2014

01

Working Paper

INSTITUTO DE POLÍTICAS Y BIENES PÚBLICOS (IPP)

COMPARING CONTINGENT VALUATION AND CHOICE MODELING USING FIELD AND EYE-TRACKING LAB DATA

JOSE L. OVIEDO
&
ALEJANDRO CAPARRÓS

CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

CCHS

INSTITUTO DE POLÍTICAS Y BIENES PÚBLICOS

INSTITUTO DE POLÍTICAS Y BIENES PÚBLICOS – CSIC

Instituto de Políticas y Bienes Públicos (CSIC-IPP)
Consejo Superior de Investigaciones Científicas
C/ Albasanz, 26-28
28037 Madrid (España)

Tel: +34 91 6022300
Fax: +34 91 3045710

http://www.ipp.csic.es

How to quote or cite this document:

Oviedo, J. L., & Caparrós, A. (2014) Comparing contingent valuation and choice modeling using field and eye-tracking lab data. Instituto de Políticas y Bienes Públicos (IPP) CSIC, Working Paper. 2014-01

Available at: digital.csic.es

Comparing contingent valuation and choice modeling using field and eye-tracking lab data

November 03, 2014

Jose L. Oviedo (*corresponding author*)
Associate Research Professor, Institute of Public Goods and Polices (IPP)
Consejo Superior de Investigaciones Científicas (CSIC)
Albasanz 26-28, 28037 Madrid, Spain. E-mail: jose.oviedo@csic.es

Alejandro Caparrós
Associate Research Professor, Institute of Public Goods and Polices (IPP)
Consejo Superior de Investigaciones Científicas (CSIC)
Albasanz 26-28, 28037 Madrid, Spain. E-mail: alejandro.caparros@csic.es

Abstract

We compare contingent valuation and choice modeling with field and eye-tracker data. Contrary to previous research, results yield significantly different structural models. These divergences remain for modified formats that minimize visual and cognitive differences between formats. We also find divergent results concerning information processing. In choice modeling, respondents devote more time to attributes, including the bid, although total answering time does not vary. Presenting several questions with varying attribute levels works for choice modeling but not for contingent valuation. Using the attribute-stimulus format does not change contingent valuation results. Dominated alternatives increase the probability of paying in choice modeling.

JEL Classification: C91, Q23, Q51
Key words: stated preferences, environmental valuation, elicitation formats, eye-tracking

1. Introduction

Contingent valuation (CV) and choice modeling (CM) are widely employed stated preference methods for environmental valuation. Although they pursue the same objective, they employ different formats and procedures to elicit responses. In CV, respondents are asked whether they are willing to pay a given amount of money to receive an environmental good; usually one overall value is obtained for a single environmental outcome. CM uses an attribute-stimulus format and provides more information, as it indicates not only whether a specific alternative is preferred to the status quo but also whether it is preferred over another alternative with different attribute levels. Hence, this method yields values for the overall environmental good and its attributes, allowing the evaluation of multiple environmental outcomes.

Theoretically, both methods should yield similar underlying preferences and willingness to pay (WTP) estimates. However, CM provides more information and hence poses a higher cognitive burden, which may imply different cognitive processes and choice decisions (Payne 1976; Breffle and Rowe 2002; Caussade et al. 2005). Alternatively, CM presents the information differently than CV does (using an attribute-stimulus format), and the way information is presented in stated choice situations has also been shown to impact respondents' decisions (Hoehn et al. 2010).

Although CM has become the preferred approach to stated preference research in disciplines such as marketing and transport economics, CV has a long-standing tradition in environmental valuation, and both methods are still used[1] (see Jacquemet et al. (2013) and Schaafsma et al. (2013) for recent applications). One of the probable reasons for the survival of both environmental valuation methods is that previous comparisons have yielded mixed and conflicting results across studies. Furthermore, while these studies provide relevant results, they also have shortcomings that require additional research, which are discussed in more detail in the next section.

Available comparison studies focus on the convergent validity of results, but they do so without analyzing information processing strategies across formats. Most research on this issue has focused solely on different variants of CM formats (e.g., Hoehn, 2010). Hensher

[1] Ranking has fallen out of use. Therefore, we refrain from presenting our results for this format within the main text, to simplify the discussion. Nevertheless, our results confirm previous studies; we obtain significant differences between ranking and contingent valuation. These results are provided in the supplemental appendix.

(2006) posits that differing complexity in CM experiments can lead to different information processing strategies that undermine the method's internal and external validity. This critique relates to attribute non-attendance (Scarpa et al. 2010) or other CM strategies that are inconsistent with random utility models.

In our paper, using a field split sample, we compare CV and CM and investigate information processing by (i) presenting the information to CV respondents using the attribute-stimulus format common in CM studies, (ii) applying to the CV format the standard CM approach of presenting several questions to the same respondent, and (iii) including dominant alternatives in the design. In addition, to further investigate information processing, we replicate these analyses in the lab using eye-tracking technology.

An eye-tracker is a screen that incorporates technology for measuring eye position, eye movement, gaze direction, and gaze points. This technology has been used to examine visual search and its relationship to decision making (Koenpfle et al. 2009, Caplin et al. 2011). An eye-tracker is promising as a tool to collect information about respondents' behavior in valuation surveys. Most research on information processing strategies in CM uses *ex post* stated responses, which is not always an accurate means to recover information about respondents' actual behavior in valuation tasks. Employing the eye-tracker, we test whether information is processed differently in CV and CM by analyzing the time devoted to viewing the attributes and bid within each format. The starting hypothesis is that, given the higher WTP values from CM observed in our field survey (and in the literature) compared to CV, CM respondents pay less attention to the bid because this format provides more information and emphasizes attribute and alternative trade-offs, whereas CV respondents pay relatively more attention to the bid because this format emphasizes the trade-off between the bid and the environmental good.

The results from the field survey show that structural models are significantly different between CV and CM. Estimates for the WTP are generally higher for the CM, although we obtain mixed results when comparing WTP estimates statistically between the two formats (probably due to having relatively few observations in some of our sub-samples).

Concerning the information processing analysis, our main results are (i) using several questions with varying attribute levels does not work for estimating attribute-based values in our field CV, whereas it does work in CM (providing evidence supporting the different habits of the CV and CM communities); (ii) presenting the information using the attribute-stimulus format does not change CV results in the field study, although eye-tracking results show that respondents pay relatively more attention to the bid when CV employs the attribute-stimulus format; (iii) the presence of a dominated alternative in CM (which is theoretically irrelevant) increases the probability of choosing the dominant alternative compared with CV; and (iv) CM respondents devoted more time to the attributes in the eye-tracking experiment, including the bid.

The remainder of the paper is organized as follows. The next section reviews the research literature. Section 3 presents the environmental good valued and the experimental details. Section 4 compares "standard" CV and CM formats, and section 5 compares modified versions of these formats to isolate fundamental differences between the two methods. Section 6 presents the conclusion.

2. Literature review

Table 1 summarizes the studies most relevant for the purposes of our comparison — those comparing CV to CM formats in split-sample designs. Regarding the parameter vector results (structural models), CV and CM converge in Christie and Azevedo (2009) and in Sikaamäki and Layton (2007) and diverge in Scarpa and Willis (2006). Regarding welfare measures (WTP values), Scarpa and Willis (2006) and Hanley et al. (1998) present mixed results and Foster and Mourato (2003) and Christie and Azevedo (2009) obtain divergent results for CV and CM. As mentioned in the introduction, although these studies provide relevant results, they all have shortcomings.

Hanley et al. (1998) use different survey procedures (mail and face-to-face) for different subsamples and test for overlapping confidence intervals to compare WTP estimates. Scarpa and Willis (2006) also use different survey procedures (face-to-face for a CM and phone for a CV), and their findings are primarily relevant for non-market goods with polarized preferences. Foster and Mourato (2003) compare CM and CV in the valuation of two nested-goods (one good and a component of the good). They conclude that CM is sensitive to scope while CV is not, but in the CM respondents were simultaneously asked about the two nested goods while in CV they were asked about either the component or the overall good. This makes it difficult to discern whether the obtained differences are due to different preferences or scope effects.

Sikaamäki and Layton (2007) compare CV to CM using flexible econometric models and homogeneous survey procedures. They obtain a statistically similar parametric model for both formats; however, in their exercise, the environmental good is characterized only by one

Table 1. Comparison studies between independent samples of contingent valuation and choice modeling.

Authors	Comparison	Design	Sample	Comparison results	
				Parameter vectors	Welfare measures
Christie and Azevedo (2009)	Single-bounded versus choice	- SB–CV: each respondent faced three programs (alternatives) in three separate CV questions. - CH: two alternatives + SQ. The three programs from CV appeared in some choice sets. Eight sets per respondent.	- SB–CV: 433 - CH: 231	- SB–CV = CH when analyzing a pooled model with the three CV questions.	- SB–CV ≠ CH when analyzing the three CV questions separately.
Sikaamäki and Layton (2007)	Double-bounded versus ranking and choice (both recoded from rating)	- DB–CV: three programs (alternatives). Two samples facing two questions each (one program common to the samples). - CR–CH: same three alternatives (those from DB-CV) + SQ. One set per respondent.	- DB–CV: 1,680 - CR–CH: 900	- SB/DB–CV = CH	- No comparison test
Scarpa and Willis (2006)	Double-bounded versus choice	- DB–CV: one alternative (one attribute). One question per respondent. Telephone survey. - CH: two alternatives (five attributes) + SQ. One set per respondent. Face-to-face survey.	- DB–CV: 628 - CH: 413	- DB–CV ≠ CH for proportion of zero bidders	- DB–CV = CH - DB–CV ≠ CH when accounting for heterogeneity
Foster and Mourato (2003)	Double-bounded versus choice	- DB–CV: two samples each asked about one of the two nested good from the CH experiment. - CH: two alternatives (two nested goods) + SQ. Three sets per respondent.	- DB–CV: 561 - CH: 290	- No comparison test	- DB–CV ≠ CH
Hanley et al. (1998)	Single-bounded versus choice	- SB–CV: one alternative (from the CH experiment). One question per respondent. - CH: two alternatives + SQ. Eight choice sets per respondent.	- SB–CV: 809 - CH: 1,480	- No comparison test	- SB–CV ≠ CH (resident sample) - SB–CV = CH (visitor sample) (overlapping test)

CV: contingent valuation; SB: single-bounded; DB: double-bounded; CR: contingent ranking; CH: choice experiment; SQ: status quo.

attribute (in addition to the bid), and CM respondents face only one valuation set, although this is not common for this method. In our paper, we increase the complexity of the question formats and information contained within them by valuing a two-attribute environmental good and asking each respondent four valuation question sets with varying attribute levels. Therefore, we work with similar parametric models and estimate attribute-based values for multiple environmental outcomes under both formats. Nevertheless, we maintain the two fundamental differences between the methods: the attribute-based visual format and the presence of at least one additional alternative in CM. It is also important to note that although Sikaamäki and Layton (2007) used a rating format that was recoded to ranking and choice in their CM experiment, we used a ranking that was recoded to choice. Caparrós et al. (2008) have shown that the recoded choice from a ranking yields the same results as a choice format. However, no similar result is available for a recoded choice from a rating.

In another comparison of CM and CV, Christie and Azevedo (2009) ask several consecutive question sets of the same respondents, valuing several attributes and multiple environmental outcomes with both methods. However, the researchers use neither fully equivalent models nor flexible econometric models, and they present three question sets in CV and eight question sets in CM. We address all of these issues in our analysis.

There are other comparison studies beyond those identified in Table 1, but we consider them less relevant for our purposes because they use within-subject comparisons between formats, which may pose endogeneity problems (Halvorsen 2000; Boxall et al. 1996; Adamowicz et al. 1998), or they compare other CV and CM formats such as open-ended questions, paired-comparisons, and ratings (Magat et al. 1988; Ready et al. 1995; Johnson and Desvouges 1997; Stevens et al. 1997 and 2000). Regardless, these studies generally obtain higher WTP values under CM formats. Cameron et al. (2002) obtain similar results from the CV and CM formats. However, their CM experiment is part of a joint CV-CM model constructed from four CV questions and a subsequent choice task, which may result in an endogeneity problem.

None of these comparison studies applied eye-tracking technology to investigate information processing strategies in the valuation tasks. Eye-trackers are relatively new in economics, with applications to game theory (Tanida and Yamagishi 2010), behavioral economics (Caplin et al. 2011), and choice modeling applied to marketing (Wedel and Pieters 2007), but we have not found eye-tracker applications for comparing information processing strategies across environmental valuation methods.

3. Survey design and application

In our study, the good valued is a stone pine reforestation program in southwest Spain. In this region, stone pine forests cover 240,000 hectares. In the last two decades, stone pine reforestation efforts have been subsidized in Spain under the EU Common Agricultural Policy framework. We decided to investigate whether social preferences, expressed through WTP, are aligned with conserving and increasing the extent of stone pine forests in southwest Spain.

In the field survey, the sample was drawn from the population of Spanish adults (>18 years of age) in 14 provinces from southwest and west Spain. The provinces were selected considering their proximity to stone pine forests to ensure that respondents would be familiar with them. A total of 750 face-to-face interviews were conducted from April to July of 2008 (with a 30% refusal rate). Respondents were provided with a booklet of information about stone pine forests in Spain and the implications of different reforestation options. Two focus groups and a pre-test were used to identify the primary attributes of a reforestation campaign in order to evaluate how the information was understood and test a preliminary design of the formats. We used open-ended CV questions in the pre-test to design the bid vector for the main survey. The pre-test was presented to 50 individuals.

The eye-tracking experiment was conducted in three locations: the University of Cádiz, a research center (CSIC) and a public administration department. Before beginning the questionnaire, the experimenter asked each respondent to follow a point through the screen for 10 seconds. This step is necessary to calibrate the screen's sensors with the respondent's eye movements in order to record gaze direction and points during the questionnaire (none of the respondents reported feeling uncomfortable, and those who asked questions did so only at the end of the questionnaire). A total of 104 face-to-face interviews were conducted in December 2010 and January 2011. As these questionnaires closely followed the design of the field questionnaires, only a few pre-tests were performed prior to the main experiment.

4. Standard formats

Respondents in the field survey and in the eye-tracking experiment faced either a discrete-choice contingent valuation question (CV sample) or a choice modeling question (CM sample) of two alternatives, plus the status quo. In the CV sample, respondents answered four double-bounded questions. In the CM sample, respondents had to complete four ranking sets in each questionnaire. We employed an experimental design 'as if' it were a choice experiment, to be able to also analyze the recoded choice from the first rank response (Caparrós et al. 2008). Thus, we analyzed the ranking and recoded choice responses. For both the CV and CM samples, respondents did not know in advance how many valuation questions they were facing.

The attributes for the experiment are: the vegetation removed, the area (hectares) covered by the reforestation, and the bid. We denote the different reforestation alternatives in the valuation questions [VEG-HEC-BID], where VEG indicates the vegetation removed with two levels (shrub (SHR) or eucalyptus (EUC)), HEC stands for the reforested hectares in thousands with four levels (20, 40, 60 and 80), and BID stands for the hypothetical payment requested from the respondent in the form of a one-time increase in taxes with four levels (€5, €20, €35 and €50).

In the CV questionnaires, alternative [SHR-40-BID] was valued in the first question, alternative [EUC-40-BID] in the second question, alternative [SHR-80-BID] in the third question, and alternative [EUC-80-BID] in the fourth question. Two HEC attribute levels (20 and 60) were not included in this design, as this would have implied a substantial number of questionnaires. The bid was randomly assigned to each question, resulting in four different questionnaire types. The design of the bid vector uses the quintiles of the WTP distribution obtained from the pre-test (Alberini (1995) shows that this is a good compromise between efficiency and information about the shape of the WTP distribution).

For the CM experiment, we constructed a main effects experimental design with 16 treatments from the 32 possible combinations of attributes, placing them in pair-wise combinations to obtain 32 sets. We then grouped these 32 sets into eight questionnaire types, where alternative [SHR-40-BID] was always in the first set, alternative [EUC-40-BID] in the second set, alternative [SHR-80-BID] in the third set, and alternative [EUC-80-BID] in the fourth set.

This yields eight alternatives having the same attribute levels shared by the CV and the CM sample and valued by a question located in the same position within the questionnaire. This design allows the comparison of preferences for a specific alternative that is located in the same position on the questionnaires in both formats. This avoids confounding order-effects when analyzing preferences for the same alternatives valued across different formats. The potential impact of order-effects in our results due to the adopted design is investigated below.

Figure 1 presents a graphical illustration of the split-sample design for these standard formats, and Annex 1 in the supplemental appendix presents examples of the valuation questions.

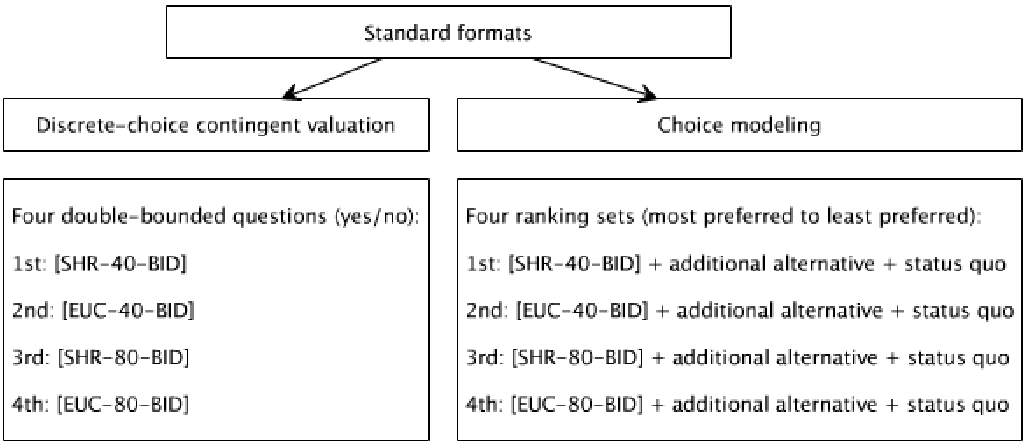

Figure 1. Split-sample survey design for the standard formats. Contingent valuation and choice modeling.

4.1 Data analysis

For the field survey, the most straightforward analysis is determining whether the same reforestation alternatives (those shared by all formats) are chosen over the status quo more or less often in the CV compared with the CM exercises. This makes it possible to identify different preferences regarding the same alternative in the same question position, without the use of econometric models. We use a *z-test* of differences between proportions to achieve this goal. Our specific design and the use of a ranking format permit this analysis. For the subsequent analysis, however, we focus on the choice answers.

We also compare the results from econometric models and their corresponding WTP values for the alternatives shared in both formats. For the CV sample, we follow Cameron (1988) for the single-bounded analysis and Hanemann et al. (1991) for the double-bounded analysis. For the CM sample, we use a conditional logit (McFadden 1974) to analyze the recoded choice data. For all these CV and CM models, we employ fixed and random parameter models (Train 2009).

Using these models, we generate empirical distributions for the individual parameters and the marginal WTP per attribute ($-\beta_{attribute}/\beta_{bid}$) through the Krinsky and Robb (1986) bootstrapping technique with 1,000 replacements. Using these empirical distributions, we estimate the WTP for a specific reforestation alternative defined by the attribute levels of the corresponding alternative. The standard deviation and the 95% confidence interval are estimated using the percentile approach (Efron and Tibshriani 1993). To test the equality of the WTP values for the same reforestation alternative obtained through different formats, we employ the complete combinatorial test (Poe et al. 2005).

We use a Likelihood Ratio test to establish whether the parameter vectors are statistically similar across formats, and therefore the valuation tasks derive from the same structural models. We follow the Likelihood Ratio test proposed by Swait and Louviere (1993), as it makes possible to test whether divergences in the results are due to differences in taste parameters or in scale parameters. This is a double stage test to test the hypothesis H_0:$(\lambda^A\beta^A) = (\lambda^B\beta^B)$, where A and B correspond to the format samples under comparison. The test separately examines hypothesis H_a:$(\beta^A) = (\beta^B)$, where the relative scale parameter is set as λ^B/λ^A, and hypothesis H_b:$(\lambda^A) = (\lambda^B)$. If H_a is rejected, H_0 is rejected and the differences derive from taste parameters. If H_a is not rejected but H_b is rejected, H_0 is rejected and the differences derive from scale parameters. If both H_a and H_b are not rejected, then H_0 is not rejected.

In the eye-tracker experiment, we analyze the time devoted to the attributes, including the bid, within each valuation format using the area of interest tool (Tobii AOI tool). This tool measures the time respondents are viewing areas containing the information about the attributes and the bid while answering the question (in "eye-tracking" terminology: total time visit duration). We estimate the percentage of this time for each attribute relative to the time employed in answering the whole valuation question in each format; that is, our focus is on the relative amount of time paid to these elements in each format. For example, in the CV question displayed in Annex 1 of the supplemental appendix, the areas of interest (AOIs) in sentence 10a are “5 euros”, “Shrubland” and “40,000 hectares” and either “20 euros” in 10a.1 or “2 euros” in 10a.2. In the CM question displayed in Annex 1, the AOIs are the corresponding areas in the rectangles of the column that describes the options in question 10a.

With the aim of identifying whether information processing strategies affect responses differently in CV and CM, we analyze how the relative time devoted to these elements and the total time spent answering the valuation questions affect the probability of a “yes” response to the alternatives in each format. We use a binary logit model (“yes” = paying for any alternative; “no” = not paying) to perform this analysis.

4.2 Field survey results

The sub-sample using the standard formats has 159 questionnaires for the CV sample and 294 for the CM sample. After removing invalid and protest responses, we have 139 respondents for the CV sample (556 useable observations) and 259 for the CM sample (1,036 useable observations).

Table 2 presents the percentages of shared alternatives choices across formats. We find significant differences (*z-test*) in eight out of 16 possible cases between CV and CM, when the latter considers the full rank. Alternative options are more frequently selected in CM and that significant differences appear for higher bids. In some cases, we also find significant differences between CV and CM when considering only the first choice. Now, the alternative options are more frequently selected in CV in most cases (Table 2). By using the full-rank in this single-alternative analysis, we ensure that we analyze preferences for the targeted alternative versus the status quo in all cases, even when the alternative is ranked second. As said previously, the subsequent analysis focuses on the first choice (ranking results for this analysis are provided in Annex 2 of the supplemental appendix).

Table 2. Percentage of choices of shared alternatives over the status quo in the standard contingent valuation and choice modeling (ranking and choice) formats. *Z-test* statistics for the difference between proportions.

Alternative	Percentage of choices of the alternatives over the status quo				*z-test* results for the difference between proportions			
	CV		CM		CR versus CV		CH versus CV	
	SB	DB	CR	CH	SB	DB	SB	DB
[SHR-40-€35]	52%	50%	68%	30%	0.089*	0.116	0.025**	0.078*
[SHR-40-€20]	67%	56%	76%	45%	0.356	0.041**	0.053*	0.271
[EUC-40-€50]	32%	11%	55%	14%	0.040**	0.014**	0.038**	0.809
[EUC -40-€20]	71%	60%	83%	61%	0.187	0.012**	0.288	0.932
[SHR-80-€5]	86%	78%	86%	68%	0.964	0.374	0.042**	0.398
[SHR-80-€50]	35%	9%	58%	28%	0.041**	0.003***	0.465	0.180
[EUC-80-€20]	63%	67%	76%	58%	0.150	0.372	0.627	0.445
[EUC-80-€50]	10%	57%	53%	27%	0.001***	0.748	0.067*	0.010**

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively.

CV: Contingent valuation

CM: Choice modeling.

SB: Single-bounded.

DB: Double-bounded.

CR: Ranking.

CH: Recoded choice.

Table 3. Regression models and willingness to pay values (with confidence intervals at 95%) for the standard contingent valuation and choice formats. Fixed and random parameter estimations.

Variable	Single-bounded contingent valuation		Double-bounded contingent valuation		Choice	
	FP	RP	FP	RP	FP	RP
Intercept / ASC_{REF}	2.219***	2.591*	1.209***	1.210***	1.556***	2.213**
	(0.197)	(1.512)	(0.222)	(0.225)	(0.142)	(0.393)
Vegetation removed (VEG) (=1 shrub; =-1 eucalyptus)	0.095	0.115	0.061	0.062	-0.096***	-0.249**
	(0.095)	(0.133)	(0.063)	(0.063)	(0.037)	(0.110)
Area (hectares/10,000) covered by the reforestation (HEC)	-0.054	-0.064	-0.021	-0.021	0.053***	0.109**
	(0.095)	(0.065)	(0.031)	(0.031)	(0.020)	(0.055)
Bid	-0.057***	-0.066*	-0.030***	-0.030***	-0.033***	-0.077***
	(0.006)	(0.037)	(0.004)	(0.004)	(0.002)	(0.017)
Standard deviation parameters						
Intercept / ASC_{REF}						
Vegetation removed (VEG) (=1 shrub; =-1 eucalyptus)		0.927		0.067		1.944***
		(2.014)		(0.913)		(0.593)
Area (hectares/10,000) covered by the reforestation (HEC)				0.001		0.252
				(0.080)		(0.104)
n	556	556	556	556	1,036	1,036
AIC	1.182	1.186	1.302	1.306	1.866	1.852
Log-likelihood	-324.776	-324.714	-720.107	-720.104	-962.514	-953.525
Alternative	Mean WTP	Mean WTP	Mean WTP	Mean WTP	Mean WTP	Mean WTP
[SHR-40]	36.82	39.20	39.71	39.82	49.67	31.65
	[30.69, 43.20]	[20.37, 54.18]	[31.75, 48.19]	[32.09, 48.69]	[43.16, 57.45]	[23.74, 44.01]
[EUC-40]	33.52	33.92	35.63	35.92	55.55	38.15
	[28.03, 39.67]	[18.26, 50.24]	[28.62, 43.99]	[28.74, 44.69]	[45.32, 67.04]	[30.88, 48.16]
[SHR-80]	32.95	37.13	36.82	36.91	55.94	37.39
	[26.94, 38.80]	[9.02, 61.39]	[29.27, 44.77]	[29.47, 45.19]	[36.65, 76.62]	[27.87, 50.47]
[EUC-80]	29.65	31.84	32.75	33.00	61.82	43.89
	[24.28, 35.29]	[9.84, 58.02]	[25.78, 40.26]	[25.16, 40.95]	[54.02, 71.18]	[34.65, 55.08]

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively; Standard errors are shown in parentheses; n: number of observations. AIC: Akaike Information criterion.

FP: Fixed parameter estimations.

RP: Random parameter estimations.

Table 3 presents the results of the regression models for CV and CM (fixed and random parameters) and the corresponding mean WTP values for the alternatives valued in the first, second, third and fourth questions in the questionnaires. For the RP models, we present the random parameters specification that yields the lowest Akaike Information Criterion score.

The single-bounded and double-bounded models do not yield significant estimates for the VEG and HEC attributes. This suggests that valuing multiple environmental outcomes by using several questions with varying attribute levels within the same questionnaire is not a desirable strategy in CV. This is not the case for the CM models, in which the parameters are significant and take the expected signs (Table 3). We obtain significant standard deviation parameters in the CM random parameter models but not in the CV ones (Table 3).

The complete combinatorial test results for WTP comparisons show significantly higher WTP values for CM using the fixed parameter estimations (Table 4). However, for the random parameter estimations, we obtain statistically similar WTP values between CV and CM in some cases.

Table 4. Complete combinatorial test for the comparison of mean WTP values for alternatives [SHR-40], [EUC-40], [SHR-80] and [EUC-80] between the standard contingent valuation and choice formats. Fixed and random parameter estimations.

Alternative	Complete combinatorial test			
	Single-bounded versus choice		Double-bounded versus choice	
	FP estimations	RP estimations	FP estimations	RP estimations
[SHR-40]	0.003***	0.204	0.034**	0.100
[EUC-40]	<0.001***	0.239	<0.001***	0.358
[SHR-80]	<0.001***	0.328	0.001***	0.492
[EUC-80]	<0.001***	0.084*	<0.001***	0.041**

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively.

FP: Fixed parameter estimations.

RP: Random parameter estimations.

The results from the Likelihood Ratio test (Table 5) show that structural models are significantly different in all cases. Differences between the single-bounded and the CM models derive only from taste parameters. Differences between the double-bounded and the

CM models come from both taste and scale parameters for the fixed parameter estimations and only from scale parameters for the random parameter estimations (Table 5).

Table 5. Likelihood ratio test for the comparison of parameter vectors (structural models) estimated with the standard contingent valuation and choice formats. Fixed and random parameter estimations.

Hypotheses	Likelihood ratio test			
	Single-bounded versus choice		Double-bounded versus choice	
	FP	RP	FP	RP
Ha: $\beta^{CV} = \beta^{CM}$	43.439	17.010	23.202	10.079
Reject Ha: $\beta^{CV} = \beta^{CM}$? (5% level)	Yes	Yes	Yes	No
Hb: $\lambda^{CV} = \lambda^{CM}$	0.478	1.444	27.288	12.420
Reject Hb: $\lambda^{CV} = \lambda^{CM}$? (5% level)	No	No	Yes	Yes
Reject H0: $\beta^{CV}\lambda^{CV} = \beta^{CM}\lambda^{CM}$? (5% level)	Yes	Yes	Yes	Yes

For the hypothesis Ha, the critical value at the 5% level for the fixed parameter estimations is 11.070 (χ^2 statistic for 5 degrees of freedom), and the critical value at the 5% level for the random parameter estimations is 14.067 (χ^2 statistic for 7 degrees of freedom).

For the hypothesis Hb, the critical value at the 5% level for both the fixed and random parameter estimations is 3.841 (χ^2 statistic for 1 degree of freedom).

FP: Fixed parameter estimations.

RP: Random parameter estimations.

Our procedure of asking a respondent several consecutive valuation questions raises the concern of order-effects. We investigated the influence of such effects by comparing subsamples in scenarios designed to remove the impact of potential order-effects. In Annex 3 of the supplemental appendix, we report the econometric models and WTP results from this analysis. The results are mostly similar to those obtained in the analysis of the complete samples. Although our CV samples do not seem to pass the scope test, the analysis of order-effects shows that the comparison results hold in most cases, even for the first questions.

4.3 Eye-tracking results

The sub-sample using the standard formats has 27 interviews for the CV sample and 26 for the CM sample. The only difference for the eye-tracking survey is that respondents faced eight ranking sets within the CM questionnaires instead of four. This design was used because respondents faced eight sets in the modified CM format in the eye-tracking survey, which is analyzed in section 5. Because the amount of information to be processed is a key issue in the eye-tracking analysis, we decided to present eight sets for the standard CM in this experiment. The comparison can still be made using the four first sets. In addition, we have the same information (eight sets presented) as in the modified CM that is analyzed below. For the analysis of the time devoted to the attributes and the bid, we provide mean values for a single valuation question using position and mean results from all question sets in each format, as well as from the first four sets in the CM format.

Table 6 shows that respondents dedicate a larger proportion of their time to the attributes, including the bid, in the CM format. These differences are significant in all cases for the mean values and in most cases for the values we obtain for each question positioned separately (*p-values* for the *t-test* of differences are reported in Annex 4 of the supplemental appendix). The small amount of attention paid to the attributes in CV could explain why the CV models from the field survey did not seem to pass the scope test.

Table 6. Eye-tracking results for the standard contingent valuation and choice modeling formats. Percentage of time devoted to the attributes and the bid relative to the time devoted to completing the whole valuation question. Mean values for all questions and for each question position.

Format	Total visit duration to (% of time relative to the time devoted to answer the whole valuation question):						Time devoted to the valuation question (seconds)	
	Attribute VEG values (%)		Attribute HEC values (%)		BID vector values (%) (including €0 in CM)			
	Mean	s. d.	Mean	s. d.	Mean	s. d.	Mean	s. d.
CV (1st question)	6.55%	4.74%	4.10%	2.54%	7.31%	2.63%	39.70	19.50
CV (2nd question)	7.44%	5.70%	4.33%	2.96%	2.73%	2.28%	33.21	21.83
CV (3rd question)	3.00%	2.26%	4.17%	3.53%	2.59%	2.38%	22.25	8.21
CV (4th question)	5.47%	3.51%	2.41%	2.54%	3.23%	3.03%	20.51	8.81
Mean (all questions)	5.62%	4.51%	3.75%	2.98%	3.96%	3.22%	28.92	17.51
CM (1st question)	5.96%	4.35%	7.13%	3.28%	5.89%	4.17%	72.68	47.35
CM (2nd question)	9.06%	4.96%	9.96%	8.15%	9.21%	6.36%	24.81	12.33
CM (3rd question)	9.83%	5.59%	7.71%	6.51%	6.51%	5.89%	23.54	17.95
CM (4th question)	10.64%	7.40%	10.86%	10.21%	10.74%	8.33%	19.51	15.56
CM (5th question)	10.93%	5.96%	15.04%	11.35%	8.77%	6.37%	16.17	10.30
CM (6th question)	13.16%	8.30%	12.92%	12.22%	12.07%	8.60%	12.66	5.90
CM (7th question)	8.01%	5.34%	13.37%	11.63%	9.55%	7.98%	13.14	8.37
CM (8th question)	10.28%	7.46%	10.52%	9.95%	10.88%	8.34%	12.50	12.48
Mean (all questions)	9.73%	6.52%	10.94%	9.77%	9.20%	7.31%	24.37	27.51
Mean (1st to 4th questions)	8.87%	5.88%	8.91%	7.53%	8.09%	6.58%	35.13	34.59

Note: The bid vector values are analyzed considering the €0 associated with the status quo as part of the bid vector in the CM sample.

s. d.: standard deviation

CV: Contingent valuation.

CM: Choice modeling.

VEG: attribute “vegetation removed”.

HEC: attribute “area (hectares) covered by the reforestation”.

Regarding the mean time (in seconds) used to answer the valuation questions, the CM questions take longer than the CV ones if we consider the four first questions, whereas the opposite occurs if we consider the eight sets presented in the CM format. However, in both cases, these differences are not significant (see Table 6 and Annex 4 of the supplemental appendix). As expected, the first questions take the most time in both formats, being significantly longer for the CM case. However, the drop in the time devoted to answer the subsequent questions is higher in CM (a 73% drop in CM versus a 48% drop in CV from the first to the fourth question). This indicates that, despite the higher load of information in CM, this information is rapidly processed in this format. As a result, by the second and subsequent questions, the time devoted to respond does not differ significantly from that in CV.

Table 7 shows the logit analysis of “yes” responses and how the time devoted to the different questions elements affect these responses. We present models for the CV (single-bounded and double-bounded) and CM formats. CM models are analyzed considering the eight sets presented and the four first sets in this sample.

We obtain that for the single-bounded answers the percentage of total time devoted to HEC increases the probability of a “yes” response. For the double-bounded answers, there is no significant explanatory variable. For the CM responses, the percentage of total time devoted to HEC and the time spent in answering the valuation question decreases the probability of a “yes” response. When analyzing only the first four questions, the effect of HEC is not significant. We also create pooled models with the single-bounded and CM answers, including as explanatory variable a dummy for the question format. These pooled models indicate a difference in the results from the different formats and no significant effect in the remaining explanatory variables.

These results reject convergent validity between these formats based on the time devoted to the attributes and how this affects the probability of paying for the environmental good. Both formats identify significant effects on the “yes” responses of the time devoted to some attributes (not in the case of the bid) and the time devoted to answer the questions, but the effects differ and even go in opposite directions. An advantage that could be attributed to CM is that respondents pay relatively more attention to the bid in this format and require the same time to respond compared to CV, even when CM contains a higher load of information.

Table 7. Logit analysis of "yes" responses (paying for the environmental good) to the standard contingent valuation and choice questions in the eye-tracking experiment. Models including all questions made to every respondent and including only the four first questions in the choice format.

Variable	Single-bounded	Double-bounded	Choice		Pooled single-bounded and choice	
	All questions	All questions	All questions	4 first questions	All questions	4 first questions
Intercept	-0.600	0.019	4.164***	4.782***	2.784***	2.505***
	(0.500)	(0.488)	(0.854)	(1.725)	(0.556)	(0.751)
% time BID	7.066	-0.036	-2.602	13.391	1.389	9.009
	(6.427)	(6.108)	(4.114)	(16.656)	(3.484)	(5.762)
% time attribute VEG	-2.359	1.839	-1.458	-6.826	0.876	1.158
	(5.439)	(5.138)	(4.600)	(11.702)	(3.248)	(4.527)
% time attribute HEC	14.727*	-9.698	-5.714**	-5.228	-2.722	6.712
	(7.804)	(7.085)	(2.593)	(9.148)	(2.425)	(5.955)
Time devoted to the valuation question (seconds)	0.013	0.019	-0.016**	-0.025**	-0.004	-0.010
	(0.014)	(0.015)	(0.007)	(0.011)	(0.007)	(0.007)
CV format					-2.208***	-2.428***
					(0.424)	(0.600)
n	108	108	208	104	316	212
AIC	1.362	1.419	0.529	0.344	0.831	0.867

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively; Standard errors are shown in parentheses; n: number of observations. AIC: Akaike Information criterion.

VEG: attribute "vegetation removed".

HEC: attribute "area (hectares) covered by the reforestation".

CV format: dummy = 1 if the response was obtained with the contingent valuation format, = 0 if the response was obtained with the choice modeling format.

5. Modified formats

The comparison using modified formats was intended to test the convergent validity when the fundamental differences between these methods are not present. Thus, a separate sample of respondents was presented with a CV question that uses an attribute-stimulus format for describing the environmental good alternative (CV_{CM} sample). In this modified CV format, we present the same information as in the standard CV; the difference resides in how this information is presented. The questionnaire design followed the one used in the standard CV.

Another sample of respondents faced a ranking of two alternatives plus the status quo, but with dominated alternatives (CM_D sample). This approximates CM to CV as the additional alternative is now irrelevant, in principle. Some sets have dominance due to the bid ($CM_{D\text{-}BID}$), implying two alternatives with the same attribute levels, except for the bid; other sets have dominance due to the HEC attribute ($CM_{D\text{-}HEC}$), implying two alternatives with the same attribute levels and bids, except for the attribute HEC.

The CM_D questionnaires included four sets with dominated alternatives, where two were $CM_{D\text{-}BID}$ sets and two were $CM_{D\text{-}HEC}$ sets. To avoid any undesirable effects from presenting these unusual sets, we alternated these sets among four sets randomly taken from the standard CM design. This meant CM_D respondents had to complete eight sets each. We included the alternatives shared between CV and CM in the sets with dominance and placed them in the third, fourth, sixth, and eighth question positions, placing the four standard sets in the remaining positions[2]. Thus, the CM_D sample has four different questionnaire types (with eight choice sets each) and eight alternatives in common with CV and CM.

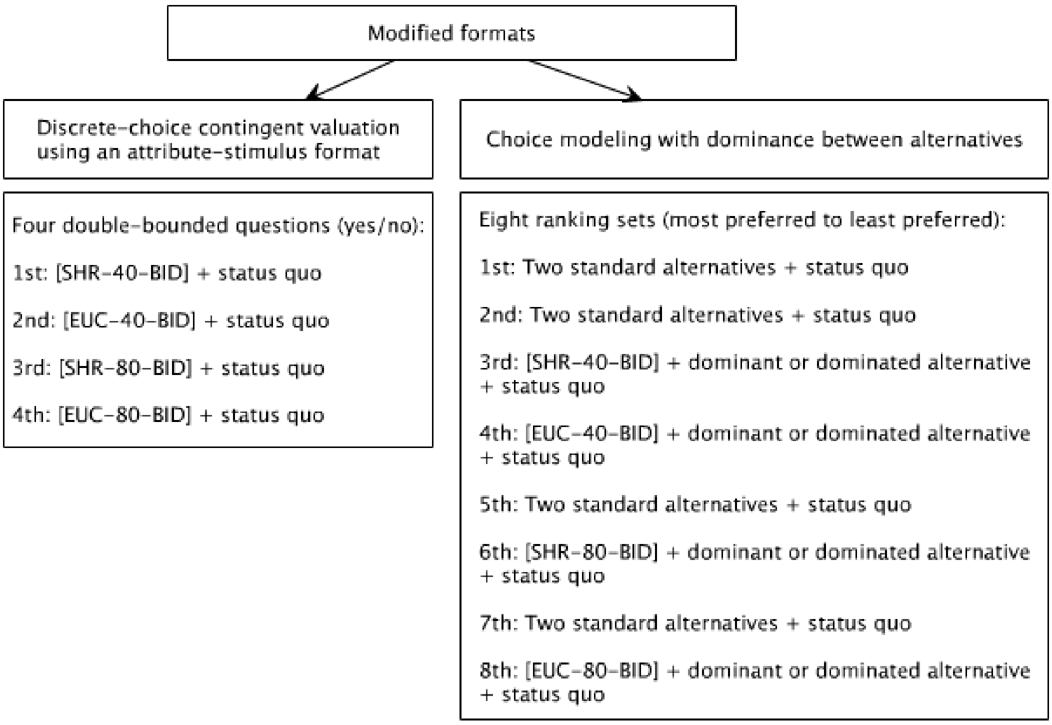

Figure 2. Split-sample survey design for the modified formats. Contingent valuation using an attribute-stimulus format and choice modeling with dominance between alternatives.

[2] Although this implies a lack of homogeneity in the question positions for the shared alternatives and a higher cognitive load for respondents to this sample, we considered this a better option than grouping all sets with dominated alternatives as the first four valuation questions because they all are unusual sets.

Figure 2 presents a graphical illustration of the split-sample design for these modified formats, and Annex 5 of the supplemental appendix presents examples of the corresponding valuation questions. The data analysis for the CV_{CM} format follows the one used in the standard CV format. For the CM_D format, we only perform the single alternative analysis (including the analysis of the full-rank) in the field survey, as the small subsamples resulting from the design presents challenges for working with econometric models. The eye-tracker data are analyzed as in the standard formats.

5.1 Field survey results

The CV_{CM} sample yielded 151 completed questionnaires that, after removing invalid and protest responses, resulted in 137 respondents (548 useable observations). A comparison of the standard and modified CV samples yields no significant differences in the single alternative analysis, except in two cases of the double-bounded format. The econometric models and WTP results show similar results between these two CV formats. Overall, we find no significant difference between the CV and the CV_{CM}. These results are presented in Annex 6 of the supplemental appendix.

The CM_D sample yielded 80 completed questionnaires that, after removing invalid and protest responses, resulted in 73 respondents (with 156 useable observations for $CM_{D\text{-}HEC}$ sets, 136 for $CM_{D\text{-}BID}$ sets, and 292 for standard sets; a total of 584 useable observations).

Table 8 reveals that the shared alternatives are more frequently preferred in all CM cases (both when we analyze the full-rank and the first choice) compared to the standard CV formats. These results hold when the comparison is made with the CV_{CM} sample (see Annex 7 of the supplemental appendix). Inconsistent answers (i.e., ranking a dominated alternative first or ranking a dominant alternative second when it is preferred over the status quo) comprise less than 10% and do not change the results.

As the CV, CV_{CM} and CM_D formats are virtually identical for the shared alternatives (the dominated alternative provides no relevant information in the CM_D sets), it is surprising these shared alternatives are chosen significantly more often in CM_D. Particularly troubling is the case for the CV_{CM}, as it replicates the CM attribute-stimulus question format. A potential problem arising from this behavior is that experimental designs may introduce bias, for example, when including dominated alternatives or when respondents' lexicographic preferences favor certain attributes, rendering some alternatives irrelevant.

Table 8. Percentage of choices of shared alternatives over the status quo in the modified choice modeling (ranking and choice) format. *Z-test* statistics for the difference between proportions with the standard contingent valuation format.

Alternative	Percentage of choices of the alternatives over the status quo			*z-test* results for the difference between proportions					
				$CM_{D\text{-}BID}$ (ranking) versus CV		$CM_{D\text{-}BID}$ (choice) versus CV		$CM_{D\text{-}HEC}$ (choice) versus CV	
	$CM_{D\text{-}BID}$ (ranking)	$CM_{D\text{-}BID}$ (choice)	$CM_{D\text{-}HEC}$ (choice)	SB	DB	SB	DB	SB	DB
[SHR-40-€20]	n. a.	74%	77%	n. a.	n. a.	0.549	0.128	0.344	0.055^{**}
[EUC-40-€20]	n. a.	76%	77%	n. a.	n. a.	0.633	0.142	0.589	0.116
[SHR-80-€50]	59%	n. a.	64%	0.060^{*}	0.004^{***}	n. a.	n. a.	0.017^{**}	0.001^{*}
[EUC-80-€50]	56%	n. a.	59%	0.001^{***}	0.927	n. a.	n. a.	0.001^{***}	0.891

Note: For the D-BID subscript samples, alternatives [SHR-40-€20] and [EUC-40-€20] are dominant and alternatives [SHR-80-€50] and [EUC-80-€50] are dominated. For the D-HEC subscript sample, all analyzed alternatives are dominant.

n. a.: not applicable because when the shared alternative is dominant, we analyze the information provided by the recoded choice, as if the dominant alternative is preferred over the status quo it should always be ranked 1st; and when the shared alternative is dominated, we analyze the information provided by the full rank, as if the dominated alternative is preferred over the status quo it should always be ranked 2nd.

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively.

CV: Contingent valuation.

SB: Single-bounded.

DB: Double-bounded.

$CM_{D\text{-}BID}$: Choice modeling with dominance between the “do something” alternatives due to the bid.

$CM_{D\text{-}HEC}$: Choice modeling with dominance between the “do something” alternatives due to the attribute “area (hectares) covered by the reforestation”.

Our survey design included an additional CM sample in which the only difference among “do something” alternatives was in the level of a single attribute. The concept was to compare CV with a CM sample that contains the minimum amount of additional information required to be a CM experiment. The comparison results are similar to those obtained with the standard formats. The valuation questions used and the results from this sample for the single alternative analysis are reported in Annex 8 of the supplemental appendix.

5.2 Eye tracking results

We performed 26 interviews for the CV_{CM} format and 26 for the CM_D format. Table 9 shows that respondents dedicate a larger proportion of their time to the attributes VEG and HEC in CM_D, with these differences being significant in most cases. However, the time dedicated to the bid is statistically similar (Annex 9 of the supplemental appendix). These results hold for mean values and for each question position separately.

When comparing CV formats, the time devoted to the bid is significantly higher in CV_{CM} (see Tables 6 and 9 and Annex 9). This may indicate that the attribute stimulus-format helps respondents pay more attention to the bid. This difference is also present among the standard CV and both CM formats, but it is not present between both CM formats. Thus, all formats using an attribute-stimulus design have in common that respondents devote a larger proportion of their answering time to the bid.

The mean time used in the valuation questions (for all presented sets and for the four first sets in CM) is not significantly different between CV_{CM} and CM_D (see Table 9 and Annex 9 of the supplemental appendix). As in the standard samples, although the first question takes significantly longer in CM_D, the drop in the time devoted to answer the subsequent questions is higher in CM_D and makes the time spent in all valuation questions similar between formats. When comparing modified and standard formats, we find that in most cases (for mean values and for each question position separately) there are no significant differences in the time devoted to answer the questions.

Table 9. Eye-tracking results for the modified contingent valuation and choice modeling formats. Percentage of time devoted to the attributes and the bid relative to the time devoted to completing the whole valuation question. Mean values for all questions and for each valuation question position.

Format	Total visit duration to (% of time relative to the time devoted to answer the whole valuation question):						Time devoted to the valuation question (seconds)	
	Attribute VEG values (%)		Attribute HEC values (%)		BID vector values (%) (including €0 in CM)			
	Mean	s. d.	Mean	s. d.	Mean	s. d.	Mean	s. d.
CV_{CM} (1st question)	1.98%	1.50%	3.09%	2.10%	5.81%	3.99%	44.42	20.41
CV_{CM} (2nd question)	6.12%	5.17%	3.42%	3.19%	7.08%	5.97%	25.18	14.52
CV_{CM} (3rd question)	4.43%	3.18%	6.32%	6.82%	9.04%	6.76%	15.24	7.64
CV_{CM} (4th question)	5.97%	5.19%	4.09%	5.24%	10.22%	7.89%	13.46	7.18
Mean (all questions)	4.63%	4.34%	4.23%	4.81%	8.04%	6.46%	24.57	18.21
CM_D (1st question)	4.80%	2.12%	7.08%	4.15%	8.04%	4.24%	62.12	26.00
CM_D (2nd question)	9.89%	6.00%	5.79%	3.75%	12.85%	19.15%	21.37	10.82
CM_D (3rd question)	10.93%	10.56%	9.07%	8.11%	8.46%	8.18%	19.01	15.09
CM_D (4th question)	10.26%	9.00%	8.33%	8.63%	10.13%	8.02%	13.44	7.57
CM_D (5th question)	10.58%	7.37%	11.30%	9.71%	7.18%	5.16%	12.71	7.65
CM_D (6th question)	11.20%	11.34%	8.98%	7.21%	9.27%	5.16%	10.99	5.84
CM_D (7th question)	11.89%	9.10%	8.34%	8.80%	8.64%	7.16%	11.63	7.72
CM_D (8th question)	8.09%	8.30%	9.71%	9.09%	10.31%	8.50%	9.68	5.73
Mean (all questions)	9.71%	8.55%	8.58%	7.75%	9.36%	9.29%	20.12	20.50
Mean (1st to 4th questions)	8.97%	7.91%	7.57%	6.58%	9.87%	11.35%	28.99	25.29

Note: The bid vector values are analyzed considering the €0 associated with the status quo as part of the bid vector in the CV_{CM} and CM_D samples.

s. d.: standard deviation.

CV_{CM}: Contingent valuation using an attribute-stimulus format.

CM_D: Choice modeling with dominance between the “do something” alternatives.

VEG: attribute “vegetation removed”.

HEC: attribute “area (hectares) covered by the reforestation”.

In Annex 9, we report the logit analysis of “yes” response to the single-bounded, double-bounded and CM answers. For both the single-bounded and double-bounded answers, the percentage of time devoted to VEG decreases the probability of a “yes” response. For the CM responses, only the time spent in answering the valuation question has a significant impact, decreasing the probability of a “yes” response. The pooled models indicate a difference in the results from the different formats and no significant effect in the remaining explanatory variables. As in the standard sample, the results point toward the rejection of convergent validity for these modified formats in this analysis.

6. Conclusions

This paper has tested convergent validity of contingent valuation and choice modeling focusing on the role of information and information processing strategies. The results obtained from a field survey and an eye-tracking experiment show that structural models and information processing strategies are different between these two formats. Our results are obtained for a two-attributes good and a two-alternatives CM design, which can be considered as the smallest departure from CV that maintains the multi-attribute nature of CM. Hence, one should expect larger differences between the two formats for goods with more attributes or for more alternatives in the CM design.

Our result that structural models are different between CV and CM is the opposite to the one obtained by Siikamäki and Layton (2007) and Christie and Azevedo (2009). The reason for the convergent validity found in Siikamäki and Layton (2007) may come from the use of a good with a single attribute; a practice that is not common in CM and that renders CV and CM tasks very similar. In Christie and Azevedo (2009), the utility functions in the CV and CM experiments were not exactly similar and they performed this test only for a subsample of the parameters from the CM model. A reason for the convergent validity they obtain may stem from the use of fixed parameters models, which tend to obtain larger variances than their flexible counterparts.

Despite our efforts to minimize the differences between the two formats, divergences remain. Presenting the CV question with the visual format (attribute-based) usually used in CM has no impact on our results. Making the additional alternative in the CM exercise irrelevant, or more precisely dominated, not only maintained the differences but increased them. Finally, the standard approach used in CM of presenting several valuation scenarios to the same respondent, which works well in our CM field survey, fails altogether in our CV field survey.

Although with the caveat discussed below, our results seem to support the use of CM. If the practitioner is interested in attribute-based values, this is hardly surprising. However, our results confirm that repeated CV questions are not reliable. More surprising is the result obtained from the eye-tracker experiment that the attention paid to the bid is not an argument in favor of using CV. Intuition suggests that respondents will pay more attention to the bid in a standard CV format and they will be able to answer faster. However, our eye-tracking results show that the opposite is true for the time devoted to the bid, which is an indicator of the attention paid, and that the time devoted to the valuation questions does not increase with the CM format.

The caveat of CM that we observed is that the presence of dominated alternatives favors choosing "do something" alternatives. This finding could be especially relevant in the presence of lexicographical preferences, which could make alternatives dominated even when this was not intended in the design.

Before concluding, we would like to mention that our eye-tracking results represent a first step in analyzing information processing strategies in stated preference formats for environmental valuation. We believe this technique offers valuable information and has a great potential for improving environmental valuation methods (e.g., when designing a field experiment one could use this technology to identify relevant attributes and to avoid non-attendance issues).

Acknowledgements

We are thankful for the comments and suggestions from Pablo Campos, Begoña Alvaréz-Farizo and participants at the 4th World Congress of Environmental and Resource Economists, Montreal-June 2010, and the 4th AERNA Conference, Faro-June 2010. Manuel Acosta, Manuel Navarro and Gloria García provided assistance with the eye-tracking experiment. We gratefully acknowledge funding from INIA (Project CPE03-001-C5), the I + D National Plan of the Ministry of Science and Innovation (Projects DYNOPAGROF and UTPCC) and the Intramural Grant 200910I130 of CSIC. Both authors share the first-authorship of the article. The usual disclaimer applies.

References

Adamowicz, W.L., P.C. Boxall, M. Williams, and J.J. Louviere. 1998. Stated preference approaches for measuring passive use values: choice experiments and contingent valuation. Am. J. Agric. Econ. 80: 65-75.

Alberini, A. 1995. Optimal design for discrete choice contingent valuation surveys: single-bound, double-bound and bivariate models. J. Environ. Econ. Manage. 28(3):287-306.

Boxall, P.C., W.L. Adamowicz, J. Swait, M. Williams, and J. Louviere. 1996. A Comparison of Stated Preference Methods for Environmental Valuation. Ecol. Econom. 18:243-253.

Breffle, W.S., and R.D. Rowe. 2002. Comparing Choice Question Formats for Evaluating Natural Resources Tradeoffs. Land Econ. 78(2): 298-314.

Cameron, T. A. 1988. A new paradigm for valuing non-market goods using referendum data: maximum likelihood estimation by censored logistic regression. J. Environ. Econ. Manage. 15:355-379.

Cameron, T.A., G.L. Poe, R.G. Either, and W.D. Schulze. 2002. Alternative Non-Market Value-Elicitation Methods: Are the Underlying Preferences the Same? J. Environ. Econ. Manage. 34:391-425.

Caparrós, A., J.L. Oviedo, P. Campos. 2008. Would you Choose your Preferred Option? Comparing Choice and Recoded Ranking Experiments. Am. J. Agric. Econ. 90(3):843-855.

Caplin, A., M. Dean, and D. Martin. 2011. Search and Satisficing. American Economic Review 101(7):2899-2922.

Caussade, S., J. D. Ortuzar, L. I. Rizzi, and D. A. Hensher. 2005. Assessing the Influence of Design Dimensions on Stated Choice Experiment Estimates. Transportation Research Part B−Methodological 39(7): 621–640.

Christie, M., and C.D. Azevedo. 2009. Testing the Consistency Between Standard Contingent Valuation, Repeated Contingent Valuation and Choice Experiments. Journal of Agricultural Economics 60(1):154-170.

Efron, B., and R.J. Tibshirani. 1993. An introduction to the bootstrap. New York: Chapman & Hall ed.

Foster, V., and S. Mourato. 2003. Elicitation Format and Sensitive to Scope: Do Contingent Valuation and Choice Experiments Give the Same Results? Environ. Res. Econom. 24:141-160.

Halvorsen, B. 2000. Comparing Ranking and Contingent Valuation for Valuing Human Lives, Applying Nested and Non-Nested Logit Models. Environ. Res. Econom. 17:1-19.

Hanemann, M., J. Loomis, and B. Kanninen. 1991. Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation. Am. J. Agric. Econ. 73(4):1255-1263.

Hanley, N., D. MacMillan, R.E. Wright, C. Bullock, I. Simpson, D. Parsisson, and B. Crabtree. 1998. Contingent Valuation Versus Choice Experiments: Estimating the Benefits of Environmentally Sensitive Areas in Scotland. J. Agr. Econ. 49(1):1-15.

Hensher, D. 2006. How Do Respondents Process Stated Choice Experiments? Attribute Consideration Under Varying Information Load. Journal of Applied Econometrics 21:861-878.

Hoehn, J.P., F. Lupi, and M. Kaplowitz. 2010. Stated Choice Experiments with Complex Ecosystem Changes: The Effect of Information Formats on Estimated Variances and Choice Parameters. J. Agr. and Resour. Econ. 35(3):568-590.

Jacquemet, N., R.V. Joule. S. Luchini, J.F. Shogren. 2013. Preference elicitation under oath. J. Environ. Econ. Manage. 65(1):110-132.

Johnson, F.R., and W.H. Desvouges. 1997. Estimating Stated Preferences With Rated-Paired Data: Environmental, Health and Employment Effects of Energy Programs. J. Environ. Econ. Manage. 34:79-99.

Koenpfle, D.T., C.F. Camerer, and J.T. Wang. 2009. Studying Learning in Games Using Eye-tracking. Journal of the European Economic Association 72(2-3):388-398.

Krinsky, I., and A.L. Robb .1986. On Approximating the Statistical Properties of Elasticities. Rev. Econom. Stat. 68(4):715-719.

Magat, W.A., W.K. Viscusi, and J. Huber. 1998. Paired Comparison and Contingent Valuation Approaches to Morbidity Risk Valuation. J. Environ. Econ. Manage. 15:395-411.

McFadden, D. 1974. Conditional Logit Analysis of Qualitative Choice Behaviour. In P. Zarembka, ed., Frontier in Econometrics. New Cork: Academia Press.

Payne, J.W. 1976. Task Complexity and Contingent Processing in Decision Making: An Information Search and Protocol Analysis. Org. Behav. and Human Performance 16:366–387.

Poe, G.L., K.L. Giraud, and J.B. Loomis. 2005. Computational Methods for Measuring the Difference of Empirical Distributions. Am. J. Agric. Econ. 87(2): 353-365.

Ready, R.C., J.C. Whitehead, and G.C. Blomquist. 1995. Contingent Valuation When Respondents Are Ambivalent. J. Environ. Econ. Manage. 29:181-196.

Scarpa, R., and K.G. Willis. 2006. Distribution of Willingness-to-pay for Speed Reduction with Non-positive Bidders: Is Choice Modeling Consistent with Contingent Valuation. Transport Rev. 26(4):451-469.

Scarpa, R., M. Thiene, and D. Hensher. 2010. Monitoring Choice Task Attribute Attendance in Nonmarket Valuation of Multiple Park Management Services: Does it Matter? Land Economics 86(4):817-839.

Schaafsma M., R. Brouwer, A. Gilbert, J. van den Bergh, and A. Wagtendonk. 2013. Estimation of Distance-Decay Functions to Account for Substitution and Spatial Heterogeneity in Stated Preference Research. Land Economics 89:514-537.

Siikamäki, J., and D.F. Layton. 2007. Discrete Choice Survey Experiments: A Comparison Using Flexible Methods. J. Environ. Econ. Manage. 53(1):122-139.

Stevens, T.H., C. Barret, and C.E. Willis. 1997. Conjoint Analysis of Groundwater Protection Programs. Agr. Resource Econ. Rev. 26(2):229-236.

Stevens, T.H., R. Belkner, D. Dennis, D. Kittredge, and C. Willis. 2000. Comparison of Contingent Valuation and Conjoint Analysis in Ecosystem Management. Ecol. Econom. 32:63-74.

Swait, J., and J.J. Louviere. 1993. The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. J. Market. Res. 30(3):305-314.

Tanida, S., and Yamagishi, T. 2010. Testing Social Preferences Through Differential Attention to Own and Partner's Payoff in a Prisoner's Dilemma Game. Letters on Evolutionary Behavioral Science 1(2):31-34

Train, K. 2009. Discrete Choice Methods with Simulation. Cambridge University Press, 2nd edition. New York, 383 p.

Wedel, M., and R. Peters. 2007. A Review of Eye-tracking research in marketing. Review of Marketing Research 4:123-147.

Supplemental Appendix to "Comparing contingent valuation and choice modeling using field and eye-tracking lab data"

Annex 1: Standard contingent valuation and choice modeling questions

Standard contingent valuation question (CV sample)

10a **(code 52)**. Would you be willing to pay **5 euros (ONLY this year)** to fund a reforestation project on land currently occupied by **SHRUBLAND** that will increase by **40,000 hectares** the surface occupied by **STONE PINES** in the southwest of Spain in the next 5 years? Keep in mind that the payment would be real and that you could not use the money for other things.

□ Yes (*p. 10a.1*) □ No (*p. 10a.2*)

10a.1 (If *Yes* to question 10a) Would you be willing to pay **20 euros**?

□ Yes □ No

10a.2 (If *No* to question 10a) Would you be willing to pay **2 euros**?

□ Yes □ No

Standard choice modeling question (CM sample)

Rank the following alternatives from the MOST PREFERRED (1) to the LEAST PREFERRED (3). Keep in mind that the payment would be real and that you could not use the money for other things.

10a.	**SET 1 (code 1)**		
	Option A	**Option B**	**Option C**
Increase in the STONE PINE surface in the southwest of Spain in the next 5 years	20,000 hectares	40,000 hectares	No reforestation
Land use where the reforestation would be carried out	Eucalyptus grove	Shrubland	
Additional taxes ONLY this year	20 euros	35 euros	0 euros
RANK THE THREE OPTIONS (A, B and C)	**OPTION A** **1ª□ 2ª□ 3ª□**	**OPTION B** **1ª□ 2ª□ 3ª□**	**OPTION C** **1ª□ 2ª□ 3ª□**

Annex 2: Results from the standard ranking format and comparison with the standard contingent valuation

Table A2.1 Regressions models and willingness to pay values (with confidence intervals at 95%) for the standard ranking format. Fixed and random parameter estimations.

Variable	Ranking	
	FP	RP
Intercept / ASC_{REF}	1.918***	2.174***
	(0.117)	(0.166)
Vegetation removed (VEG) (=1 shrub; =-1 eucalyptus)	-0.065**	-0.076***
	(0.031)	(0.024)
Area (hectares/10,000) covered by the reforestation (HEC)	0.026	0.035**
	(0.016)	(0.015)
Bid	-0.035***	-0.039***
	(0.002)	(0.002)
Standard deviation parameters		
Intercept / ASC_{REF}		1.176***
		(0.301)
Vegetation removed (VEG) (=1 shrub; =-1 eucalyptus)		
Area (hectares/10,000) covered by the reforestation (HEC)		0.070**
		(0.035)
n	1,036	1,036
AIC	2.971	2.963
Log-likelihood	-1,535.269	-1,529.260
Alternative	Mean WTP	Mean WTP
[SHR-40]	55.20	57.49
	[50.91, 60.19]	[50.14, 65.75]
[EUC-40]	59.00	61.34
	[54.53, 64.18]	[53.71, 69.77]
[SHR-80]	58.18	61.09
	[52.82, 64.24]	[53.44, 69.45]
[EUC-80]	61.97	64.95
	[56.22, 68.51]	[57.13, 73.28]

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively; Standard errors are shown in parentheses; n: number of observations. AIC: Akaike Information criterion.

FP: Fixed parameter estimations.

RP: Random parameter estimations.

Table A2.2 Complete combinatorial test for the comparison of mean WTP values for alternatives [SHR-40], [EUC-40], [SHR-80] and [EUC-80] between the standard contingent valuation and ranking formats. Fixed and random parameter estimations.

Alternative	Complete combinatorial test			
	Single-bounded versus ranking		Double-bounded versus ranking	
	FP estimations	RP estimations	FP estimations	RP estimations
[SHR-40]	<0.001***	0.024**	0.002***	0.002***
[EUC-40]	<0.001***	0.014**	<0.001***	<0.001***
[SHR-80]	<0.001***	0.029**	<0.001***	<0.001***
[EUC-80]	<0.001***	0.020**	<0.001***	<0.001***

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively.

FP: Fixed parameter estimations.

RP: Random parameter estimations.

Annex 3: Exploring order-effects

To investigate the influence of order-effects, we compare subsamples in two scenarios designed to remove the impact of potential order-effects: (i) assuming that order-effects are present in all formats, we analyze the results from the 1st question presented in the questionnaire (we denote these subsamples as SB1, DB1, CR1 and CH1) − in this case, the WTP analysis focuses exclusively on alternative [SHR-40]; (ii) assuming that order-effects are present only in the CV formats, we compare the results for alternative [SHR-40] from the 1st question in the CV subsample (SB1 and DB1 subsamples) with those for alternative [SHR-40] from the complete CM samples (CR and CH subsamples). The scenario where order-effects are present in the CM but not in the CV format is unlikely, and we do not consider it.

Table A3.1 presents the models and WTP values from the subsamples of questions/sets located in the 1st position, and Tables A3.2 and A3.3 present the complete combinatorial test results for the WTP comparisons between formats in the two scenarios described above. Random parameter models do not converge for the CV samples that only analyze the 1st question, and we only use fixed parameter models in these cases. For the CR1 and CH1 subsamples, random parameter models are presented.

The WTP values for alternative [SHR-40] differ between the CV subsamples that analyze the 1st question (SB1 and DB1 subsamples) and the CR1 and CR subsamples. These WTP values are statistically similar in two cases when compared to CH1 and CH. Thus, the comparative results from these order-effect free subsamples are not very different from the WTP comparison results for the standard formats.

Table A3.1 Regression models and willingness to pay values (with confidence intervals at 95%) obtained from the first question in the standard contingent valuation and choice modeling formats. Fixed parameter estimations for the contingent valuation and fixed and random parameter estimations for the choice format.

	Contingent valuation		Choice			
Variable	SB1	DB1	CR1		CH1	
	FP	FP	FP	RP	FP	RP
Intercept / ASC_{REF}	2.077^{***}	2.029^{***}	2.205^{***}	2.801^{***}	1.487^{***}	1.490^{***}
	(0.407)	(0.267)	(0.263)	(0.470)	(0.317)	(0.322)
Vegetation removed (VEG) (=1 shrub; =-1 eucalyptus)			-0.102^{*}	-0.130^{*}	-0.132^{*}	-0.133^{*}
			(0.061)	(0.075)	(0.075)	(0.075)
Area (hectares/10,000) covered by the reforestation (HEC)			0.037	0.106	0.148^{**}	0.148^{**}
			(0.060)	(0.077)	(0.072)	(0.072)
Bid	-0.052^{***}	-0.061^{***}	-0.041^{***}	-0.052^{***}	-0.042^{***}	-0.042^{***}
	(0.012)	(0.006)	(0.006)	(0.008)	(0.007)	(0.006)
Standard deviation parameters						
Intercept / ASC_{REF}				1.302^{*}		
				(0.698)		
Area (hectares/10,000) covered by the reforestation (HEC)				0.404^{***}		0.027
				(0.102)		(0.212)
n	139	139	259	259	259	259
AIC	1.165	2.636	3.015	2.975	1.921	1.929
Log-likelihood	-78.963	-183.201	-386.406	-386.406	-244.848	-244.840
Alternative	Mean WTP	Mean WTP	Mean WTP	Mean WTP	Mean WTP	Mean WTP
[SHR-40]	40.68	33.06	55.03	59.95	46.42	46.74
	[32.65, 51.84]	[27.64, 38.11]	[47.44, 65.72]	[49.02, 72.74]	[37.56, 57.73]	[37.34, 58.88]

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively; Standard errors are shown in parentheses; n: number of observations. AIC: Akaike Information criterion.

SB1: Single-bounded for the first question in the standard contingent valuation.

DB1: Double-bounded for the first question in the standard contingent valuation.

CR1: Ranking for the first set in the standard choice modeling.

CH1: Recoded choice for the first set in the standard choice modeling.

FP: Fixed parameter estimations.

RP: Random parameter estimations.

Table A3.2 Complete combinatorial test for the comparison of mean WTP values for alternative [SHR-40] obtained from the first question in the standard contingent valuation and choice formats. Fixed parameter estimations for the contingent valuation and fixed and random parameter estimations for the choice format.

Alternative	Complete combinatorial test							
	SB1 (FP estimations) versus				DB1 (FP estimations) versus			
	CR1	CH1	CR1	CH1	CR1	CH1	CR1	CH1
	FP estimations		RP estimations		FP estimations		RP estimations	
[SHR-40]	0.019**	0.199	0.009***	0.199	<0.001***	0.005***	<0.001***	0.004***

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively.

SB1: Single-bounded for the first question in the standard contingent valuation.

DB1: Double-bounded for the first question in the standard contingent valuation.

CR1: Ranking for the first set in the standard choice modeling.

CH1: Recoded choice for the first set in the standard choice modeling.

FP: Fixed parameter estimations.

RP: Random parameter estimations.

Table A3.3 Complete combinatorial test for the comparison of mean WTP values for alternative [SHR-40] obtained from the first question in the standard contingent valuation format and from the complete sample in the standard choice format. Fixed parameter estimations for the contingent valuation and fixed and random parameter estimations for the choice format.

Alternative	Complete combinatorial test							
	SB1 (FP estimations) versus				DB1 (FP estimations) versus			
	CR1	CH1	CR1	CH1	CR1	CH1	CR1	CH1
	FP estimations		RP estimations		FP estimations		RP estimations	
[SHR-40]	0.012**	0.069*	0.009***	0.090*	<0.001***	<0.001***	<0.001***	0.359

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively.

SB1: Single-bounded for the first question in the standard contingent valuation.

DB1: Double-bounded for the first question in the standard contingent valuation.

CR: Ranking.

CH: Recoded choice.

FP: Fixed parameter estimations.

RP: Random parameter estimations

Annex 4: Results of the *p-values* for eye tracker comparison tests between the contingent valuation and choice modeling standard formats.

Table A4.1 Results of the t-test (*p-values*) for the differences between the standard contingent valuation and choice modeling formats in the percentage of time devoted to the attributes and the bid relative to the time devoted to completing the whole valuation question, and in the time devoted to completing the whole valuation question. Results for mean values for all questions and for each question position.

Format	Total visit duration to (% of time relative to the time devoted to answer the whole valuation question):			Time devoted to the valuation questions (seconds)
	Attribute VEG values	Attribute HEC values	BID vector values (including €0 in CM)	
	p-values	*p-values*	*p-values*	*p-values*
CV vs CM (1st questions)	0.6401	0.0005***	0.1472	0.0024***
CV vs CM (2nd questions)	0.2759	0.0023***	<0.0001***	0.0903*
CV vs CM (3rd questions)	<0.0001***	0.0191**	0.0035***	0.7413
CV vs CM (4th questions)	0.0027***	0.0003***	0.0001***	0.7742
CV vs CM (all questions)	0.0107**	0.0012***	0.0020***	0.4787
CV vs CM (1st to 4th questions)	0.0290**	0.0027***	0.0066***	0.4176

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively.

CV: Contingent valuation.

CM: Choice modeling.

VEG: attribute "vegetation removed".

HEC: attribute "area (hectares) covered by the reforestation".

Annex 5: Modified contingent valuation and choice modeling questions

Contingent valuation question using an attribute-stimulus format (CV_{CM} sample)

From the following two alternatives, please mark the ONE THAT YOU WOULD CHOOSE (ONLY ONE). Keep in mind that the payment would be real and that that you could not use the money for other things.

10a.	**SET 1 (code 52)**	
	OPTION A	**OPTION B**
Increase in the STONE PINE surface in the southwest of Spain in the next 5 years	40,000 hectares	No reforestation
Land use where the reforestation would be carried out	Shrubland	
Additional taxes ONLY this year	5 euros	0 euros
Please, mark ONLY ONE OPTION	□ **OPTION A (*q. 10a.1*)**	□ **OPTION B (*q. 10a.2*)**

10a.1 (If OPTION A was marked in q. 10a) Which option would you choose if the amount to be paid were **20 euros**?

Additional taxes ONLY this year	20 euros	0 euros
Please, mark ONLY ONE OPTION	□ **OPTION A**	□ **OPTION B**

10a.2 (If OPTION B was marked in q. 10a) Which option would you choose if the amount to be paid were **2 euros**?:

Additional taxes ONLY this year	2 euros	0 euros
Please, mark ONLY ONE OPTION	□ **OPTION A**	□ **OPTION B**

Choice modeling question with dominance between alternatives due to the bid ($CM_{D\text{-}BID}$ sample)

Rank the following alternatives from the MOST PREFERRED (1) to the LEAST PREFERRED (3). Keep in mind that the payment would be real and that you could not use the money for other things.

10c.	**SET 3 (code 35)**		
	Option A	**Option B**	**Option C**
Increase in the STONE PINE surface in the southwest of Spain in the next 5 years	40,000 hectares	40,000 hectares	No reforestation
Land use where the reforestation would be carried out	Shrubland	Shrubland	
Additional taxes ONLY this year	20 euros	50 euros	0 euros
RANK THE THREE OPTIONS (A, B and C)	**OPTION A** **1ª□ 2ª□ 3ª□**	**OPTION B** **1ª□ 2ª□ 3ª□**	**OPTION C** **1ª□ 2ª□ 3ª□**

Choice modeling question with dominance between alternatives due to the attribute "area (hectares) covered by the reforestation" ($CM_{D\text{-}HEC}$ sample)

Rank the following alternatives from the MOST PREFERRED (1) to the LEAST PREFERRED (3). Keep in mind that the payment would be real and that you could not use the money for other things.

10c.	**SET 3 (code 36)**		
	Option A	**Option B**	**Option C**
Increase in the STONE PINE surface in the southwest of Spain in the next 5 years	20,000 hectares	40,000 hectares	No reforestation
Land use where the reforestation would be carried out	Shrubland	Shrubland	
Additional taxes ONLY this year	20 euros	20 euros	0 euros
RANK THE THREE OPTIONS (A, B and C)	**OPTION A** **1ª□ 2ª□ 3ª□**	**OPTION B** **1ª□ 2ª□ 3ª□**	**OPTION C** **1ª□ 2ª□ 3ª□**

Annex 6: Results from the modified contingent valuation and comparison with the standard contingent valuation.

Table A6.1. Percentage of choices of shared alternatives over the status quo in the modified contingent valuation format. *Z-test* statistics for the difference between proportions with the standard contingent valuation format. The percentages of choices of shared alternatives over the status quo from the standard contingent valuation are reported in Table 2 in the main text.

Alternative	Percentage of choices of the alternatives over the status quo		*z-test* results for the difference between proportions	
	CV_{CM}		CV versus CV_{CM}	
	SB	DB	SB	DB
[SHR-40-€35]	69%	50%	0.170	1.000
[SHR-40-€20]	67%	61%	1.000	0.681
[EUC-40-€50]	38%	54%	0.645	0.040**
[EUC-40-€20]	69%	59%	0.855	0.911
[SHR-80-€5]	86%	94%	0.962	0.189
[SHR-80-€50]	52%	47%	0.205	0.040**
[EUC-80-€20]	47%	73%	0.165	0.611
[EUC-80-€50]	17%	70%	0.432	0.393

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively.

SB: Single-bounded.

DB: Double-bounded.

SB_{CM}: Single-bounded using an attribute-stimulus format.

DB_{CM}: Double-bounded using an attribute-stimulus format.

Table A6.2 Regression models and willingness to pay values (with confidence intervals at 95%) for the modified contingent valuation format. Fixed and random parameter estimations.

Variable	Modified contingent valuation (CV_{CM})			
	Single-bounded		Double-bounded	
	FP	RP	FP	RP
Intercept / ASC_{REF}	2.086***	2.194***	1.225***	1.228***
	(0.346)	(0.800)	(0.223)	(0.231)
Vegetation removed (VEG) (=1 shrub; =-1 eucalyptus)	0.206**	0.219*	0.119*	0.120*
	(0.093)	(0.118)	(0.063)	(0.064)
Area (hectares/10,000) covered by the reforestation (HEC)	-0.077*	-0.078	-0.039	-0.039
	(0.047)	(0.054)	(0.031)	(0.032)
Bid	-0.046***	-0.048***	-0.022***	-0.023***
	(0.006)	(0.017)	(0.003)	(0.004)
Standard deviation parameters				
Intercept / ASC_{REF}				
Vegetation removed (VEG) (=1 shrub; =-1 eucalyptus)		0.416		0.090
		(2.056)		(1.255)
Area (hectares/10,000) covered by the reforestation (HEC)		0.054		0.015
		(0.118)		(0.070)
n	548	548	548	548
AIC	1.236	1.242	1.313	1.317
Log-likelihood	-334.651	-334.511	-715.799	-715.772
Alternative	Mean WTP	Mean WTP	Mean WTP	Mean WTP
[SHR-40]	43.79	44.59	53.14	53.11
	[36.22, 52.73]	[35.40, 60.26]	[42.25, 66.95]	[42.68, 66.48]
[EUC-40]	34.67	35.11	42.45	42.44
	[27.97, 42.56]	[25.30, 48.45]	[33.10, 54.48]	[32.81, 54.41]
[SHR-80]	36.88	37.20	46.09	46.04
	[29.51, 44.65]	[27.40, 48.42]	[35.48, 58.30]	[36.04, 59.28]
[EUC-80]	27.76	27.72	35.40	35.37
	[20.78, 34.52]	[17.07, 36.97]	[26.23, 45.80]	[26.21, 46.23]

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively; Standard errors are shown in parentheses; n: number of observations.

AIC: Akaike Information criterion.

CV_{CM}: Contingent valuation using an attribute-stimulus format.

FP: Fixed parameter estimations

RP: Random parameter estimations

Table A6.3 Complete combinatorial test for the comparison of mean WTP values for alternatives [SHR-40], [EUC-40], [SHR-80] and [EUC-80] between the modified contingent valuation and the standard contingent valuation formats. Fixed and random parameter estimations.

Alternative	Complete combinatorial test			
	CV versus CV_{CM}			
	FP estimations		RP estimations	
	Single-bounded	Double-bounded	Single-bounded	Double-bounded
[SHR-40]	0.087*	0.030**	0.153	0.030**
[EUC-40]	0.405	0.146	0.438	0.162
[SHR-80]	0.211	0.092*	0.314	0.088*
[EUC-80]	0.343	0.346	0.394	0.363

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively.

CV: Contingent valuation.

CV_{CM}: Contingent valuation using an attribute-stimulus question format.

FP: Fixed parameter estimations.

RP: Random parameter estimations.

Annex 7: Single-alternative analysis comparison between the contingent valuation and the choice modeling (CM_D) modified formats.

Table A7.1 *Z-test* statistics for the difference between proportions of choices of shared alternatives over the status quo with the modified contingent valuation and choice formats. The percentages of choices of shared alternatives over the status quo for the modified contingent valuation format are reported in Table 6.1 in Annex 6, and for the modified choice format are reported in Table 8 in the main text.

Alternative	*z-test* for the difference between proportions					
	CV_{CM} versus $CM_{D\text{-}BID}$ (ranking)		CV_{CM} versus $CM_{D\text{-}BID}$ (choice)		CV_{CM} versus $CM_{D\text{-}HEC}$ (choice)	
	Single-bounded	Double-bounded	Single-bounded	Double-bounded	Single-bounded	Double-bounded
[SHR-40-€20]	n. a.	n. a.	0.531	0.292	0.323	0.156
[EUC-40-€20]	n. a.	n. a.	0.509	0.129	0.464	0.106
[SHR-80-€50]	0.572	0.430	n. a.	n. a.	0.305	0.243
[EUC-80-€50]	0.001***	0.304	n. a.	n. a.	0.001***	0.407

Note: For the D-BID subscript samples, alternatives [SHR-40-€20] and [EUC-40-€20] are dominant and alternatives [SHR-80-€50] and [EUC-80-€50] are dominated. For the D-HEC subscript sample, all analyzed alternatives are dominant.

n. a.: not applicable.

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively.

CV_{CM}: Contingent valuation using an attribute-stimulus format.

$CM_{D\text{-}BID}$: Choice modeling with dominance between the "do something" alternatives due to the bid.

$CM_{D\text{-}HEC}$: Choice modeling with dominance between the "do something" alternatives due to the attribute "area (hectares) covered by the reforestation".

Annex 8: Valuation question and single alternative analysis for the modified choice modeling format where the only difference between "do something" alternatives is in the level of a single attribute.

Choice modeling question where the only difference between "do something" alternatives is in the level of a single attribute (CM_{FIX} sample)

Rank the following alternatives from the MOST PREFERRED (1) to the LEAST PREFERRED (3). Keep in mind that the payment would be real and that you could not use the money for other things.

10c.	**SET 3 (code 33)**		
	Option A	**Option B**	**Option C**
Increase in the STONE PINE surface in the southwest of Spain in the next 5 years	40,000 hectares	40,000 hectares	No reforestation
Land use where the reforestation would be carried out	Shrubland	Eucalyptus grove	
Additional taxes ONLY this year	35 euros	20 euros	0 euros
RANK THE THREE OPTIONS (A, B and C)	**OPTION A** **1ª□ 2ª□ 3ª□**	**OPTION B** **1ª□ 2ª□ 3ª□**	**OPTION C** **1ª□ 2ª□ 3ª□**

Table A8.1 Percentage of choices of shared alternatives over the status quo in the choice modeling format where the only difference between "do something" alternatives is in the level of a single attribute (CM_{FIX}). *Z-test* statistics for the difference between proportions with the standard and modified contingent valuation formats. The percentages of choices of shared alternatives over the status quo for the standard contingent valuation are reported in Table 2 in the main text, and for the modified contingent valuation are reported in Table 6.1 in Annex 6.

Alternatives	Percentage of choices of the alternatives over the status quo		*z-test* results for comparison between proportions							
			CV versus CM_{FIX}				CV_{CM} versus CM_{FIX}			
			Ranking		Choice		Ranking		Choice	
	CM_{FIX} (ranking)	CH_{FIX} (choice)	Single-bounded	Double-bounded	Single-bounded	Double-bounded	Single-bounded	Double-bounded	Single-bounded	Double-bounded
[SHR-40-€35]	88%	40%	0.001***	0.001***	0.310	0.453	0.035*	0.001***	0.012**	0.423
[EUC-40-€50]	77%	26%	0.001***	0.001***	0.555	0.323	0.001***	0.088*	0.267	0.054*
[SHR-80-€5]	91%	54%	0.409	0.126	0.002***	0.077*	0.438	0.746	0.002***	0.004***
[EUC-80-€20]	91%	74%	0.001***	0.005***	0.244	0.506	0.001***	0.029**	0.001***	0.954

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively.

CV: Contingent valuation.

CV_{CM}: Contingent valuation using an attribute-stimulus format.

CM_{FIX}: Choice modeling where the only difference between "do something" alternatives is in the level of a single attribute.

Annex 9: Results of the *p-values* for eye tracker comparison tests between the contingent valuation and choice modeling modified formats, and logit analysis of the "yes" responses to the contingent valuation and choice questions in these samples.

Table A9.1 Results of the *t-test* (*p-values*) for the differences between the modified contingent valuation and choice modeling formats in the percentage of time devoted to the attributes and the bid relative to the time devoted to completing the whole valuation question, and in the time devoted to completing the whole valuation question. Results for mean values for all questions and for each question position.

Format	Total visit duration to (% of time relative to the time devoted to answer the whole valuation question):			Time spent devoted to the valuation questions (seconds)
	Attribute VEG values	Attribute HEC values	BID vector values (including €0 in CM)	
	p-values	*p-values*	*p-values*	*p-values*
CV_{CM} vs CM_D (1^{st} question)	<0.0001***	<0.0001***	0.0560*	0.0089***
CV_{CM} vs CM_D (2^{nd} question)	0.0190**	0.0178**	0.1532	0.2898
CV_{CM} vs CM_D (3^{rd} question)	0.0055***	0.1923	0.7795	0.2627
CV_{CM} vs CM_D (4^{th} question)	0.0414**	0.0384**	0.9661	0.9894
CV_{CM} vs CM_D (all questions)	0.0104**	0.0196**	0.5546	0.4114
CV_{CM} vs CM_D (1^{st} to 4^{th}	0.0187**	0.0425**	0.4789	0.4741
CV_{CM} vs CV (1^{st} question)	<0.0001***	0.1215	0.1147	0.3945
CV_{CM} vs CV (2^{nd} question)	0.3815	0.2872	0.0015***	0.1203
CV_{CM} vs CV (3^{rd} question)	0.0660	0.1594	<0.0001***	0.0022***
CV_{CM} vs CV (4^{th} question)	0.6869	0.1486	0.0002***	0.0024***
CV_{CM} vs CV (all questions)	0.4200	0.6667	0.0064***	0.3800
CV_{CM} vs CM (1^{st} question)	<0.0001***	<0.0001***	0.9444	0.0086***
CV_{CM} vs CM (2^{nd} question)	0.0420**	0.0006***	0.2203	0.9215
CV_{CM} vs CM (3^{rd} question)	<0.0001***	0.4574	0.1561	0.0374**
CV_{CM} vs CM (4^{th} question)	0.0115**	0.0047***	0.8187	0.0808*
CV_{CM} vs CM (all questions)	0.0018***	0.0034***	0.5462	0.9755
CV_{CM} vs CM (1^{st} to 4^{th} questions)	0.0049***	0.0106**	0.9794	0.1767
CV vs CM_D (1^{st} question)	0.0898*	0.0032***	0.4554	0.0009***
CV vs CM_D (2^{nd} question)	0.1338	0.1235	0.0129**	0.0162
CV vs CM_D (3^{rd} question)	0.0009***	0.0078***	0.0015***	0.3402
CV vs CM_D (4^{th} question)	0.0163**	0.0022***	0.0003***	0.0028***
CV vs CM_D (all questions)	0.0368**	0.0056***	0.0088***	0.0997*
CV vs CM_D (1^{st} to 4^{th} questions)	0.0665*	0.0107**	0.0163**	0.9915
CM vs CM_D (1^{st} question)	0.2287	0.9559	0.0706*	0.3253
CM vs CM_D (2^{nd} question)	0.5864	0.0232**	0.3643	0.2911
CM vs CM_D (3^{rd} question)	0.6412	0.5071	0.3298	0.3301
CM vs CM_D (4^{th} question)	0.8694	0.3400	0.7882	0.0820*
CM vs CM_D (5^{th} question)	0.8549	0.2087	0.3270	0.1761
CM vs CM_D (6^{th} question)	0.4794	0.1649	0.1624	0.3105
CM vs CM_D (7^{th} question)	0.0677*	0.0855*	0.6680	0.5029
CM vs CM_D (8^{th} question)	0.3224	0.7605	0.8083	0.3024
CM vs CM_D (all questions)	0.9900	0.3392	0.9458	0.5303
CM vs CM_D (1^{st} to 4^{th} questions)	0.9592	0.4946	0.4920	0.4683

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively.

CV: Standard contingent valuation; CV_{CM}: Contingent valuation using an attribute-stimulus format; CM: Standard choice modeling; CM_D: Choice modeling with dominance between alternatives; VEG: attribute "vegetation removed"; HEC: attribute "area (hectares) covered by the reforestation".

Table A9.2 Logit analysis of "yes" responses (paying for the environmental good) to the modified contingent valuation and choice questions in the eye-tracking experiment. Models including all questions made to every respondent and including only the four first questions in the choice format.

Variable	Single-bounded	Double-bounded	Choice		Pooled single-bounded and choice	
	All questions	All questions	All questions	4 first questions	All questions	4 first questions
Intercept	1.128	1.116*	3.707***	4.699**	3.603***	4.336***
	(0.688)	(0.601)	(0.924)	(2.350)	(0.683)	(1.053)
% time BID	3.629	5.425	-2.044	-3.119	0.062	0.868
	(4.335)	(3.965)	(2.727)	(3.597)	(2.867)	(3.634)
% time attribute VEG	-15.339***	-14.862***	9.997	83.372	-3.116	-8.253**
	(5.875)	(5.700)	(8.484)	(89.233)	(3.167)	(3.974)
% time attribute HEC	4.608	3.779	-2.952	118.854	0.717	4.087
	(5.982)	(5.589)	(5.660)	(88.288)	(3.663)	(4.880)
Time spent on the valuation questions (seconds)	0.021	-0.015	-0.028**	-0.153*	-0.006	-0.009
	(0.017)	(0.012)	(0.011)	(0.094)	(0.010)	(0.012)
CV_{CM} format					-2.079***	-2.675***
					(0.489)	(0.751)
n	104	104	206	102	310	206
AIC	1.043	1.298	0.338	0.175	0.610	0.702

Asterisks (e.g.,*,**,***) denote significance at the 10%, 5%, and 1% levels, respectively; Standard errors are shown in parentheses; n: number of observations. AIC: Akaike Information criterion.

VEG: attribute "vegetation removed".

HEC: attribute "area (hectares) covered by the reforestation".

CV_{CM} format: dummy = 1 if the response was obtained with the modified contingent valuation format, = 0 if the response was obtained with the modified choice modeling format.