recreation activities are not available (i.e., picnicking or mushrooming). In other words, driving on the one hand, and mushrooming or picnicking on the other, are on average perceived as substitutes, not as complements. A similar pattern is observed for mushrooming and picnicking. The WTP for mushrooming and picnicking are positive and relatively high when driving is forbidden. However, the WTP for both attributes becomes not statistically different from zero when driving is allowed.

The choice of design may therefore have relevant policy implications. For example, based on the main effects model, it appears that people are willing to pay a high amount to forbid car driving in the forest. The WTP for driving is close to -16 EUR in Models 3 (Table 4) and 7 (Table 5). However, the negative WTP associated with driving alone is relatively small, with the interactions specification equaling -4.9 EUR (Model 1) and -5.3 EUR (Model 5). There is therefore a relatively low average disutility associated with driving alone.

A larger disutility related to Drive comes from the two-way interactions with Picnic and Mushroom. Since the interdependencies between Mushroom, Picnic and Drive are not accounted for in the main effects IUF, the lower value is attributed to Drive, Picnic and Mushroom. In other words, whenever Drive is included along with Picnic or Mushroom in the same alternative, the disutility associated with Drive in this setting will be higher due to the negative impact of the interaction effect. This will in turn lead to negatively biased parameters for Mushroom, Picnic and Drive.

3.2. Results based on the main effect design. As in the previous section, two forms of IUF are assumed. Even though the design is tailored for the main effects specification, the estimation of two-way interactions is still possible. There are high correlations (> 0.7) between some interactions and the main effects; meaning that the identification of the parameters may be difficult in this case. The MNL results are reported in Table 4, while the RPL results are shown in Table 5. The standard errors for the main effects design and interaction IUF are larger than for the interaction design and interaction IUF. As a consequence, Picnic, Drive, Erosion, and Picnic×Drive, which were highly significant in Models 1 and 5, become insignificant in Models 2 and 6. There are also substantial differences in WTP values between the two designs for Picnic when driving is forbidden, and for Drive when picnicking is allowed and mushrooming forbidden. The mean WTP for these attributes based on the main effects design lies outside the 95% confidence intervals calculated for the interaction design and model.

As in Section 3.1, the inclusion of interaction effects significantly influences the main effects parameter estimates. The pattern is similar to the results obtained from the data based on the interaction design, and is therefore not discussed here. However, the statistical significance of the parameter estimates and the resulting WTP estimates is lower. This is in line with findings reported by Yu et al. (2006).

Conclusion

It is current practice in choice modeling to assume linear in parameters IUF in both the design and modeling stage, probably because interaction terms are believed to be small or even insignificant relative to main effects (Louviere et al., 2000). However, a case study has been presented here in which the interaction parameters are statistically significant and large in comparison with the main effects. The use of a partial log-likelihood analysis shows that the overall contribution of two-way interactions to the LL value is 21% in this study.

Forest policies based on the main effect results would be significantly different from those accounting for interactions. For instance, the disutility of allowing car driving in the new forests is more than three times higher if estimated from a main effects model instead of an interactions model, due to the fact that driving imposes a negative externality on other recreational activities, such as mushroom picking and picnicking.

If other case studies find similar results, the role of interactions may appear more prominent. Sometimes it is difficult to identify in focus groups or in a pilot study whether interactions are relevant. In these cases, incorporating interactions into both the design stage and the model estimation may be a safe strategy to adopt.

References

1. Alvarez-Farizo, B., Gil, J.M., Howard, B.J. (2009). Impacts from restoration strategies: assessment through valuation workshops, *Ecological Economics*, 68, pp. 787-797.
2. Bateman, I.J., Munro, A. (2009). Household versus individual valuation: What's the difference? *Environmental and Resource Economics*, 43, pp. 119-135.
3. Bliemer, M.C.J., Rose, J.M. (2005). Efficiency and sample size requirements for stated choice studies, Working Paper, Institute of Transport and Logistics Studies, Sydney.
4. Bliemer, M.C.J., Rose, J.M. (2006). Designing stated choice experiments: state-of-the art, 11th International Conference on Travel Behaviour Research, Kyoto.
5. Bliemer, M.C.J., Rose, J.M. (2010). Construction of experimental designs for mixed logit models allowing for correlation across choice observations, Transp. Res.: Part B: Methodological, 44, pp. 720-734.
6. Bliemer, M.C.J., Rose, J.M., Hensher, D.A. (2009). Efficient stated choice experiments for estimating nested logit models, Transp. Res.: Part B: Methodological 43, pp. 19-35.
7. Boyle, K., J., Ozdemir, S. (2009). Convergent validity of attribute based, choice questions in stated preference studies, *Environmental and Resource Economics*, 42, pp. 247-264.
8. Brownstone, D., Train, K. (1999). Forecasting new product penetration with flexible substitution patterns, *Journal of Econometrics*, 89, pp. 109-129.
9. Campbell, D., Hutchinson, W., Scarpa, R. (2008). Incorporating discontinuous preferences into the analysis of discrete choice experiments, *Environmental and Resource Economics*, 41, pp. 401-417.
10. Carlsson, F., Martinsson, P. (2003). Design techniques for stated preference methods in health economics, *Health Economics*, 12, pp. 281-294.
11. Carson, R.T., Louviere, J.J., Wasi, N. (2009). A cautionary note on designing discrete choice experiments: a comment on Lusk and Norwood's 'effect of experiment design on choice-based conjoint valuation estimates', *American Journal of Agricultural Economics*, 91, pp. 1056-1063.
12. Chattopadhyay (2009). The random expenditure function approach to welfare in RUM: the case of hazardous waste clean-up, *Resource Energy Economics*, 31, pp. 58-74.
13. Colombo, S., Angus, A., Morris, J., Parsons, D.J., Brawn, M., Stacey, K., Hanley, N. (2009). A comparison of citizen and 'expert' preferences using an attribute-based approach to choice, *Ecological Economic*s, 68, pp. 2834-2841.
14. Dawes, R.M., Corrigan, B. (1974). Linear models in decision making, *Psychological Bulletin*, 81, pp. 95-106.
15. Ferrini, S., Scarpa, R. (2007). Designs with a priori information for nonmarket valuation with choice experiments: a Monte Carlo study, *Journal of Environmental Economics and Management*, 53, pp. 342-363.
16. Haaijer, R., Kamakura, W., Wedel, M. (2001). The 'no choice' alternative in conjoint choice experiments, *International Journal of Market Research*, 43, pp. 93-106.
17. Hensher, D.A., Rose, J.M., Greene, W.H. (2005). Applied Choice Analysis: A Primer, Cambridge University Press, New York.
18. Hoyos, D., Mariel, P., Fernández-Macho, J. (2009). The influence of cultural identity on the WTP to protect natural resources: some empirical evidence, *Ecological Economics*, 68, pp. 2372-2381.
19. Huber, J., Swerina, K. (1996). The importance of utility balance in efficient choice designs, *Journal of Marketing Research*, 33, pp. 307-317.

20. Kanninen, B.J. (2002). Optimal design for multinomial choice experiments, *Journal of Marketing Research*, 39, pp. 214-227.
21. Kontoleon, A., Yabe, M. (2003). Assessing the impacts of alternative 'opt-out' formats in choice experiment studies: consumer preferences for genetically modified content and production information in food, *Journal of Agricultural Policy*, 5, pp. 1-43.
22. Lancsar, E., Louviere, J., Flynn, T. (2007). Several methods to investigate relative attribute impact in stated preference experiments, *Social Science and Medicine*, 64, pp. 1738-1753.
23. Lehtonen, R., Sarndal, C.E., Veijanen, A. (2003). The effect of model choice in estimation for domains, including small domains, *Surveys in Mathematics*, 29, pp. 33-44.
24. Louviere, J.J. (1988). Analysing individual decision making: metric conjoint analysis, Sage, Newbury Park.
25. Louviere, J.J., Hensher, D.A., Swait, J.D. (2000). Stated choice methods: analysis and applications, Cambridge University Press, Cambridge.
26. Lusk, J.L., Norwood, B. (2005). Effect of experimental design on choice-based contingent valuation estimates, *American Journal of Agricultural Economics*, 87, pp. 771-785.
27. McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior, in: P. Zarembka (Eds.), Frontiers in Econometrics, Academic Press, New York, pp. 105-142.
28. Ministerio de Medio Ambiente (1996). Segundo Inventario Nacional, 1986-1996, Dirección General de Conservación de la Naturaleza, Ministerio de Medio Ambiente, Madrid.
29. Mogas, J., Riera, P., Bennett, J. (2006). A comparison of contingent valuation and choice modeling with second-order interactions, *Journal of Forest Economics*, 12, pp. 5-30.
30. Rose, J.M., Bliemer, M.C.J. (2008). Stated preference experimental design strategies, in: D. A. Hensher, and K. J. Button (Eds.), Handbook of Transport Modelling. Elsevier, Oxford, pp. 151-180.
31. Samuelson, W., Zeckhauser, R. (1988). Status quo bias in decision making, *Journal of Risk Uncertainty*, 1, pp. 7-59.
32. Sandor, Z., Wedel, M. (2002). Profile construction in experimental choice designs for mixed logit models, *Market Science*, 21, pp. 430-444.
33. Sandor, Z., Wedel, M. (2005). Differentiated bayesian conjoint choice designs, *Journal of Marketing Research*, 55, pp. 210-218.
34. Scarpa, R., Philippidis, G., Spalatro, F. (2005). Product-country images and preference heterogeneity for Mediterranean food products: A discrete choice framework, *Agribusiness* 21, pp. 329-349.
35. Scarpa, R., Rose, J.M. (2008). Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why, *Australian Journal of Agricultural and Resource Economics*, 52, pp. 253-282.
36. Train, K. (1999). Halton sequences for mixed logit, Working Paper, University of California, Berkeley.
37. Train, K. (2003). Discrete choice methods with simulation, Cambridge University Press, Cambridge.
38. Yu, J., Goos, P., Vanderbroek, M. (2006). The importance of attribute interactions in conjoint choice design and modeling, Working Paper, Catholic University of Leuven, Leuven.