

Tough and easy choices: Testing the influence of utility difference on stated certainty-in-choice in choice experiments

Thomas Hedemark Lundhede¹, Søren Bøye Olsen², Jette Bredahl Jacobsen¹

& Bo Jellesmark Thorsen^{1,3}

DRAFT MANUSCRIPT, UNDER SUBMISSION; PLEASE DO NOT QUOTE WITHOUT PERMISSION

Abstract: Respondents in Stated Preference studies may be uncertain about their preferences for the good presented to them. We reformulate Wang's hypothesis [35] on respondent uncertainty into a Choice Experiment context and hypothesize that respondents' stated certainty in choice increase with the utility difference between the alternative chosen and the best alternative to that. We test this hypothesis using data from two independent Choice Experiments. In modelling respondents' self-reported certainty in choice, we find evidence that the stated level of certainty increases significantly as utility difference in choice sets increases. In addition, stated certainty is higher for men than for women, and increases with income. A learning effect may reduce stated certainty as choice set number increases. We find the latter to be true for a study of a somewhat generic good (main nature types in Denmark), but not for a study of a more specific good (specific national parks).

Keywords: Stated preferences, choice experiment, environmental valuation, respondent certainty, learning effect, utility balance

¹ Division of Economics, Policy and Management Planning, Forest & Landscape, Life Science Faculty, University of Copenhagen, Rolighedsvej 23, 1958 Frederiksberg C, Denmark

² Corresponding author: sobo@life.ku.dk, telephone (+45)35333643, fax: (+45)35336801. Environmental Economics and Rural Development Division, Institute of Food and Resource Economics, Life Science Faculty, University of Copenhagen, Rolighedsvej 25, 1958 Frederiksberg C, Denmark

³ **Acknowledgements:** The authors acknowledge the financial support from the Ministry of Environment and the Environmental Assessment Institute, which made possible the data collection for the surveys used. Thomas Hedemark Lundhede and Jette Bredahl Jacobsen further acknowledge the support of the Danish Research Councils making possible further work.

1. Introduction

Stated Preference methods are widely applied in environmental valuation in spite of much critique and a continuous struggle with various methodological challenges and issues. In recent years, the Choice Experiment method has been applied in an increasing number of environmental valuation cases. It has been argued that this method may be less susceptible to such problems as embedding, scale insensitivity and other methodological flaws often associated with the Contingent Valuation (CV) method [19,1]. In this paper we focus on the issue of respondent certainty when stating preferences in Choice Experiments.

It is a key assumption in stated preference methods that the respondents know their own true preferences and are able to assess the exact utility they derive from the good presented to them. Respondents are therefore assumed to answer valuation questions in a way that precisely reflects their true preferences for the good being assessed [12]. However, responding to a hypothetical question about willingness to pay for a non-market good is a difficult task for most people. They may be uncertain as to the exact value of the good for them, about the meaning of words and sentences applied in the survey instrument or interview, or simply unfamiliar with relating to the good in monetary terms. This gives rise to preference uncertainty as formalized by [22]. Throughout the last two decades, preference uncertainty has been recognized as a problem when conducting stated preference studies. Different approaches have been taken, primarily in CV studies, to take account of this uncertainty [24,2]. The presence of preference uncertainty of the respondent may pose several problems for the estimation of welfare measures from stated preference studies. As pointed out by [22], even if parameter estimates are consistent and unbiased, faulty estimates on variance may lead to biased value inference.

In this paper we investigate causes for post-decisional stated uncertainty in Choice Experiments. We pursue in particular the argument of Wang [35] that one would expect uncertainty to be high for indicated payments close to the individual's true WTP, and low for payments distinctly smaller or larger than the true WTP. Wang's point remains a hypothesis which has only tentatively been examined in the literature [31], whereas other factors like prior knowledge and environmental attitude have been examined more thoroughly [4,18,31]. There seems to have been little attention devoted to analysing the effects of survey instrument and survey design on respondent uncertainty in a Choice Experiment context. This is unfortunate, as this type of research is needed for improving methods in order to reduce respondent uncertainty and avoid biased value inference.

It is sometimes recommended in the choice set design literature, that in order to derive the largest possible amount of information from respondents' choices, one needs to aim for a high level of utility balance in choice sets, i.e. that utility for the alternatives in each choice set are close to each other [15,20]. This recommendation obviously may clash with the problem of uncertain responses and the handling of these if individuals' uncertainty increases, the closer in utility level the alternatives in a choice set are, i.e. the hypothesis proposed by Wang [35] and implicitly by Li and Mattson [22]. This calls for approaches to examine respondent uncertainty at choice set level – and across more alternatives than in a CV question.

We use two independent Choice Experiment studies where respondents report their perceived certainty in choice following *each* choice set evaluated. We focus on the underlying data generating processes that potentially lead to preference uncertainty in the Choice Experiment responses. We aim to identify significant determinants of preference uncertainty, especially those which could potentially be influenced by the researcher. In particular we focus on how the utility difference

across alternatives in a choice set may influence respondents' stated certainty level. Furthermore, we evaluate other survey design factors as well as respondent characteristics which might affect the perceived certainty in choice.

In order to test the Wang-hypothesis we conduct a two-step analysis. In the first step we estimate a mixed logit error component model based on respondent choices without taking into account the stated uncertainty. This produces an estimated indirect utility function, which is then utilized to assign an aggregate utility measure to each alternative in the choice sets. We estimate utility functions for each respondent, and use these to construct a measure of utility difference, defined as the difference in utility between the chosen alternative and the best alternative to that in terms of the calculated utility. In the second step of the analysis we estimate a probit model with the self-reported certainty level serving as the response variable. We include the utility difference variable created in the first step as an explanatory variable, in order to enable an explicit test of the hypothesis that respondent certainty increases with utility difference.

We find that the utility difference does indeed explain a significant part of respondent certainty, and hence we find support for Wang's hypothesis [35] in a rather subtle test of this compared to earlier work. We do, however, also find that other design related and respondent specific factors have significant impact on respondent certainty and we provide interpretations of these findings too.

The paper is organised as follows. First we take a closer look at the theoretical and empirical works on respondents' certainty, which allows us to formulate some a priori expectations concerning our analyses. Section two outlines the econometric models and theory applied. In section three we give a brief introduction to the two Choice Experiments studies, which provided the data for this study.

The results are presented in section four. In section five we discuss the results and we conclude in section six.

2. Theory and methods

2.1 Respondent uncertainty about own stated preferences

It is a key assumption in stated preference methods that respondents are able to precisely assess the utility they derive from the good presented to them. Hence, respondents can answer any given valuation question with absolute certainty [12]. There are numerous arguments why this assumption may not be valid. In recognition of this several CV studies have tried to obtain a measure of, or an expression for, the degree of certainty perceived by the respondent when answering valuation questions, and take account of this in estimations. The approaches taken to that end can roughly be split into two groups. The first approach is to have respondents choose among different answers to the payment question, which explicitly incorporate some level of certainty, e.g. ‘Don’t Know’ [35], ‘I will definitely pay’ or ‘I most probably will not pay’ [2,30,36]. 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 [9,22,24], either in the form of a numeric scale or text statements. The latter approach has the advantage that it does not interfere directly with the valuation task, yet it too hinges on some degree of researcher interpretation concerning the stated certainty in order to incorporate it in estimations.

Samnaliev et al. [31] summarizes four assumptions or hypotheses regarding stated certainty, which we briefly present and discuss here. One hypothesis, adapted from [32], 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. If such behaviour dominates, we should find a fairly constant level of certainty across alternatives and questions – but in fact stated certainty varies systematically with the bid [24,35]. A second and related hypothesis is that certainty levels may be susceptible to strategic behaviour, e.g. respondents may exaggerate certainty along with stated WTP [31]. However, any strategic behaviour in relation to stated WTP should usually be screened for, and consequently most strategic answers should be excluded from further analysis on that account. Thus, 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 reducing hypothetical bias is assumed [9]. A fourth hypothesis is formulated by Wang [35] and implied in [22]. This hypothesis maintains the assumption that respondents are rational, truth-telling, non-strategic and honest. Wang, however, breaks away from the widely held assumption in most CV studies that each respondent holds a single ‘true’ value in mind for the good being valued. Instead, Wang hypothesizes that rather than a single true value, each respondent holds a value distribution due to uncertainty about their own true utility function. Therefore, respondents may be quite uncertain as to their answer (‘Yes’/‘No’) in a referendum CV study when the offered bid price is close to their maximum WTP, as measured by the mean of the value distributions – but quite certain when very different from their maximum WTP. In other words, the choice becomes tough when the utilities associated with the ‘Yes’ and the ‘No’ option are close to each other, but when the difference in utility increases, the choice becomes relatively easier. Thus, the difference in utility between the alternatives implied by ‘Yes’ and ‘No’ is assumed to determine stated certainty in choice. While [24], [31] and [35] find indications that support this hypothesis, a formal test has yet to be made.

All of the above literature exclusively treats the issue of certainty in a CV context. How can we transfer approaches and knowledge gained into a Choice Experiment context? In Choice Experiments, respondents usually evaluate more than one alternative version of the environmental change, potentially up against a status quo alternative. In both of the two surveys used in the present paper, respondents are faced with the status quo alternative along with two policy-generated alternatives, a total of three alternatives per choice set. Thus, when respondents are asked to state their uncertainty about the choice of alternative *ex-post*, this uncertainty may relate to the choice between the alternative they chose and any of the two competing alternatives. However, it seems reasonable to assume that the respondents' uncertainty is related to the choice between the alternative chosen, and the best alternative to that. Thus, we define the utility difference between these two alternatives as the one relevant for testing Wang's hypothesis. The implication is that if the alternative chosen has the highest utility, the utility difference is positive, but if the chosen alternative is not the one with the highest utility in the choice set, it will be negative.

As the respondents make repeated choices in the Choice Experiment studies investigated here, there may be a learning effect through the sequence of choice sets, which may influence respondents' stated certainty in choice. As respondents evaluate more choice sets, initial instability of preferences may decrease [3,6,8,13], preferences may be constructed or at least completed [5], and any starting point bias should decrease [21]. It is found in [11,26] that respondents who are certain in their choices are less susceptible to starting point bias than less certain respondents. These observations suggest that the choice set number should be an important predictor of respondents' stated certainty in choice.

2.2 The random parameter error component logit model

The Choice Experiment method relies on McFadden's [27] 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 observation of utility can only be made imperfectly, the random utility model is the foundation for estimation. For a choice between a status quo alternative and two policy alternatives the utility model can formally be described as:

$$U(kin) = \begin{cases} V(ASC, x_{kin}, \beta_i) + \varepsilon_{kin} & \text{if } k = 1 \text{ (status quo);} \\ V(x_{kin}, \beta_i, \beta_{fixed}, \sigma_i) + \varepsilon_{kin} & \text{if } k = 2, 3; \end{cases} \quad (1)$$

The term $U(kin)$ is the i 'th individual's true utility function for the good described by alternative k in choice situation n . V is the observed indirect utility function, which depends on x_{kin} , the vector of variables that we can observe. Here, the vector of utility parameters β and the error component, σ , which is common for the two experimentally designed alternatives, are parameters to be estimated. The term ε_{kin} is stochastic in the sense that its variation cannot be observed by the analyst.

Assuming that ε_{ki} is IID extreme value distributed, the Mixed Logit probabilities of individual i choosing alternative k 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 generalized to allow for repeated choices by the same respondent, i.e. a panel structure, by letting k_n be a sequence of alternatives, one for each choice occasion [34]. If the utility coefficients vary over people but are constant over the N choice occasions for each individual, the resulting probabilities can be specified as:

$$\Pr(kin) = \int \left(\prod_{n=1}^N \left[\frac{\exp(\beta'_i x_{k,in})}{\sum_j \exp(\beta'_i x_{jn})} \right] \right) \phi(\beta|b, W) d\beta \quad (2)$$

where $\phi(\beta|b, W)$ is the distribution function for β , in the present case a normal distribution, with mean b and covariance W . This is also referred to as the Random Parameter Logit. The analyst chooses the appropriate distribution for each parameter in β . This standard setup is applied to the Choice Experiment data analysed here.

2.3 A probit model of certainty in choice incorporating utility difference

In this second step, we use the estimated model in (2), to calculate the expected aggregate utility of each alternative for each individual and then calculate the expected utility difference, UD , between the alternative chosen, k , and the best alternative to that (either l or m), i.e. for each choice set:

$$UD = E(u_{ki}(x_{ki}, \epsilon_{ki})) - \max\{E(u_{li}(x_{li}, \epsilon_{li})), E(u_{mi}(x_{mi}, \epsilon_{mi}))\} = \hat{\beta}'_i x_{ki} - \max\{\hat{\beta}'_i x_{li}; \hat{\beta}'_i x_{mi}\} \quad (3)$$

That is, the utility of each alternative is calculated by multiplying the estimated utility weights with the corresponding attribute levels. This utility difference is used as an explanatory variable in a standard random effects binary probit model to identify explanatory variables for the self-reported certainty in choice. The binary probit model is an index model specified as the conditional probability that a binary response variable will take the value one:

$$P(y = 1 | x) = G(x' \beta) \equiv p(x) \in [0,1] \quad (4)$$

Here, y is the binary response variable, x is a vector of explanatory variables, β is a parameter vector, and G is a function mapping the linear index $x'\beta$ into the response probability. In our analyses, the response variable, y , takes the value one if respondents report a ‘certain’ or ‘very certain’ choice on the certainty scale presented after each choice set (see appendix A and B), and zero otherwise. The utility difference variable constructed in step one enters the model as an explanatory variable in x . Remaining explanatory variables in x have been chosen on account of *a priori* expectations of their potential impact on certainty in choice. The probit model assumes that the unobservable error terms are normally distributed with mean zero.

3. Data description

The data originates from two independent Choice Experiment surveys. One that examines preferences for reducing the impact of new motorways on different types of nature, and one that examines preferences for new national parks in Denmark.

3.1 The Motorway survey

The hypothetical scenario shown to respondents was based on the assumption that 100 kilometres of new motorways are to be built in Denmark during the next ten years. The scenario describes that the exact location of these stretches of motorway through the countryside can be decided upon with more or less consideration for different nature areas.

Three different types of nature were identified and chosen as attributes in the Choice Experiment design. They 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 through taxes, lending credibility to this payment vehicle. The attributes and their assigned levels are summed up in table 1.

>>>TABLE 1 ABOUT HERE<<<

The attribute levels were assigned to alternatives using an experimental design and paired into choice sets of 3 alternatives. As a full factorial design comprised 162 alternatives, a D-optimal fractional factorial design consisting of 18 choice sets was identified [25]. To minimize the number of dominating and non-causal alternatives, the initially identified efficient design was subjected to the manual swapping procedure suggested by [15]. The sample was split into three groups, so that each respondent only had to answer six choice sets. 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 agricultural areas) and two experimentally designed improvement alternatives with an associated price. Respondents were instructed to choose which location (alternative) they would prefer in each of a series of six different choice sets. Following each choice set, they were asked to state how certain they were of the choice just made. An example of a choice set is displayed in appendix A. The data consists of 595 responses resulting in a total of 3,570 choice observations⁴. In a total of 66.1 percent of the choices, respondents state that they are certain or very certain. Looking at demographic characteristics, the gender distribution in the obtained sample of respondents is found to be representative of the target population which is Danes in general. However, age-groups below 25 and above 55, groups with relatively low income, and groups with short educations are all slightly under-represented in the sample.

⁴ For a more thorough description of the survey and a full version of the used questionnaire, the reader is referred to [29].

3.2 The National Park survey

Denmark is about to establish its first national parks following a long political process and ongoing public debate. As part of the participatory process surrounding this, a valuation 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 added as a price variable to trade against. The attributes are shown in table 2. Each respondent was only presented with four of the seven locations, allocated by a cyclic design of four groups. 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 blocked 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 data consists of 636 responses resulting in 4,866 choice observations⁵. In a total of 74.5 percent of the choices, respondents state that they are certain or very certain. Compared to the target population (Danes), the sample is a little under-represented in age-groups below 35 and above 65; and also in short educations; whereas both the lowest and the highest income groups are under-represented. There is no significant difference on gender.

⁵ For a full description of the data and analysis, see [16,17]

>>>TABLE 2 ABOUT HERE<<<

4. Results and analyses

4.1 The estimated mixed logit error component models

The Motorway survey

Table 3 displays the results obtained from the random parameter error component logit model. All attribute parameters, except income tax, are assumed to be normally distributed in the population. All parameter estimates are significant except for the coefficient and standard deviation of the alternative specific constant (ASC). The significant error component, σ_{12} , indicates that utilities across the two experimentally designed alternatives are correlated. Furthermore, this result suggests a difference between the perception of the status quo alternative and that of the two experimentally designed alternatives. Even though the insignificant ASC would reject a status quo effect, the significant error component supports the presence of a status quo effect, but in the stochastic rather than the deterministic component of utility. As expected, the nature attributes' mean estimates are all of a negative sign indicating that on average respondents experience a decrease in utility when one km of motorway is placed through the specific types of nature. Likewise, the income tax parameter estimate has a negative sign as would be expected. The estimated standard deviations of the random parameters are highly significant, revealing a considerable degree of heterogeneity in the respondents' preferences for the three types of nature attributes. The adjusted pseudo-R² value as well as the likelihood ratio test reveals that the model fits the data well [10,25].

>>>TABLE 3 ABOUT HERE<<<

The national parks

Table 4 displays the results for the national parks data as obtained from the random parameter error component logit model. Like in the previous estimation all parameters except the income tax are assumed to be normally distributed around their mean. Most of the parameter estimates are significant at the 0.05 level or less except the mean location dummy for Nordsjælland and the mean parameter for increased amount of walking and biking paths. The error component, σ_{12} , is significant, indicating correlation between the two policy alternatives. The dummy for location Læsø is excluded and is confounded with the ASC. Notice that the negative ASC can be interpreted as a WTP for a benchmark national park and thus the WTP for each of the other location attributes as add-ons hereto [16]. The location of a national park in Lille Vildmose offers the highest WTP among locations although none of the location attributes are significantly different from each other. As expected the generic attributes all have a positive sign and the income tax parameter has the expected negative sign. The estimated standard deviations in the population of the random parameters show a higher degree of heterogeneity in respondents' preferences for the location of a national park compared to the generic attributes. This indicates a larger agreement in the population concerning what to put in a national park, as compared to where to put the national park. The test statistics indicate that the model fits the data well.

>>>TABLE 4 ABOUT HERE<<<

4.2 Modelling certainty in choice and testing the Wang hypothesis

Using the estimated utilities from tables 3 and 4, we apply equation (3) to estimate the expected utility difference for each individual and for each choice set. Subsequently, this utility difference

enters the probit model in equation (4) as an explanatory variable. Note that this procedure implies that the utility difference is measured with the error implied by estimation. We ignore this inaccuracy in the following as we are not concerned with the actual parameter estimates or prediction, but only in the potential of the variable for affecting the likelihood of individuals stating that they are certain or very certain of their choice. Table 5 displays the results obtained from estimating the probit model of possible determinants of the likelihood that respondents report themselves as being certain or very certain of their choice.

>>>TABLE 5 ABOUT HERE<<<

For both surveys the goodness of fit indicates that the probit model provides a reasonably good description of the dataset. Secondly, both models reveal that the utility difference variable constructed in the first step of the analysis does indeed have a highly significant impact on respondents' self-reported certainty level; in both cases a t-value of around 8 is observed. The positive sign of the utility difference variable is as expected, i.e. the larger the utility difference as defined in (3), the higher the probability that the respondents state that they are certain or very certain of their choice – and vice versa. We also note that in the Motorway survey other variables also have a significant bearing on the likelihood that the respondent feels certain about their choice (income, gender and notably, choice set number), whereas only income level is significant in the National Park survey. The random effects parameter, σ , is also significant indicating that the stated certainty is correlated across choice sets within the individual respondent's choices.

5. Discussion

5.1 Testing the adapted Wang-hypothesis: Tough choices cause uncertainty

Among reasons for stated uncertainty in Stated Preference studies, we have focused on the hypothesis formulated by Wang [35] and also implied in the work of [22]. This hypothesis maintains the assumption that respondents are rational, truth-telling, non-strategic and honest, but also assumes that respondents may for different reasons assess the value of the environmental change with some degree of uncertainty. As a consequence they will feel uncertain about their stated answers, when the choices are tough. Tough in the sense that there is a small difference between the utility of the money they are asked to give up, and the good offered, or a small utility difference between the two best alternatives in a choice set.

This hypothesis is particularly interesting for three reasons: First of all it is testable as we have shown here. Secondly, if the hypothesis is well-founded and the researcher has an *a priori* expectation about the relative preferences for the environmental goods in question, it is possible to reduce the number of highly uncertain responses by omitting the choice sets or questions which are potentially tough. Finally, it may be possible to use the utility difference as a basis for capturing variance heterogeneity across choice sets *ex post* – in turn improving model performance and inference.

In this paper, we focus on the first issue of interest and develop an approach to test the hypothesis as this is the prerequisite for the potential gains from the other reasons for taking an interest in Wang's hypothesis. While [24,31,35] find indications that support this hypothesis, a formal test has to the authors' knowledge not been made so far. The attribute-based approach of Choice Experiments caused us to reformulate the original hypothesis of Wang [35] slightly and introduce the utility

difference as defined in (3). This in turn allows us to undertake a considerably more explicit and subtle test of the Wang hypothesis than other analyses so far presented in the literature.

Using data from two independent Choice Experiment surveys we find that the utility difference has a clear impact on the probability of respondents reporting that they are certain or very certain of their choice – *ex post*. The larger the utility difference between the alternative chosen and the best of the remaining alternatives, the more confident the respondent is likely to be in their choice. The intuition is straightforward and indeed the pattern found is also highly significant. This is strong evidence for the hypothesis of Wang [35], which has been a more or less implicit assumption in much work on respondent uncertainty. The confirmation of the Wang hypothesis also allows us to reject the hypothesis that respondents use the certainty question to show consistency – rather there is a more subtle pattern in how the stated certainty varies with choice sets. We note that we have used an estimate of the respondent and choice set specific utility difference (in (3)) for this analysis, but we have not explicitly taken into account that this is of course a variable measured with error. Nevertheless, the results are so clear that this approximation is unlikely to affect the conclusions drawn.

As pointed out by [22], preference uncertainty in choices or answers will at the very least imply a risk of biased and invalid inference in Stated Preference studies. One approach to cope with this is the explicit modelling of this uncertainty. An approach pursued in CV studies by several authors, e.g. [2,22,24], but so far not in Choice Experiments. The results of this study points to the utility difference in choice sets as a key vehicle to model variance effects of uncertainty in choice. A second approach to deal with potential preference uncertainty is to avoid presenting respondents with choices that are likely to be very tough, i.e. where the utility difference between the best

alternatives is likely to be small. Thus, one could formulate the recommendation that some degree of utility imbalance should exist in a significant part of the choice sets to reduce potential problems arising from respondent uncertainty. Note that such an approach will in fact counter the recommendation on obtaining utility balance in choice sets set forth by [15,20]. A high degree of utility balance in choice sets will indeed increase statistical efficiency of the sample based estimates, *provided* that respondents have a high degree of preference *certainty* and can in fact make the tough choices. If this is not the case, then a high degree of utility balance will reduce the performance and inference quality. The results of this study suggest the latter to be the more likely, even though it has to be balanced against the possibility of lexicographic choices if the difference is too large (c.f. [33]).

5.2 Familiarity, learning effects and respondent certainty

It is a well-documented observation that as respondents evaluate more choice sets, initial instability of preferences may decrease [3,6,8,13]. It is straightforward to formulate the hypothesis that such an effect should cause respondent certainty to increase with the choice set number. To test this hypothesis, we also included a choice set number variable in the estimated models shown in table 5. The results are somewhat mixed. In the Motorway survey, we find that the probability of a respondent stating that they are certain or very certain increases significantly with increasing choice set number. In the National Park survey, the parameter for choice set number has the expected positive sign, but it is not significantly different from zero at any conventional significance criteria. While these results show that at least for the Motorway survey, it seems likely that some degree of learning effect is taking place, the difference across the surveys suggest that learning is not a given thing – at least within the limits of the surveys here. One reason could be differences in respondents' familiarity or prior information about the good. List [23] finds increased experience

with a good to result in increased preference consistency; a result confirmed by [6]. This may help to explain why the Motorway questionnaire was more generic than the National park questionnaire. And indeed we find that the respondents in the former were more uncertain than in the latter, indicating that the case initially is more demanding.

It should however be noticed that none of the surveys were utterly imaginary. In fact, at the time of the Motorway survey, it was clearly related to a public debate on the planning of a Motorway through different nature types around the city of Silkeborg, even though that case was not exactly mentioned in the questionnaire. Similarly, the content and possible locations of one or more new national parks in Denmark had been much debated in the media before the launch of the valuation study, lending credibility to the scenarios presented as also documented by [16]. In fact, five parks have been or are currently being established.

A last potential source for the difference in stated certainty could be the survey mode – the national park data was collected through a postal questionnaire and thus respondents could go back and forth between the choice sets providing themselves with an overview before starting the response effort, whereas the Motorway data was sampled from a panel and collected through the internet with a barrier to go back. However, this is not likely to be a major reason here as it is contradicted by [28] who finds that internet panel sample respondents generally state higher certainty in choice than mail sample respondent do, and the study in [28] is based on the same questionnaire as the Motorway survey results presented here.

5.3 Respondent dependent uncertainty

Turning to the variables reflecting characteristics of the respondents, we find in the Motorway data set, that male respondents have a higher probability of reporting a 'certain' choice than female respondents. The same tendency is present in the National park data set, but not significant.

A priori, we would expect that being knowledgeable would reduce preference uncertainty and hence increase certainty for all utility differences, but prior knowledge as well as educational level are found to have no significant effect on the probability of respondents stating that they are certain about their choice. In fact, all parameters even have a somewhat surprising negative sign. Thus, being knowledgeable – either of the good or in general – has no direct effect on stated certainty in the two surveys once the effect of utility difference has been accounted for.

In both data sets there is a significant effect of respondents' household income, with the causal effect being that higher income levels increase the probability of respondents reporting a 'certain' choice. As all respondents are faced with the same payment intervals, it could be an indication that especially poorer income groups use the certainty question to scale down WTP to something more appropriate for their income level. It could also reflect that as WTP increases with income, the more wealthy respondents are less often uncertain as to the choice between the status quo and the best alternative.

5.4 A possible caveat and future research

In the present study we have not investigated the difference in specific attribute levels on stated uncertainty. We therefore assume that respondents assess the value of each attribute with equal certainty. This is reasonable so far as we keep the main assumption when using choice experiments:

That there is unlimited (at least within the considered range) substitutability between attributes. When these assumptions do not hold it may lead to discontinuous preferences and thus lexicographic ordering [7,14,33]. It would be interesting to analyse the impact of discontinuous preferences in relation to stated uncertainty, but this is beyond the scope of the present paper.

6. Conclusion

It is widely recognised that respondents in Stated Preference surveys may be uncertain about their own preferences, especially in non-market good cases. Such preference uncertainty may lead to uncertain responses. In a CV context, Wang [35] formulated the hypothesis that respondents may for different reasons assess the value of the environmental change with some degree of uncertainty. As a consequence they will feel uncertain about their answers or choices when the questions posed are tough in the sense that the requested WTP is close to their true but uncertainly assessed.

The attribute-based approach of Choice Experiments caused us to slightly reformulate the original hypothesis of Wang [35] and introduce the utility difference defined as the difference in utility between the alternative chosen and the best of the remaining alternatives in the choice set. While [24,31,35] find indications that support Wang's hypothesis, a formal test has to the authors' knowledge not been made so far, and certainly not in a Choice Experiment context.

Using data from two independent Choice Experiment surveys, we find that the utility difference indeed has a clear impact on the probability of respondents reporting that they are certain or very certain of their choice. The larger the utility difference between the alternative chosen and the best of the remaining alternatives, the more confident the respondent is likely to be in their choice. The intuition is straightforward and indeed the pattern found is also very strong. This evidence supports

Wang's hypothesis which has been a more or less implicit assumption in much work on respondent uncertainty.

The result provides a potential for further improvement of Choice Experiments when valuing non-market goods. One avenue ahead is the explicit modelling of respondent uncertainty, and the results of this study points to the utility difference in choice sets as a key vehicle to model variance heterogeneity effects of uncertainty across choice sets. Another, perhaps more crude avenue ahead is to avoid presenting respondents with choice sets that are likely to be very tough, i.e. where the utility difference is small. One could formulate the recommendation that some degree of utility imbalance should characterise a significant part of the choice sets to reduce potential problems arising from respondent uncertainty. This would be at odds with the recommendation of obtaining utility balance in choice sets as argued by [15,20]. A high degree of utility balance in choice sets will indeed increase statistical efficiency of the sample based estimates, *provided* that respondents have a high degree of preference *certainty* and can in fact make the tough choices. If this is not the case, then a high degree of utility balance will reduce the performance and inference quality. The results of this study suggest the latter to be the more likely.

References

- [1] W. Adamowicz, P. Boxall, M. Williams, J. Louviere, Stated Preference Approaches for Measuring Passive Use Values: Choice Experiments and Contingent Valuation, *Am. J. Agr. Econ.* 80 (1998) 64-75.
- [2] A. Alberini, K. Boyle, M. Welsh, Analysis of contingent valuation data with multiple bids and response options allowing respondents to express uncertainty, *J. Environ. Econ. Manag.* 45 (2003) 40-62.
- [3] I. Bateman, D. Burgess, W. G. Hutchinson, D. I Matthews, Learning design contingent valuation (LDCV): NOAA guidelines, preference learning and coherent arbitrariness, *J. Environ. Econ. Manag.* 55 (2008) 127-141
- [4] R.P. Berrens, H. Jenkins-Smith, A. Bohara, C.L. Silva, Further Investigation of Voluntary Contribution Contingent Valuation: Fair Share, Time of Contribution, and Respondent Uncertainty, *J. Environ. Econ. Manag.* 44 (2006) 144-168.
- [5] J. Braga, C. Starmer, Preference Anomalies, Preference Elicitation and the Discovered Preference Hypothesis, *Environ. Res. Econ.* 32 (2005) 55-89.
- [6] T.C. Brown, D. Kingsley, G. L. Peterson, N. E. Flores, A. Clarke, A. Birjulin, Reliability of individual valuations of public and private goods: Choice consistency, response time, and preference refinement, *J. Pub. Econ.* 92 (2008) 1595-1606.
- [7] D. Campbell, Identification and analysis of discontinuous preferences in discrete choice experiments, Paper presented at EAERE 16th Annual Conference, Gothenburg, Sweden, June 25-28 2008, 14 pp.
- [8] F. Carlsson, P. Martinsson, Do Hypothetical and Actual Marginal Willingness to Pay Differ in Choice Experiments, *J. Environ. Econ. Manag.* 41 (2001) 179-192.

- [9] A. Champ, R. C. Bishop, T. C. Brown, D. W. McCollum, Using Donation Mechanisms to Value Nonuse Benefits from Public Goods, *J. Environ. Econ. Manag.* 33 (1997) 151-162.
- [10] T. Domencich, D. McFadden, *Urban Travel Demand: A Behavioural Analysis*, North-Holland Publishing Company, (1975), Amsterdam.
- [11] E. Flachaire, G. Hollard, Starting point bias and respondent uncertainty in dichotomous choice contingent valuation surveys, *Res. Energy Econ.* 29 (2007) 183-194.
- [12] W.M. Hanemann, Welfare evaluations in contingent valuation experiments with discrete responses, *Am. J. Agr. Econ.* 66 (1984) 332-341.
- [13] N. Hanley, J. F. Shogren, Is Cost-Benefit Analysis Anomaly-Proof? *Environ. Res. Econ.* 32 (2005) 13-24.
- [14] D.A. Hensher, Joint estimation of process and outcome in choice experiments and implications for willingness to pay, *J. Transport Econ. Pol.* 42 (2008), 297-322
- [15] J. Huber, K. Zwerina, The Importance of Utility Balance in Efficient Choice Designs, *J. Market. Res.* 33 (1996) 307-317
- [16] J.B. Jacobsen, B. J. Thorsen, Where to put a national park and what to put in it? An a priori study of the willingness-to-pay for national parks. Working paper, Forest & Landscape, University of Copenhagen (2008), Frederiksberg, 23 pp.
- [17] J.B. Jacobsen, B.J. Thorsen, J.H. Boiesen, S. Anthon, J. Tranberg, Værdisætning af syv mulige nationalparker i Danmark [Valuation of seven potential national parks in Denmark]. Workingpaper 28, Forest & Landscape, University of Copenhagen, (2006) Frederiksberg. 63 pp.
- [18] B.S. Jorgensen, G.J. Syme, B. E. Nancarrow, The role of uncertainty in the relationship between fairness evaluations and willingness to pay, *Ecol. Econom.*, 56 (2006) 104-124
- [19] D. Kahnemann, J.L Knetsch, Valuing Public Goods: The Purchase of Moral Satisfaction, *J. Environ. Econ. Manag.* 22 (1992) 57-70.

- [20] W. Kuhfeld, Marketing Research Methods in SAS. Experimental Design, Choice, Conjoint and Graphical Techniques, SAS Institute Inc., (2004), online:
<URL:<http://support.sas.com/techsup/technote/ts694.pdf>>.
- [21] J. Ladenburg, S. B. Olsen, Gender-Specific Starting Point Bias in Choice Experiments: Evidence from an Empirical Study, *J. Environ. Econ. Manag.* 56 (2008) 275-285.
- [22] C. Li, L. Mattsson, Discrete Choice under Preference Uncertainty: An Improved Structural Model for Contingent Valuation, *J. Environ. Econ. Manag.* 28 (1995) 256-269.
- [23] J. List, Does Market Experience Eliminate Market Anomalies? *Quart. J. Econ.* 118 (2003) 41-71.
- [24] J. Loomis, E. Ekstrand, Alternative approaches for incorporating respondent uncertainty when estimating willingness to pay: the case of the Mexican spotted owl, *Ecol. Econ.* 27 (1998) 29-41.
- [25] J. Louviere, D. A. Hensher, J. Swait, *Stated Choice Methods. Analysis and Applications* University Press (2000) Cambridge UK.
- [26] S. Luchini, V. Watson, Does respondent uncertainty explain framing effects in double bounded contingent valuation? GREQAM, (2008), Working Paper no. 2008-6.
- [27] D. McFadden, Conditional logit analysis of qualitative choice behaviour. In: P. Zarembka *Frontiers in Econometrics.* Academic Press Inc. New York (1974) 105-142.
- [28] S. Olsen, Choosing between internet and mail survey modes for choice experiment surveys considering non-market goods – A word of caution. *Environ. Res. Econ.* (forthcoming 2009).
- [29] S. Olsen, J. Ladenburg, M. L. Petersen, U. Lopdrup, A. S. Hansen, A. Dubgaard, *Motorways versus Nature - A Welfare Economic Valuation of Impacts*, Report from FOI and IMV, (2005), Copenhagen.
- [30] S. Ready, J. C. Whitehead, G.C. Blomquist, Contingent Valuation When Respondents Are Ambivalent, *J. Environ. Econ. Manag.* 29 (1995) 181-196.

- [31] M. Samnaliev, T. H. Stevens, T. More, A comparison of alternative certainty calibration techniques in contingent valuation, *Ecol. Econ.* 57 (2006) 507-519.
- [32] N. Schwarz, S. Sudman, *Answering Questions. Methodology for Determining Cognitive and Communicative Processes in Survey Research.* Jossey-Bass Publishers, (1996), San Francisco.
- [33] K. Sælensminde, Inconsistent choices in Stated Choice data; Use of the logit scaling approach to handle resulting variance increases, *Transportation* 28 (3) (2001) 269-296
- [34] K. Train, *Discrete Choice Methods with Simulation* Cambridge University Press, (2003), Cambridge.
- [35] H. Wang, Treatment of "Don't-Know" Responses in Contingent Valuation Surveys: A Random Valuation Model, *J. Environ. Econ. Manag.* 32 (1997) 219-232.
- [36] M.P. Welsh, G. L. Poe, Elicitation Effects in Contingent Valuation: Comparisons to a Multiple Bounded Discrete Choice Approach, *J. Environ. Econ. Manag.* 36 (1998) 170-185.

Appendix A – Example of a choice set question for the Motorway data

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 question for the national park

data

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

(Mark one)

(The money have 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 cho ... *Very certain* *Certain* *Uncertain* *Very uncertain* *Don't know*

Table 1

Attribute levels used in the Motorway survey

Attribute (type of nature)	Level (km new motorway through nature area)
Forest	0 km, 5 km, 10 km
Wetland	0 km, 2.5 km, 5 km
Heath/pastoral area	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	(0 DKK), 100 DKK, 200 DKK, 400 DKK, 700 DKK, 1100 DKK, 1600 DKK

^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 80 kilometres. 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: 100 DKK \approx 17.2 \$ as per 01.02.2009

Table 2

Attributes presented 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	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

Table 3

The basic model of step 1 for the Motorway data: The random parameter error component logit model

Attribute		Estimate	Std. err.	t-value		WTP (95%CI)
ASC	- Mean	-0.3967	0.215	-1.85	N.S.	197.78
	- Standard deviation	1.0433	1.3426	0.78	N.S.	(-9.45 - 405.01)
Forest	- Mean	-0.1685	0.0147	-11.44	***	84.03
	- Standard deviation	0.2227	0.0169	13.2	***	(69.82 - 98.24)
Wetland	- Mean	-0.2182	0.0228	-9.58	***	108.79
	- Standard deviation	0.2469	0.0306	8.06	***	(86.27 - 131.31)
Heath	- Mean	-0.0937	0.0195	-4.8	***	46.71
	- Standard deviation	0.1966	0.0303	6.49	***	(27.48 - 65.94)
Additional income tax	- Mean fixed	-0.0020	0.00007	-27.29	***	
σ_{12}		3.0545	0.5273	5.79	***	
# observations		3570				
# respondents		595				
Log Likelihood at convergence		-2754.2				
Adj. McFadden's pseudo-R ²		0.2968				

Note: Simulations are based on 1000 Halton draws. *** indicates significance at the 0.001 level; N.S. non-significance at the 0.05 level. 95% confidence intervals for WTP are estimated using the Krinsky-Robb method with 1000 draws. By questionnaire construction and data coding, the nature type parameter estimates are negative, thus in the calculation of WTP these estimates are multiplied by -1.

Table 4

The basic model of step 1 for the National park data: The random parameter error component logit model

Attribute		Estimate	Std. err.	t-value		WTP (95%CI)
ASC/Læsø	- Mean	-2.1764	0.1927	-11.3	***	-1055.38
	- Standard deviation	0.8421	0.4914	1.714	N.S.	(-1225.99 - -885.77)
Location Møn	- Mean	0.4164	0.1464	2.84	**	201.91
	- Standard deviation	1.3871	0.1847	7.51	***	(66.70 - 337.12)
Location Thy	- Mean	0.3273	0.1491	2.2	*	158.7
	- Standard deviation	1.2276	0.1773	6.92	***	(17.09 - 300.31)
Location Nordsjælland	- Mean	-0.0834	0.1788	-0.47	N.S.	-40.45
	- Standard deviation	1.923	0.191	10.07	***	(-210.03 - 129.13)
Location Mols Bjerge	- Mean	0.4393	0.1649	2.67	**	213.04
	- Standard deviation	1.6654	0.1929	8.63	***	(57.59 - 368.49)
Location Lille Vildmose	- Mean	0.7757	0.1383	5.61	***	376.13
	- Standard deviation	1.3832	0.1695	8.16	***	(242.78 - 509.48)
Location Vadehavet	- Mean	0.6734	0.1513	4.45	***	326.53
	- Standard deviation	1.6322	1.7496	9.33	***	(185.30 - 467.76)
Nature preservation	- Mean	0.1239	0.03414	3.63	***	60.09
	- Standard deviation	0.3321	0.0522	6.36	***	(27.87 - 92.31)
Effort for animal/plants	- Mean	1.1269	0.0808	13.95	***	546.44
	- Standard deviation	0.8459	0.1126	7.51	***	(474.56 - 618.32)
Walking and biking paths	- Mean	0.1469	0.0879	1.67	N.S.	71.22
	- Standard deviation	0.8294	0.1327	6.25	***	(-12.88 - 155.32)
Price	-Mean (fixed)	-0.00206	0.00006	-33.40	***	
σ_{12}		2.6206	0.2582	10.15	***	
# observations		4866				
# respondents		636				
Log Likelihood at convergence		-3736.32				
Adj. McFadden's pseudo-R ²		0.2995				

Note: Simulations are based on 1000 Halton draws. *** indicates significance at the 0.001 level; ** at the 0.01 level; * at the 0.05 level; N.S. non-significance at the 0.05 level. 95% confidence intervals for WTP are estimated using the Krinsky-Robb method with 1000 draws.

Table 5

Determinants of certainty in choices - Random effect binary probit model

Parameter	Motorway survey				National parks survey			
	Estimate	Std. err.	t-value		Estimate	Std. err.	t-value	
Utility difference	0.1716	0.0212	8.11	***	0.1003	0.0126	7.97	***
Choice set number	0.0515	0.0164	3.14	**	0.0089	0.0100	0.88	N.S.
Age group	-0.0224	0.0700	-0.32	N.S.	0.0041	0.0057	0.71	N.S.
Male	0.7276	0.1696	4.29	***	0.2340	0.1440	1.62	N.S.
Educational level	-0.0897	0.1076	-0.83	N.S.	-0.0676	0.0839	-0.81	N.S.
Income level	0.2992	0.1118	2.68	**	0.1462	0.0706	2.07	*
Knowledge of good	-0.0542	0.1947	-0.28	N.S.	-0.1745	0.1571	-1.11	N.S.
Constant	-0.6481	0.4891	-1.33	N.S.	0.8154	0.3566	2.29	N.S.
σ	0.7820	0.0184	42.52	***	0.7191	0.0192	37.39	***
# observations	3570				4866			
LL	-1567.8				-2006.0			
McFadden's pseudo-R ²	0.298				0.269			

*** indicates significance at the 0.001 level; ** at the 0.01 level; * at the 0.05 level; N.S. non-significance at the 0.05 level