Journal of Choice Modelling 47 (2023) 100401

Contents lists available at ScienceDirect

Journal of Choice Modelling

journal homepage: www.elsevier.com/locate/jocm

Cost vector effects in discrete choice experiments with positive status quo cost

Heini Ahtiainen [a], Eija Pouta [a], Wojciech Zawadzki [b], Annika Tienhaara [a,*]

[a] Natural Resources Institute Finland (Luke), Latokartanonkaari 9, FI-00790, Helsinki, Finland
[b] Department of Economic Sciences, University of Warsaw, Długa 44/50, 00-241, Warsaw, Poland

ARTICLE INFO

Keywords:
Discrete choice experiments
Cost vector
Choice behaviour
Preferences
Willingness to pay
Agri-environmental policies

ABSTRACT

An important component of the design phase of a discrete choice experiment (DCE) is formulating the cost vector, which specifies the costs of the alternatives and enables the calculation of marginal willingness to pay (WTP) estimates. If the cost vector affects choice behaviour, welfare estimates may depend on the choice of the cost vector, which leads to problems with the validity and reliability of DCE results. We employ a split-sample design to examine cost vector effects on choice behaviour and WTP estimates. Our data come from a DCE on agri-environmental policies to a nationally representative sample in Finland. We provide additional insights compared to previous research by including four cost vectors with otherwise identical surveys and experimental designs and a positive cost for the status quo alternative, with cost levels for policy alternatives both below and above the status quo cost. We obtain some evidence that the cost vector affects choice behaviour, as the proportion of status quo choices is larger with higher cost vectors. Both absolute and relative cost levels matter for choices. The marginal WTP estimates are highest in the sub-sample with the largest range cost vector that has cost levels both below and above the status quo cost. We suggest more careful pre-testing of the cost levels compared to current practices to determine a plausible range of cost levels to produce valid welfare estimates.

1. Introduction

An important component of the design phase of a discrete choice experiment (DCE) is formulating the attribute levels, including the cost vector, which specifies the costs of the alternatives and allows the calculation of monetary welfare estimates (marginal willingness to pay, WTP) for changes in attributes and their combinations. Most DCE studies have employed a single cost vector, chosen based on pre-testing of attribute levels or insights from previous valuation studies.

In general, the attribute levels in a DCE should be realistic and plausible to respondents, and useful for decision-making (Johnson et al., 2013). For the cost vector, this implies that the levels should be credible to avoid possible biases and protest behaviour (Johnston et al., 2017), thus refraining from presenting implausibly high or low costs for the good. Credibility of the attribute levels can and should be evaluated as part of pre-testing, and this is particularly important for the cost attribute. However, there is no clear guidance on the practical construction of the cost levels, and thus DCE studies may employ different strategies in determining the levels.

This should not be a problem if respondents adhere to the theoretical assumptions of having stable and well-defined preferences, as in that case the cost vector should not affect the observed choices and WTP. However, there is empirical evidence from DCEs that

* Corresponding author.
E-mail address: annika.tienhaara@luke.fi (A. Tienhaara).

https://doi.org/10.1016/j.jocm.2023.100401
Received 18 March 2021; Received in revised form 13 January 2023; Accepted 20 January 2023
Available online 2 February 2023

various features of the choice context can have an impact on the subsequent estimates, such as the range of attribute levels and inclusion of a cost attribute (Mørkbak et al., 2010; Luisetti et al., 2011; Pedersen et al., 2011; Glenk et al., 2019). In addition, studies in behavioural economics have observed the provision of arbitrary anchors or value cues to affect preferences (Ariely et al., 2003; Tversky and Kahneman, 1974), which could occur particularly when respondents are uncertain about their valuations (Glenk et al., 2019). For example, in DCEs on less familiar environmental goods and services, respondents may use the cost levels in the valuation task as information on the good's value and anchor their responses to the costs presented (Mitchell and Carson 1989; Kragt 2013). If this is the case, welfare estimates may be affected by the choice of the cost vector.

Cost vector design has been the focus of several contingent valuation studies (e.g., Cooper and Loomis 1992; Kanninen and Kriström 1993; Vossler et al., 2004), and several previous studies have used split samples to investigate the impact of different cost vectors on choice behaviour in the DCE setting (Ryan and Wordsworth 2000; Hanley et al., 2005; Carlsson and Martinsson 2008; Mørkbak et al., 2010; Kragt 2013; Liesivaara and Myyrä 2014; Svenningsen and Jacobsen 2018; Glenk et al., 2019). These studies have examined effects on parameter estimates and marginal WTP, as well as on systematic choice behaviour, such as the proportion of status quo choices and protest responses, the proportion of respondents choosing the lowest or highest cost alternative in all choice sets, and bid acceptance at different cost levels. The findings of these studies have been mixed.

Ryan and Wordsworth (2000) examined the effect of attribute levels on WTP estimates. For five of the six attributes, including cost, there were no significant differences in the parameter estimates across the two split samples, but unit WTP estimates were significantly different for four attributes. However, no difference was found in the WTP for a programme (combination of attributes). Hanley et al. (2005) used a split-sample design to compare two cost vectors, reporting no significant impact on preferences or WTP estimates. Carlsson and Martinsson (2008) observed significantly higher marginal WTP estimates for attributes when a scaled-up cost vector was used. Examining the effect of changes in the maximum level of the cost attribute, Mørkbak et al. (2010) demonstrated significant differences in preferences and WTP estimates, with an increased level of the maximum cost inducing higher WTP estimates for all attributes. Kragt (2013) investigated the effect of different cost vectors on choices, allowing for individual preference and scale heterogeneity, finding, for the most part, no significant difference between the treatments. Liesivaara and Myyrä (2014) used three cost vectors and found WTP to be sensitive to the cost vector employed, with the high vector producing higher welfare estimates. Svenningsen and Jacobsen (2018) tested the impact of different survey designs, including changes in cost range, on hypothetical bias. The results did not allow for isolating the effect of varying the cost vector, but they were able to conclude that the observed hypothetical bias was likely caused by differences in either the elicitation format (hypothetical versus actual payment) or the cost range, instead of changing the payment vehicle. Glenk et al. (2019) examined the effect of cost vectors on choice behaviour, also considering the role of respondents' income. Their results indicated that in addition to affecting WTP estimates, cost vectors can influence decision strategies via differences in attribute non-attendance and the proportion of respondents always choosing the lowest cost policy alternative. They also found weak evidence that cost vector effects on WTP estimates can be different for lower-versus higher-income respondents. Besides this and studies on the factors affecting the propensity for anchoring (e.g. Furnham and Boo 2011), there are only few investigations how respondent and other characteristics impact systematic choice behaviour.

The effect of the inclusion of a cost attribute on choice behaviour has been examined similarly in health and environmental valuation with mixed findings. In the health care context, several studies have found that including the cost attribute had no significant effect on preferences for the other attributes (Bryan et al., 1998; Essers et al., 2010; Sever et al., 2019; Genie et al., 2021). However, Aravena et al. (2014) observed larger standard errors and lower model fit and Sever et al. (2019) and Genie et al. (2021) an increased response error variance (lower scale) when a cost attribute was included. Both Carlsson et al. (2007) and Van Zanten et al. (2016) found the inclusion of a cost attribute to affect the observed preferences and ranking of the other attributes, and in Pedersen et al. (2011), adding a cost attribute had a significant effect on the parameter estimates in a forced choice context.

A few additional studies have provided information on anchoring effects in DCE. Ladenburg and Olsen (2008) tested two versions of instructional choice sets, presented before the actual choice sets, which differed in the costs of alternatives. They observed differences in preferences and WTP estimates for female respondents and considered this to provide evidence of starting point bias in choice experiments. Similarly, Meyerhoff and Glenk (2015) observed instructional choice sets to affect WTP, particularly when low quality improvements are combined with high prices in the presented sets, and vice versa. In their case, these extreme combinations resulted in lower overall WTP for the attributes, which they argue could potentially be explained by reduced credibility of the entire DCE. In an experiment, Ariely et al. (2003) showed the provision of arbitrary cues of values (i.e., social security numbers) to affect willingness to pay for familiar consumer products.[1]

Choice behaviour may be affected by many (unobservable) factors besides the cost vector and other aspects of the choice context, which could lead to differences in observed choices and WTP across sub-samples. However, as illustrated by the review of empirical evidence, cost vector effects are often observed, and thus it is important to provide more systematic evidence on when and how they may occur. Despite the central status of the cost vector in DCEs and determining welfare impacts, there is yet no consistent evidence on the effect of the cost vector on choice behaviour, resulting in relatively few practical recommendations for cost vector construction that would have wide applicability. If WTP estimates for sub-samples exposed to different cost vectors diverge and there are no other observable characteristics behind this divergence, the scale and range of the costs could be driving the differences. This can have major implications for the validity of DCE results and their use for policy support.

We contribute to the literature by providing additional insights into cost vector effects in the DCE setting using four split samples

[1] Glenk et al. (2019) provide a comprehensive overview of previous research.

and otherwise identical surveys and experimental designs. Our data come from a DCE on agri-environmental policies and their environmental benefits, administered to a representative sample of the national population in Finland. The survey elicited preferences for four environmental attributes: traditional rural biotopes and endangered species, agricultural landscape, climate and water quality, with the payment vehicle being a tax paid by all taxpayers. Our aim was to provide welfare estimates to support discussions on future agri-environmental policies, while examining cost vector effects in DCE.

This setting allows us to test the impacts of the scale of the cost vector on choices and WTP. Furthermore, the status quo alternative is associated with a positive cost to reflect the current situation of financing the Finnish agri-environmental policy, and thus cost levels both below and above the reference level costs are included in two of the cost vectors. We contribute to the literature by providing information on cost vector effects on systematic choice behaviour and welfare estimates when the status quo alternative has a positive cost and there are cost levels lower than the status quo cost for the policy alternatives, previously lacking in the literature. We test the differences across the cost vector sub-samples in a systematic and varied way and are able to rule out observable differences in socio-demographic factors across the samples that could lead to observed differences.

The paper is organized as follows. Section 2 describes the cost vector design and hypotheses to be tested. Section 3 presents the econometric models and section 4 the design and implementation of the choice experiment. Results are provided in section 5, and discussion and conclusions in section 6.

2. Cost vectors and hypotheses

2.1. Cost vectors and split-sample design

We examined cost vector effects in a DCE focused on the benefits from agricultural environments in Finland, with the aim to support policy discussions on a transfer towards results-based agri-environmental policies. The survey aimed at eliciting preferences and willingness to pay for four environmental attributes that could be affected by changes in agri-environmental policy: traditional rural biotopes and endangered species, agricultural landscape, climate effects and water quality effects. The payment vehicle was specified as a tax per individual to be paid for the next 10 years by all Finnish taxpayers. The cost vector scale and range effects were investigated using four versions differing only in the cost levels presented (Table 1). Apart from the varying cost vector, the choice context, design and the levels of the attributes were identical between the four versions. The sub-samples are referred to as A(5–300), B(5–500), C(40–300) and D(40–500), which in each case indicate the lowest and highest level of the cost vector.

The construction of the cost vectors proceeded as follows. We first decided to employ four cost vectors in the final survey. The construction of the cost levels within the vectors relied on earlier valuation studies on the agricultural environment in Finland, the current and potential future costs of the Finnish agri-environmental programme, and the results of the pilot survey.

The pilot included a DCE with cost vector A(5–300) (with 6 levels), followed by an open-ended contingent valuation (CV) question on WTP for the best possible alternative. In the DCE, acceptance of alternatives costing €300 was 14%. In the CV, the mean WTP for the best scenario was 80€ and the median 50€, but few respondents expressed WTP as high as 700€. Considering these pilot results, we increased the highest cost to €500 in two sub-samples (B(5–500) and D(40–500)). The clear increase from the second highest to the highest cost level at the upper end of the cost vector follows the studies by Svenningsen and Jacobsen (2018) (67% increase), Pedersen et al. (2011) (167% increase) and Mørkback et al. (2010) (85% increase) and aimed to avoid choice uncertainty and reduce error variances (Svenningsen and Jacobsen 2018).

The cost of the status quo alternative was always €40 per year. This was the approximated annual cost of the current Finnish agri-environmental programme per individual collected through taxes, based on expert judgement. The lowest level of the cost vector was set to the current cost level (approximately €40) in two of the four sub-samples (C(40–300), D(40–500)). In the two other sub-samples (A(5–300), B(5–500)), the lowest level was specified to be below the current level but still reasonable, i.e. €5 per year. Choosing policy alternatives costing less than €40 would result in cost savings for the respondent. Lower than current costs are justified, as more effective targeting of measures in the results-based scheme could decrease the total costs of the agri-environmental policy and reduce the tax burden in the future, while increasing the environmental benefits (VainioTienhaara et al., 2021; Niskanen et al., 2021).

The combinations of the lowest and highest costs defined the cost ranges for the 4 vectors (A(5–300), B(5–500), C(40–300), D(40–500)). First, using the CV results from the pilot, five intermediate bid levels were constructed for sub-sample A(5–300) by setting the median WTP (€60) in the middle of the cost vector. For the remaining cost vectors, the intermediate levels were defined by keeping the relative differences between the cost levels approximately the same as in sub-sample A(5–300). Thus, the range between the lowest and the highest cost differed across the vectors, but the relative differences between the costs were the same. The cost vectors for the policy alternatives had seven levels.

Our setting with six cost vector pairs enables capturing various cost vector effects on choices and welfare estimates. Pairs A(5–300)–

Table 1
Cost vectors employed in the choice experiment study.

Cost vector	Levels (€)
A(5–300)	5, 20, 40, 60, 80, 100, 300
B(5–500)	5, 30, 60, 100, 130, 160, 500
C(40–300)	40, 50, 70, 90, 110, 130, 300
D(40–500)	40, 60, 90, 130, 160, 190, 500

B(5–500) and C(40–300)–D(40–500) can be considered as a standard comparison of low and high scale cost vectors, where the lowest cost level is the same but the other cost levels in B and D have been scaled up compared to A and C, respectively. Furthermore, sub-samples A and B include cost levels that are lower than the status quo cost, and thus pairs A(5–300)–C(40–300) and B(5–500)–D (40–500) can be examined to identify the influence of having lower cost levels for the policy alternatives than the status quo cost. Various impacts from the inclusion of cost decreases for policy alternatives are possible, such as on the proportion of status quo or lowest cost choices. There could also be asymmetric sensitivity for cost increases and decreases.

2.2. Hypotheses

We focus on four hypotheses to investigate possible cost vector effects on choice behaviour identified in previous literature and pre-defined before the implementation of our study (Hanley et al., 2005; Kragt 2013; Glenk et al., 2019). The starting point in the investigation is that features of the choice context, such as the cost vector, should not influence choices or WTP if preferences are well-defined. However, empirical studies have found higher cost vectors to lead to more frequent systematic choice behaviour due to differences in decision and information processing strategies (Hanley et al., 2005; Glenk et al., 2019). Further, people may use the cost levels as information on the reasonable payment for or actual value of the good, thus anchoring their choices to the presented levels of the cost (Frykblom and Shogren 2000; Kragt 2013), and may be more concerned about relative than absolute costs (Kragt 2013; Glenk et al., 2019).

We test four hypotheses, both as joint hypotheses across the sub-samples and as pairwise comparisons. First, we examine cost vector effects on systematic choices of the status quo, lowest cost and highest cost alternatives (hypotheses 1 and 2). Second, we investigate the impacts of the cost vector on respondents' sensitivity to absolute and relative cost levels, as well as on marginal WTP estimates (hypotheses 3 and 4). All hypotheses are nondirectional to maintain consistent presentation, as we have no strong theory-based prior expectations of the direction of the effects of the cost vectors for any of the hypotheses.

Our hypotheses are the following.

Hypothesis 1. There are no significant differences in the proportion of status quo choices across cost vectors.

P_{SQ} denotes the proportion of status quo (SQ) choices.

The testing of this hypothesis is divided into two parts. In part 1a, we test the difference in the systematic choice of the status quo across the cost vector treatments by examine the proportion of status quo choices in all choice tasks (P_{SQall}). In part 1b, we examine differences in the proportion of respondents choosing the status quo in any of the choice tasks (P_{SQany}).

We are particularly interested in the differences between 1) pairs C(40–300) and D(40–500), which have the same lowest cost level (€40), but D is of a higher scale overall compared to C, and 2) pairs A(5–300)–C(40–300) and B(5–500)–D(40–500), which allow us to examine whether the presence of a lower cost level than the status quo for the policy alternatives affects the proportion of SQ choices.

$H0_{1a}$: $P_{SQall}A = P_{SQall}B = P_{SQall}C = P_{SQall}D$
$H1_{1a}$: $P_{SQall}A \neq P_{SQall}B \neq P_{SQall}C \neq P_{SQall}D$
$H0_{1b}$: $P_{SQany}A = P_{SQany}B = P_{SQany}C = P_{SQany}D$
$H1_{1b}$: $P_{SQany}A \neq P_{SQany}B \neq P_{SQany}C \neq P_{SQany}D$

Hypothesis 2. There are no significant differences in always choosing the lowest/highest cost level across cost vector treatments.

Let P_L and P_H denote the proportion always choosing the lowest and the highest cost, respectively.

Our interest lies in comparing 1) pairs A(5–300)–B(5–500) and C(40–300)–D(40–500), which have the same lowest bid but B and D are overall higher in scale and have a higher highest cost level, as well as 2) pairs A(5–300)–C(40–300) and B(5–500)–D(40–500), which have the same highest cost level but differ in overall scale and in the lowest cost levels.

$H0_2$: $P_LA = P_LB = P_LC = P_LD$ and $P_HA = P_HB = P_HC = P_HD$
$H1_2$: $P_LA \neq P_LB \neq P_LC \neq P_LD$ and $P_HA \neq P_HB \neq P_HC \neq P_HD$

Hypothesis 3. Respondents are sensitive to absolute cost levels.

The testing of this hypothesis is divided into two parts. In 3a, we test whether there are significant differences in the proportion of alternatives with the nth cost level selected across the cost vector sub-samples. In 3b, we test for significant differences in the choice proportion of the same absolute value of cost across the sub-samples. P_n denotes the proportion of choosing the nth cost level and P_V the proportion of choosing the absolute cost value.

Here, we focus on testing those situations where the nth cost level differs in absolute value across sub-samples (3a) and where the absolute cost levels are the same across sub-samples (3b).

$H0_{3a}$: $P_nA = P_nB = P_nC = P_nD$
$H1_{3a}$: $P_nA \neq P_nB \neq P_nC \neq P_nD$
$H0_{3b}$: $P_VA = P_VB = P_VC = P_VD$

$H1_{3b}$: $P_VA \neq P_VB \neq P_VC \neq P_VD$

Hypothesis 4. There are no significant differences in coefficient estimates (4a) or marginal WTP for the environmental attributes (4b) across the cost vector sub-samples.

Let β denote the coefficient for the environmental attributes j, and WTP the marginal willingness to pay. Here, we are interested in each pairwise comparison.

$H0_{4a}$: $\beta_{jA} = \beta_{jB} = \beta_{jC} = \beta_{jD}$
$H1_{4a}$: $\beta_{jA} \neq \beta_{jB} \neq \beta_{jC} \neq \beta_{jD}$
$H0_{4b}$: $WTP_A = WTP_B = WTP_C = WTP_D$
$H1_{4b}$: $WTP_A \neq WTP_B \neq WTP_C \neq WTP_D$

3. Econometric methods

Hypotheses 1, 2 and 3 were tested using two-sided chi[2] and two-sample proportion tests. These allow examination of whether there are significant differences in the proportion of status quo choices, systematic lowest/highest cost level choices and choosing the nth cost level across the cost vector sub-samples.

To analyse hypothesis 4, respondents' choices were modelled using the mixed logit (MXL) model, which relaxes the independence of irrelevant alternatives (IIA) assumption and accommodates heterogeneity in individual preferences. The MXL model allows for preference heterogeneity through random parameters that have both a mean and a standard deviation, picking up preference variation across individuals in terms of both unconditional taste heterogeneity (random heterogeneity) and individual characteristics (conditional heterogeneity) (Hensher et al., 2015; Hensher and Greene, 2003; Revelt and Train, 1998; Train, 2003).

In the MXL model, respondent *i*'s utility associated with choosing alternative *j* in the choice task *t* is represented by the general utility expression:

$$U_{jti} = (\beta + \eta_i)' X_{jti} + \varepsilon_{jti} \tag{1}$$

where β is a vector of mean attribute parameters in the population, η_i is a vector of person-specific deviations from the mean, X_{jti} is a vector of explanatory variables, and ε_{jti} is a stochastic error term assumed to have an independent and identical (IID) extreme value type 1 distribution. The analyst must impose a distribution for η_i, the most common being normal, lognormal, triangular and uniform (Hensher et al., 2015).

Models were estimated using Stata 13. The environmental attributes followed a normal distribution and the cost attribute a lognormal distribution to limit the coefficient to negative values and assure finite moments of the WTP distribution (Daly et al., 2012). As the lognormal distribution is defined only over positive real values, cost was multiplied with −1 to constrain the random coefficient to be negative. The cost variable was specified as continuous, and the environmental attributes were dummy-coded.

As sub-samples A and B included cost levels below and above the status quo cost (40€), we included an interaction term equalling the product of the cost level (continuous) and cost being lower than the status quo cost (binary). The aim was to examine whether the sensitivity to costs varies between cost increases and decreases.[2] The interaction term was specified as random and followed a normal distribution.

The models were estimated using 2000 Halton draws with 2000 draws as burn-in. Halton draws were chosen as they have been found to outperform pseudorandom draws (Czajkowski and Budziński 2019; Bhat 2001; Train 2000).[3] Consistency of the model results was tested by estimating each sub-sample model with 500 and 5 000 draws. The signs, significance and magnitudes of the coefficients were robust to changes in the number of draws. To examine hypothesis 4a on differences in preferences for environmental attributes, both sub-sample-specific and pooled models including two sub-samples were estimated. Marginal WTP estimates and confidence intervals were calculated using the delta method in Stata with the *nlcom* command.[4] For hypothesis 4b, differences in the WTP distributions across sub-samples were tested using the complete combinatorial method, which calculates every possible difference between two empirical distributions (Poe et al., 2005).

4. Choice experiment survey design and implementation

The survey data were collected using an Internet survey during the spring of 2016, with the sample drawn from the probability-

[2] We also examined the option of including two cost variables in the A and B sub-sample models, one for cost levels lower and one for higher than the status quo cost, but this approach led to convergence problems and was thus discarded.

[3] Scrambled Sobol draws, as well as scrambled or shuffled Halton draws have been found to perform the best in terms of simulation error (e.g. Czajkowski and Budziński 2019; Bhat 2003). We tested both shuffled (Nlogit) and scrambled (Stata) Halton draws but had issues with model convergence. Thus, we have continued to use regular Halton draws in the estimation.

[4] As cost is lognormal, the marginal WTP for an attribute is calculated as $-1 * \mu_{attribute} / mean\ cost$. *Mean cost* can be calculated as $\left(-1 * exp\left(\mu + \frac{\sigma^2}{2}\right)\right)$, where μ is the mean coefficient and σ is the standard deviation coefficient for $-cost$.

based online panel of a private survey company Taloustutkimus. Their online panel comprises over 30 000 respondents who have been recruited using random sampling to represent the Finnish population (Taloustutkimus 2017). A pilot survey of 202 people was conducted before the final survey. Altogether, 2066 respondents completed the final survey, corresponding to a response rate of 25%. The survey employed a split-sample design with four different cost vectors. Respondents were randomly assigned to one of the four sub-samples.

The choice experiment was framed around new agri-environmental policies producing environmental benefits (for details, see Tienhaara et al., 2020). The attributes were selected based on the Common International Classification of Ecosystem Services (CICES) using a literature review, analysis of previous survey data, evaluation by stakeholders, as well as expert judgement by agricultural and environmental economists and ecologists. The selection process is described in more detail in Tienhaara et al. (2021). The final four environmental attributes were traditional rural biotopes and endangered species, typical agricultural landscape, climate effects and water quality effects. Table 2 presents the attributes together with their descriptions and levels, as well as the abbreviations of the variables used in the results section (marked in parentheses in the first column). Status quo levels are marked with (SQ). We aimed to enhance consequentiality of the responses by describing the funding mechanism of the agri-environmental programme and stating that the information from the choice tasks would help decision-makers revise the programme (Johnston et al., 2017). Further, truthful preference revelation and reduction in the number of strategic responses was invoked by a standard reminder of the additional costs people would incur based on choosing certain alternatives and by including a follow-up question on certainty of the choices, both potential techniques to mitigate hypothetical bias (Penn and Hu 2018).

The design consisted of 36 choice sets divided into six blocks. These six blocks and 36 choice tasks were used to generate four versions of the design by changing only the cost levels. Thus, the designs were otherwise identical across the sub-samples. Each respondent was presented with six choice tasks, each comprised of three alternatives: the status quo alternative, described as maintaining the current policy, and two policy alternatives with higher levels of environmental benefits compared to the current state. The order of the choice sets within a block was randomized. The status quo alternative corresponded to the current agri-environmental policy in Finland and was identical across the choice tasks.[5] An example of a choice task is presented in Table 3.

The experimental design was a Bayesian D-efficient design using Ngene (v. 1.0.2), taking 500 Halton draws for the prior parameter distributions and using the parameter estimates obtained from the pilot study (see, e.g., Rose and Bliemer 2009). Zero priors were used in the pilot study. The final design was optimized for the MNL model and for D-efficiency and interactions were not included in the design. A Bayesian prior was used for the plants in cultivation attribute and fixed priors were used for all the other attributes.[6] According to the S-estimate, sample size requirement was 360 respondents in each sub-sample (A, B, C and D). The S-estimate informs about the minimum number of observations necessary to obtain statistically significant parameter estimates at the 95% confidence level.

5. Results

5.1. Sample characteristics

Each cost vector sub-sample included a little over 500 respondents, and in total there were 2066 respondents to the survey (Table 4). The sample was divided equally across the four treatments, i.e., the number of respondents who were contacted in each treatment was the same. Respondents were randomly recruited to the survey versions, both for the cost vector sub-samples and the choice experiment blocks. The share of the total number of respondents was the same across the treatments. All six blocks included the lowest cost level and four blocks out of six the highest cost level. Thus, approximately 67% of the respondents faced both the lowest and the highest cost in the choice experiment in all sub-samples.

Table 5 presents the descriptive statistics for the four samples. We examined whether there were significant differences in respondent characteristics and views across the samples with the chi^2 test (gender, education level, living environment, second home environment, CE question follow-ups), ANOVA (household size, income, response time) and the Kruskal-Wallis test (age group) and found no significant differences. Thus, the possible differences in the results across sub-samples do not appear to be driven by identifiable differences in the socio-demographic or other characteristics of the respondents.[7]

[5] The lower price of the policy alternatives compared to the status quo resulted in the status quo being a dominated alternative in some choice tasks of treatments A(5–300) and B(5–500), as policy alternatives with higher environmental attribute levels and the same or a lower cost than the status quo alternative were available. This allowed us to check how often the status quo was chosen as a dominated alternative. Approximately 20% of the alternatives/choice tasks in A and B had a lower than status quo cost, and status quo was chosen as a dominated alternative in 9% of the cases. Potential explanations for these seemingly irrational choices could be altruism and fairness considerations, as well as concern for farmers' income.

[6] The attribute for plants in agricultural landscape was not significant in the pilot study. However, a similar DCE was conducted for farmers in the same project (Tienhaara et al., 2020) and the attribute was significant in the analysis of the farmer data. Thus, the attribute was kept in the citizen survey in order to keep the citizen and farmer DCEs similar and enable comparison of citizens' WTP and farmers' WTA for agricultural ecosystem services. As the pilot did not provide a prior for the plant attribute, a Bayesian prior was used for this attribute for the final design.

[7] Examination of sample characteristics by blocks for each sub-sample is provided in Appendix A. There are only few significant differences in respondents' characteristics across blocks: age in sub-sample A, household size and response time in B, age and share having a second home in agricultural environment in C, and age and income in D.

Table 2
Attributes of conservation programmes and their levels (attribute abbreviations marked in parentheses in the first column).

Attribute	Description	Levels
Traditional rural biotopes and endangered species (traditional)	Mowed or grazed seminatural grasslands (meadows, leas, pastures) can provide a habitat for several endangered species.	Current area, 0 species protected (SQ) 30% increase in area, 100 species protected 60% increase in area, 200 species protected
Typical agricultural landscape (animals) (plants)	Grazing animals and crops grown in open fields affect the diversity of the landscape.	Grazing animals: seldom seen (SQ), seen often during the summer season, seen often during the summer and unfrozen season Plants in cultivation:3 species (SQ): 4 species: 5 species
Climate effects: decrease in current emissions (climate)	Agricultural greenhouse gas emissions contribute to climate change. Emissions can be reduced by various cultivation practices and capturing greenhouse gasses.	0% (SQ) 10% 30%
Water quality effects: Share of surface waters in good or excellent status (%) (waterqual)	About half of the nutrients leaching to waters are from fields. This is affected by the amount of fertilizers used, cultivation practices and annual weather conditions.	60% (SQ) 70% 80%
Cost	Cost for taxpayers, €/year during 2017–2026. Cost of status quo alternative: €40	A: €5, 20, 40, 60, 80, 100, 300 B: €5, 30, 60, 100, 130, 160, 500 C: €40, 50, 70, 90, 110, 130, 300 D: €40, 60, 90, 130, 160, 190, 500

Table 3
Example of a choice task.

	Current programme	Alternative X	Alternative Y
Traditional rural biotopes and endangered species	Current area,0 species protected	Area is increased 60%,200 species protected	Area is increased 30%,100 species protected
Typical agricultural landscape, Grazing animals	Seldom seen	Seen often during the summer season	Seen often during the summer season
Plants in cultivation	3 species	5 species	4 species
Climate effects, Decrease in current emissions	0%	0%	30%
Water quality effects, Share of surface waters with a good or excellent status	60%	80%	60%
Cost/taxpayer/year, during 2017–2026	€40	€70	€130
I would choose	*	*	*

Table 4
Number of respondents and response rates.

Cost vector	Number of respondents (observations)	Share of total respondents (%)	Share of respondents facing the entire cost vector range (both lowest and highest cost) (%)
A(5–300)	520 (3 120)	25.2	67.1
B(5–500)	514 (3 084)	24.9	66.9
C(40–300)	516 (3 096)	25.0	67.0
D(40–500)	516 (3 096)	25.0	66.9
Total	2066 (12 396)		

5.2. Systematic differences in choice behaviour

In the first part of the analysis, we tested whether there were differences in the proportion of status quo choices (hypothesis 1), systematic highest and lowest cost choices (hypothesis 2) and in sensitivity to absolute cost levels (hypothesis 3) across the four cost vector sub-samples using two-sided chi^2 and two-sample proportion tests.

Hypothesis 1. There are no significant differences in the proportion of status quo choices across cost vector sub-samples.

The findings in Table 6 show significant differences in choosing the status quo (SQ) alternative across the cost treatments, both in choosing the status quo in all tasks ($chi^2 = 8.5$, $p = 0.037$) (hypothesis 1a) and in any of the tasks ($chi^2 = 52.7$, $p < 0.001$) (hypothesis 1b). According to the two-sample proportion tests, there is a significant difference in the proportion of respondents choosing the SQ alternative in all tasks between treatments A(5–300) and D(40–500), with less systematic SQ choices in cost vector A ($z = 2.5$, $p = 0.011$). It is worth noting that cost vector A has two policy cost levels that are below the cost for the status quo alternative. This could explain the observed difference, aside from cost vector D being higher overall. However, cost vector B(5–500) also has lower cost levels than the status quo cost, and exhibits the second largest proportion of status quo choices. The other significant difference in the proportion of systematic SQ choices is across sub-samples C(40–300) and D(40–500) ($z = 2.2$, $p = 0.026$), where there are less

Table 5
Respondent characteristics in the sub-samples.

Characteristic	Average				P-value
	A(5–300)	B(5–500)	C(40–300)	D(40–500)	
Women (%)	46.5	40.5	45.0	45.5	0.210
Age group (1–5; 1 = 18–24 years old, 5 = 65–74 years old)	3.7	3.8	3.7	3.7	0.685
Household size (people)	2.34	2.33	2.27	2.32	0.726
University level education (%)	28.3	26.8	27.3	24.8	0.639
Individual monthly gross income (€)	2 701	2 697	2 746	2 725	0.962
Response time (minutes)[a]	21.3	21.8	22.8	22.5	0.478
Agricultural living environment (%)	25.4	26.3	24.8	22.9	0.633
Second home in agricultural environment (%)	30.0	31.3	30.2	32.2	0.864
Easy to answer CE questions (%)	49.4	44.9	46.9	47.9	0.534
Understood CE questions (%)	77.7	76.3	75.8	76.2	0.894
Certain of CE choices (%)	56.5	57.0	54.3	55.2	0.807

[a] Response times over 180 min (34 respondents) were removed from the examination, as these were not considered to reflect actual response times. Differences across samples were not significant even with full samples.

Table 6
Proportion of respondents choosing the status quo alternative, highest cost level and lowest cost level. Detailed test results are provided in Appendix B.

Cost vector	Proportion of respondents choosing SQ in all choice tasks (%)	Proportion of SQ choices in any of the choice tasks (%)	Proportion of respondents choosing the highest cost in all choice tasks (%)	Proportion of respondents choosing the lowest cost in all choice tasks (%)
A(5–300)	4.6	17.7	3.9	10.2
B(5–500)	6.8	23.7	3.7	8.4
C(40–300)	5.0	19.5	4.1	10.5
D(40–500)	8.5	23.8	3.5	13.2
Total	6.2	21.2	3.8	10.6

systematic SQ choices in cost vector C. Both of these sub-samples have the SQ cost level as the lowest cost level for the policy alternatives, but D is higher overall compared to C.

There is a significant difference in the proportion of the SQ alternative chosen in any of the choice tasks between most of the cost vector pairs. Two-sample proportion tests reveal significant differences in the proportion of SQ choices between the following cost vector pairs: A(5–300)–B(5–500) ($z = -5.8$, $p < 0.001$), A(5–300)–C(40–300) ($z = 1.8$, $p = 0.07$), A(5–300)–D(40–500) ($z = -5.9$, $p < 0.001$), B(5–500)–C(40–300) ($z = 4.0$, $p < 0.001$) and C(40–300)–D(40–500) ($z = 4.1$, $p < 0.001$). The status quo alternative is chosen the least often in sub-sample A(5–300), followed by C(40–300). The proportion of SQ choices is higher in sub-sample D(40–500) than in sub-sample C(40–300), both of which have the same lowest cost level (€40), but otherwise D has higher cost levels than C. This result is similar to Glenk et al. (2019), who found a larger number of SQ choices in the average and high-cost vectors compared to the low vector. There is no significant difference between B(5–500) and D(40–500), where the highest bid is the same for both treatments, but B has bids that are smaller than the SQ cost. Detailed test results are provided in Appendix B.[8]

Considering the results, it appears that a higher cost vector and the presence of higher bids in the vector increases the proportion of systematic SQ choices, as well as choosing SQ in any of the choice tasks. Thus, we have some evidence to reject hypothesis 1 and conclude that there appear to be significant differences in the proportion of SQ choices across sub-samples. However, there is no consistent evidence that the presence of a lower cost for the policy alternative than the status quo cost level affects the proportion of SQ choices.

Hypothesis 2. There are no significant differences in always choosing the lowest/highest cost level across cost vector sub-samples.

Second, we examined the systematic choice of the highest and lowest cost alternative across cost vector sub-samples (Table 6).[9] The differences in the proportion of respondents systematically choosing the highest cost alternative in all choice tasks is not significant across cost vector sub-samples according to the chi^2 and two-sample proportion tests ($chi^2 = 0.3$, $p = 0.968$). The proportion of lowest cost choices differs significantly only across treatments B(5–500) and D(40–500) ($z = 2.5$, $p = 0.01$), with significantly fewer

[8] Examining only those respondents who faced the highest cost level (either 300 or 500€) in the choice experiment produced similar results. For systematic SQ choices, the only significant difference was between sub-samples A and D at the 10% level ($z = 1$, $p = 0.052$). For choosing SQ in any of the choice tasks, significant differences were found for pairs: A(5–300)–B(5–500) ($z = -2.9$, $p = 0.004$), A(5–300)–C(40–300) ($z = -1.7$, $p = 0.09$), A(5–300)–D(40–500) ($z = -3.5$, $p < 0.001$), and C(40–300)–D(40–500) ($z = 1.9$, $p = 0.06$).

[9] As the cost of the status quo alternative is positive, the status quo alternative is also included in this examination.

respondents always choosing the lowest cost in treatment B(5–500). Vectors B and D have the same maximum cost level, but B includes two bids that are lower than the SQ cost. Thus, there is no consistent evidence that offering lower bid levels (having a lower scale in the cost vector) reduces the propensity to systematically choose the lowest cost alternative, particularly as the differences between A and B or C and D are not significant (detailed test results are provided in Appendix B). Thus, we cannot provide evidence to reject hypothesis 2, which states that there are no significant differences in always choosing the lowest/highest cost level across cost vectors.[10]

Hypothesis 3. Respondents are sensitive to absolute cost levels.

Table 7 presents the proportions of the cost level choices for the policy alternatives. As expected, acceptance rates decline as the cost increases in all treatments. However, acceptance rates are still relatively high, being around 25% for the highest cost level of €300 (A, C) and around 20% for the highest cost level of €500 (B, D).

We examined the importance of absolute and relative cost levels by testing whether there is a difference in the choice proportions for 1) a specific cost level but different absolute values of cost (hypothesis 3a) and 2) for the same absolute value of cost but different cost levels (hypothesis 3b) using a two-sample test of proportions for each cost vector pair. We find some evidence that the absolute cost level matters for choices. Choice proportions for the highest cost levels are significantly lower for the cost vectors B(5–500) and D (40–500), which have €500 as the highest cost, compared to A(5–300) and C(40–300), which have €300 as the highest cost (detailed results are provided in Appendix B). In several cases, there is a significant difference in the proportion of choices across cost vectors, with a lower proportion of choices for higher absolute cost levels in vectors B, C and D compared to A. The results show that the proportion of lowest cost level choices is significantly smaller in treatments A and B (€5) than C and D (€40). This probably results from having several low cost levels in sub-samples A and B, including lower cost levels for the policy alternatives than the status quo. There is some evidence to support hypothesis 3 and reject $H0_3$, which states that respondents are insensitive to absolute cost levels.

Testing the choice proportions for the same absolute value of cost (but different cost level) across the cost vectors reveals that there are significant differences in the majority of the cases (detailed results are provided in Appendix B). In general, cost levels of a specific value are chosen more often in those cost vectors that have overall higher cost levels (e.g., the choice proportion for cost level €60 is 43% in A and 55% in C). Thus, relative cost levels also seem to affect choices. This result is in accordance with the findings of Kragt (2013), Luisetti et al. (2011) and Glenk et al. (2019), who suggest that relative costs may be more important than absolute costs.

Fig. 1 presents the bid acceptance for the four cost vectors at different cost levels for the policy alternatives (cumulative probability of choosing the alternative with a specific bid level when presented to the respondent). It confirms the finding that higher cost levels are accepted less often than lower levels within each sub-sample (Kragt, 2013; Glenk et al., 2019).

All in all, there is no consistent evidence that the cost vector affects the employment of systematic decision strategies. The results, for the most part, support the rejection of hypothesis 1, as the proportion of SQ choice appears to differ across cost vectors. This can partly be explained by having a positive cost for the SQ alternative and including lower than SQ costs for the policy alternatives. There is no consistent evidence to reject hypothesis 2. Our data provide limited evidence that the cost vector affects the employment of systematic decision strategies for lowest cost level choices, but there is no similar support for systematic highest cost choices. There is some evidence to support hypothesis 3, as we find some evidence that both the absolute and relative cost level matters for choices. Choice proportions are lower for higher costs, but in some cases, specific cost values are chosen more often in higher cost vectors.

5.3. Model results

To test hypothesis 4a, we estimated two sets of mixed logit models: separate models for each cost vector treatment and pooled models including two treatments with interactions.

Sub-sample specific mixed logit model results are presented in Table 8. Overall, the coefficients are significant and of the expected sign. The coefficient for the cost attribute is positive as the cost is lognormal, and significant for all treatments. For sub-samples A and B, an interaction term (*costint*) was included to reveal potential asymmetric sensitivity to cost increases and decreases, specified as the product of the cost level and a binary variable indicating that the cost was lower than the status quo cost. The interaction term is significant in the B model. As the mean cost for sub-sample B is −0.051, the size of interaction term (0.024) is considerable, and the sign of the interaction term implies that sensitivity to costs is lower for cost decreases than for increases.[11] Although the interaction term is not significant in sub-sample A, its standard deviation is large and statistically significant, suggesting a significant effect of a cost decrease relative to status quo cost for some part of the sample.

Alternative-specific constants are specified for the policy alternatives separately. For sub-samples C(40–300) and D(40–500), they indicate that respondents have a tendency to choose policy alternatives over the status quo. All agri-environmental attributes are consistently significant and positive, except for the plants in cultivation attribute and the lower level of the grazing animals attribute. In general, larger changes in the environmental attributes have a higher impact on utility, as expected. Standard deviations indicate significant individual preference heterogeneity for most of the attributes in all sub-samples.

The pooled mixed logit models use the data from two cost vector sub-samples to enable pairwise comparisons (Table 9). The models

[10] Robustness of the test results for hypotheses 1 and 2 was examined with a latent class clustering model that grouped respondents based on their systematic status quo, highest and lowest bid choices. The model results were similar to the test results.

[11] Note that as cost is lognormally distributed and lognormal distribution is limited to the positive domain, the variables cost and costint enter the model as negatives of the original variables, i.e. as mcost = −1 * cost and mcostint = −1 * costint. The mean cost is −0.051 without the interaction term, and −0.027 with the interaction term. See footnote 4 for the calculation of mean cost.

Table 7
Cost acceptance proportions (number of alternatives with the cost level in parentheses).

Proportion chosen (%)				
Cost level	A(5–300)	B(5–500)	C(40–300)	D(40–500)
1	52.54 (1045)	52.14 (1028)	60.89 (1033)	56.26 (1031)
2	54.06 (862)	52.10 (858)	54.55 (858)	55.27 (863)
3	44.62 (874)	41.57 (854)	37.91 (860)	39.09 (857)
4	43.24 (865)	39.46 (854)	43.79 (861)	35.28 (856)
5	33.72 (866)	26.64 (852)	27.24 (859)	27.17 (854)
6	33.37 (863)	31.93 (858)	28.47 (857)	29.32 (863)
7	24.16 (865)	20.49 (864)	24.88 (864)	20.74 (868)

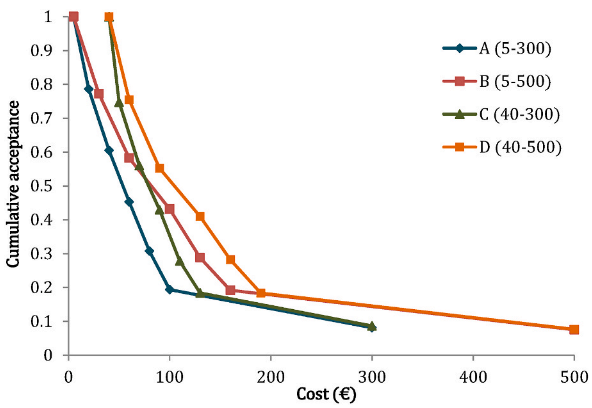

Fig. 1. Bid acceptance at different costs for the four sub-samples.

allow the coefficient means to vary across the cost vector treatments by specifying interaction terms between each variable and the treatment. Table 10 summarizes information on the significant interactions in the pooled models. The pairwise comparisons indicate that the main differences across treatments are in the cost coefficient and the ASCs. The interactions for environmental attributes are, in general, not significant (hypothesis 4a). The only exception is the pair B(5–500)–D(40–500). For D, respondents derive higher utility from the policy alternatives, but lower utility from the environmental attributes compared to B. Combined with the higher sensitivity to costs in D, this results in lower welfare estimates for the environmental attributes compared to B. One possible explanation for the differences is that while both cost vectors have the same highest bid, cost vector B has smaller bids than the cost of the status quo alternative at the lower end, while the lowest bid in D is the status quo cost (€40). Thus, policy alternatives with higher levels of environmental attributes and a lower cost compared to the status quo in B, while in D, respondents have been able to choose policy alternatives at the same cost as the status quo alternative. This may have resulted in the choice behaviour and differences observed. However, similar findings are not true for the pair A(5–300)–C(40–300), which has the same setup (the same highest bid, with A having lower bids than the status quo cost).

Except for one cost vector pair, there appears to be no consistent evidence that coefficient estimates for the environmental attributes would be influenced by the cost vector, i.e., hypothesis 4a cannot be rejected. However, for several pairs, there are significant differences in the cost coefficients and thus the sensitivity to costs across treatments. Cost vector B(5–500) produces the lowest sensitivity to costs compared to all other cost vectors and, as a consequence, is likely to lead to the highest marginal WTP estimates.

5.4. *Welfare estimates*

Marginal willingness to pay estimates (Table 11) from the sub-sample-specific mixed logit models (presented in Table 8) provide additional insights into the possible cost vector effects and enable the testing of hypothesis 4b. The WTP estimates were calculated using the delta method with the *nlcom* command in Stata. The significance of the differences in marginal WTP across the cost vectors was tested using the complete combinatorial method, which calculates every possible difference between the two empirical distributions and enables testing whether the WTP distribution from one sub-sample is higher than the other (Poe et al., 2005). The tests were conducted in a pairwise manner for all six combinations using a user-written command *poetest* in Stata (Sagebiel 2022), and results are presented in Table 12.

The marginal WTP estimates for the environmental attributes are significantly higher in treatment B compared to the other sub-samples. This applies both when the WTP estimates in sub-sample B are based on increases and decreases in costs for the policy alternatives. Although the marginal WTPs are considerably higher when based on a decrease in costs, the estimates are less often significant, and when significant, only at the 10% level. The confidence interval of the WTP estimates based on cost decreases is wider

Table 8
Mixed logit model results. Cost is log-normally distributed. Variables are significant at the *** 1%, ** 5% and * 10% level.

Cost vector	A(5–300)		B(5–500)		C(40–300)		D(40–500)	
Variable	Coeff.	Z-statistic	Coeff.	Z-statistic	Coeff.	Z-statistic	Coeff.	Z-statistic
cost	−4.430***	−22.01	−4.766***	−29.60	−4.115***	25.2	−4.346***	32.53
costint	−0.024	−1.57	−0.024***	−2.74				
ASC2	0.253	−0.50	−0.504	−1.18	2.112***	6.22	2.506***	6.32
ASC3	−0.803	−1.53	−1.200***	−2.80	1.367***	4.08	1.852***	4.82
traditional30	1.512***	−5.65	1.352***	−6.65	0.659***	4.53	0.417***	2.66
traditional60	1.568***	−5.23	1.696***	−7.09	0.709***	4.32	0.838***	4.65
animals1	0.397*	−1.70	0.355*	−1.94	0.164	1.12	−0.020	0.12
animals2	0.886***	−3.35	1.001***	−4.83	0.679***	4.29	0.323*	1.87
plants4	−0.08	−0.39	−0.176	−1.04	−0.202	1.61	−0.277*	1.92
plants5	0.062	−0.31	0.134	−0.77	0.058	0.46	0.009	0.06
climate10	1.084***	−4.14	1.001***	−4.73	0.652***	4.20	0.263	1.53
climate30	1.224***	−4.27	0.961***	−4.32	0.598***	3.60	0.552***	3.14
waterqual70	0.894***	−3.53	0.765***	−3.78	0.292**	1.98	0.347**	2.11
waterqual80	1.354***	−5.06	1.073***	−5.29	0.679***	4.41	0.597***	3.73
Standard deviations								
cost	2.038***	−10.08	1.891***	−12.77	2.104***	13.69	1.957***	14.78
costint	0.115***	−4.27	0.078***	−5.23				
ASC2	2.443***	−8.18	1.916***	−8.89	−0.841***	3.91	1.267***	5.78
ASC3	2.284***	−7.39	2.223***	−9.18	1.286***	7.77	1.653***	8.45
traditional30	2.282***	−6.92	1.373***	−5.35	1.115***	5.62	1.251***	5.49
traditional60	3.205***	−7.58	2.203***	−6.85	−1.082***	4.61	1.282***	5.02
animals1	1.549***	−4.87	0.12***	−0.22	0.892***	4.00	0.944***	3.69
animals2	2.323***	−6.70	1.365***	−5.40	−0.942***	4.48	1.283***	5.78
plants4	1.641***	−4.29	0.938***	−3.08	−0.382	0.98	0.997***	4.00
plants5	1.719***	−5.69	1.735***	−6.63	0.759***	3.62	0.775***	3.12
climate10	1.589***	−4.07	1.083***	−3.98	0.697***	2.63	1.145***	4.77
climate30	2.243***	−6.20	1.812***	−6.16	1.305***	6.46	1.131***	4.00
waterqual70	−1.094***	−2.89	−0.828**	−2.13	0.707***	2.58	0.538	1.35
waterqual80	1.173***	−2.67	1.363***	−4.84	−0.840***	3.82	−0.534	1.38
Model statistics								
N (observations)	9 360		9 252		9 288		9 288	
Log-likelihood	−2 579.16		−2 577.19		−2 467.58		−2 459.54	
BIC	5 414.36		5 410.098		5 172.7		5 156.635	
AIC	5 214.32		5 210.385		4 987.152		4 971.086	

than for increases, due to the smaller number of observations and a large standard deviation of the cost decrease interaction term in the model. Taking this into consideration, we use the estimates based on cost increases for comparisons across sub-samples.

Comparison of A and B indicates that higher cost levels lead to higher marginal WTP estimates when lower than status quo costs are present in the cost vector for the policy alternatives. The same is not observed when comparing WTP estimates based on C and D. Sub-sample B leading to higher WTPs than C and D could potentially result from differences in whether lower cost levels than the status quo are available, rather than other differences in the cost vectors. The welfare impact of the alternatives not explained by the attributes differs between B and both C and D - WTP for the second policy alternative (ASC3) is negative in B but positive in C and D. This result is surprising, as the only difference in the policy alternatives across sub-samples stems from the cost and B has lower cost levels than the status quo. Considering all of the above, the results provide some evidence for rejecting hypothesis 4b on the similarity of marginal WTPs, for the case where cost decreases from selecting policy alternatives compared to the status quo are possible.

Table 13 presents welfare estimates for two policy scenarios: moderate, indicating a one-step improvement in all environmental attributes and high, indicating two-step improvement in all attributes. Cost vector B produces the highest average welfare estimates, followed by A, D and C. The estimates from B are 2–3 times larger compared to C. The confidence intervals are quite wide and, in most cases, overlapping, which suggests that there might not be a significant difference in the policy scenario welfare estimates. It is surprising that the scenario welfare estimates are lower than the lowest (status quo) cost for C. These results would indicate that respondents would need reductions in current payments even though there were improvements in the agri-environmental attributes. This is likely influenced by the log-normal cost parameter, which often results in lower marginal mean WTP estimates than other distributions (see e.g. Glenk et al., 2019).[12]

6. Discussion and conclusions

This paper examines cost vector effects in discrete choice experiments. The empirical study focuses on changes in agri-

[12] Using a non-random cost parameter results in larger marginal WTP estimates for the environmental attributes and some additional significant differences in the marginal WTP estimates between sub-samples, but the overall results on the relative magnitude of the WTP estimates and significance of differences are robust to the assumption on the cost parameter.

Table 9
Pooled mixed logit model results. Cost is log-normally distributed. Variables are significant at the *** 1%, ** 5% and * 10% level.

Variable	A(5–300)–B(5–500) (base model = A)		A(5–300)–C(40–300) (base model = A)		A(5–300)–D(40–500) (base model = A)		B(5–500)–C(40–300) (base model = B)		B(5–500)–D(40–500) (base model = B)		C(40–300)–D(40–500) (base model = C)	
	Coefficient	Z-statistic	Coefficient	Z-statistic	Coefficient	Z-statistic	Coefficient	Z-statistic	Coefficient	Z-statistic	Coefficient	Z-statistic
cost*sub-sample	−0.003**	2.11	0.003	1.42	−0.001	0.42	0.003*	1.66	0.0001	0.12	−0.002*	1.86
ASC2*sub-sample	−0.971*	1.65	2.034***	3.47	2.426***	3.98	2.640***	4.92	2.959***	5.3	0.36	0.67
ASC3*sub-sample	−0.828	1.43	1.939***	3.35	2.516***	4.17	2.400***	4.51	2.941***	5.29	0.488	0.9
traditional30*sub-sample	0.117	0.47	−0.251	1.00	−0.649**	2.47	−0.311	1.38	−0.703***	2.98	−0.364	1.59
traditional60*sub-sample	0.430	1.44	−0.397	1.35	−0.449	1.48	−0.684**	2.54	−0.736***	2.65	0.038	0.14
animals1*sub-sample	0.039	0.15	−0.017	0.07	−0.333	1.23	−0.164	0.7	−0.415*	1.68	−0.316	1.29
animals2*sub-sample	0.348	1.25	0.170	0.62	−0.345	1.21	−0.153	0.61	−0.641**	2.5	−0.465*	1.88
plants4*sub-sample	−0.033	0.14	−0.271	1.22	−0.263	1.14	−0.225	1.11	−0.207	0.98	−0.012	0.06
plants5*sub-sample	0.132	0.57	−0.005	0.02	−0.007	0.03	−0.176	0.83	−0.182	0.85	−0.02	0.1
climate10*sub-sample	0.231	0.83	−0.013	0.05	−0.525*	1.87	−0.124	0.5	−0.629**	2.44	−0.393	1.61
climate30*sub-sample	0.146	0.51	−0.142	0.5	−0.399	1.39	−0.225	0.86	−0.42	1.6	−0.148	0.59
waterqual70*sub-sample	0.166	0.63	−0.218	0.84	−0.219	0.83	−0.349	1.45	−0.39	1.6	0.005	0.02
waterqual80*sub-sample	0.106	0.40	−0.063	0.24	−0.325	1.25	−0.159	0.67	−0.436*	1.82	−0.214	0.92
cost	−4.394***	33.2	−4.378***	29.47	−4.269***	33.82	−4.568***	35.04	−4.487***	40.16	−4.023***	35.84
ASC2	0.46	1.12	0.41	1.03	0.385	0.95	−0.5	1.36	−0.464	1.25	2.274***	5.99
ASC3	−0.379	0.93	−0.425	1.08	−0.479	1.18	−1.113***	3	−1.126***	2.98	1.437***	3.81
traditional30	1.137***	6.27	1.098***	5.96	1.124***	6.04	1.137***	6.86	1.183***	6.92	0.753***	4.64
traditional60	1.230***	5.74	1.229***	5.93	1.302***	6.13	1.489***	7.65	1.575***	7.88	0.782***	4.27
animals1	0.345*	1.93	0.264	1.48	0.316*	1.73	0.412**	2.52	0.440***	2.61	0.207	1.25
animals2	0.718***	3.66	0.719***	3.72	0.699***	3.51	0.982***	5.46	1.022***	5.61	0.792***	4.42
plants4	−0.044	0.28	−0.049	0.32	−0.038	0.24	−0.046	0.32	−0.08	0.53	−0.276*	1.92
plants5	0.071	0.43	0.062	0.4	0.074	0.47	0.2	1.32	0.18	1.19	0.039	0.28
climate10	0.775***	3.92	0.822***	4.24	0.789***	3.93	0.895***	5.04	0.918***	4.98	0.727***	4.13
climate30	0.890***	4.31	0.890***	4.4	0.961***	4.65	0.914***	4.79	0.994***	5.2	0.687***	3.78
waterqual70	0.685***	3.64	0.619***	3.36	0.645***	3.47	0.771***	4.39	0.806***	4.61	0.351**	2.11
waterqual80	1.031***	5.36	0.983***	5.32	1.004***	5.39	1.042***	5.99	1.073***	6.22	0.798***	4.69
Standard deviations												
cost	1.709***	14.78	2.162***	15.01	1.838***	14.72	2.127***	17.88	1.824***	17.36	1.939***	16.95
ASC2	1.984***	12.81	1.625***	11.22	1.884***	11.57	1.485***	10.96	1.659***	11.56	1.085***	6.5
ASC3	1.991***	12.64	1.689***	10.96	1.926***	11.67	1.753***	12.19	1.926***	13.09	1.647***	11.39
traditional30	1.431***	8.1	1.609***	9.47	1.631***	9.15	1.244***	7.67	1.266***	7.48	1.360***	8.08
traditional60	2.227***	10.6	2.015***	9.8	2.143***	9.66	1.711***	9.31	1.762***	9.16	−1.378***	7.22
animals1	0.788***	3.37	1.236***	6.42	−1.237***	6.17	0.618**	2.23	0.868***	4.15	−1.156***	6.63
animals2	1.585***	9.3	1.556***	9.04	1.739***	10.13	1.262***	7.99	1.291***	7.78	1.208***	7.54
plants4	1.052***	4.68	1.029***	5.23	1.138***	5.9	−0.715***	3.43	1.043***	5.59	−0.872***	4.88
plants5	1.442***	8.57	1.160***	7.03	1.239***	7.27	1.246***	8.26	1.185***	7.47	0.975***	6.27
climate10	1.279***	6.99	1.147***	6.29	1.269***	6.78	1.060***	6.12	1.184***	6.53	0.980***	5.47
climate30	1.714***	8.61	1.566***	8.31	1.511***	7.43	1.583***	8.99	1.320***	7.14	−1.356***	7.7
waterqual70	0.811***	3.14	1.095***	4.89	−0.720***	2.69	0.880***	4.07	0.531*	1.72	−0.827***	4.26
waterqual80	1.308***	6.89	1.172***	5.93	−0.842**	2.54	1.023***	5.79	0.894***	4.25	0.881***	4.87
Model statistics												
N (observations)	18 612		18 648		18 648		18 540		18 540		18 576	
Log-likelihood	−5 188.35		−5 077.83		−5 058.61		−5 088.65		−5 068.01		−4 918.82	
BIC	10760.13		10539.17		10500.72		10560.58		10519.29		10 221	
AIC	10454.7		10233.66		10195.22		10255.3		10214.01		9 915.648	

Table 10
Significant interactions in the pooled models. Interactions are significant at the *** 1%, ** 5% and * 10% level.

Interaction coefficient	A(5–300)– B (5–500) (base model = A)	A(5–300)–C (40–300) (base model = A)	A(5–300)–D (40–500) (base model = A)	B(5–500)–C (40–300) (base model = B)	B(5–500)–D (40–500) (base model = B)	C(40–300)–D (40–500) (base model = C)
cost	**			*		*
ASC2	*	***	***	***	***	
ASC3		***	***	***	***	
traditional30			**		***	
traditional60				***	***	
animals1					*	
animals2					**	
plants4						
plants5						
climate10			*		**	
climate30						
waterqual70						
waterqual80					*	

Table 11
Mean willingness to pay estimates from mixed logit models per person and year in 2019 euros (95% confidence intervals in brackets). NS = not significant at the 90% level.

	A(5–300)	B(5–500) – increase in cost	B(5–500) – decrease in cost	C(40–300)	D(40–500)
ASC2	NS	NS	NS	14.1 [5.5; 22.8]	28.5 [13.9; 43.1]
ASC3	NS	−23.6 [-42.8; −4.5]	NS	9.2 [2.7; 15.6]	21.1 [8.9; 33.2]
traditional30	15.9 [5.4; 26.4]	26.6 [12.7; 40.5]	50.5 [-5.9; 106.8]	4.4 [1.3; 7.5]	4.7 [0.6; 8.9]
traditional60	16.5 [5.5; 27.5]	33.4 [16.6; 50.1]	63.4 [-6.3; 133.0]	4.7 [1.3; 8.1]	9.5 [3.7; 15.4]
animals1	NS	7.0 [-0.9; 14.9]	NS	NS	NS
animals2	9.3 [1.9; 16.8]	19.7 [8; 31.4]	37.4 [-4.6; 79.4]	4.5 [1.3; 7.8]	3.7 [-0.5; 7.8]
plants1	NS	NS	NS	NS	−3.1 [-6.6; 0.3]
plants2	NS	NS	NS	NS	NS
climate10	11.4 [3.5; 19.3]	19.7 [8.3; 31.1]	37.4 [-5.3; 80.1]	4.4 [1.3; 7.5]	NS
climate30	12.9 [3.7; 22]	18.9 [7.3; 30.5]	35.9 [-5.0; 76.8]	4.0 [0.9; 7.1]	6.3 [1.5; 11.1]
waterqual70	9.4 [2.2; 16.6]	15.0 [4.9; 25.1]	28.6 [-4.4; 61.6]	2.0 [-0.2; 4.1]	3.9 [-0.1; 8.0]
waterqual80	14.3 [4.5; 24]	21.1 [9.2; 33.0]	40.1 [-4.7; 84.9]	4.5 [1.4; 7.7]	6.8 [2.2; 11.4]

Table 12
Differences in willingness to pay distributions across cost vectors. Differences are significant at the *** 1%, ** 5% and * 10% level, and - indicates a nonsignificant difference. A result is not reported if the WTP estimate from either of the sub-sample models is nonsignificant.

	A(5–300)–B (5–500)	A(5–300)–C (40–300)	A(5–300)–D (40–500)	B(5–500)–C (40–300)	B(5–500)–D (40–500)	C(40–300)–D (40–500)
ASC2						–
ASC3				*** (C higher)	*** (D higher)	* (D higher)
traditional30	*** (B higher)	** (A higher)	–	*** (B higher)	** (B higher)	–
traditional60	*** (B higher)	* (A higher)	–	*** (B higher)	–	–
animals1						
animals2	***	–	* (A higher)	*** (B higher)	*** (B higher)	–
plants1						
plants2						
climate10	*** (B higher)	–		** (B higher)-		
climate30	** (B higher)	–	–	** (B higher)	* (B higher)	–
waterqual70	** (B higher)	** (A higher)	–	*** (B higher)	* (B higher)	–
waterqual80	*** (B higher)	* (A higher)	–	*** (B higher)	* (B higher)	–

environmental policies and their benefits in Finland, i.e. provision of environmental goods that are at present financed with tax revenues. We used otherwise identical surveys and experimental designs, but four different cost vectors: A(5–300), B(5–500), C (40–300) and D(40–500). Choosing the status quo alternative entails a (constant) positive cost (40€) to the respondent based on the current costs of the policy, and thus sub-samples A and B included also cost levels below the status quo cost. We examined differences in both systematic choice behaviour and welfare estimates across the sub-samples.

Our data provide some evidence that the cost vector affects the employment of systematic decision strategies, in line with previous studies (Kragt 2013; Glenk et al., 2019). The provision of lower cost levels in the cost vector (lower scale) may reduce the propensity to systematically choose the status quo alternative, as the proportion of status quo choices was higher in sub-sample D(40–500) than in sub-sample C(40–300), both of which have the same lowest cost level, but the other cost levels are higher in D. This result is similar to

Table 13
Welfare estimates for selected policy scenarios per person and year in 2019 euros (95% confidence intervals in brackets). In the calculation, mean of ASC2 and ASC3 coefficients is used and nonsignificant attribute levels are excluded. For sub-sample B, WTPs based on cost increases are used.

Scenario	Description	A(5–300)	B(5–500)	C(40–300)	D(40–500)
Moderate scenario	30% increase in traditional rural biotopes; 100 endangered species protected; grazing animals often seen in the summer; 4 plant species in cultivation; 10% decrease in GHG emissions; 70% of surface waters in good quality	40.9 [13.4; 68.4]	56.5 [27.1; 85.9]	16.7, [7.4; 25.9]	36.4, [20.0; 52.8]
High scenario	60% increase in traditional rural biotopes; 200 endangered species protected; grazing animals often seen in the summer and unfrozen season; 5 plant species in cultivation; 30% decrease in GHG emissions; 80% of surface waters in good quality	53.0 [18.9; 87.0]	81.3 [41.9; 120.7]	29.5 [13.0; 46.0]	51.0 [27.9; 74.2]

Glenk et al. (2019), who found a larger number of status quo choices in the average and high cost vectors compared to the low vector. However, there is no consistent evidence that the presence of cost decreases for policy alternatives affects the proportion of SQ choices. There was no significant difference in the proportion of respondents always choosing the highest cost alternative across treatments, and no difference in the proportion of lowest cost choices for treatments that had the same lowest cost. Further, we found no consistent evidence that having a lower scale cost vector that offers lower bid levels reduces the propensity to systematically choose the lowest cost alternative. This result is in contrast to Glenk et al. (2019), who found a significant difference in the proportion of respondents always choosing the lowest cost alternative across cost treatments.

Our results indicate that both the absolute and relative cost levels matter for choices. As expected, the proportion of the highest cost level choices was lower when the absolute cost was higher. Also for several other cost levels, a specific level was chosen less frequently when it was higher in absolute terms. We found that the proportion of lowest cost level choices was significantly smaller in treatments that had several low absolute cost levels, including those lower than the status quo cost. Consistently with earlier studies (Luisetti et al., 2011; Kragt 2013; Glenk et al., 2019), relative cost levels appear to affect choices, as the likelihood of choosing cost levels of a specific absolute value was higher in those cost vectors that have overall higher cost levels.

The results of the mixed logit models indicated that there were no significant differences across the environmental attribute coefficients across treatments, but the sensitivity to cost differed. Cost sensitivity was the lowest in B(5–500), and B produced the highest marginal WTP and scenario welfare estimates. The differences in the marginal WTP distributions for the attributes were significant between cost vector B and the other vectors, but not between A, C and D. Cost vector B has the largest range and includes cost levels also below the status quo cost. The findings also indicate that WTP estimates from B that are based on cost decreases compared to the status quo cost are larger than those based on cost increases, and the difference is considerable. This is likely one element contributing to the observed differences in WTP. The results provide some support for avoiding cost vectors with a notably large range, as that is found to lead to significant differences in marginal WTP estimates compared to the other cost vectors. The results from previous research examining the effects of the scale of the cost vector on welfare estimates have been mixed. Hanley et al. (2005) and Kragt (2013) found nonsignificant differences in marginal WTP for most or all attributes, while Carlsson and Martinsson (2008) observed significant differences in marginal WTPs between cost treatments for most attributes. Glenk et al. (2019) also found significant differences in marginal WTP after excluding respondents who systematically chose the cheapest non-status quo alternative, with higher cost vectors producing higher WTP estimates.

Examination of cost vector effects is relevant for assessing the ability of DCEs to provide reliable and valid welfare estimates to support policy decisions. A single study cannot provide guidelines that apply to all situations, but it can provide some insights into the construction of cost levels. Our examination suggests that it may be possible to reduce systematic choice behaviour by using a cost vector that includes several lower cost levels, and that both relative and absolute costs should be considered in the construction of the cost vector. According to our findings, inclusion of higher cost levels in the cost vector may, in some cases, increase the WTP estimates, as can potentially a large range of cost levels. This emphasizes the importance of selecting the cost levels, noted already in the CV setting by Kanninen and Kriström (1993). A standard procedure is to test the cost levels of the DCE in a pilot study imitating the features of the final study. We argue that this might not be sufficient, and suggest that more efforts should be placed on careful pre-testing of the cost levels to find a realistic and appropriate cost levels to produce valid welfare estimates. Although a two-step procedure is often applied in DCEs, its more rigorous implementation would be warranted, with careful literature review, qualitative pre-testing and larger sample sizes for the pilot study, combined with a clear understanding of what the changes to the cost vector may imply to the results based on existing literature on cost vector effects.

In two treatments (A and B), there were policy alternatives that had a lower cost and higher levels of environmental attributes compared to the status quo. This was justified, as a shift to a result-based policy could lower the total costs of environmental good provision, while at the same time decreasing the ambition level would not correspond well with the actual aims and development needs of the Finnish agri-environmental policy. Although this enabled the examination of cost vector effects in a setting where the status quo alternative has a positive cost and policy alternatives may result in cost decreases, it also brings about limitations. In comparisons across cost vectors with and without cost decreases, the observed impacts may have been confounded with the effect of having lower than the status quo cost levels. However, with four cost vectors, comparisons could be made for cost vector pairs either with or without lower than status quo cost levels, with no confounding effects. In addition, many of our results could be observed for several cost vector pairs.

Our approach implied that there were choice tasks where the status quo alternative could be dominated by a policy alternative. In those cases, respondents still had to make trade-offs in the choice between the two policy alternatives. The status quo was chosen only in few cases as a dominated alternative, and overall, the results are as expected with no apparent problems in the data or model results. This implies that the inclusion of alternatives with lower than status quo cost did not thwart the examination.

After this investigation, the evidence on cost vector effects remains mixed. Further research is needed to obtain a more comprehensive understanding of cost vector effects in different study settings and geographic areas. A systematic examination of the occurrence of cost vector effects in existing stated preference literature, e.g. using meta-analysis, could provide further insights on the factors that influence whether effects are observed and contexts where they are more likely. In future case studies on cost vector effects, systematic examination of the various alternative bid distributions would be needed, as both the cost range and the distribution of bids within that range may be important. An interesting future topic would also be to provide further evidence on cost vector effects when the status quo alternative has a positive cost and policy alternatives imply both cost decreases and increases.

Author statement

Heini Ahtiainen: Conceptualization, Methodology, Investigation, Software, Formal analysis, Writing - Original Draft.
Eija Pouta: Conceptualization, Methodology, Investigation, Supervision, Writing – review & editing.
Wojciech Zawadzki: Software, Formal analysis, Writing – review & editing.
Annika Tienhaara: Conceptualization, Methodology, Investigation, Writing – review & editing.

Declaration of competing interest

There are no conflicts of interest.

Data availability

Data will be made available on request.

Acknowledgements

The authors gratefully acknowledge the funding from the Development Fund for Agriculture and Forestry (MAKERA) by the Finnish Ministry of Agriculture and Forestry (MAAESP project, 1798/312/2014) and the Academy of Finland (VABARO project, grant number 310205). We thank Julian Sagebiel for providing his user-written Stata code “poetest2” for testing the differences in the marginal WTP distributions, and Mikołaj Czajkowski for guidance on modelling. We would also like to thank three anonymous reviewers for their valuable comments on the manuscript.

Appendix A. Sample characteristics by blocks

Table A1
Sub-sample A(5–300)

Characteristic	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	P-value
Number of respondents (observations)	85 (1 530)	85 (1 530)	86 (1 548)	86 (1 548)	87 (1 566)	91 (1 638)	
Share of respondents in sub-sample (%)	16.3	16.3	16.5	16.5	16.7	17.5	
Women (%)	49.4	48.2	38.4	48.8	46.0	48.4	0.700
Age group (1–5; 1 = 18–24 years old, 5 = 65–74 years old)	3.6	3.6	3.5	3.9	3.8	3.8	0.0996
Household size (people)	2.4	2.6	2.3	2.4	2.1	2.3	0.0532
University level education (%)	25.9	29.4	30.2	30.2	24.1	29.7	0.9323
Individual monthly gross income (€)	2 656	2 721	2 669	2 631	2 716	2 810	0.9841
Response time (minutes)*	20	21	23	19	22	22	0.4779
Agricultural living environment (%)	28	28	21	26	26	23	0.858
Second home in agricultural environment (%)	28	26	26	29	39	32	0.390
Easy to answer CE questions (%)	44	56	43	47	53	54	0.337
Understood CE questions (%)	74	81	78	76	80	77	0.868
Certain of CE choices (%)	53	64	56	56	53	58	0.737

* Response times over 180 min (34 respondents) were removed from the examination, as these were not considered to reflect actual response times. Differences across samples were not significant even with full samples.

Table A2
Sub-sample B(5–500)

Characteristic	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	P-value
Number of respondents (observations)	85 (1 530)	85 (1 530)	86 (1 548)	86 (1 548)	87 (1 566)	91 (1 638)	
Share of respondents in sub-sample (%)	16.3	16.3	16.5	16.5	16.7	17.5	
Women (%)	37.6	38.8	39.1	41.2	41.4	44.7	0.952

(*continued on next page*)

Table A2 (*continued*)

Characteristic	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	P-value
Age group (1–5; 1 = 18–24 years old, 5 = 65–74 years old)	3.9	3.8	3.9	3.8	3.7	3.6	0.4583
Household size (people)	2.1	2.7	2.3	2.2	2.3	2.4	0.0134
University level education (%)	30.6	21.2	32.2	21.2	33.3	22.4	0.4108
Individual monthly gross income (€)	2 726	2 794	2 868	2 606	2 744	2 441	0.5946
Response time (minutes)*	21	19	21	23	20	26	0.0264
Agricultural living environment (%)	24	27	28	27	18	34	0.311
Second home in agricultural environment (%)	31	27	34	34	32	29	0.898
Easy to answer CE questions (%)	45	46	45	44	37	54	0.376
Understood CE questions (%)	78	73	69	79	77	82	0.387
Certain of CE choices (%)	51	51	59	66	57	59	0.326

* Response times over 180 min (34 respondents) were removed from the examination, as these were not considered to reflect actual response times. Differences across samples were not significant even with full samples.

Table A3
Sub-sample C(40–300)

Characteristic	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	P-value
Number of respondents (observations)	85 (1 530)	85 (1 530)	86 (1 548)	86 (1 548)	87 (1 566)	91 (1 638)	
Share of respondents in sub-sample (%)	16.3	16.3	16.5	16.5	16.7	17.5	
Women (%)	47.1	45.9	37.9	46.5	48.8	47.1	0.760
Age group (1–5; 1 = 18–24 years old, 5 = 65–74 years old)	4.0	3.8	3.6	3.8	3.7	3.5	0.0751
Household size (people)	2.1	2.4	2.3	2.3	2.4	2.5	0.5355
University level education (%)	23.5	22.4	28.7	27.9	23.3	23.0	0.9596
Individual monthly gross income (€)	3 085	2 738	2 710	2 785	2 651	2 388	0.1833
Response time (minutes)*	21	23	22	23	22	25	0.7741
Agricultural living environment (%)	28	27	26	16	20	31	0.205
Second home in agricultural environment (%)	34	24	25	43	30	25	0.049
Easy to answer CE questions (%)	44	44	47	38	52	56	0.192
Understood CE questions (%)	74	74	71	74	80	80	0.658
Certain of CE choices (%)	46	55	59	51	58	56	0.534

* Response times over 180 min (34 respondents) were removed from the examination, as these were not considered to reflect actual response times. Differences across samples were not significant even with full samples.

Table A4
Sub-sample D(40–500)

Characteristic	Block 1	Block 2	Block 3	Block 4	Block 5	Block 6	P-value
Number of respondents (observations)	85 (1 530)	85 (1 530)	86 (1 548)	86 (1 548)	87 (1 566)	91 (1 638)	
Share of respondents in sub-sample (%)	16.3	16.3	16.5	16.5	16.7	17.5	
Women (%)	37.6	50.0	47.1	44.7	42.0	48.2	0.615
Age group (1–5; 1 = 18–24 years old, 5 = 65–74 years old)	3.9	3.4	3.7	3.6	3.6	3.9	0.0561
Household size (people)	2.2	2.3	2.1	2.2	2.4	2.4	0.8742
University level education (%)	31.8	30.2	26.4	22.4	22.7	30.6	0.594
Individual monthly gross income (€)	3 250	2 387	2 589	2 518	2 741	3 000	0.0043
Response time (minutes)*	22	22	23	24	22	22	0.9940
Agricultural living environment (%)	29	21	16	27	17	27	0.162
Second home in agricultural environment (%)	35	37	28	32	28	33	0.728
Easy to answer CE questions (%)	45	42	48	45	55	53	0.496
Understood CE questions (%)	78	78	70	73	76	82	0.510
Certain of CE choices (%)	64	49	52	54	57	56	0.492

* Response times over 180 min (34 respondents) were removed from the examination, as these were not considered to reflect actual response times. Differences across samples were not significant even with full samples.

Appendix B. Supplementary test results

Table B1
Test results on choice behaviour across cost vector pairs

Tested behaviour	Z-statistic					
	A(5–300)–B (5–500)	A(5–300)–C (40–300)	A(5–300)–D (40–500)	B(5–500)–C (40–300)	B(5–500)–D (40–500)	C(40–300)–D (40–500)
Systematic SQ choices	−1.52	−0.32	−2.54	1.20	−1.04	−2.23
SQ choices in any tasks	−5.84	−1.81	−5.94	4.04	−0.09	−4.14

(*continued on next page*)

Table B1 (*continued*)

Tested behaviour	Z-statistic					
	A(5–300)–B (5–500)	A(5–300)–C (40–300)	A(5–300)–D (40–500)	B(5–500)–C (40–300)	B(5–500)–D (40–500)	C(40–300)–D (40–500)
Systematic highest cost choices	0.13	−0.18	0.31	−0.31	0.18	0.49
Systematic lowest cost choices	1.01	−0.14	−1.50	−1.15	−2.49	−1.35

Table B2

Test results on the differences in proportions of the cost level chosen across cost vector pairs

Cost level		Z-statistic				
	A(5–300)–B(5–500)	A(5–300)–C(40–300)	A(5–300)–D(40–500)	B(5–500)–C(40–300)	B(5–500)–D(40–500)	C(40–300)–D(40–500)
1	0.18	−3.84	−1.70	−4.01	−1.87	−2.14
2	0.81	−0.20	−0.51	−1.02	−1.32	0.30
3	1.28	2.84	2.33	1.55	1.05	0.50
4	1.59	−0.23	3.38	−1.82	1.79	−3.60
5	3.19	2.92	2.95	−0.28	−0.24	−0.03
6	0.64	2.20	1.82	1.56	1.18	0.39
7	1.84	−0.35	1.71	−2.18	−0.13	−2.06

Table B3

Test results on the differences in proportions of the absolute cost chosen across cost vector pairs

Cost (€)	Z-statistic					
	A(5–300)–B(5–500)	A(5–300)–C(40–300)	A(5–300)–D(40–500)	B(5–500)–C(40–300)	B(5–500)–D(40–500)	C(40–300)–D(40–500)
40		−7.10	−5.06			
60	0.70		−5.00		−5.68	
100	−2.62					
130				−0.85	−3.86	3.02
160					2.16	

References

Aravena, C., Martinsson, P., Scarpa, R., 2014. Does money talk?—the effect of a monetary attribute on the marginal values in a choice experiment. Energy Econ. 44, 483–491.

Ariely, D., Loewenstein, G., Prelec, D., 2003. Coherent arbitrariness": stable demand curves without stable preferences. Q. J. Econ. 118 (1), 73–106.

Bhat, C.R., 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. Transp. Res. Part B Methodol. 35 (7), 677–693.

Bhat, C.R., 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transp. Res. Part B Methodol. 37 (9), 837–855.

Bryan, S., Buxton, M., Sheldon, R., Grant, A., 1998. Magnetic resonance imaging for the investigation of knee injuries: an investigation of preferences. Health Econ. 7 (7), 595–603.

Carlsson, F., Frykblom, P., Lagerkvist, C.J., 2007. Preferences with and without prices-does the price attribute affect behavior in stated preference surveys? Environ. Resour. Econ. 38 (2), 155–164.

Carlsson, F., Martinsson, P., 2008. How much is too much? An investigation of the effect of the number of choice sets, context dependence and the choice of bid vectors in choice experiments. Environ. Resour. Econ. 40 (2), 165–176.

Cooper, J., Loomis, J., 1992. Sensitivity of willingness-to-pay estimates to bid design in dichotomous choice contingent valuation models. Land Econ. 68 (2), 211–224.

Czajkowski, M., Budziński, W., 2019. Simulation error in maximum likelihood estimation of discrete choice models. J. Choice Model. 31, 73–85.

Daly, A., Hess, S., Train, K., 2012. Assuring finite moments for willingness to pay in random coefficient models. Transportation 39 (1), 19–31.

Essers, B.A., van Helvoort-Postulart, D., Prins, M.H., Neumann, M., Dirksen, C.D., 2010. Does the inclusion of a cost attribute result in different preferences for the surgical treatment of primary basal cell carcinoma? Pharmacoeconomics 28 (6), 507–520.

Frykblom, P., Shogren, J.F., 2000. An experimental testing of anchoring effects in discrete choice questions. Environ. Resour. Econ. 16 (3), 329–341.

Furnham, A., Boo, H.C., 2011. A literature review of the anchoring effect. J. Soc. Econ. 40 (1), 35–42.

Genie, M.G., Ryan, M., Krucien, N., 2021. To pay or not to pay? Cost information processing in the valuation of publicly funded healthcare. Soc. Sci. Med. 276, 113822.

Glenk, K., Meyerhoff, J., Akaichi, F., Martin-Ortega, J., 2019. Revisiting cost vector effects in discrete choice experiments. Resour. Energy Econ. 57, 135–155.

Hanley, N., Adamowicz, W., Wright, R.E., 2005. Price vector effects in choice experiments: an empirical test. Resour. Energy Econ. 27 (3), 227–234.

Hensher, D.A., Greene, W.H., 2003. The mixed logit model: the state of practice. Transportation 30, 133–176.

Hensher, D.A., Rose, J.M., Greene, W.H., 2015. Applied Choice Analysis, second ed. Cambridge University Press, Cambridge, UK.

Johnson, F.R., Lancsar, E., Marshall, D., Kilambi, V., Mühlbacher, A., Regier, D.A., Bresnahan, B.W., Kanninen, B., Bridges, J.F., 2013. Constructing experimental designs for discrete-choice experiments: report of the ISPOR conjoint analysis experimental design good research practices task force. Value Health 16 (1), 3–13.

Johnston, R.J., Boyle, K.J., Adamowicz, W., Bennett, J., Brouwer, R., Cameron, T.A., Hanemann, W.M., Hanley, N., Ryan, M., Scarpa, R., Tourangeau, R., 2017. Contemporary guidance for stated preference studies. J. Assoc. Environ. Res. Economic. 4 (2), 319–405.

Kanninen, B.J., Kriström, B., 1993. Sensitivity of willingness-to-pay estimates to bid design in dichotomous choice valuation models: comment. Land Econ. 69 (2), 199–202.

Kragt, M.E., 2013. The effects of changing cost vectors on choices and scale heterogeneity. Environ. Resour. Econ. 54 (2), 201–221.

Ladenburg, J., Olsen, S.B., 2008. Gender-specific starting point bias in choice experiments: evidence from an empirical study. J. Environ. Econ. Manag. 56 (3), 275–285.

Liesivaara, P., Myyrä, S., 2014. Willingness to pay for agricultural crop insurance in the northern EU. Agric. Finance Rev. 74 (4), 539–554.

Luisetti, T., Bateman, I.J., Turner, R.K., 2011. Testing the fundamental assumption of choice experiments: are values absolute or relative? Land Econ. 87 (2), 284–296.
Meyerhoff, J., Glenk, K., 2015. Learning how to choose—effects of instructional choice sets in discrete choice experiments. Resour. Energy Econ. 41, 122–142.
Mitchell, R.C., Carson, R.T., 1989. Using Surveys to Value Public Goods: the Contingent Valuation Method. Resources for the Future, Washington, D.C. United States.
Mørkbak, M.R., Christensen, T., Gyrd-Hansen, D., 2010. Choke price bias in choice experiments. Environ. Resour. Econ. 45 (4), 537–551.
Niskanen, O., Tienhaara, A., Haltia, E., Pouta, E., 2021. Farmers' heterogeneous preferences towards results-based environmental policies. Land Use Pol. 102 https://doi.org/10.1016/j.landusepol.2020.105227.
Pedersen, L.B., Kjær, T., Kragstrup, J., Gyrd-Hansen, D., 2011. Does the inclusion of a cost attribute in forced and unforced choices matter? Results from a web survey applying the discrete choice experiment. J. Choice Model. 4 (3), 88–109.
Penn, J.M., Hu, W., 2018. Understanding hypothetical bias: an enhanced meta-analysis. Am. J. Agric. Econ. 100 (4), 1186–1206.
Poe, G., Giraud, K., Loomis, J., 2005. Computational methods for measuring the difference of empirical distributions. Am. J. Agric. Econ. 87 (2), 353–365.
Ryan, M., Wordsworth, S., 2000. Sensitivity of willingness to pay estimates to the level of attributes in discrete choice experiments. Scot. J. Polit. Econ. 47 (5), 504–524.
Revelt, D., Train, K., 1998. Mixed logit with repeated choices: households' choices of appliance efficiency level. Rev. Econ. Stat. 80 (4), 647–657.
Rose, J.M., Bliemer, M.C., 2009. Constructing efficient stated choice experimental designs. Transport Rev. 29 (5), 587–617.
Sagebiel, J., 2022. Poetest Stata codes. Available at: https://github.com/sagebiej/poetest_stata. (Accessed 26 November 2022). Updated 26.11.2022.
Sever, I., Verbič, M., Sever, E.K., 2019. Cost attribute in health care DCEs: just adding another attribute or a trigger of change in the stated preferences? J. Choice Model. 32, 100135.
Svenningsen, L.S., Jacobsen, J.B., 2018. Testing the effect of changes in elicitation format, payment vehicle and bid range on the hypothetical bias for moral goods. J. Choice Model. 29, 17–32.
Taloustutkimus, 2017. Internet panel. Available at: http://www.taloustutkimus.fi/in-english/products_services/internet_panel/. (Accessed 25 January 2017).
Tienhaara, A., Ahtiainen, H., Pouta, E., 2021. Heterogeneous preferences of citizens towards agricultural ecosystem services: the demand relevance in a choice experiment. Ecosyst. People 17 (1), 114–127.
Tienhaara, A., Haltia, E., Pouta, E., Arovuori, K., Grammatikopoulou, I., Miettinen, A., Koikkalainen, K., Ahtiainen, H., Artell, J., 2020. Demand and supply of agricultural ES: towards benefit-based policy. Eur. Rev. Agric. Econ. 47 (3), 1223–1249.
Train, K., 2000. Halton Sequences for Mixed Logit. Department of Economics, UCB, 2000.
Train, K., 2003. Discrete choice methods with simulation. Cambridge University Press, United Kingdom.
Tversky, A., Kahneman, D., 1974. Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty. Science 185 (4157), 1124–1131.
Vainio, A., Tienhaara, A., Haltia, E., Hyvönen, T., Pyysiäinen, J., Pouta, E., 2021. The legitimacy of result-oriented and action-oriented agri-environmental schemes: a comparison of farmers' and citizens' perceptions. Land Use Pol. 107, 104368.
van Zanten, B.T., Koetse, M.J., Verburg, P.H., 2016. Economic valuation at all cost? The role of the price attribute in a landscape preference study. Ecosyst. Serv. 22, 289–296.
Vossler, C.A., Poe, G.L., Welsh, P., Ethier, R.G., 2004. Bid design effects in multiple bounded discrete choice contingent valuation. Environ. Resour. Econ. 29 (4), 401–418.