Analysis

Embedding effects in choice experiment valuations of environmental preservation projects

Jette Bredahl Jacobsen [a,*], Thomas Hedemark Lundhede [a], Louise Martinsen [b], Berit Hasler [b], Bo Jellesmark Thorsen [a]

[a] Forest & Landscape and Center for Macroecology, Evolution and Climate, University of Copenhagen, Rolighedsvej 23, DK-1958 Frederiksberg C, Denmark
[b] Aarhus University, National Environmental Research Institute, Frederiksborgvej 399, PB 358, DK-4000 Roskilde, Denmark

ARTICLE INFO

ABSTRACT

This paper addresses the question, whether attribute values derived from a CE study of one nature restoration and preservation project depend on the overall scale of nature preservation activities in which it is embedded. A split-sample CE study was undertaken in which a particular nature preservation project was evaluated in three plausible and strictly different embedding contexts. Respondents' attention was drawn to a varying number of forthcoming substitute preservation projects to be implemented prior to the one in question. Results show that while WTP for the project in focus is sensitive to the presentation of substitute projects as such, there does not seem to be any clear and unidirectional relationship between WTP and the number of substitute projects. Furthermore, effects vary across the project's attributes.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

Environmental preservation projects are seldom undertaken one at a time. In Denmark, many such projects have been debated and evaluated in recent years. This includes a large public debate and several studies of the potential costs and benefits of establishing a number of national parks (Wilhjelmudvalget, 2001; Larsen et al., 2008; Jacobsen and Thorsen, 2010). For several projects, key attributes of each environmental change have been evaluated using Choice Experiments (CE). For example, the benefits of increasing the scope of the individual project have been investigated, with attributes like recreational access or biodiversity preservation (Jacobsen et al., 2008; Jacobsen and Thorsen, 2010).

However, as each project or environmental improvement evaluated is likely to be only one of several such efforts, an obvious question asked by policy-makers is: How do the environmental values derived from one project depend on the overall scale of planned (Danish) nature preservation activities and the extent of existing nature areas? This concern relates to the issue of embedding effects, where the valuation of a good depends on the context in which it is embedded, and notably on the presence of substitutes in this context. Thus, the present study builds on the literature on embedding in Contingent Valuation (CV) (Kahnemann and Knetsch, 1992; Carson and Mitchell, 1995; Randall and Hoehn, 1996), on the substitution effect (Arrow

et al., 1993; Carson and Mitchell, 1995; Randall and Hoehn, 1996; Carson et al., 2001) and on the sub-additivity property (Hanemann, 1994; Bateman et al., 1997).

However, the approach taken here differs from these earlier embedding effect analyses in applying the CE technique and, more importantly, in focusing on the value of one specific project embedded in different contexts, rather than on the value of different combinations of goods. In the recent decade, the CE method has become popular due to its flexibility in terms of allowing for simultaneous valuation of several attributes. Only a few studies have explored the role of substitution in a CE context. Rolfe et al. (2002) and Jacobsen and Thorsen (2010) analysed substitution effects by including the substitutes in the alternatives of the choice sets. This approach allows respondents to consider substitutes, but it comes at a cost in terms of increased complexity of the experimental design and the information to be conveyed to respondents. The challenge grows as the number of potential substitutes increase.

This study tested embedding effects caused by the presence of substitute projects. It was done by varying the status quo level in a CE context valuing a proposed environmental and cultural heritage restoration and preservation project. The different degrees of substitution were implemented by asking different respondents to assume a different number of other, already planned, nature preservation projects being performed first. The CE approach allowed us to evaluate the presence of a substitution effect at attribute level for the project, in turn providing additional information of relevance to environmental policy and management. Three versions of the survey were developed and used on three samples of respondents. The

* Corresponding author. Tel.: +45 35331746.
E-mail addresses: jbj@life.ku.dk (J.B. Jacobsen), thlu@life.ku.dk (T.H. Lundhede), lom@dmu.dk (L. Martinsen), bh@dmu.dk (B. Hasler), bjt@life.ku.dk (B.J. Thorsen).

treatment in the three surveys varied in only one aspect: a text reminder informing about a number of forthcoming potential substitute projects (none, two or seven) to be implemented prior to the project in question. All other factors were held constant in the choice sets for the three samples, inclusive of the experimental design.

The substitute projects stated in the text reminder referred to a contemporary public debate on the establishment of a number of national parks in Denmark (Jacobsen and Thorsen, 2010). This provided high credibility to the scenarios. The text reminder was shown before each and every choice set that the respondent had to consider. This corresponds with the common recommendation that instruction reminders be "…forcefully and directly prior to the main valuation question …" (Arrow et al., 1993, p. 33), as is now commonplace in CE studies (e.g. Verbic and Slabe-Erker, 2009; Jacobsen and Thorsen, 2010). Compared to the reminders applied in most of the CV literature, it is a new feature to remind respondents of substitutes that are not already in place, but represent other environmental improvement projects that *will be* undertaken. Following the approach of Diamond (1996), respondents were asked to assume that they had already received a specified number of substitutes and paid an – unspecified – amount for it through income taxes. As already pointed out by Carson and Mitchell (1995) and Carson et al. (2001), respondents may find such a set-up quite complex. Thus, to evaluate their ability to grasp the overall context in which the valuation project is embedded, detailed follow-up questions were included.

The results clearly showed embedding effects on willingness to pay (WTP), which proved sensitive to the presentation of substitute projects as such, but nevertheless did not depend convincingly on the number of projects. Furthermore, effects varied across the attributes of the evaluated project.

The remainder of the paper is organised as follows. In Section 2, the relevant literature is reviewed and hypotheses presented. Section 3 describes the method, and Section 4 presents the environmental project in focus, questionnaire design, and data collection method. Results are presented in Section 5, and in Section 6 we discuss various implications. We provide a brief conclusion in Section 7.

2. Literature and Hypotheses

2.1. Literature on Embedding Effects and Sensitivity to Substitutes in Environmental Valuation

Though subject to substantive criticism by, e.g. Smith (1992), Carson and Mitchell (1995), and Randall and Hoehn (1996), the study of Kahnemann and Knetsch (1992) remains influential in putting focus on the issue of embedding effects. This is the phenomenon that the WTP people state for a composite good may differ from the aggregate WTP obtained if sub-sets of the good are valued independently. Several arguments have been given that such effects are to be expected and rational. Carson and Mitchell (1995) and Randall and Hoehn (1996) argued that economic theory would predict the value of a particular good to be progressively smaller the later in a sequence of goods it is assessed, *if* these goods are substitutes. Rollins and Lyke (1998) argued that people may have sharply declining marginal utility of an environmental good after a certain quantity of the good has been provided, causing a saturation effect on WTP for larger bundles. Finally, embedding effects may be due to sub-additivity or attribute-splitting (Hanemann, 1994; Bateman et al., 1997), namely when respondents value the sum of marginal changes in several goods higher than the same change in a larger good. This can be explained by diminishing marginal rate of substitution (Hanemann, 1994) and is likely to occur in sequential settings. Powe and Bateman (2003), and in particular Powe and Bateman (2004), argued that the political credibility of large scale projects may be a reason for variation in value with varying sequence.

Numerous empirical analyses exist of various forms of embedding effect. Powe and Bateman (2003) distinguished between sequencing effects in inclusive and exclusive lists, where in the former, goods are presented as additions (or subtractions) to the list, and in the latter as alternatives to elements within the list. They argued that in an inclusive list, which is the most frequently studied, sequencing is to be expected for public goods as their amount varies throughout the sequence. For exclusive lists, sequencing effects are not to be expected if the goods are independent. Again, this independency can be questioned for public goods and the authors do indeed find sequencing effects. Veisten et al. (2004) also analysed sequencing effect empirically and found evidence of it in a CV study of endangered species. However, if the alternative goods are not substitutes, the good in question may be perceived increasingly scarce in a relative sense, and the embedding effect may even be to raise the WTP. Hailu et al. (2000) provided an example of this where for some goods they found a complementary effect when valued in a multi-programme context. McDaniels et al. (2003) find that embedding problems are reduced when respondents are allowed to discuss and structure their preferences in small groups, and they consider this an improvement of the quality of preference judgement.

Other studies focus more on identifying contexts where substitution causes embedding effects. Whitehead and Blomquist (1995) found that substitution effects occurred when reminders or budget constraints were introduced (WTP was reduced), a result that Cummings et al. (1994), Loomis et al. (1994), and Kotchen and Reiling (1999) did not find. Nevertheless, Kotchen and Reiling (1999) found that information improved the efficiency of the WTP estimates. Sælensminde (2003) found more mixed results in a valuation study of non-marketed goods related to transport. He found that a reminder of substitutes and budget constraints significantly reduced the value of a two-goods package, but had no influence on the valuation of a three-goods package. Neill (1995) found no substitution effects when reminders were introduced, only when respondents were directly valuing the substitutes.

Besides the sequencing or presence of substitutes also the nature of the good may have an impact on WTP Many environmental valuation studies focus on public goods, but even for goods of a private character the setting may determine the sensitivity to sequencing and substitutes. Clark and Friesen (2008) analysed order effects for very familiar goods and found that goods with a large scope were affected less by sequencing than goods with smaller scope. This might indicate that substitutes are more easily available for smaller than for larger goods. Clark and Friesen (2008) also found that order effects were not present for donations but for private consumption, even when the good valued was the same. Dupont (2003) analysed embedding and found ordering effects to be larger for passive than for active users in a CV of recreational activities.

The literature reviewed above is focused on testing embedding with CV applications at the project level, but few authors have explored embedding in a CE context. Substitution issues between projects are rarely considered, but examples do exist. Rolfe et al. (2002) found substitution effects in a CE experiment, where they included the substitute among the alternatives in the choice sets. Jacobsen and Thorsen (2010) found varying degree of substitution effects when evaluating trade-offs between different sites for national parks. External (to the choice set) substitution analyses are few, but e.g. Loureiro and Ojea (2008) found no substitution effect in a CE where one split sample, valuing the preservation of a specific bird population, was told that the bird was common in other countries.

Closely linked and often confounded with embedding is the issue of scope sensitivity. It refers to the extent to which WTP is sensitive to changes in the quantity or quality of the environmental good subject to valuation (Hanemann, 1994; Veisten et al., 2004). Like for embedding, there may be rational reasons for lack of sensitivity to scope. Amiran and Hagen, (2010) argue that depending on the shape

of the utility function, sensitivity to scope may be more or less pronounced. If utility functions are directionally bounded, the degree of sensitivity can be arbitrarily small. The CE technique has by construction often an implicit internal scope test, if the attributes have more than two levels (e.g. Campbell, 2007; Jacobsen et al., 2008). Several studies addressing sensitivity to scope look at characteristics and presentation of the good for explanations (e.g. Czajkowski and Hanley, 2009; Ojea and Loureiro, 2009). The issue of scope is sometimes difficult to separate entirely from the issue of embedding (Czajkowski and Hanley, 2009).

The approach taken in this study to test for embedding effects differed from the perhaps more common approaches in the CV literature, as we focused on the valuation of several attributes of a specific nature preservation project, when this is presented as a stand-alone project to be evaluated, or as one out of a number of forthcoming nature preservation projects.

2.2. Hypothesis Formulation

The test designed for this study was not a classic embedding test. Simply stated, an embedding effect is present when the valuation of a composite good with parts A and B does not show perfect additivity, $WTP(A+B) \neq WTP(A) + WTP(B)$. The typical embedding effect is that the WTP for either part of the good is larger than when evaluated embedded in the overall good. The test designed here also has some relations to the scope literature, but is not a scope test. A typical scope test is of the form $WTP(A+B) > WTP(A)$. Of course, as several attributes of project A are valued through a CE, there is an internal scope test of the attributes of A. The scope effects at the level of a series of nature projects are not directly evaluated.

Instead the test was formalised as differences in WTP for project A, when evaluated as an addition to a series of forthcoming substitute projects, $WTP(A) > WTP(A|B) > WTP(A|B,C)$. More specifically, the test performed was:

$$WTP_{s=0,x^m} > WTP_{s=2,x^m} > WTP_{s=7,x^m} \qquad (1)$$

where s refers to the number of substitute projects in the sample (0, 2 or 7) highlighted prior to each choice set, and x^m is the specific relevant attribute among the M attributes characterising the good, x. Assuming that the project is in fact a substitute with respect to attribute m, we expect the WTP to decrease the more substitute projects are available. The null hypothesis of no effect is:

$$WTP_{s=0,x^m} = WTP_{s=2,x^m} = WTP_{s=7,x^m} \qquad (2)$$

In a split-sample CE context it is possible to test these hypotheses at attribute level, using standard econometric approaches.

3. Method

In choice models, a scale parameter that is inversely proportional to the standard deviation of the error in each sample is embedded in the estimated coefficients. This scale parameter may vary between our splits and invalidate tests for identical preferences across splits, if pooled. We used Biogeme (Bierlaire, 2003) to test for variations in the scale parameter, and finding no significant variation we are able to proceed specifying the model for the pooled data sets.

A model in WTP-space was specified (Train and Weeks, 2005; Scarpa et al., 2008; Thiene and Scarpa, 2009).[1] A set of variables was defined, capturing the value of project Store Åmose in the situation with no substitutes. Subsequently, two sets of additional difference

variables were defined, relating to the situation with two or seven forthcoming substitute projects. Thus, the utility function for individual i, for attribute m in project alternative j for the k'th choice situation can be specified as:

$$U_{ijk} = -\lambda_i p_{ijk} + c_i x_{ijk} + a_2 b_{2i} x_{ijk} + a_7 b_{7i} x_{ijk} + \varepsilon_{ijk} \qquad (3)$$

where $a_2 = 1$ when at least two national parks are presented in the k'th choice situation and zero otherwise, and $a_7 = 1$ when seven national parks are presented and zero otherwise. p_{ijk} is the price, λ_i, c_i, b_{2i}, b_{7i}[2] are parameters to be estimated, and ε_{ijk} is the remaining i.i.d. extreme value error term. Under the null hypothesis in Eq. (2) $b_{2i} = b_{7i} = 0$.

Letting V_{ijk} represent the deterministic part of U_{ijk}, then the probability of individual i choosing a sequence of choices, $y_i = (y_{i1} \ldots y_i K_i)$, conditional on β_i (being either c_i, b_{2i} or b_{7i}), is given by (Revelt and Train, 1998; Thiene and Scarpa, 2009):

$$P(y_i | \beta_i) = \prod_{k=1}^{k=K} \frac{exp(V_{ink}(\beta_i))}{\sum_j exp\left(V_{ijk}(\beta_i)\right)}, \forall n \neq j \qquad (4)$$

The integral of $P(y_i | \beta_i)$ over all β_i weighted by the density is the unconditional probability or the mixed logit choice probabilities:

$$P_i(y_i) = \int P(y_i | \beta_i) f(\beta_i | \theta) d\beta_i \qquad (5)$$

where f is the distributional function for parameter β_i and θ is the parameters of the distribution. The model was estimated using Nlogit version 4.3 and some sub-analyses, using Biogeme (Bierlaire, 2003). In Nlogit, the procedure for generalised mixed models was used to allow for scale heterogeneity. The utility function can be written as (Fiebig et al., 2010):

$$U_{ijk} = (\beta \sigma_i) x_{ijk} + \varepsilon_{ijk} \qquad (6)$$

where

$$\sigma_i = exp(\overline{\sigma} + \tau \varepsilon_i) \qquad (7)$$

Thus, the model included a scale parameter, σ, and allowed for scale heterogeneity (expressed through τ).

Follow-up questions regarding respondents' motives for choices, in particular if they chose none of the alternatives, were included as is standard in the literature (e.g. Giraud et al., 1999; Jorgensen et al., 2001; Foster and Mourato, 2003; Bateman et al., 2004; Heberlein et al., 2005). Furthermore, we included follow-up questions asking respondents to state their perception of the embedding context in which they just answered the choice questions.

4. Survey Design and Data Collection

4.1. The Choice Set and the Store Åmose Project in Focus

The area known as Store Åmose is situated in Western Zealand in Denmark and was originally one of the largest wet moorland areas in the eastern part of the country. Following intensive draining, much of the area has been converted into agricultural uses. The survey respondents were informed the following:

- The proposed project could involve the abandonment of agriculture to restore a larger coherent nature area, and additional measures could be undertaken to enhance biological diversity, to

---

[1] We are grateful to an anonymous reviewer for several constructive suggestions for modelling and testing our hypothesis.

[2] Notice that these parameters are defined as the ratio of the parameter to the price coefficient, μ_i, in preference space, $\lambda_i = \alpha_i / \mu_i$ and $c_i = \beta_i / \mu_i$, and correspondingly for b_{2i} and b_{7i}, cf. Thiene and Scarpa (2009).

re-establish meadows and wetlands, and to improve recreational opportunities.

- In addition, the area is rich in archaeological artefacts buried in the topsoil and deeper.
- The project could preserve this cultural heritage through restoring the original hydrological and top soil conditions in selected parts, which would halt the current oxidation of the artefacts.
- Note that while some of these measures will be undertaken jointly to some extent, specific and independent measures are needed to secure high provision levels of any of the above goods.

Four attributes reflecting these environmental goods (see Table 1) were used to describe two alternatives and a status quo option in each choice set. The status quo represents the land use and preservation status expected in the case of no project. The status quo situation implies continued loss of artefacts, maintenance of the current low level of biodiversity, and restricted level of accessibility, and thus the lowest level of all attributes. Most Danish nature preservation projects of this size are financed by public funds. Therefore, income tax was used as the payment vehicle. It was highlighted that all taxpayers should expect to contribute, to discourage respondents from free-riding, and respondents were told that the stated amounts would be additional to current tax payments.

Alternatives dominating the status quo (e.g. improvements with zero payment) were included. The combination of alternatives into choice sets was performed, using a design where d-efficiency was sought for a multinomial logit model by implying a modified Fedorov candidate set search algorithm (Kuhfeld, 2004). The input was a fractional factorial design containing 36 alternatives, corresponding to 18 choice sets. The choice sets were grouped into two blocks, such that each respondent was presented with nine choice sets. The resulting design, with zero priors and adding a status quo, had a d-error of 0.107 when evaluated ex-ante as continuous variables with levels corresponding to lists from 0 to 6. When evaluated with actual levels of price and area, the d-error drops to 0.001064. The same

design was used across the three split-samples. Ex-post, the d-error was 0.000344 for a logit model with dummy coding for biodiversity and ancient artefacts, and without effects of including national parks, i.e. without b_{2i} and b_{7i} in Eq. (3) (Scarpa and Rose, 2008).

4.2. Setting the Embedding Context Using the National Park Projects as Potential Substitutes

For some years prior to this study, the Danish public and political decision-makers debated the potential of establishing national parks in Denmark. The process started in 2001, when a committee pointed out the need for larger coherent nature areas (Wilhjelmudvalget, 2001). Seven pilot projects were nominated as possible national park candidates and their potentials assessed in various ways (Larsen et al., 2008; Jacobsen and Thorsen, 2010). Some of them are now established or will be so, and a law has been passed to rule their administration. However, when the survey was performed, it was not yet decided how many national parks would be implemented in the short run. The aim of the national park projects is to create large coherent areas of a natural or semi-natural character to secure nature preservation and improve possibilities for recreation. To a smaller extent the projects also contribute to the preservation of cultural heritage. It is therefore safe to assume that the proposed national park projects can in general be regarded as substitutes for the *Store Åmose* area. This is certainly true in respect of biodiversity and recreation, but less so in respect of the cultural heritage element, which is more prominent in *Store Åmose* compared with the national parks.

All nature projects, and especially the national parks, have been discussed in the media prior to the launch of the survey. According to Jacobsen et al. (2006) more than 70% of the population was aware of the national parks debate, and almost 20% kept themselves informed about it. This widespread awareness was also reflected in the focus group interviews, performed prior to the survey, and revealed that the fairly concise information given in the questionnaire on the national parks was perceived as sufficient.

A split-sample approach was used, where the questionnaire[3] in each of three split samples differed only in the following aspects: In two split samples, respondents were asked to assume that it had recently been politically decided to establish large nature areas in Denmark, including two and seven national parks, respectively. Respondents should assume that these parks were to be established prior to the possible preservation of *Store Åmose* with which the choice sets were solely concerned. The description outlined how these new national parks would provide large coherent nature areas, enhancing the biological diversity and the recreational access.

In one of the introductory warm-up questions, the respondents were asked to point out three of the seven potential national park projects as their most preferred. This was done in order to make them form preferences about such nature projects and awareness of the substitutes. In the samples that included national parks as substitutes, the context of embedding was furthermore accentuated in key places. The price attribute had the following text attached to it (here in the sample with two national parks):

> "For each alternative, the annual tax payment increase to your household due to the nature preservation project is stated. Note that this payment would be in addition to your current tax payment, and **in addition** to any extra tax payment your household would have to make towards the establishment of the two forthcoming national parks."

Table 1
The attributes used in the CE. Description as given to the respondents.

Attribute	Levels
Size of the protected area Biodiversity	Six levels between 230 and 1,750 hectares **Small diversity**: A large number of animals, but distributed among a smaller number of ordinary species. Vulnerable and rare species are threatened by extinction because of dry conditions and cultivation. **Some diversity:** Many animals distributed among a larger number of ordinary species, including small birds. Vulnerable and rare plants are only threatened by extinction in a few places. **Large diversity**: Many ordinary and rare animal species, especially bird species. Rare species are protected against dry conditions and cultivation.
Preservation of the ancient artefacts	**Continued devastation:** Important artefacts, including internationally unique places of sacrifice and of residence, will be demolished within in a number of years. **Reduced devastation:** The speed of devastation is significantly reduced for some of the most important artefacts, but devastation is not brought to an end. **Preservation now and in the future:** The artefacts will be protected within the soil now and in the future, so that they can be excavated and placed on view in the future.
Public access	**Restricted access** by a few tracks and paths. **Extended access** to a larger part of the area by a larger system of paths and tracks
Extra payment in annual tax (household level)	Six levels of: EUR 0, 10, 34, 61, 115, and 203.

---

[3] For a full questionnaire please contact the corresponding author. For review purpose the questionnaire is attached.

The previous questions, where you had to choose between alternatives, concerned payment for

nature restoration. Which of the following statements is correct?

(mark one)

The payment only concerned the restoration of Store Åmose .... ❑

The payment concerned 2 national parks .... ❑

The payment concerned 2 national parks and Store Åmose......................... ❑

The payment concerned nature in general .... ❑

I don't know.. ❑

Fig. 1. Debriefing question concerning the embedded context. The question is related to the split-sample where respondents were informed about two national park projects to be implemented prior to the Store Åmose project. A similar question was given to the split-sample related to seven parks.

At the beginning of the section on choice sets, a reminder of substitutes was added:

*[Please bear in mind that] "you should assume that it **has been** decided to establish two national parks, which will in turn contribute to the preservation of Danish nature. The payment for Store Åmose would therefore be in addition to your - and your household's - **future** payment for two national parks."*

Furthermore, before each choice set a reminder was formulated like this in the cases with substitutes:

"Assume that the two mentioned national parks would be established whatever alternative you choose for the preservation of Store Åmose. Your choice and the extra income tax therefore only concern the next nature preservation project, namely Store Åmose."

A number of debriefing questions were asked to identify protest behaviour, and in particular debriefing questions were designed to follow up on the embedding test, cf. Fig. 1.

4.3. Pre-testing and Data Collection Procedure

Data were collected through an internet-based questionnaire managed by the survey institute GALLUP, using a sample from their internet panel of approximately 35,000 individuals, representative of the population. For each version of the questionnaire, responses from approximately 350 respondents were received. The survey institute invited a number of panel members to participate in surveys, and re-invited until a desired amount and representative mix of responses had been obtained. The resulting response rate was somewhat above 50% for this way of collecting data. The use of an internet-based questionnaire made it possible to prevent respondents from browsing back to check their choice sets when answering the follow-up questions regarding which embedding context they thought they replied to. Prior to the launch of the survey, three focus group interviews were performed to test and improve the questionnaire. Specific attention was paid to understanding of the presentation of substitutes.

5. Results

Respondents were invited to answer the questionnaire in December 2006, and the survey continued until the required number of responses had been returned. The three samples have a slight underrepresentation of respondents in the group above 60 years of age and, thereby, also in the low income group compared with the population in Denmark. However, in relation to the findings of this

study, there is no statistically significant difference in age, gender and income between the three sample splits.

Follow-up questions were used to check motivations for different responses and to eliminate some answers. These included serial non-respondents (von Haefen et al., 2005) - respondents who chose the status quo alternative in all nine choice sets - who explained this with 'I don't want to pay more tax' or 'I don't know why', rather than with answers of the type 'I could not afford the amounts mentioned' (see Table 2). Respondents who stated 'My payment concerned nature in general' were also removed from the sample, because they were considered not to have related to the issue at hand, and their answers are likely to reflect warm glow (Andreoni, 1990).

The proportions of serial non-respondents and status quo choices can be an indication of complexity (Boxall et al., 2009). The proportions of serial non-respondents were not statistically different between the three split-samples, whereas homogeneity in terms of proportions of status quo choices was rejected.

As a debriefing question, the respondents were asked about their perception of what was the good valued in the choice experiment, see Fig. 1. Table 2 shows that in the sub-samples with substitutes, almost 20% of the respondents had an incorrect post-choice perception of the good valued. Furthermore, between 8% and 11% were not able to recall the embedding context and answered 'I don't know'. These groups of respondents were also excluded, and we are left with the 70-75% of the respondents who understood the embedding context correctly. This exclusion of 25-30% of the respondents is considerable, but we estimated a series of preference space models, where the above groups of respondents were excluded in a step-wise manner. The exclusions did not affect the pattern of results and the conclusions of this paper, but only increased the efficiency of estimates.[4]

Table 3 shows the estimation result of a model corresponding to Eq. (3). Limitations of how many variables the software (Nlogit 4.3) could handle for this type of model made some simplifications necessary. The 'Biodiversity' attribute showed an almost linear pattern across levels in initial estimations, and therefore the variable was coded with levels 0, 1, 2. 'Ancient artefacts' showed insignificant differences between the two levels and they were hence merged into one dummy variable. Limitations of the software when handling many variables caused us to exclude also the variable corresponding to parameter b_7 for the 'Area' attribute, as this was not significantly different from zero. Excluding b_7 for 'Biodiversity' instead of 'Area' gave similar results for the other parameters. Alternative specific constants were highly insignificant and are also excluded. Parameter distribution was only specified for c, jointly for all sub-samples, as

---

[4] Based on a simple probit regression we found no significant effect of the exclusion of respondents in relation to age, gender or income, and only a moderate effect of education.

Table 2
Sample size reduction.

	Only Store Åmose (Sample 0)	Store Åmose and 2 additional nature areas (Sample 2)	Store Åmose and 7 additional nature areas (Sample 7)
Initial number of respondents	**360**	**355**	**359**
- Serial non-respondents with protest behaviour (I don't want to pay more tax, I don't know why I chose the status quo in all 9 choice situations)	−16	−11	−18
	4.4%	3.1%	5.1%
- Mental account (My payment concerned nature in general)	−36	−27	−21
	10.0%	7.6%	5.8%
- The payment concerned 2(7) national parks	n/a	−24	−19
		6.8%	5.3%
- The payment concerned 2(7) national parks and Store Åmose	n/a	−43	−49
		12.1%	13.6%
- Don't know (what the payment concerned)	−34	−39	−29
	9.4%	11.0%	8.1%
Reduced sample	**274**	**211**	**223**
Serial non-respondents in total	27	28	33
	7.5%	7.9%	9.2%
Status quo choices in total	672	746	647
	6.9%	7.8%	6.7%

specifying further distributions resulted in unstable models, probably due to multi-collinearity and to limitations of the estimation procedures.

Table 3 shows that all main effect parameters, c_i, are significant and with the expected sign. The parameters with highest WTP are 'Biodiversity' and 'Ancient Artefacts', but also the 'Area' attribute is important, considering that the value estimated is per 100 ha and the largest area increase suggested was 1750 ha. 'Access' to a larger system of roads and paths appears to have a negative value, corresponding to a WTA measure. The large standard deviations show that there is quite some difference in the population's WTP.

When two or more national parks were presented as being established beforehand, we see that the WTP decreases for all attributes and so does the WTA for access. All these parameters are significant. Thus, we can reject the null hypothesis from Eq. (2) of no effect between the first and the second term. Also note that the

Table 3
Parameter estimates for a model in WTP space corresponding to Eq. (3). All values in DKK. The variable for biodiversity is continuously coded with values 0, 1, 2, corresponding to no change, some biodiversity increase, and large biodiversity, respectively. The two levels for ancient artefacts are merged.

Main effects, c_i:	Estimate	t-value
Tax	1.0	Fixed
Biodiversity	162.5	17.5
- Std. Dev.	209.8	22.4
Ancient Artefacts	380.0	23.1
- Std. Dev.	410.7	18.9
Access	−47.0	4.3
- Std. Dev.	81.1	3.7
Area (per 100 ha)	19.7	14.8
- Std. Dev.	23.5	16.5
Adjustments to Main Effects by Substitutes		
+2 or more national parks b_{2i}:		
Biodiversity	−31.3	4.9
Ancient Artefacts	−117.0	10.3
Access	−28.0	4.2
Area (per 100 ha)	−6.6	10.4
+7 national parks b_{7i}:		
Biodiversity	−2.4	0.3
Ancient Artefacts	25.1	−1.8
Access	17.4	−2.1
Tau scale	1.7	39.5
Sigma(i)	0.6	0.6
Log likelihood	−4490.8	
Chi^2	5019	
McFadden pseudo R^2	0.36	

smallest reduction in WTP in relative terms is on the 'Biodiversity' attribute.

The effect of referring to the forthcoming establishment of five more national parks, and hence a total of seven, is captured in the parameter set b_7. These results are more mixed. For 'Biodiversity', the sign is as expected, but the further decrease in WTP is not significant. For 'Ancient Artefacts', the sign is now positive, and the parameter is almost significant at the 5% level, suggesting a small correction to the decreased WTP effect captured in the b_2 set. For 'Access', the WTA now decreases and thus reduced the effects captured in the b_2 set. Thus, we cannot reject the null hypothesis of equal WTP for biodiversity, when referring to seven instead of two national parks to be established prior to Store Åmose, and marginally this is also true for 'Ancient Artefacts'. For the 'Access' attribute, it seems that the introduction of further substitutes reduces the demand for compensation, and hence the WTA.

It might be expected, at least for use-related attributes like access, that values would depend on respondents' distance to Store Åmose and to the potential national parks, respectively. Various ways of incorporating distance in the model, including respondents' distance from home to Store Åmose, distance to the nearest substitute national park, relationships between distances etc., were examined. None of the attempts gave explanations.

6. Discussion

6.1. Test for Embedding Effects

The core question addressed by this paper is: Do the attribute values derived from a CE study of one preservation project depend on substitute projects to be undertaken beforehand? Provided these projects are considered substitutes in respect of one or more of the attributes, theory would suggest that 'Yes, they should' (Carson and Mitchell, 1995; Randall and Hoehn, 1996). An empirical analysis of this question was conducted by setting up a split-sample CE, where one particular project was evaluated in three plausible and strictly different embedding contexts.

The overall results suggest that WTP measures derived from CE studies are sensitive to the context in which the valuation is embedded; in this case whether two or more national parks will be established beforehand. However, WTP seems fairly insensitive to the number of national parks being presented. Hence there is only a weak tendency of decreasing WTP (WTA) when the valuation is performed in a context with seven parks forthcoming as opposed to two.

Based on the policy processes of the study and on the focus group interviews, the national parks should represent good substitutes for

Store Åmose in all attributes, except perhaps for those concerning the preservation of ancient artefacts. The focus of most of the proposed national parks is heavily on biodiversity and recreation – although preservation of historic cultural values does feature in some national parks. Jacobsen and Thorsen (2010) find that preservation of biodiversity is the primary good that people expect and want from the national parks. Nevertheless, significant decreases in WTP are found for both 'Biodiversity' and 'Ancient Artefacts', when presenting the valuation exercise in the context of two forthcoming national parks. The decrease is in the range of 20-30%, largest for 'Ancient Artefacts', which is clearly a non-trivial effect for policy-makers. However, as the number of forthcoming projects is increased, there is little further change in the WTP for these attributes. Thus, the effect could be interpreted as being mainly one of reminding respondents of the overall context of nature preservation plans.

'Access' exhibits a negative WTP. This was also found by Lundhede et al. (2010) in another study of *Store Åmose* with the same attributes, and their interpretation was that respondents perceived extended access to reduce the value of biodiversity preservation. The same could be the case here, and this might also explain the increase in requested compensation when more substitutes are introduced: As more, and perhaps closer and better, recreational sites are available, the value of access to *Store Åmose* is reduced even further and traded against the preservation of biodiversity and artefacts.

The fact that the null hypothesis of equal WTP is most clearly rejected across attributes when only two parks are presented (i.e. rejection of similar results between the split samples with 0 and 2 substitutes), suggests that respondents may simply have used the information on the establishment of national parks as a reminder of likely substitutes to be forthcoming. Thus, they ignored the specific number of national parks. An alternative explanation could be that the scenario of establishment of seven national parks lacked credibility from a respondent point of view (cf. Powe and Bateman, 2003). However, focus groups did not find a lack of credibility, which also aligns well with the debate at the time, where interest groups in fact called for more and different areas to be considered. Indeed today, four years later, five national parks have been decided, of which four are among the original seven pilot projects. Discussion on further increases in the number of national parks is ongoing.

In this study, a WTP space model was applied. The same hypotheses were also investigated in preference space models, but with less clear results. Thus, allowing for heterogeneous price sensitivity among respondents, as in the WTP space model, seems to better reveal the trade-offs that respondents are making. This finding is similar to Thiene and Scarpa (2009), and it relies on the assumption that respondents consider their trade-offs in terms of WTP.

6.2. Possible Caveats and Potential for Improvements

As reported, no significant differences were found in the scale parameter between the sub-samples. This indicated that respondents did not find the questionnaire more difficult to answer in the samples where national park projects were included. Thus, the quite massive extra information provided about the national parks did not confuse respondents about their preference for *Store Åmose* – even if they perhaps not fully understood the implications of the embedding context. This is supported by the finding of no statistical difference in proportions of serial non-respondents between the three split samples. The statistical difference in proportions of status-quo could indicate otherwise, however, the difference in absolute terms is rather small (see Table 2).

There may be ways that would provide respondents a better understanding of the embedding context. As shown in Table 2, a rather large proportion of the sample respondents did not understand the context as intended. It was deliberately not suggested to

respondents what the implementation of two or seven national parks might cost in the end. In principle, this should matter nothing or very little to their valuation of the *n*'th project (a small effect only through the imagined effect on their budget, cf. Bateman et al. (1997)) and suggesting cost would most likely imply a severe risk of bias through an anchoring effect. However, it creates an uncertainty with regard to the total future payment of protection when more substitutes are introduced. This might be a reason why the WTP for *Store Åmose* with seven substitutes is not fully consistent with WTP for *Store Åmose* with two substitutes.

7. Concluding Remarks

The contribution of this study is the attempt to provide an empirical answer to the following highly policy-relevant question: Do the values derived from a CE study of one preservation project depend on the number of substitute projects to be undertaken prior to project in focus? A split-sample CE was developed, where one particular nature preservation project was evaluated in three plausible and strictly different embedding contexts, in which respondents were alerted to forthcoming substitute nature preservation projects. This set-up was endowed with unusual credibility to respondents as the implementation of the first Danish national parks was discussed at the time of the experiment (Jacobsen and Thorsen, 2010).

The overall results suggest that WTP measures derived from CE studies are sensitive to the context in which the valuation is embedded, most clearly for the presence of substitute projects *per se*, and less so for the more specific number of substitute projects. The observed decrease in WTP was in the range of 20-30% of WTP for core attributes.

The problem of embedding in environmental valuation studies has great relevance to policy. In many political contexts, decision-makers must decide on a sequence of several environmental preservation projects, and thus it is an important question whether benefit estimates of a given project vary depending on the extent to which the project is evaluated as the (only) forthcoming project or as the *n*'th project. Environmental valuation research must be able to answer this question. The present study has shown that while this is probably not an easy task, it does not appear to be a hopeless task.

Acknowledgements

References

Amiran, E.Y., Hagen, D.A., 2010. The scope trials: Variation in sensitivity to scope and WTO with directionally bounded utility functions. J. Env. Econ. Man. 59, 293–301.
Andreoni, J., 1990. Impure altruism and donations to public goods: A theory of warm-glow of giving. Econ. J. 100, 464–477.
Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Roy, R., Schuman, H., 1993. Report of the NOAA Panel on Contingent Valuation. http://www.darp.noaa.gov1993.
Bateman, I.J., Munro, A., Rhodes, B., Starmer, C., Sugden, R., 1997. Does part-whole bias exist? An experimental investigation. Econ. J. 107, 322–332.
Bateman, I.J., Cole, M., Cooper, P., Georgiou, S., Hadley, D., Poe, G.L., 2004. On visible choice sets and scope sensitivity. J. Environ. Econ. Manage. 47, 71–93.
Bierlaire, M., 2003. BIOGEME: A free package for the estimation of discrete choice models. Proceedings of the 3 rd Swiss Transportation Research Conference, Ascona, Switzerland.
Boxall, P., Adamowicz, W.L., Moon, A., 2009. Complexity in choice experiments: choice of the status quo alternative and implications for welfare measurement. Aust. J. Agric. Resour. Econ 53 (4), 503–551.
Campbell, D., 2007. Willingness to Pay for Rural Landscape Improvements: Combining Mixed Logit and Random-Effects Models. J. Agric. Econ. 58, 467–483.
Carson, R.T., Mitchell, R.C., 1995. Sequencing and Nesting in Contingent Valuation Surveys. J. Environ. Econ. Manage. 28, 155–173.

Carson, R., Flores, N.E., Meade, N.F., 2001. Contingent Valuation: Controversies and Evidence. Environ. Res. Econ. 19, 173–210.

Clark, J., Friesen, L., 2008. The causes or order effects in contingent valuation surveys: An experimental investigation. J. Environ. Econ. Manage. 56, 195–206.

Cummings, R.G., Ganderton, P.T., McGuckin, T., 1994. Substitution effects in CVM values. Am. J. Agric. Econ. 76, 205–214.

Czajkowski, M., Hanley, N., 2009. Using Labels to Investigate Scope Effects in Stated Preference Methods. Environ. Resour. Econ. 44, 521–535.

Diamond, P.A., 1996. Testing the Internal Consistency of Contingent Valuation Surveys. J. Environ. Econ. Manage. 30, 337–347.

Dupont, D.P., 2003. CVM Embedding Effects When There Are Active Potentially Active and Passive Users of Environmental Goods, Environ. Resour. Econ. 25, 319–341.

Fiebig, D.G., Keane, M.P., Louviere, J., Wasi, N., 2010. The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. Market. Sci. 29, 393–421.

Foster, V., Mourato, S., 2003. Elicitation Format and Sensitivity to Scope. Environ. Resour. Econ. 24, 141–160.

Giraud, K.L., Loomis, J.B., Johnson, R.L., 1999. Internal and external scope in willingness-to-pay estimates for threatened and endangered wildlife. J. Environ. Manage. 56, 221–229.

Hailu, A., Adamowicz, W.L., Boxall, P.C., 2000. Complements, Substitutes, and Budget Constraints and Valuation: Application of a Multi-Program Environmental Valuation Method. Environ. Resour. Econ. 16, 51–68.

Hanemann, W.M., 1994. Valuing the Environment Through Contingent Valuation. J. Econ. Perspect. 8, 19–43.

Heberlein, T.A., Wilson, M.A., Bishop, R.C., Schaeffer, N.C., 2005. Rethinking the scope test as a criterion for validity in contingent valuation. J. Environ. Econ. Manage. 50, 1–22.

Jacobsen, J.B., Boiesen, J.H., Thorsen, B.J., Strange, N., 2008. What's in a Name? The use of Quantitative Measures vs. 'Iconised' Species when Valuing Biodiversity. Environ. Resour. Econ. 39, 249–263.

Jacobsen, J.B., Thorsen, B.J., Boiesen, J.H., Anthon, S., Tranberg, J., 2006. Værdisætning af syv mulige nationalparker i Danmark [Valuation of seven potential national parks in Denmark], Arbejdsrapport 28. Skov og Landskab, KVL, Frederiksberg. 63 pp.

Jacobsen, J.B., Thorsen, B.J., 2010. Preferences for site and environmental functions when selecting forthcoming national parks. Ecol. Econ. 69, 1532–1544.

Jorgensen, B.S., Wilson, M.A., Heberlein, T.A., 2001. Fairness in the contingent valuation of environmental public goods: attitude toward paying for environmental improvements at two levels of scope. Ecol. Econ. 36, 133–148.

Kahnemann, D., Knetsch, J.L., 1992. Valuing Public Goods: The Purchase of Moral Satisfaction. J. Environ. Econ. Manage. 22, 57–70.

Kotchen, M.J., Reiling, S.D., 1999. Do reminders of substitutes and budget constraints influence contingent valuation estimates? Another comment (1999). Land Econ. 75, 478–482.

Kuhfeld, W.F., 2004. Marketing Research Methods in SAS. Experimental Design, Choice, Conjoint, and Graphical Techniques, SAS 9.1. Edition, TS-694. SAS Institute Inc., Cary, NC, USA. 782 pp.

Larsen, F.W., Petersen, A.H., Strange, N., Lund, M.P., Rahbek, C., 2008. A quantitative analysis of biodiversity and the recreational value of potential national parks in Denmark. Environ. Manage. 41, 685–695.

Loomis, J.B., Gonzalez-Caban, A., Gregory, R., 1994. Do Reminders of Substitutes and Budget Constraints Influence Contingent Valuation Estimates? Land Econ. 70, 499–506.

Loureiro, M.L., Ojea, E., 2008. Valuing local endangered species: The role of intra-species substitutes. Ecol. Econ. 68, 362–369.

Lundhede, T.H., Olsen, S.B., Jacobsen, J.B., Thorsen, B.J., 2009. Handling respondent uncertainty in Choice Experiments: Evaluating recoding approaches against explicit modelling of uncertainty. J. Choice Model. 2, 1–30.

McDaniels, T.L., Gregory, R., Arvai, J., Chuenpagdee, R., 2003. Decision structuring to alleviate embedding in environmental valuation. Ecol. Econ. 46, 33–46.

Neill, H.R., 1995. The Context for Substitutes in CVM Studies: Some Empirical Observations. J. Environ. Econ. Manage. 29, 393–397.

Ojea, E., Loureiro, M.L., 2009. Valuation of wildlife: Revising some additional considerations for scope tests. Contemp. Econ. Policy 27, 236–250.

Powe, N., Bateman, I.J., 2003. Ordering effects in nested top–down and bottom–up contingent valuation designs. Ecol. Econ. 45, 255–270.

Powe, N., Bateman, I.J., 2004. Investigating insensitivity to scope: a split-sample test of perceived scheme realism. Land Econ. 80, 258–271.

Randall, A., Hoehn, J.P., 1996. Embedding in Market Demand Systems. J. Environ. Econ. Manage. 30, 369–380.

Revelt, D., Train, K., 1998. Mixed logit with repeated choices: Households' choices of appliance efficiency level. Rev. Econ. Stat. 80, 647–657.

Rollins, K., Lyke, A., 1998. The Case for Diminishing Marginal Existence Values. J. Environ. Econ. Manage. 36, 324–344.

Rolfe, J., Bennett, J., Louviere, J., 2002. Stated values and reminders of substitute goods: Testing for framing effects with choice modelling. Aust. J. Agric. Resour. Econ. 46, 1–20.

Scarpa, R., Rose, J.M., 2008. Designs efficiency for nonmarket valuation with choice modelling: how to measure it, what to report and why. Aust. J. Agric. Res. Econ. 52, 253–282.

Scarpa, R., Thiene, M., Train, K., 2008. Utility in willingness to pays pace: a tool to address confounding random scale effects in destination choice to the Alps. Am. J. Agric. Econ. 90, 994–1010.

Smith, V.K., 1992. Arbitrary Values, Good Causes, and Premature Verdicts. J. Environ. Econ. Manage. 22, 71–89.

Sælensminde, K., 2003. Embedding effects in valuation of non-market goods. Transp. Pol. 10, 59–72.

Thiene, M., Scarpa, R., 2009. Deriving and Testing Efficient Estimates of WTP Distributions in Destination Choice Models. Environ. Resour. Econ. 44, 379–395.

Train, K., Weeks, M., 2005. Discrete choice models in preference space and willingness-to-pay space. In: Scarpa, R., Alberini, A. (Eds.), Applications of simulation methods in environmental and resource economics. Springer Publisher, Dordrecht, The Netherlands, pp. 1–16. Chapter 1,.

Veisten, K., Hoen, H.F., Navrud, S., Strand, J., 2004. Scope insensitivity in contingent valuation of complex environmental amenities. J. Environ. Manage. 73, 317–331.

Verbic, M., Slabe-Erker, R., 2009. An econometric analysis of willingness-to-pay for sustainable development: A case study of the Volcji Potok landscape area. Ecol. Econ. 68, 1316–1328.

Von Haefen, R.H., Massey, D.M., Adamowicz, W.I., 2005. Serial Nonparticipation in Repeated Discrete Choice Models. Am. J. Agric. Econ. 87 (4), 1061–1076.

Whitehead, J.C., Blomquist, G.C., 1995. Do Reminders of Substitutes and Budget Constraints Influence Contingent Valuation Estimates? Comment. Land Econ. 71, 541–543.

Wilhjelmudvalget, 2001. En rig natur i et rigt samfund [A rich nature in a rich society]. Wilhjelmudvalget, Skov & Naturstyrelsen [Danish Forest and Nature Agency], Copenhagen, Denmark. . 121 pp.