

Lundhede, Thomas Hedemark; Olsen, Søren Bøye; Jacobsen, Jette Bredahl;
Thorsen, Bo Jellesmark

Article

Handling respondent uncertainty in choice experiments: Evaluating recoding approaches against explicit modelling of uncertainty

Journal of Choice Modelling

Provided in Cooperation with:

Journal of Choice Modelling

Suggested Citation: Lundhede, Thomas Hedemark; Olsen, Søren Bøye; Jacobsen, Jette Bredahl; Thorsen, Bo Jellesmark (2009) : Handling respondent uncertainty in choice experiments: Evaluating recoding approaches against explicit modelling of uncertainty, Journal of Choice Modelling, ISSN 1755-5345, University of Leeds, Institute for Transport Studies, Leeds, Vol. 2, Iss. 2, pp. 118-147

This Version is available at:

<https://hdl.handle.net/10419/66802>

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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



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

Handling respondent uncertainty in Choice Experiments: Evaluating recoding approaches against explicit modelling of uncertainty

Thomas Hedemark Lundhede^{1,*} Søren Bøye Olsen^{2,†}
Jette Bredahl Jacobsen^{1,‡} Bo Jellesmark Thorsen^{1,§}

¹ Division of Economics, Policy and Management Planning, Forest & Landscape, University of Copenhagen, Rolighedsvej 23, DK-1958 Frb.C., Denmark

² Environmental Economics and Rural Development Division, Institute of Food and Resource Economics, University of Copenhagen, Rolighedsvej 25, DK-1958 Frb.C., Denmark

Received 23 April 2009, revised version received 30 June 2009, accepted 16 September 2009

Abstract

Using data from two Choice Experiment environmental valuation surveys we investigate several different ways of handling respondent uncertainty. In both surveys respondents are asked to state their certainty of choice after each single choice set. We evaluate three different recoding-of-answers methods adapted from the Contingent Valuation literature. Furthermore, we evaluate two models which directly capture the effect of respondent uncertainty by parameterization of a scale function. In one model the scale parameter is a function of respondents' stated uncertainty level. In the other it is a function of respondent and choice set characteristics found to be significant determinants of stated uncertainty. All approaches are compared to a benchmark utility model.

While some of the recoding approaches apparently reduce noise in the data, they generally have no significant effect on attribute Willingness-To-Pay (WTP) estimates, and standard errors on WTP estimates tend to increase. The explicit modelling of the scale parameter using stated uncertainty reveals that unobserved variation decreases as certainty of choice increases. While the model performance does not improve much, this approach offers a structurally and intuitively much more appealing way of accounting for uncertainty in choices in CE.

Keywords: Environmental valuation, stated certainty in choice, recoding, scaling approach, variance heterogeneity, random parameter error component logit.

* T: +45-35331535, F: +45-35331508, thlu@life.ku.dk

+ Corresponding author, T: +45-35333643, F: +45-35336801, sobo@life.ku.dk

‡ T: +45-35331746, F: +45-35331508, jbj@life.ku.dk

§ T: +45-35331700, F: +45-35331508, bjt@life.ku.dk



Licensed under a Creative Commons Attribution-Non-Commercial 2.0 UK: England & Wales License
<http://creativecommons.org/licenses/by-nc/2.0/uk/>

1 Introduction

It is usually a key assumption in stated preference methods that respondents' are able to assess without any error the utility they may derive from the good presented to them, and hence can answer any valuation question with absolute certainty (Hanemann 1984). There are numerous arguments why this assumption may not be valid, and recognising this several Contingent Valuation (CV) studies have investigated the causes of respondent uncertainty, as well as ways to handle it (e.g. Loomis and Ekstrand 1998; Alberini *et al.* 2003). However, so far this issue has been given little attention in Choice Experiments (CE) studies, which have in recent years been increasingly used for environmental valuation. While it may be perfectly reasonable for respondents to feel uncertain about their stated responses and choices (Wang 1997; Li and Mattson 1995), the problem is that failing to accommodate for such uncertainty in the modelling of data may bias, if not the valuation estimates, then at least the variance estimates and hence the inference and conclusions made (Li and Mattson 1995).

In this paper, we present an analysis of different ways to handle uncertain responses in CE surveys. We use data from two environmental valuation CE surveys in which respondents stated their certainty of choice after each choice set. For both data sets, the effects on model performance and Willingness-To-Pay (WTP) of each way of handling uncertain choices are compared with a benchmark model in which respondent uncertainty is ignored. Inspired by the way stated uncertainty has been handled in CV studies, we estimate three models where we recode the data sets according to different ways of interpreting uncertain answers, taking the hypotheses summarised by Samnaliev *et al.* (2006) as points of departure. We also estimate two models in which respondent uncertainty in choice is handled directly through the explicit modelling of the scale parameter: an approach parallel to that applied in a CV study by Alberini *et al.* (2003). In the first of these we account for variation in scale as a function of respondents' stated level of uncertainty in each choice set. In the second, we model the scale parameter as a function of specific variables found to affect respondents' stated uncertainty (Lundhede *et al.* 2009a). While some of the recoding methods reduce noise in the data, the effect on the WTP estimates for the environmental attributes is generally insignificant and not unidirectional. The explicit modelling of the scale parameter using stated uncertainty reveals that a higher degree of certainty of choice is equivalent to a lower degree of unobserved variation. Thus, this approach holds the promise of obtaining a higher degree of precision in the estimation of the environmental attribute parameters and their estimation errors.

The remainder of the paper is organised as follows. In Section 2 we provide an overview of the handling of uncertainty in the CV literature. We also describe the applied econometric models and the five different ways of handling uncertain answers. In Section 3 we present the surveys providing the data. Section 4 contains the results and in Section 5 we discuss the results in terms of effects on WTP estimates and model performance. We summarise with a few concluding remarks in Section 6.

2 Theory and Methods

2.1 Handling uncertain answers in Contingent Valuation surveys

Several studies have tried to obtain an expression for the degree of uncertainty felt by the respondent when answering CV questions. The approaches taken can roughly be classified into two groups. The first approach is to make respondents choose among answers to the payment question, which explicitly incorporates some level of uncertainty, e.g. ‘Don’t Know’ (Wang 1997), ‘I will definitely pay’ or ‘I most probably will not pay’ (Ready *et al.* 1995; Welsh and Poe 1998; Alberini *et al.* 2003). The second approach is to have respondents first answer the payment question (‘Yes’/‘No’) and then state their degree of certainty regarding the answer just provided, either in the form of a numeric scale or as text statements (Li and Mattsson 1995; Champ *et al.* 1997; Loomis and Ekstrand 1998; Johannesson *et al.* 1999; Blumenschein *et al.* 2008). There are pros and cons of both approaches. For example, the latter has the advantage that it does not directly interfere with the valuation task, yet it hinges to some degree on researcher interpretation concerning the stated certainty in order to handle it in estimations. We will take this approach, asking respondents to indicate their certainty *post* choice. This stated certainty may of course differ from true certainty, which limits how far one can make conclusions on true certainty from the observed stated certainty. However, our focus here is the evaluation of different approaches to using respondents’ stated certainty in choice to improve estimations, which is a relevant objective even if does not coincide perfectly with the unobserved true certainty in choice. How the researcher handles respondents’ stated certainty will depend on what is assumed to be the reasons for the stated certainty. Samnaliev *et al.* (2006) summarise four such assumptions or hypotheses, which we briefly present and discuss here in a slightly different order and wording. One hypothesis (adapted from Schwarz and Sudman 1996) is that certainty levels indicated by respondents will reflect only their attempt to appear consistent in answers: Once they have chosen ‘Yes’ or ‘No’, they indicate some degree of certainty to signal consistency. The main objection to this hypothesis is that if such behaviour dominates, we should find a fairly constant level of certainty across alternatives – but in fact stated certainty varies systematically with, e.g. the bid (Wang 1997; Loomis and Ekstrand 1998).

A second and related hypothesis is that certainty levels may be susceptible to protesting and strategic behaviour such as respondents exaggerating certainty along with stated WTP (Samnaliev *et al.* 2006). This sort of behaviour is usually screened for in quality studies and samples. While it may be a source of noise, it should not be dominant. A third hypothesis concerning preference uncertainty is that when respondents are allowed to state uncertainty, they use this option to scale down their stated WTP, i.e. an asymmetric effect on WTP reducing hypothetical bias is assumed (Champ *et al.* 1997). A fourth hypothesis is formulated by Wang (1997) and implied in Li and Mattsson (1995). This hypothesis maintains the assumption that respondents are rational, truth-telling and non-strategic, but may for different reasons assess the value of the environmental change with some degree of uncertainty and, therefore, they may be quite uncertain as to their answer (‘Yes’/‘No’) when the bid price is close to their maximum WTP – but quite certain when very different from it. The implication is that the response itself may be subject to error which in a CE setting

translates into a probability that the respondent does not choose the utility maximising alternative.

If the third assumption is preferred, it is relevant to apply some sort of asymmetric method to correct for the signal sent by respondents stating uncertainty about their 'Yes'. A direct recoding of these answers from 'Yes' to 'No' has been applied by several authors (e.g. Champ *et al.* 1997; Welsh and Poe 1998), whereas Loomis and Ekstrand (1998) propose an asymmetric uncertainty model incorporating stated uncertainty levels of 'Yes' answers into the likelihood function. Not surprisingly, the first approach implies an often dramatic downward adjustment of WTP estimates, whereas the effect on WTP is somewhat less with the second approach. If the fourth assumption is preferred, it follows that respondents can also be uncertain about voting 'No'. This calls for a symmetric approach and several studies have suggested ways to incorporate uncertainty for all responses directly in the likelihood function, be the uncertainty level implied by the chosen answer (as in Wang 1997 and in Alberini *et al.* 2003), or stated on some sort of scale post decision (Li and Mattson 1995; Loomis and Ekstrand 1998). As noted by both Loomis and Ekstrand (1998) and Alberini *et al.* (2003) the symmetric approach tends to increase the estimated WTP, even if it also provides a better performing model.

All of the above literature exclusively treats the issue of uncertainty in CV studies. The transfer of these approaches to the data obtained in CE studies is slightly complicated by the fact that respondents usually evaluate more than one alternative version of the environmental change, potentially against a status quo alternative as in the present surveys. Thus, if a respondent states uncertainty about a choice, there will be as many possible alternative answers as there are alternatives left in the choice set. Next, we elaborate on the way in which we treat the stated certainty levels in the different approaches.

2.2 Handling Uncertain Answers in Choice Experiment surveys

In the following, we maintain the common assumption in CE that respondents attend to and evaluate all attributes presented in the choice set, and that there is full substitutability between attributes within the ranges presented to respondents. If these assumptions do not hold, it may lead to discontinuous preferences and lexicographic ordering. Recent studies have investigated biases in respondents' processing of the different attributes in CE studies and their use of heuristics in response, as an underlying cause of uncertainty and heterogeneity (Campbell 2008; Hensher 2008), and found that such biases can be severe. These findings are important and may have bearings also on the present study, if respondents are in fact not processing all attributes as assumed. This limitation should be kept in mind as we proceed to address our objective here: to evaluate the usefulness of stated certainty at the choice set level in censoring of responses or as a basis for explicitly modelling heterogeneity in scale.

2.2.1 Three Recoding Approaches

We evaluate the effect of three different ways of recoding respondent choices: Uncertain choices are either i) eliminated from the sample, ii) asymmetrically recoded in the sense that a choice reported to be uncertain is recoded as a choice of the status-quo alternative, or iii) symmetrically recoded such that an uncertain choice is recoded as a choice of the best alternative different from the one chosen. 'The best alternative'

is assessed in terms of the utility to each respondent of each alternative, using the parameters from the benchmark model in which stated uncertainty is ignored. This means that if respondents according to the expected utility have chosen the best alternative among the possible alternatives in the choice set and have reported the choice as uncertain, we recode their answer into a choice of the second best alternative. Correspondingly, if they have chosen the second or third best alternative and have reported uncertainty about their choice, we recode the choice as a choice of the best alternative in the choice set. In other words, the symmetric recoding should reflect uncertain respondents' most likely choice if they had chosen differently.

As pointed out in Section 2.1, the recoding approach hinges on the researcher deciding on an interpretation of what is an uncertain answer. Based on the construction of the scales in which the respondents have reported certainty, see Appendices A and B, we assume that a certain choice is one where the stated certainty level is either 'Certain' or 'Very certain'. All other response categories are interpreted as uncertain¹.

2.2.2 Uncertainty and Variation in the Scale Parameter

An approach used in CE to take into account differences in unexplained variation between groups is scaling, which makes use of the fact that embedded in all random utility choice models are scale parameters. Since utility in itself is an ordinal measure, it has no absolute scale. The logit scaling approach introduced by Bradley and Daly (1994) allows for differences in the amount of unexplained variance across different types of data. If one part of a data set has more unexplained variance than the other, and this is not recognised in the model, it may lead to biased model parameters and the model may wrongly predict changes in choice probabilities (Bradley and Daly 1994). The importance of handling variations in scale across data types has been stressed repeatedly in studies investigating the merits of joint analysis of Stated Preference and Revealed Preference data (Bradley and Daly 1994; Hensher and Bradley 1993; Adamowicz *et al.* 1994; Adamowicz *et al.* 1997; Hensher *et al.* 1999; Brownstone *et al.* 2000; Whitehead *et al.* 2008). In such cases it is crucial that the variations in the scale parameter across data set are adequately taken into account, if the underlying taste parameters are to be reliably estimated. More generally, variation in scale is likely to be an integral part of the behavioural and decision-making processes reflected in the response patterns of stated preference studies. To uncover these, several studies have focused on the scale parameters dependence on variables like choice complexity, effort and fatigue (Bradley and Daly 1994; Swait and Adamowicz 2001; Hensher *et al.* 2001; DeShazo and Fermo 2002; Sælensminde 2001), as well as demographic variables of the respondents (Louviere and Hensher 2001; Scarpa *et al.* 2003; Hu *et al.* 2006). In this paper, we use respondents' explicitly stated certainty in choice in two models: One in which variation in scale is linked directly to the stated certainty, and one in which scale is modelled as a function of other variables found to correlate with stated certainty.

¹ The recoding approach may hinge on the distinction that the researcher makes between a certain and an uncertain answer. In the motorway data the level 'Neither certain nor uncertain' has been categorised as an uncertain answer. For the purpose of testing the influence of the used distinction, we also estimated the models where 'Neither certain nor uncertain' was categorised as a certain answer. Results showed that this specific distinction had no noteworthy effect on the conclusions drawn

2.3 The Econometric Models

2.3.1 The Random Parameter Error Component Logit Model

The random parameter error component logit model relies on McFadden's (1974) random utility model, where the utility of a good is described as a function of its attributes, and people choose among complex goods by evaluating their attributes. Since utility can only be imperfectly observed, the random utility model is the basis for estimation. In a specific case, where a respondent, i , faces a choice between a status quo and two management alternatives, the utility, U , of these j alternatives in the n 'th choice occasion can be described in the following way:

$$U_{ijn} = \begin{cases} V(ASC, x_{ijn}, \tilde{\beta}_i, \beta) + \varepsilon_{ijn} & \text{if } j = 1 \text{ (statusquo);} \\ V(x_{ijn}, \tilde{\beta}_i, \beta, \sigma_i) + \varepsilon_{ijn} & \text{if } j = 2,3; \end{cases} \quad (1)$$

Here the indirect utility, V , is a function of the vector of explanatory variables, x_{ijn} , including characteristics of the individual, the alternative and the choice situation, as well as the vectors of individual-specific random parameters, $\tilde{\beta}_i$, and fixed parameters, β . Following Scarpa *et al.* (2005) an Alternative Specific Constant (ASC) is specified for the status quo alternative in order to capture the systematic component of a potential status quo effect. Furthermore, an error component additional to the usual Gumbel-distributed error term is incorporated in the model to capture any remaining status quo effects in the stochastic part of utility. This error component, σ_i , which is implemented as an individual-specific zero-mean normally distributed random parameter, is assigned exclusively to the two non-status quo alternatives. By specifying a common error component across these two alternatives, a correlation pattern in utility over these alternatives is induced. Thus, it captures any additional unexplained variance associated with the cognitive effort of evaluating two experimentally designed hypothetical scenarios relative to a status quo scenario (Greene and Hensher 2007; Scarpa *et al.* 2007; Ferrini and Scarpa 2007; Scarpa *et al.* 2008).

Assuming that ε_{ijn} is IID extreme value distributed, the probability of individual i choosing alternative k out of j alternatives can be defined by the Conditional Logit model:

$$\Pr(ikn) = \frac{\exp^{\lambda_{ikn}\beta'x_{ikn}}}{\sum_j^J \exp^{\lambda_{ikn}\beta'x_{ijn}}} \quad (2)$$

where β' is a vector of all betas, λ is the scale parameter which is typically normalised to 1, and the ASC and error terms from eq. (1) are left out for simplicity. Following Train (2003), the Mixed Logit probabilities can be described as integrals of the standard conditional logit function evaluated at different β 's with a density function as the mixing distribution. Furthermore, this specification can be generalised to allow for

repeated choices by the same respondent, i.e. a panel structure, by letting k be a sequence of alternatives, one for each choice occasion, $k = \{k1, \dots, kN\}$. Thus, the utility coefficients vary over people but are constant over the N choice occasions for each individual. If the density, as in this paper, is specified to be normal, the probabilities of the model become:

$$\Pr(ik) = \int \left(\prod_{n=1}^N \left[\frac{\exp^{\beta_i' x_{ikn}}}{\sum_j \exp^{\beta_i' x_{ijn}}} \right] \right) \phi(\beta | b, W) d\beta \tag{3}$$

where $\phi(\beta | b, W)$ is the distribution function for β , with mean b and covariance W . The analyst chooses the appropriate distribution for each parameter in β . For simplicity, λ is normalised to unity.

2.3.1 The Scaling Approach

In our first four models estimated, i.e. the benchmark model ignoring uncertainty and the three models based on recoded samples, the econometric model just described was used. In the fifth and sixth models, we expand the model with the logit scaling approach (see Louviere *et al.* 2000). Since utility is an ordinal measure, the scale of utility has to be normalised and usually this is done by normalising the variance of the error term (Ben-Akiva and Lerman 1985; Train 2003). Assumptions concerning the distribution of the unobserved part of utility are required in any random utility model for instance by assuming that the error terms are IID Gumbel distributed as above. This implies that the scale of utility is normalised. It can be shown that the Gumbel scale parameter, λ , is inversely proportional to the standard deviation of the random component up to a constant of $\pi^2/6 \approx 1.3$ (Ben-Akiva and Lerman 1985).

In the benchmark and recoding models of this paper, we use the common normalisation of λ to unity (Train 2003; Scarpa *et al.* 2003), implying that unexplained variance is assumed uniform across responses. In Models 5 and 6 we relax this assumption. A general parameterisation of the scale function is (Hensher *et al.* 2005):

$$\lambda_{ikn} = \exp(\gamma_w Z_{iknw}) = \frac{\pi}{\sqrt{\text{var}(\mu_{ikn})} \sqrt{6}} \tag{4}$$

where Z_{iknw} is a vector of covariates associated with the individual, the choice set and the alternative (the elements in Z are indexed by w), and γ_w is a row-vector of the corresponding scale function parameters. The exponential form of the scale function ensures nonnegative estimates of model variance, as the scale is inversely proportional to the standard deviation of the unobserved component, μ_{ikn} .

In the context of preference uncertainty, it makes sense to assume that uncertain choices will exhibit a greater degree of unobserved variability in choices than certain choices, and thus a lower degree of estimation precision (Hole 2006). In order to incorporate this, in Model 5 we parameterise scale as a function of respondents' stated certainty level for each choice set. Hence, Z consists of the different levels of stated certainty in choice, with one parameter fixed to unity for identification purposes. In this way, stated certainty is explicitly accounted for in the parametric model as a source of unobserved variability in choices. This specification allows us to use a panel

specification, as we can index each respondent and choice combination, capturing the way respondents' switch between uncertainty-groups through the course of the choice sets.

We recognise that most CE surveys do not or may not want to ask respondents to state their experienced uncertainty after each single choice set. Therefore, in a sixth model, we model scale as a function of variables found to determine the respondents' stated certainty levels. This is based on the findings from Lundhede *et al.* (2009a) where respondents' stated certainty in choice is modelled using an ordered probit model. Among other things, they find evidence that stated certainty is higher for men than for women, that certainty increases with income, and that a learning effect may increase stated certainty as the choice set number increases.

3 Data

The data originate from two CE surveys: one that examined preferences for reducing the impact of new motorways on different types of nature and one that examined the preferences for the establishment of national parks.

3.1 The Motorway Survey

The hypothetical scenario was that 100 kilometres of new motorways were to be built in Denmark during the next ten years. The scenario described that the exact location of these stretches of motorway can be decided on with more or less consideration of potential encroachment on nature areas.

Three different types of nature were identified and chosen as attributes in the CE design. The three attributes were 'forest', 'wetland', and 'heath'. To enable estimation of WTP, a price attribute was defined in terms of an extra annual income tax on the household. In Denmark, the building of motorways is financed over taxes, lending credibility to this payment vehicle. The attributes and their assigned levels are summed up in Table 1.

Table 1: Attribute Levels used in the Motorway Survey

Attribute	Levels
Forest	0 km, 5 km, 10 km
Wetland	0 km, 2.5 km, 5 km
Heath/pastoral land	0 km, 2.5 km, 5 km
Arable land ^a	80 km, 82.5 km, 85 km, 87.5 km, 90 km, 92.5 km, 95 km, 97.5 km, 100 km
Annual extra tax payment per household ^b	DKK 0, 100, 200, 400, 700, 1100, 1600

^a As the total stretch of motorway was fixed at 100 kilometres, a fourth supplementary attribute, 'arable land', was introduced to account for the location of the remaining kilometres summing up to 100 in total. This attribute functioned as an accumulation attribute, its level being determined by the other attribute levels. Thus, due to perfect correlation, it was not included in the experimental choice set design and it is not included in the parametric modelling of preferences.

^b Note: DKK 100 \approx € 13.4

The attribute levels were assigned to alternatives using an experimental design and paired into choice sets of three alternatives. As a full factorial design comprised 162 alternatives, a D-optimal fractional factorial design consisting of 18 choice sets was identified (Louviere *et al.* 2000). To minimise the number of dominating and non-causal alternatives, the initially identified efficient design was subjected to the manual swapping procedure suggested by Huber and Zwerina (1996). A choice set consisted of three alternatives: the zero-priced status quo alternative (the motorway would be placed through 10 kilometres of forest, 5 kilometres of wetland, 5 kilometres of heath, and 80 kilometres of arable areas) and two experimentally designed improvement alternatives. The respondent sample was split into three groups, so that each respondent only had to answer six choice sets, and respondents were instructed to choose which alternative they would prefer in each of the choice sets. An example of a choice set is displayed in appendix A along with the associated question on certainty in choice. The dataset which was collected using an internet-based questionnaire consists of 595 responses, resulting in a total of 3,570 answered choice sets². In only 37 out of these did respondents answer 'Don't know' or fail to respond to the certainty follow-up question. The distribution of responses is displayed in Table 3. The majority of the choices (66.0 percent) were classified as either 'certain' or 'very certain' choices.

3.2 The National Park Survey

Denmark is establishing its first national parks following a long political process and public debate. As part of the participatory process included, a study was performed to evaluate preferences for different environmental attributes of national parks as well as seven potential sites. Respondents were asked to evaluate choice sets in which the 'Location' of the new national park was one attribute along with four generic attributes of the parks, namely 'Extra initiatives for special plant and animals', 'Extra effort for general nature protection', 'Increased amount of walking and biking paths'. The establishment, nature protection efforts and management of the national parks will be paid for over the general taxes in Denmark, and thus 'Extra income tax per year and household' was the fourth generic attribute. The attributes and levels are shown in Table 2.

Each respondent was given only four of the seven locations, allocated by a cyclic design of seven groups. For each group the same combinations of attribute levels were used. The attribute levels were assigned to alternatives by a fractional factorial design and resulted in an orthogonal, balanced experimental design of 32 choice sets consisting of two alternatives and a status quo (no national park). The choice sets were grouped into 4 blocks of 8. No choice sets were eliminated from the design, i.e. also alternatives with zero payment for a national park occurred. An example of a choice set is shown in Appendix B. Each respondent replied to 8 choice sets. The dataset which was collected using self-administered paper-and-pencil questionnaires consists of 636 responses resulting in 4,866 choice observations³. Compared with the target population (Danes), the sample is a little under-represented in age-groups below 35 and above 65 years; and also in short educations. Also both the lowest and the highest

² For a more thorough description of the survey and a full version of the questionnaire used, the reader is referred to Olsen *et al.* (2005).

³ For a full description of the data and analysis, see Jacobsen *et al.* (2006) and Jacobsen and Thorsen (2009).

Table 2: Attributes Levels used in the National Park Survey

Attribute	Levels
Location	None, Læsø, Møn, Thy, Nordsjælland, Mols Bjerger, Lille Vildmose, Vadehavet
Nature preservation	No extra effort, Little extra effort, Some extra effort, Large extra effort
Extra effort for specific animals and plants	No, Yes (with indication of which species for the given location)
Walking and biking path	No increased amount of path, Increased amount of path
Extra income tax per year per household	DKK 0, 50, 100, 200, 400, 700, 1500, 2000

income groups are under-represented. There is no significant difference on gender. The distribution of responses to the certainty groups is displayed in Table 3. With a total of 73.6 percent of the choices being classified as either ‘certain’ or ‘very certain’, the respondents in this survey seem to express a generally higher degree of certainty about their choices than the respondents in the survey considering motorways⁴.

4 Results and Analyses

4.1 The Benchmark Model

Assumptions concerning the distribution of random parameters, i.e. the density function specified in eq. (3), are necessary. The true distribution is unknown, so in principle any distribution could be applied (Carlsson *et al.* 2003; Hensher and Greene

Table 3: Distributions of Stated Certainty

Certainty group	Percentage of responses	
	Motorway	National park
Very uncertain	3.1	2.7
Uncertain	9.8	21.0
Neither certain nor uncertain	20.0	-
Certain	45.2	49.5
Very certain	20.8	24.1
Don't know	1.0	2.7

⁴ Considering that Olsen (2009) find evidence of internet respondents generally stating higher certainty in choice than mail respondents, it is slightly surprising to find the opposite here. However, the two surveys differ in a number of other aspects than survey mode (i.e. the good being evaluated, the number of choices per respondent, the presentation of the certainty question) which might explain this result. We refrain from elaborating further on this topic here, as it is beyond the scope of this paper.

2003). In the present analyses, all parameters except the ones for tax payment and the alternative specific constant (ASC) are assumed to be normally distributed random parameters to allow for preference heterogeneity. This choice is based on experience from other Danish valuation studies, where the preferences for several similar attributes have been found to vary considerably across different groups of respondents. Extended access rights and facilities can have a negative value for some respondent groups, see e.g. Jacobsen *et al.* (2008) and Lundhede *et al.* (2009b), probably due to concerns for protection and conservation in nature areas. Likewise, Jacobsen and Thorsen (2009) find that preferences for national park sites differ greatly across respondents – some sites may have a negative value for respondents in some regions. Also, the value of increased wildlife and protection of different habitats is not necessarily (equally) positive for all (Jacobsen *et al.* 2009). As the explicit modelling of these patterns in preference variation, e.g. through the use of latent-class models, is not the focus of the paper, we have chosen the normal distribution as a reasonable way of taking into account possible taste variation. The ASC and tax parameters are treated as fixed rather than random parameters, even though keeping tax fixed implies that the marginal utility of money is fixed over the population. This is done to avoid a number of potentially severe problems associated with specifying a random price parameter (see e.g. Meijer and Rouwendal 2006; Hensher *et al.* 2005; Hess *et al.* 2005; Train and Sonnier 2005; Hensher and Greene 2003; Train 2003; Train and Weeks 2005). Tables 4 and 5 display the results obtained from the random parameter error component estimation of our benchmark model, where no measures have been taken to account for certainty in choice. All models are based on simulations using 1000 Halton draws. In Tables 10 and 11 the resulting WTP estimates are shown alongside the estimates from the other models.

For the motorway data it is evident from the t-values that all parameter estimates are significant with ASC being the only exception. As expected, the nature attribute mean estimates all have a negative sign, indicating that on average respondents experience a diminishing utility when one km of motorway is placed through the specific types of nature. Likewise, the price parameter estimate has a negative sign as would be expected. The impact of heath on utility is significantly lower than that of forest and wetland areas. The estimated standard deviations of the random parameters are highly significant, revealing a considerable degree of heterogeneity in the respondents' preferences for the three nature type attributes. The calculated probabilities of obtaining reversed sign suggest that this heterogeneity translates into between 19 and 31 percent of respondents generally being indifferent or even in favour of placing motorways through these types of landscape⁵. The adjusted pseudo- R^2 value as well as the likelihood ratio test reveal that the model generally fits the data very well (Domencich and McFadden 1975; Louviere *et al.* 2000).

⁵ This finding confirms results from focus group interviews where a minor share of the participants exhibited strong preferences for driving through scenic landscapes.

Table 4: The Benchmark Model for the Motorway Data (Model 1).

Attribute		Estimate	(<i>t</i> -value)	Prob. rev. sign
ASC		-0.3356	(-1.60)	
Forest ^a	- Mean	-0.1641	(-11.31)	0.23
	- St.dev.	0.2200	(13.12)	
Wetland ^a	- Mean	-0.2161	(-9.50)	0.19
	- St.dev.	0.2473	(8.14)	
Heath ^a	- Mean	-0.0923	(-4.88)	0.31
	- St.dev.	0.1824	(5.96)	
Price		-0.0020	(-27.17)	
Sigma		3.1437	(14.73)	
Num. observations		3570		
Num. respondents		595		
LL _{zero}		-3922		
LL _{Full}		-2756		
LRI (pseudo- R^2)		0.297		
Adj. LRI		0.296		
LR test		2331.9		

^a Due to the chosen coding of the landscape type variables, the utility estimates are associated with an increase in the number of kilometres of motorway going through the landscape types. Hence, the negative estimates do not imply that people generally dislike these landscape types, but rather the opposite. Accordingly, the calculated probabilities of reversed sign denote the probability of obtaining negative preferences for the landscape types. In the calculation of WTP estimates presented in Table 10, this coding approach is taken into account by multiplying by -1.

For the national park data in Table 5 we see that most parameter estimates are significant and with a positive sign. The exceptions are the parameter estimates for location Nordsjælland and for 'Walking and biking paths' which are not significantly different from zero. However, this is a result of highly heterogeneous preferences that more or less split the sample in two equally sized portions with opposing preferences. These results are in line with findings in Jacobsen *et al.* (2008) and Lundhede *et al.* (2009b). The numerically high negative ASC corresponds to a high WTP for the establishment of a park *per se*, cf. discussion in Jacobsen and Thorsen (2009). As for the motorway data, the estimated standard deviations of the random parameters are highly significant. Thus, a considerable degree of heterogeneity is established and this is underlined by the generally high probabilities of sign reversal. The adjusted pseudo- R^2 value as well as the likelihood ratio test indicates that the model fits the data very well.

We use the estimated utility weights implied by the parameters shown in Tables 4 and 5 to calculate the expected utility of each alternative in each choice set for each respondent, cf. eq.(1) and the attribute definitions in Tables 1 and 2. This is used for the symmetric recoding approach applied in Model 4 as explained in Section 2.2.1.

Table 5: The Benchmark Model for the National Park Data (Model 1).

Attribute		Estimate	(<i>t</i> -value)	Prob. rev. sign
ASC		-2.1833	(-10.99)	
Location Møn	- Mean	0.3555	(2.49)	0.40
	- St.dev.	1.3458	(7.20)	
Location Thy	- Mean	0.2978	(2.02)	0.40
	- St.dev.	1.1871	(6.91)	
Location Nordsjælland	- Mean	-0.1146	(-0.66)	0.52
	- St.dev.	1.8776	(10.16)	
Location Mols Bjerge	- Mean	0.4415	(2.76)	0.39
	- St.dev.	1.5742	(8.26)	
Location Lille Vildmose	- Mean	0.8141	(5.92)	0.28
	- St.dev.	1.3630	(8.17)	
Location Vadehavet	- Mean	0.6014	(3.95)	0.35
	- St.dev.	1.6092	(10.09)	
Nature preservation	- Mean	0.1216	(3.77)	0.33
	- St.dev.	0.2835	(4.89)	
Effort for animal/plants	- Mean	1.1291	(14.02)	0.09
	- St.dev.	0.8242	(7.46)	
Walking and biking paths	- Mean	0.1615	(1.85)	0.42
	- St.dev.	0.8086	(5.85)	
Price		-0.0020	(-33.49)	
Sigma		2.7036	(14.68)	
Num. observations		4866		
Num. respondents		636		
LL _{zero}		-5346		
LL _{Full}		-3742		
LRI (pseudo- R^2)		0.300		
Adj. LRI		0.298		
LR test		3205.8		

4.2 Estimating Models on Recoded Data

The results of re-estimating the benchmark models using the three different recoded data sets are shown in Tables 6 and 7⁶. Due to potential differences in the scale parameters confounded within the estimated beta parameters, the attribute parameter estimates are not directly comparable across models. Section 4.4 will present a more appropriate comparison of attribute preferences by comparing attribute WTP estimates. It is however evident from Tables 4 to 7 that the signs of estimated attribute parameters are identical across the four models. One exception is that of the ASC mean estimate. For both datasets, the asymmetric recoding in Model 3 results in a sign reversal for this parameter. Specifically it changes from a negative to a positive sign, indicating that the asymmetric recoding entails a marked increase in the utility

⁶ In order to avoid information overload, the calculated probabilities of reversed sign are not reported in Tables 6 to 9, as these were not found to differ much from the benchmark models.

Table 6: The Results of Elimination (Model 2), Asymmetric Recoding (Model 3) and Symmetric Recoding (Model 4) for the Motorway Data.

Attribute	Model 2 (Elimination)		Model 3 (Asymmetric recoding)		Model 4 (Symmetric recoding)		
	Estimate	(<i>t</i> -value)	Estimate	(<i>t</i> -value)	Estimate	(<i>t</i> -value)	
ASC	-0.3453	(-0.98)	2.3454	(11.15)	-0.1848	(-1.19)	
Forest	- Mean	-0.2211	(-10.45)	-0.1718	(-10.39)	-0.1369	(-11.93)
	- St.dev.	0.2758	(10.47)	0.2144	(11.70)	0.1685	(13.20)
Wetland	- Mean	-0.3169	(-9.56)	-0.2678	(-10.59)	-0.1968	(-10.50)
	- St.dev.	0.2221	(4.96)	0.1824	(4.88)	0.1563	(5.33)
Heath	- Mean	-0.1003	(-3.65)	-0.1060	(-4.89)	-0.0966	(-6.08)
	- St.dev.	0.2060	(5.13)	0.1266	(3.17)	0.0782	(1.89)
Price		-0.0022	(-20.46)	-0.0019	(-22.80)	-0.0017	(-23.67)
Sigma		4.4494	(11.22)	3.1785	(15.20)	2.1447	(16.28)
Num. observations		2358		3570		3570	
Num. respondents		505		595		595	
LL _{zero}		-2591		-3922		-3922	
LL _{Full}		-1614		-2414		-3003	
LRI (pseudo- R^2)		0.377		0.384		0.234	
Adj. LRI		0.375		0.384		0.233	
LR test		1951.4		3015.6		1838.1	

associated with the status quo situation. Across all models, the estimated standard deviations suggest a significant amount of heterogeneity in preferences.

For Model 2 based on elimination the log-likelihood and the adjusted pseudo- R^2 measures suggest a better model fit than in the original model and sample, reflecting that eliminating uncertain responses reduces noise in data. This tendency is evident in both datasets. Similarly, in Model 3 based on asymmetric recoding, the log-likelihood and the adjusted pseudo- R^2 measures suggest a better model fit than in the benchmark model and sample, suggesting that much noise has been eliminated. This is however only the case for the motorway data, whereas for the national park data the model fit appears to be lower for the recoded data set than for the benchmark data. For Model 4 based on a symmetric recoding of the choice data, the log-likelihood and the adjusted pseudo- R^2 measures indicate a markedly lower model fit than the benchmark for both datasets. This suggests that the symmetric recoding actually increases the amount of noise in the data. Note that as the dataset changes across Models 1 to 4 due to the applied recoding, the likelihood and pseudo- R^2 values are strictly speaking not directly comparable across Models 1 to 4. Hence, the differences in model fit can only be interpreted in terms of the models' ability to explain the resulting data. The differences in model fit do not tell us anything about whether the underlying assumption behind one model is more true than the assumptions behind other models.

Table 7: The Results of Elimination (Model 2), Asymmetric Recoding (Model 3) and Symmetric Recoding (Model 4) for the National Park Data

Attribute	Model 2 (Elimination)		Model 3 (Asymmetric recoding)		Model 4 (Symmetric recoding)		
	Estimate	(<i>t</i> -value)	Estimate	(<i>t</i> -value)	Estimate	(<i>t</i> -value)	
ASC	-2.5647	(-9.45)	0.1356	(0.89)	-1.5371	(-11.58)	
Location Møn	- Mean	0.3729	(2.20)	0.2399	(1.73)	0.2499	(2.25)
	- St.dev.	1.3991	(5.94)	1.0958	(5.56)	0.6828	(3.90)
Location Thy	- Mean	0.3624	(2.04)	0.2512	(1.70)	0.1682	(1.37)
	- St.dev.	1.2046	(5.47)	1.1048	(5.98)	0.9033	(5.46)
Location Nordsjælland	- Mean	-0.1607	(-0.76)	-0.2232	(-1.21)	0.0076	(0.06)
	- St.dev.	1.8749	(7.61)	1.7239	(9.00)	1.1754	(7.28)
Location Mols Bjerger	- Mean	0.5371	(2.58)	0.3043	(1.82)	0.3045	(2.30)
	- St.dev.	1.9009	(7.22)	1.6082	(8.38)	1.1501	(7.11)
Location Lille Vildmose	- Mean	0.8953	(5.20)	0.6157	(4.36)	0.5389	(4.71)
	- St.dev.	1.4776	(6.99)	1.1567	(6.54)	0.8528	(5.26)
Location Vadehavet	- Mean	0.7891	(4.42)	0.5489	(3.80)	0.4733	(4.19)
	- St.dev.	1.7096	(8.42)	1.2999	(8.04)	0.8247	(5.21)
Nature preservation	- Mean	0.1186	(3.06)	0.0990	(3.19)	0.0652	(2.42)
	- St.dev.	0.2566	(3.25)	0.2148	(3.12)	0.2271	(4.37)
Effort for animal/plants	- Mean	1.2476	(12.86)	0.9009	(11.59)	0.8294	(13.44)
	- St.dev.	0.8424	(5.92)	0.8465	(7.47)	0.6436	(6.74)
Walking and biking paths	- Mean	0.1330	(1.22)	0.0568	(0.72)	-0.0025	(-0.04)
	- St.dev.	0.8082	(4.57)	0.5806	(3.55)	0.4608	(3.13)
Price	- Mean	-0.0021	(-29.63)	-0.0018	(-32.28)	-0.0014	(-33.69)
	- St.dev.	0.8082	(4.57)	0.5806	(3.55)	0.4608	(3.13)
Sigma		3.0839	(11.45)	2.5009	(19.32)	0.1060	(15.96)
Num. observations		3581		3581		3581	
Num. respondents		579		579		579	
LLzero		-3934		-5346		-5346	
LLFull		-2663		-4006		-4137	
LRI (pseudo- R^2)		0.323		0.251		0.226	
Adj. LRI (Adj.pseudo- R^2)		0.321		0.249		0.224	
LR test		2542.9		2680.2		2417.4	

4.3 Explicit Modelling of Certainty in Choice Through use of Scaling

Tables 8 and 9 present the results obtained from the random parameter error component models, where we have also allowed for a parametric representation of the scale variation as explained in Section 2.2.2. We estimate the scale function models using Biogeme (Bierlaire 2003).

Table 8: Results from modelling scale variation in the motorway data set using stated certainty (Model 5) or determinants of stated certainty (Model 6)

Attribute	Model 5		Model 6		
	Estimate	(<i>t</i> -value)	Estimate	(<i>t</i> -value)	
ASC	-0.1300	(-2.45)	-0.2970	(-1.91)	
Forest	- Mean	-0.0399	(-4.54)	-0.0884	(-2.33)
	- St.dev.	0.0572	(4.62)	0.1300	(2.32)
Wetland	- Mean	-0.0594	(-4.54)	-0.1230	(-2.35)
	- St.dev.	0.0591	(4.10)	0.1450	(2.21)
Heath	- Mean	-0.0222	(-3.40)	-0.0539	(-2.08)
	- St.dev.	0.0311	(3.32)	0.0869	(1.90)
Price		0.0496	(4.82)	0.1170	(2.34)
Sigma		0.6730	(4.69)	1.3200	(2.52)
Scale function parameters^a					
Don't know	0.77	(-0.97)	-	-	
Uncertain	0.76	(2.27)	-	-	
Neither/nor	1.01	(0.05)	-	-	
Certain	1.48	(-0.61)	-	-	
Very certain	1.71	(3.20)	-	-	
Income medium	-	-	0.88	(-1.24)	
Income high	-	-	1.11	(1.17)	
Male	-	-	1.02	(0.27)	
Choice set number	-	-	-0.09	(-3.24)	
Num. observations	3570		3570		
Num. respondents	595		595		
LL _{zero}	-3922		-3922		
LL _{Full}	-2717		-2740		
LRI (pseudo- <i>R</i> ²)	0.307		0.301		
Adj. LRI	0.304		0.298		
LR test	2410.3		2364.2		

^a The scale function parameters are dummy variables estimated relative to the base levels which are normalised to 1. In Model 5, the normalised level is the 'very uncertain'-responses. In Model 6, the normalised base levels are 'income low' and 'female'. Hence, the *t*-test values reported for the associated parameters test the hypothesis of the estimated scale function parameter being equal to 1. The choice set number variable is, however, entered as a continuous variable, so the relevant null hypothesis for this variable is the parameter being equal to zero.

Table 9: Results from Modelling Scale Variation in the National Park Data set using Stated Certainty (Model 5) or Determinants of Stated Certainty (Model 6)

Attribute		Model 5		Model 6	
		Estimate	(<i>t</i> -value)	Estimate	(<i>t</i> -value)
ASC	- Mean	-0.6074	(-4.53)	-0.2670	(-7.61)
Location Møn	- Mean	0.1045	(2.97)	0.0354	(1.98)
	- St.dev.	0.3243	(1.85)	0.1580	(4.94)
Location Thy	- Mean	0.1120	(4.78)	0.0368	(2.09)
	- St.dev.	0.3191	(3.13)	0.1000	(2.86)
Location Nordsjælland	- Mean	-0.0322	(3.85)	-0.0104	(-0.49)
	- St.dev.	0.5827	(4.09)	0.2530	(-6.93)
Location Mols Bjerge	- Mean	0.1552	(2.85)	0.0564	(2.85)
	- St.dev.	0.4201	(4.20)	0.1630	(5.17)
Location Ll. Vildmose	- Mean	0.2397	(2.34)	0.1010	(4.98)
	- St.dev.	0.3625	(3.71)	0.1610	(5.55)
Location Vadehavet	- Mean	0.1861	(-0.65)	0.0810	(3.91)
	- St.dev.	0.4448	(4.44)	0.2140	(6.47)
Nature preservation	- Mean	0.0326	(1.80)	0.0150	(3.78)
	- St.dev.	0.0496	(2.65)	0.0197	(1.39)
Effort for animal/plants	- Mean	0.3156	(2.38)	0.1280	(8.65)
	- St.dev.	0.1736	(3.78)	-0.0975	(5.36)
Walking and biking paths	- Mean	0.0457	(3.30)	0.0205	(1.92)
	- St.dev.	0.1513	(4.30)	0.0825	(4.13)
Price		-0.0006	(-4.96)	-0.0003	(-9.47)
Sigma		0.8271	(4.73)	-0.3600	(-8.76)
<u>Scale function parameters^a</u>					
Don't know		0.68	-1.06	-	-
Uncertain		1.01	0.05	-	-
Certain		1.22	1.07	-	-
Very certain		1.21	1.00	-	-
Income medium		-	-	1.22	2.29
Income high		-	-	1.08	0.97
Male		-	-	0.93	-1.07
Choice set number		-	-	-0.01	-0.42
Num. observations		4866		4866	
Num. respondents		636		636	
LL _{zero}		-5346		-5346	
LL _{Full}		-3752		-3767	
LRI (pseudo- <i>R</i> ²)		0.298		0.295	
Adj. LRI		0.293		0.291	
LR test		3187.3		3157.6	

^a The scale function parameters are dummy variables estimated relative to the base levels which are normalised to 1. In Model 5, the normalised certainty level is 'Very uncertain'. In Model 6, the normalised base levels are 'Income low' and 'female'. Hence, the *t*-test values reported for the associated parameters test the hypothesis of the estimated scale function parameter being equal to 1. The choice set number variable is, however, entered as a continuous variable, so the relevant null hypothesis for this variable is the parameter being equal to zero.

In Model 5, for both data sets we see that all attribute parameter estimates are significant and of the expected sign when compared with Model 1. The estimated scale parameters provide a test of the hypothesis of equal error variance across certainty level groups. For both data sets the base level for the scale function is the group of 'very uncertain'-responses, which is normalised to 1. For the motorway data the 'don't know' and 'uncertain' groups of responses obtain estimates below 1, indicating higher error variance for these choices, but the t-test values cannot reject the hypothesis of equality of error variance across these two groups and the base. The 'neither certain nor uncertain' group is estimated with a parameter just above one and is also not significantly different from the base group of 'very uncertain' answers. Turning to the 'certain' and 'very certain' responses, these groups obtain significantly higher scale function parameter estimates than the base and the other estimated scale parameters. For instance, in the motorway data the scale parameter of the 'very certain' responses relative to the base is $\exp(1.71) \approx 5.5$, which corresponds to a variance of about 0.05 for the assumed Gumbel error term. The base group error variance is about 0.22. Hence, the error variance in the 'very certain' group is more than one fourth of that of the base. The error variance of the 'uncertain' group is 0.36, i.e. seven times higher. For the national park data, the scale function parameter of the group 'don't know' is the only one to be smaller than the base level. The groups of 'certain' and 'very certain' answers are significantly different from the 'don't know' group of responses, but none of the estimated scale parameters are significantly different from the base of 'very uncertain' answers. The fit of Model 5 is better than in Model 1 for the motorway data, but worse for the national park data. Contrary to the recoding models, direct comparisons of model fit across Models 1, 5 and 6 are valid as the datasets underlying these models are identical. Models 5 and 6 simply make use of more of the information available in the datasets.

In Model 6, the scale function contains several of the variables found to influence respondents' stated certainty in choice (Lundhede *et al.* 2009a). As can be seen, the results are less encouraging. The model fit of the motorway data is worse, though still marginally improved when compared with Model 1. For the national park data the model fit in terms of adjusted pseudo- R^2 is again slightly worse. The scale function parameters do not suggest any evidence of gender differences in any of the data sets. For the different income groups we find that medium and high income groups do not differ significantly from the low income base group in the motorway data. However, the high income group does obtain a significantly higher scale function parameter (and hence a lower error variance) than the medium income group. For the national park data, the highest scale parameter is for the middle income group, and it is the only one that is significantly different from the basis. The effect of the choice set number on scale is significant in the motorway model while insignificant in the national park model. However, as the parameter in both cases is negative it indicates a weak evidence of a possible fatigue effect (Bradley and Daly 1994). The evidence is seen in spite of the fact that our surveys had sequences of only six, respectively eight, choices per respondent whereas Bradley and Daly applied up to 16 choices per respondent.

4.4 Comparison of WTP Estimates

In Tables 10 and 11, we compare the WTP estimates across all six models along with their confidence intervals and changes in these – for both data sets. As can be seen, looking across models for both data sets, most of the WTP estimates show only minor or insignificant differences for most environmental attributes, when compared with

Model 1 – or each other. The largest differences are found for the asymmetric recoding in Model 3. More specifically, the WTP for the ASC increases drastically in both datasets, whereas remaining attributes are only subject to minor and largely insignificant changes. This result is not surprising considering that a large amount of improvement choices have been recoded to status quo choices which should lead to an increase in the estimated utility associated with for the status quo situation.

Consider also the relative width of the confidence intervals around the attribute level WTP estimates, as compared with the width of the confidence interval of the benchmark in Model 1. Note that relative to the benchmark the recoding Models 2 and 3 generally show increases in the width of the WTP confidence intervals, whereas the tendencies for Models 4 to 6 are less clear.

Table 10: Comparison of WTP Estimates and Confidence Intervals for the Motorway Data

	Model 1 Benchmark model	Model 2 Uncertain eliminated	Model 3 Asymmetric recoding	Model 4 Symmetric recoding	Model 5 Certainty level in scale	Model 6 Respondent variables in scale
ASC	-169 (-373;35)	-157 (-469;155)	1251 (1004;1498)	-106 (-279;67)	-262 (-454;-70)	-254 (-330;-177)
	-	+53%	+21%	-15%	-6%	-62%
Forest	82 (69;96)	101 (82;119)	92 (75;109)	78 (66;91)	80 (65;96)	76 (69;82)
	-	+30%	+22%	-13%	+12%	-55%
Wetland	109 (86;131)	144 (116;172)	143 (116;170)	113 (91;134)	120 (95;144)	105 (95;115)
	-	+25%	+19%	-6%	+9%	-56%
Heath	46 (28;65)	46 (21;70)	57 (34;79)	55 (37;74)	45 (25;65)	46 (38;54)
	-	+31%	+23%	-2%	+8%	-57%
# observations	3570	2358	3570	3570	3570	3570
# respondents	595	505	595	595	595	595
LL _{zero}	-3922	-2591	-3922	-3922	-3922	-3922
LL _{Full}	-2756	-1614	-2414	-3003	-2717	-2740
LRI (pseudo- R^2)	0.297	0.377	0.384	0.234	0.307	0.301
Adj. LRI	0.296	0.375	0.384	0.233	0.304	0.298
LR test	2331.9	1951.4	3015.6	1838.1	2410.3	2364.2

Note: Intervals in parenthesis are 95% confidence intervals obtained using the Krinsky-Robb procedure (Krinsky and Robb 1986; 1990) with 10,000 replications. Below these we show the percentage increase in the width of the confidence interval for each model compared with the benchmark in Model 1.

Table 11: Comparison of WTP Estimates and Confidence Intervals for the National Park Data

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Benchmark model	Uncertain eliminated	Asymmetric recoding	Symmetric recoding	Certainty level in scale	Respondent variables in scale
ASC	-1078 (-1256;-899)	-1234 (-1470;-997)	75 (-90;240)	-1065 (-1244;-886)	-1063 (-1278;-848)	-1068 (-1258;-878)
	-	+33%	-7%	0%	+20%	+6%
Location Møn	175 (39;312)	179 (21;338)	133 (-17;282)	173 (23;323)	183 (39;327)	142 (3;280)
	-	+16%	+9%	+10%	+6%	+2%
Location Thy	147 (5;289)	174 (7;342)	139 (-21;299)	117 (-50;283)	196 (42;350)	147 (10;284)
	-	+18%	+13%	+17%	+8%	-4%
Location Nordsjælland	-57 (-226;113)	-77 (-280;125)	-123 (-326;79)	5 (-182;193)	-56 (-238;126)	-42 (-211;128)
	-	+20%	+19%	+11%	+7%	0%
Location Mols Bjerge	218 (63;373)	258 (61;456)	168 (-15;351)	211 (30;392)	272 (106;438)	226 (77;374)
	-	+28%	+18%	+17%	+7%	-4%
Location Ll. Vildmose	402 (269;534)	431 (269;593)	341 (188;494)	373 (218;529)	419 (271;567)	404 (264;544)
	-	+22%	+15%	+17%	+12%	+6%
Location Vadehavet	297 (151;442)	380 (213;546)	304 (148;459)	328 (175;481)	326 (167;484)	324 (174;474)
	-	+15%	+7%	+5%	+9%	+3%
Nature preservation	60 (29;91)	57 (21;94)	55 (21;88)	45 (9;82)	57 (24;90)	60 (30;90)
	-	+18%	+8%	+17%	+6%	-3%
Effort for animal/plants	557 (486;628)	600 (516;684)	498 (416;580)	575 (493;656)	552 (475;630)	512 (437;587)
	-	+18%	+15%	+14%	+8%	+5%
Walking and biking paths	80 (-5;164)	64 (-38;166)	31 (-54;117)	-2 (-96;92)	80 (-6;166)	82 (1;163)
	-	+21%	+1%	+12%	+2%	-4%
# obs	4866	3581	3581	3581	4866	4866
# resp	636	579	579	579	636	636
LLzero	-5346	-3934	-5346	-5346	-5346	-5346
LLFull	-3742	-2663	-4006	-4137	-3752	-3767
LRI	0.300	0.323	0.251	0.226	0.298	0.295
Adj. LRI	0.298	0.321	0.249	0.224	0.293	0.291
LR test	3205.8	2542.9	2680.2	2417.4	3187.3	3157.6

Note: Intervals in parenthesis are 95% confidence intervals obtained using the Krinsky-Robb procedure (Krinsky and Robb 1986; 1990) with 10,000 replications. Below these we show the percentage increase in the width of the confidence interval for each model compared with the benchmark in Model 1.

5 Discussion

We have adapted and evaluated three recoding approaches from the CV literature as well as two approaches to capture respondent uncertainty explicitly as systematic variations in the scale parameter. We discuss the merits and problems associated with the different approaches in turn.

5.1 The Effect of Elimination of Uncertain Choices and Symmetric and Asymmetric recoding

The effect of elimination of uncertain responses on WTP in the motorway data is a considerable increase for the 'Forest' and the 'Wetland' attributes. For the national park data, effects are more mixed, but not significant at the attribute level. These observations do not lend support to the hypothesis of Champ *et al.* (1997) that respondents use the certainty question to scale down their stated WTP. If this was the case, we should have seen significant decreases in WTP following elimination of choices that respondents 'regret'. On this censored data set, the estimated models show a better data fit in terms of pseudo- R^2 , in spite of the much lower number of observations. This shows that much noise is eliminated when eliminating uncertain answers, but not that the underlying assumption of uncertainty is the true one. Furthermore, confidence intervals of WTP increase quite a lot. In conclusion, not much is gained by elimination – observations and information are lost and WTP may slightly increase, but the differences in terms of for instance confidence intervals of WTP-measures give little justification for eliminating uncertain choices.

The effect of the asymmetric recoding is a dramatic increase in the WTP estimate for the ASC in both data sets. This is not surprising considering that all uncertain answers are recoded as a choice of the status quo alternative and, hence, the status quo alternative would appear relatively more attractive on the recoded dataset. A similar effect is well-documented for CV studies (Welsh and Poe 1998; Loomis and Ekstrand 1998; Alberini *et al.* 2003). For the other parameters we see two different albeit weak reactions. In the motorway data we see a small but mostly insignificant increase in the WTP, whereas for the national park data we see a small but insignificant decrease in all other attributes. Considering the large increase in the marginal WTP for the ASC and the only minor changes in marginal WTP for the quality attributes, the total WTP for a change from the current situation is markedly decreased by the asymmetric recoding. This result is in accordance with findings from the CV literature (Little and Berrens 2004; Blumenschein *et al.* 2008). For the motorway data, we note that a considerable amount of noise is removed with this recoding approach, resulting in a better model fit to the recoded data, but this pattern is different for the national park data, where the model fit is reduced for the recoded data. This probably reflects the strong preferences for picking a national park, any alternative to status quo, as also reflected in the high negative parameter for ASC in the benchmark model. Thus, for the national park data the asymmetric recoding seems at odds with the respondents' underlying preferences for expressing support to a national park establishment. Again we see that the recoding tends to increase the width of confidence intervals of WTP estimates.

The effect of symmetric recoding on WTP is insignificant. The approach here is not directly comparable to the symmetric uncertainty model of, e.g. Loomis and Ekstrand (1998). The recoding here is inevitably more involved as it concerns a CE

with two alternatives to the status quo. Furthermore, the approach we suggest is also dependent on the estimated benchmark model. In most symmetric uncertainty models in the CV literature (e.g. Loomis and Ekstrand 1998 and Alberini *et al.* 2003), it is found that taking into account symmetric uncertainty increases the WTP estimates. Our finding of low impact on WTP estimates indicates that the symmetric approach to recoding we have evaluated here mainly recodes choices between alternatives that are already quite close in utility level. This confirms the findings in Lundhede *et al.* (2009a) that uncertainty partly originates from a high level of utility balance between alternatives. Thus, the symmetric recoding increases the noise in the model, which suggests that respondents have on average picked the alternative that best fits their preferences, even in the choice situations where they are uncertain. This is supported by the fact that elimination and asymmetric recoding hardly affect the WTP estimates at attribute level.

5.2 Models Explicitly Modelling Uncertainty as Scale Variations

Recoding implies that the researcher changes the preferences stated by the respondents according to what the researcher considers the best alternative answer or choice. As such it hinges on the researcher's interpretation of stated certainty in choice and how answers are grouped into certain and uncertain answers. Therefore it is intuitively more appealing to use the scaling approach where the stated uncertainty is explicitly parameterised in the model through an exponential formulation of the indirect utility function. However, Sælensminde (2001) notes that this approach could pose an ethical problem. As certain answers are weighted higher than uncertain answers due to a higher scale parameter, it may imply putting less weight on choices obtained from specific demographic groups. Considering the findings of Lundhede *et al.* (2009a), using the scaling approach in the present case would indeed result in putting less weight on for instance choices obtained from female respondents than from male respondents. Still, compared to the recoding approaches, which to a much higher degree entail a considerable down-weighting of uncertain responses, with the elimination of uncertain responses equivalent to a weight of zero being the most severe, the scaling approach would seem the preferred method.

In Model 5, cf. Tables 8 and 9, we have modelled the scale parameter as a function of stated certainty, essentially providing six (five) different scale parameter estimates for each of the six (five) certainty statements and hence groupings of choice sets. The overall effect on WTP is negligible. For the motorway data, the model performance as measured by pseudo- R^2 is slightly better than for the benchmark model in Table 4, but the same is not the case for the national park data (Table 5).

There is strong and significant variation in scale across the choice sets in the different stated certainty categories. There are large differences in the unexplained error variances across the certainty categories – sometimes to an order of 10. This implies two things: First, that mapping respondents' certainty of choice, as we have done here, may enable explicit modelling of scale variation to improve model performance. In the present surveys, WTP estimates at attribute level remain largely unaffected as compared to the benchmark. For some attributes the confidence intervals are tighter than in the benchmark model whereas the opposite is the case for other attributes. There is however a tendency that Model 5 generally produces tighter confidence intervals than the recoding models. This implies a stronger and more reliable inference, as also predicted by Li and Mattsson (1995). Secondly, the results stress that reducing uncertainty in response is an important issue, which could be

treated already in the design phase of CEs as for instance suggested by Lundhede *et al.* (2009a).

For both data sets, 'don't know' has the lowest scale parameter. It seems reasonable that the group of 'don't know' responses has the highest degree of error variance, as not knowing how certain one is would indicate a high degree of uncertainty to the whole questionnaire as such, and consequently a relatively high error variance as compared to at least knowing your degree of uncertainty, whether certain or uncertain. The results support this.

In the sixth model, cf. Tables 8 and 9, we have tried to capture variation in scale using respondent and choice set variables as proxies for the stated certainty. These proxies are variables⁷ that have been found to correlate with stated certainty levels in Lundhede *et al.* (2009a). The model performs slightly better than the benchmark model for the motorway data and slightly worse for the national park data, and there is no overall effect on the WTP measures. However, for none of the two datasets does Model 6 perform as good as Model 5. For both data sets we find that none of the variables are significant in determining the variations in scale. The result is that we recommend eliciting respondents' perceived certainty in choice directly instead of using proxies for stated certainty.

6 Concluding Remarks

This paper is, to our knowledge, the first to specifically address and evaluate different ways of handling respondents' stated choice uncertainty in CE. We adapt and evaluate three recoding approaches from CV and evaluate two approaches to capture respondent uncertainty explicitly as systematic variations in the scale parameter.

Recoding uncertain choices either by eliminating them or asymmetrically recode them as status quo choices may reduce the noise in the data sets, but the effects on WTP measures and the confidence intervals of these are ambiguous and in most cases insignificant. A symmetric recoding approach, where uncertain choices are recoded to the best alternative choice, did not affect WTP measures significantly, but in fact increased the noise in data. A shortcoming of all three recoding approaches is that the researcher essentially has to change the preferences stated by the respondents in some way or leave out information. Whether either is a valid approach, depends on largely unrevealed causes for respondent uncertainty. Consequently, recoding approaches do not seem to be a satisfactory way to handle uncertain answers unless specific behavioural hypotheses are present.

As an alternative we evaluate the use of stated respondent certainty in choice as a basis for modelling scale variation across the responses and data, and find that indeed scale varies greatly and significantly with the certainty in choice. Specifically, the higher the level of certainty in choice is, the lower is the amount of unexplained variance. Again, we find no significant differences in WTP compared with a benchmark model ignoring respondent certainty in choice, but explicitly modelling variation in scale reduces the unexplained variance considerably and offers a

⁷ Lundhede *et al.* (2009a) find that a key determinant of stated certainty is the level of utility balance within choice set. Unfortunately, model identification is not possible when we include this variable in model 6, probably due to correlation between this utility balance variable and the attribute parameters. If this problem can be solved, we have reason to suspect that model 6 will perform much better.

structurally and intuitively appealing way of accounting for uncertainty in choices in CE surveys.

While our contribution in the present paper sheds some light on particular ways of accounting for preference uncertainty, it is obvious that much research still remains to be done in this area. Recently, there has been an increasing focus on the behavioural decision processes that lead to observed behaviour which deviates from that of the traditionally assumed economic man. We believe that delving further into this topic of preference uncertainty offers a line of research which can contribute to the great challenge of developing more plausible behavioural models of choice. Specifically, asking more detailed choice set level follow-up questions aiming to identify underlying behavioural mechanisms and reasons for stating an (un-)certain answer seems like a promising way forward. Furthermore, there would seem to be an interesting research area in the potential linkage between the behavioural implications of stated certainty in choice and the very recent and rapidly expanding literature on attribute processing (see e.g. Hensher and Rose 2009). One research question in this direction could be to establish a mapping between certainty statements and the number of strategies derived from heuristics that individuals might invoke when processing information in choice experiments, preferably at the choice set level.

Acknowledgements

We acknowledge the helpful comments of two anonymous reviewers on an earlier version of the paper, and we thank Riccardo Scarpa for suggesting the use of scale modelling. We also thank participants at the EAERE annual conference in Amsterdam 2009 for useful comments. We further acknowledge the financial support from the Ministry of Environment, which facilitated the data collection for the two surveys. Finally, Thomas Hedemark Lundhede and Jette Bredahl Jacobsen acknowledge the support of the Danish Research Councils, making possible further work.

Appendix A: Example of a Choice Set used in the Motorway Survey

Q.1: Which of the following locations for the future motorways would you prefer?

Remember to imagine that the nature areas, which you visit most often, *will* be affected.

	Alternative 0	Alternative 1	Alternative 2
Number of kilometres through:			
– Forest	10 km	0 km	10 km
– Wetland	5 km	0 km	5 km
– Heath/common	5 km	5 km	2.5 km
– Arable land	80 km	95 km	82.5 km
Annual extra payment	0 DKK	200 DKK	100 DKK
I prefer ... (tick one):	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Q.15A: How certain are you of your choice?

It's ok to be uncertain – Your reply will be no less valuable for that reason!

(tick one)

- Very uncertain
- Uncertain.....
- Neither certain nor uncertain.....
- Certain
- Very certain.....
- Don't know.....

Appendix B: Example of a Choice Set used in the National Park Survey

Do you prefer Choice 1, Choice 2 or No national park?

(Mark one)

(The money has to be taken from your normal budget, and you will therefore have less money available for other things)

	Choice I	Choice II	No national park
Location of the national park	Thy	Mols Bjerge	
Nature preservation	Little extra effort	Some extra effort	
Extra initiatives for special plants and animals	Yes (Crane and red deer)	Yes Butterflies and barn owl	
Paths	No more paths	More paths	
Yearly extra income tax for your household	700 kr.	50 kr.	0 kr.

Choose only one of the possibilities ...

How certain were you of your choice? *Very certain* *Certain* *Uncertain* *Very uncertain* *Don't know*
 ...

References

- Adamowicz, W., Louviere, W. and Williams, M., 1994. Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities. *Journal of Environmental Economics and Management*, 26, 271-292.
- Adamowicz, A., Swait, J. and Boxall, P., 1997. Perceptions versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Valuation. *Journal of Environmental Economics and Management*, 32, 65-84.
- Alberini, A., Boyle, K. and Welsh, M., 2003. Analysis of contingent valuation data with multiple bids and response options allowing respondents to express uncertainty. *Journal of Environmental Economics and Management*, 45, 40-62.
- Ben-Akiva, M. and Lerman, S. R., 1985. *Discrete Choice Analysis. Theory and Application to Travel Demand*. The MIT Press, Cambridge.
- Bierlaire, M., 2003. BIOGEME: A free package for the estimation of discrete choice models. Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.
- Blumenschein, K., Blomquist, G. C., Johannesson, M., Horn, N. and Freeman, P., 2008. Eliciting willingness to pay without bias: evidence from a field experiment *The Economic Journal*, 118, 114-137.
- Bradley, M. and Daly, A., 1994. Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data. *Transportation*, 21, 167-184.

- Brownstone, D., Bunch, D. S. and Train, K., 2000. Joint Mixed Logit Models of Stated and Revealed Preferences for Alternatives Fuel Vehicles. *Transportation Research Part B*, 34, 315-338.
- Campbell, D., 2008. Identification and analysis of discontinuous preferences in discrete choice experiments. Paper presented at EAERE 16th Annual Conference, Gothenburg, Sweden, June, 25-28, 14 pp.
- Carlsson, F., Frykblom, P. and Liljenstolpe, C., 2003. Valuing wetland attributes: an application of choice experiments. *Ecological Economics*, 47, 95-103.
- Champ, P. A., Bishop, R. C., Brown, T. C. and McCollum, D. W., 1997. Using Donation Mechanisms to Value Nonuse Benefits from Public Goods. *Journal of Environmental Economics and Management*, 33, 151-162.
- DeShazo, J.R. and Fermo, G., 2002. Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. *Journal of Environmental Economics and Management*, 44, 123-143
- Domencich, T. and McFadden, D., 1975. Urban Travel Demand: A Behavioral Analysis North-Holland Publishing Company, Amsterdam.
- Ferrini, S. and 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, 342-363.
- Greene, W. H. and Hensher, D. A., 2007. Heteroscedastic control for random coefficients and error components in mixed logit. *Transportation Research E: Logistics and Transportation Review*, 43, 610-623.
- Hanemann, W. M., 1984. Welfare evaluations in contingent valuation experiments with discrete responses. *American Journal of Agricultural Economics*, 66, 332-341.
- Hensher, D. A., 2008. Joint Estimation of Process and Outcome in Choice Experiments and Implications for Willingness to Pay. *Journal of Transport Economics and Policy*, 42, 297-322.
- Hensher, D. A. and Bradley, M., 1993. Using stated response choice data to enrich revealed preference discrete choice models. *Marketing Letters*, 4, 139-151.
- Hensher, D. A., Louviere, J. J., Swait, J., 1999. Combining sources of preference data. *Journal of Econometrics*, 89, 197-221.
- Hensher, D. A., Stopher, P. R. and Louviere, J. J., 2001. An exploratory analysis of the effect of numbers of choice sets in designed choice experiments: An airline choice application. *Journal of Air Transport Management*, 7, 373-379.
- Hensher, D. A. and Greene, W. H., 2003. The mixed logit model: The state of practice. *Transportation*, 30, 133-176.
- Hensher, D. A., Rose, J. M. and Greene, W. H., 2005. Applied Choice Analysis: A Primer. Cambridge University Press, Cambridge, UK.
- Hensher, D. A. and Rose, J. M., 2009. Simplifying choice through attribute preservation or non-attendance: Implications for willingness to pay. *Transportation Research Part E: Logistics and Transportation Review*, 45, 583-590.
- Hess, S., Bierlaire, M. and Polak, J., 2005. Estimation of value of travel-time savings using Mixed Logit models. *Transport Research A: Policy and Practice*, 39, 221-236
- Hole, A. R., 2007. A comparison of approaches to estimating confidence intervals for willingness to pay measures. *Health Economics*, 16, 827-840.

- Hu, W., Adamowicz W. L. and Veeman, M. M., 2006. Labeling Context and Reference Point Effects in Models of Food Attribute Demand. *American Journal of Agricultural Economics*, 88, 1034-1049.
- Huber, J. and Zwerina, K., 1996. The Importance of Utility Balance in Efficient Choice Designs. *Journal of Marketing Research*, 33, 307-317.
- Jacobsen, J. B., Thorsen, B. J., Boiesen, J. H., Anthon, S. and Tranberg, J., 2006. Værdisætning af syv mulige nationalparker i Danmark [Valuation of seven potential national parks in Denmark]. Working paper 28, Forest & Landscape, University of Copenhagen, Denmark, 63 pp.
- Jacobsen J. B., Boiesen, J. H., Thorsen, B. J. and Strange, N., 2008. What's in a Name? The use of Quantitative Measures vs. 'Iconised' Species when Valuing Biodiversity. *Environmental and Resource Economics*, 39, 249-263.
- Jacobsen, J. B. and Thorsen, B. J., 2009. Where to put a national park and what to put in it? An a priori study of the willingness-to-pay for national parks. Manuscript under submission.
- Jacobsen, J. B., Lundhede, T. H. and Thorsen B. J., 2009: Are Economists Valuing Biodiversity at Gunpoint? Evidence of Decreasing Willingness to Pay for Wildlife Population Increases beyond Species Survival. Manuscript under submission.
- Johannesson, M., Blomquist, G. C., Blumenschein, K., Johansson, P. O., Liljas, B. and O'Connor, R. M., 1999. Calibrating Hypothetical Willingness to Pay Responses. *Journal of Risk and Uncertainty*, 18, 21-32.
- Krinsky, I. and Robb, A. L., 1986. On approximating the statistical properties of elasticities. *Review of Economics and Statistics*, 68, 715-719.
- Krinsky, I. and Robb, A. L., 1990. On approximating the statistical properties of elasticities: a correction. *Review of Economics and Statistics*, 72, 189-190.
- Li, C. and Mattsson, L., 1995. Discrete Choice under Preference Uncertainty: An Improved Structural Model for Contingent Valuation. *Journal of Environmental Economics and Management*, 28, 256-269.
- Little, J. and Berrens, R., 2004. Explaining Disparities between Actual and Hypothetical Stated Values: Further Investigation Using Meta-Analysis. *Economics Bulletin*, 3, 1-13.
- Loomis, J. and Ekstrand, E., 1998. Alternative approaches for incorporating respondent uncertainty when estimating willingness to pay: the case of the Mexican spotted owl. *Ecological Economics*, 27, 29-41.
- Louviere, J. J., Hensher, D.A. and Swait, J., 2000. Stated Choice Methods. Analysis and Applications University Press, Cambridge, UK.
- Louviere, J. J. and Hensher, D. A., 2001. Combining Sources of Preference Data. In: Hensher, D. A. ed: Travel Behavior Research: The Leading Edge. Oxford Pergamon Press, pp. 125-144.
- Lundhede, T.H., Olsen, S.B., Jacobsen, J.B. and Thorsen, B.J., 2009a. Tough and easy choices: Influence of utility difference on self-reported certainty levels. Paper presented at the Envecon Conference, London, United Kingdom, March 20, 30 pp.
- Lundhede, T., Hasler, B. and Bille, T., 2009b. Exploring Preferences for Hidden Archaeological Artifacts – a case from Denmark. Manuscript under submission.
- McFadden, D., 1974. Conditional logit analysis of qualitative choice behaviour. In: Zarembka, P. Frontiers in Econometrics. Academic Press Inc., New York. , pp. 105-142.

- Meijer, E. and Rouwendal, J., 2006. Measuring welfare effects in models with random coefficients. *Journal of Applied Econometrics*, 21, 227-244.
- Olsen S. B., 2009. Choosing between internet and mail survey modes for choice experiment surveys considering non-market goods. *Environmental and Resource Economics*, DOI: 10.1007/s10640-009-9303-7
- Olsen, S. B., Ladenburg, J., Petersen, M. L., Lopdrup, U., Hansen, A. S. and Dubgaard, A., 2005. Motorways versus Nature - A Welfare Economic Valuation of Impacts. Report from FOI and IMV, Copenhagen, Denmark, 230 pp.
- Ready, R. C., Whitehead, J. C. and Blomquist, G. C., 1995. Contingent Valuation When Respondents Are Ambivalent. *Journal of Environmental Economics and Management*, 29, 181-196.
- Samnaliev, M., Stevens, T. H. and More, T., 2006. A comparison of alternative certainty calibration techniques in contingent valuation. *Ecological Economics*, 57, 507-519.
- Scarpa, R., Ruto, E. S. K., Kristjanson, P., Radeny, M., Drucher, A. G. and Rege, J. E. O., 2003. Valuing indigenous cattle breeds in Kenya: an empirical comparison of stated and revealed preference value estimates. *Ecological Economics*, 45, 409-426.
- Scarpa, R., Ferrini, S. and Willis, K., 2005. Performance of Error Component Models for Status-Quo Effects in Choice Experiments. In: Scarpa, R., Alberini, A. eds: Applications of Simulation Methods in Environmental and Resource Economics. The Economics of Non-Market Goods and Resources 6. Springer, Dordrecht, The Netherlands, pp 247-273.
- Scarpa, R., Willis, K. and Acutt, M., 2007. Valuing externalities from water supply: status quo, choice complexity, and individual random effects in panel kernel logit analysis of choice experiments. *Journal of Environmental Planning and Management*, 50, 449-466.
- Scarpa, R., Thiene, M. and Marangon, F., 2008. Using flexible taste distributions to value collective reputation for environmentally friendly production methods. *Canadian Journal of Agricultural Economics*, 56, 145-162.
- Schwarz, N. and Sudman, S., 1996. Answering Questions. Methodology for Determining Cognitive and Communicative Processes in Survey Research. Jossey-Bass Publishers, San Francisco.
- Swait, J. and Adamowicz, W., 2001. The Influence of Task Complexity in Consumer Choice: A Latent Class Model of Decision Strategy Switching. *The Journal of Consumer Research*, 28, 135-148.
- Sælensminde, K., 2001. Inconsistent choices in Stated Choice data: Use of the logit scaling approach to handle resulting variance increases. *Transportation*, 28, 269-296.
- Train, K., 2003. Discrete Choice Methods with Simulation Cambridge University Press, Cambridge.
- Train, K. and Sonnier, G., 2005. Mixed Logit with Bounded Distributions of Correlated Partworths. In: Scarpa, R., Alberini, A. eds., Applications of Simulation Methods in Environmental and Resource Economics. The Economics of Non-Market Goods and Resources 6. Springer, Dordrecht, The Netherlands, pp. 117-134.
- Train, K. and Weeks, M., 2005. Discrete Choice Models in Preference Space and Willing-to-pay Space. In: Scarpa R, Alberini A eds: Applications of Simulation Methods in Environmental and Resource Economics. The Economics of Non-Market Goods and Resources 6. Springer, Dordrecht, The Netherlands, pp. 1-16

- Wang, H., 1997. Treatment of "Don't-Know" Responses in Contingent Valuation Surveys: A Random Valuation Model. *Journal of Environmental Economics and Management*, 32, 219-232.
- Welsh, M. P., and Poe, G. L., 1998. Elicitation Effects in Contingent Valuation: Comparisons to a Multiple Bounded Discrete Choice Approach. *Journal of Environmental Economics and Management*, 36, 170-185.
- Whitehead J. C., Pattanayak, S. K., van Houtven, G. L. and Gelso, B. R., 2008. Combining revealed and stated preference data to estimate the nonmarket value of ecological services: An assessment of the state of the science. *Journal of Economic Surveys*, 22, 872-908.