ECOLOGICAL ECONOMICS 59 (2006) 247–255

available at www.sciencedirect.com

ScienceDirect

www.elsevier.com/locate/ecolecon

METHODS

A pilot test of a new stated preference valuation method: Continuous attribute-based stated choice

Richard Ready, Ann Fisher, Dennis Guignet, Richard Stedman, Junchao Wang*

Department of Agricultural Economics and Rural Sociology, Pennsylvania State University, University Park, PA 16802, United States

ARTICLE INFO

Article history:
Received 31 August 2005
Received in revised form
3 November 2005
Accepted 4 November 2005
Available online 4 January 2006

Keywords:
Nonmarket valuation
Attribute based stated choice
Global climate change

ABSTRACT

A new stated preference nonmarket valuation technique is developed. In an interactive computerized survey, respondents move continuous sliders to vary levels of environmental attributes. The total cost of the combination of attributes is calculated according to a preprogrammed cost function, continuously updated and displayed as respondents move the sliders. Each registered choice reveals the respondent's marginal willingness to pay for each of the attributes. The method is tested in a museum exhibit on global climate change. Two construct validity tests were conducted. Responses are sensitive to the shape of the cost function in ways that are consistent with expectations based on economic theory. Implied marginal willingness to pay values were similar to those estimated using a more traditional paired comparisons stated choice format. However, responses showed range effects that indicate potential cognitive biases.

1. Introduction

The first study using surveys to elicit values for environmental goods was by Robert Davis (1963). Since then, several different formats have been used in contingent valuation. Early studies used value elicitation formats that generate continuous measures of maximum willingness to pay (WTP). These include open-ended questions, payment cards, and iterative bidding. In each case, a specific improvement in environmental quality is described to survey respondents. Respondents are told that the improvement is feasible, but would entail a cost to households like theirs. In an open-ended value elicitation question, the respondent is asked to state the largest amount of money she would pay in return for the environmental improvement. In the payment card approach, the respondent is presented with a printed list of dollar amounts, and is asked to choose the largest printed amount she would pay. In iterative bidding, the respondent is presented with a single dollar amount (the bid) and asked whether she would pay that amount in return for the improvement in environmental quality. The bid is then iteratively adjusted up or down depending on the respondent's answers, until the respondent's switch point (where the response switches from "no" to "yes") is identified.

While these approaches are efficient in the sense that they provide a precise WTP measure for each survey respondent, there are concerns about their validity. The iterative bidding approach has been shown to exhibit starting point effects, where the switch point is influenced by the first bid presented (Boyle et al., 1985; Holmes and Kramer, 1995). Payment card responses have in some cases shown range effects, where the stated WTP is influenced by the range of dollar amounts printed on the card, though that result is not universal (Dubourg et al., 1997; Rowe et al., 1996). Because open-ended valuation questions do not mimic everyday purchase decisions, they are thought to be unfamiliar and unrealistic to

* *Corresponding author*. Tel.: +1 814 863 5575; fax: +1 814 865 3746.
E-mail address: rready@psu.edu (R. Ready).

0921-8009/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.ecolecon.2005.11.004

respondents (Arrow et al., 1993). All three approaches have the potential for strategic responses, depending on the respondent's understanding of how the research results will be used (Carson et al., 1999).

Starting with Bishop and Heberlein (1979), dichotomous choice (DC) or referendum formats were increasingly used during the 1980s and early 1990s. In DC contingent valuation, respondents are presented with only one bid, with no follow-up. Respondents state either that they would pay the bid amount or that they would not pay the bid amount. Bid amounts are randomly assigned to respondents, allowing identification of the distribution of WTP across respondents. DC contingent valuation has several purported advantages over continuous elicitation formats. First, it is thought to be more familiar to respondents because of its similarity to everyday purchase decisions. Second, starting point and range effects do not arise, though there is a concern over potential yea-saying, where respondents give positive responses to DC questions without considering the size of the bid (Boyle et al., 1998; Holmes and Kramer, 1995; Ready et al., 1996). Third, the DC format can be incentive compatible, so that the respondent's dominant strategy is to truthfully reveal preferences (Hoehn and Randall, 1987). One drawback of the DC format is that it does not generate a precise measure of WTP for each respondent. Instead, each DC response generates a discrete datum that reveals only whether WTP is greater than or less than the posed bid.

A limitation of both continuous and discrete contingent valuation is that they can value only one change in environmental quality at a time. If environmental quality involves multiple attributes, then a discrete change in each attribute is described. If the analyst wants to evaluate a different change in environmental quality, one that involves a larger or smaller discrete change or a different combination of changes in multiple attributes, then multiple surveys or survey versions are required.

Starting in the early 1990s there has been a shift in valuation practice toward the use of attribute-based stated choice approaches (Holmes and Adamowicz, 2003). In a typical stated choice format, the respondent chooses from a set of two or more options; each option differs in its levels of the environmental quality attributes and in the cost to the respondent. Based on the stated choices, marginal utility for each of the attributes is estimated using a random utility model. With these estimated marginal utilities, it is possible to calculate WTP for any proposed change in environmental quality. A limitation of the stated choice format is that it too generates discrete data, in the sense that each stated choice reveals only ordinal preferences among the options presented in that choice. Many stated choices from each respondent and a large sample size of respondents are needed to identify marginal utility for each of the attributes.

The purpose of this study is to develop a new elicitation method, called the continuous attribute-based stated choice method (CABSCM), that values attributes of environmental quality but collects continuous data. In an interactive computer-based survey, respondents choose levels of each attribute. The total cost of the package of attributes is continuously updated according to a cost function determined by the researcher. Each respondent's choice reveals his or her marginal WTP for each attribute. In a large sample ($n=14{,}100$) pilot test, internal validity tests are conducted for the CABSCM. In a second, small sample test, responses to the CABSCM are compared to responses from a more traditional discrete attribute-based stated choice method survey (DABSCM).

2. Theoretical foundation

In a typical DABSCM survey, the respondent makes a series of choices among sets of options. In each choice, a given option, j, is defined by the levels of the attributes of environmental quality that would result if that option is chosen, $\mathbf{a}_j=\mathbf{a}_j^1, \ldots, \mathbf{a}_j^M$, and in the cost to the respondent, c_j. The choice data is typically analyzed using multinomial logit regression. The respondent's utility from option j is assumed to follow a form given by

$$U_j = V(\mathbf{a}_j, c_j) + \varepsilon_j = \beta'\mathbf{a}_j - \gamma c_j + \varepsilon_j \tag{1}$$

where $V(\mathbf{a}_j, c_j)$ is the deterministic component of utility, and ε_j is an option-specific error term distributed according to a type II extreme value distribution.

The respondent is assumed to choose the option that gives the highest utility among those available in a given choice set. The probability that option j is chosen from a set of options 1, …, j, …, J is then

$$\Pr\{j\} = \Pr\{U_j \geq U_i \text{ for all } i \varepsilon 1, \ldots J\} = \frac{\exp(\beta'\mathbf{a}_j - \gamma c_j)}{\sum_{i \varepsilon J} \exp(\beta'\mathbf{a}_i - \gamma c_i)} \tag{2}$$

The parameters of the utility function, β and γ, are estimated using maximum likelihood techniques. The marginal willingness to pay for a given attribute, $\mathbf{a}^m$, is then the ratio of the marginal utility for that attribute to the marginal utility of income, β^m/γ. The total willingness to pay for a discrete change that affects multiple attributes of environmental quality, $\Delta\mathbf{a}$, is given by $\boldsymbol{\beta}'\Delta\mathbf{a}/\gamma$.

The information needs to estimate a utility function from DABSCM data are high. As in DC contingent valuation, each choice response reveals only partial information about preferences. To estimate marginal utilities requires a sample of many respondents, each facing several choices. Typically each respondent is asked to make eight or more choices. When designing a DABSCM survey, care must be taken to construct a set of choices that identifies the parameters β and γ including any attribute interactions that may be of interest. When the number of attributes is large, and interactions are of interest, it may be necessary to construct an experimental design that includes more than eight choices. In such cases, the choices may be divided into subsets, so that different respondents face a different set of choices.

The resulting parameters are aggregates across all respondents, and calculated willingness-to-pay values can be interpreted as median estimates for the respondent population. The marginal utilities can be individualized to some extent by interacting attribute levels or cost with measurable characteristics of the individual respondents. Alternatively, a random parameters or latent class model can approximate variability in preferences within the population (Greene and Hensher, 2003). However, even with these statistical approaches, individual WTP is not directly measured. Rather, it is inferred from population parameters.

In contrast, the CABSCM generates unique marginal WTP values for each attribute for each respondent. The preliminary material in a CABSCM survey is similar to what would be included in a DABSCM survey. The attributes are described, and motivation is given for how they might vary independently and how that might affect the respondent's budget. In a CABSCM survey, the respondent is then presented a computer screen with continuous sliders for each attribute in the design.

Fig. 1 shows one of the slider screens used in this study. This particular task presents respondents with the possibility that society might invest in planting trees to sequester carbon dioxide, or in improved energy efficiency in buildings to reduce carbon dioxide emissions. On the screen, respondents can choose any level of tree planting from 0 to 13 million acres, and any level of building efficiency improvement from 0% to 55%. As sliders A and B are moved, slider C (monthly cost per household) moves with them according to a preprogrammed cost function. The total reduction in net US carbon dioxide emissions that would result if the respondent's choices were implemented is given as well. When the respondent is satisfied with the set of attributes and their cost, she enters her choices. A popup screen repeats the total cost, and asks the respondent if she is sure she is willing to pay that much money. The response is recorded when the respondent answers "yes" to the "are you sure" question.

The theoretical foundation for this design differs from that for the DABSCM. Here, it is assumed that each respondent, i, has a utility function $U^i(\mathbf{a}, c)$. Given a cost function, $c=f(\mathbf{a})$, the respondent solves the problem

$$\max_{\mathbf{a}} \quad U^i(\mathbf{a}, f(\mathbf{a})) \tag{3}$$

A necessary first order condition for an interior solution to this maximization is

$$-\frac{\partial U^i/\partial \mathbf{a}_j}{\partial U^i/\partial c} = \partial f/\partial \mathbf{a}_j \tag{4}$$

The left hand side of Eq. (4) is the respondent's marginal WTP for attribute j, while the right hand side is the marginal cost at the combination of attributes chosen.

This choice is shown graphically for one attribute in Fig. 2. The cost function is designed such that marginal costs increase as the level of the attribute increases. This curvature in the cost function increases the chance of an interior solution, even in cases where indifference curves are relatively straight. Three indifference curves are shown, drawn to reflect the fact that trees generate positive utility while costs generate disutility. The respondent's optimal choice occurs where an indifference curve is tangent to the cost curve.

In practice, the respondent is presented with all of the attribute sliders at the bottom of their ranges. She pushes each attribute's slider up as long as her marginal WTP for that attribute is larger than the marginal cost of providing more of the attribute. At the final quantity/cost chosen, marginal WTP is exactly equal to the marginal cost of providing the attribute. Because we know the marginal cost of provision for each attribute, we know the respondent's marginal WTP for each attribute.

Each choice in a CABSCM survey, then, precisely identifies the respondent's marginal WTP for each attribute in the choice, assuming an interior choice. Corner solutions, where

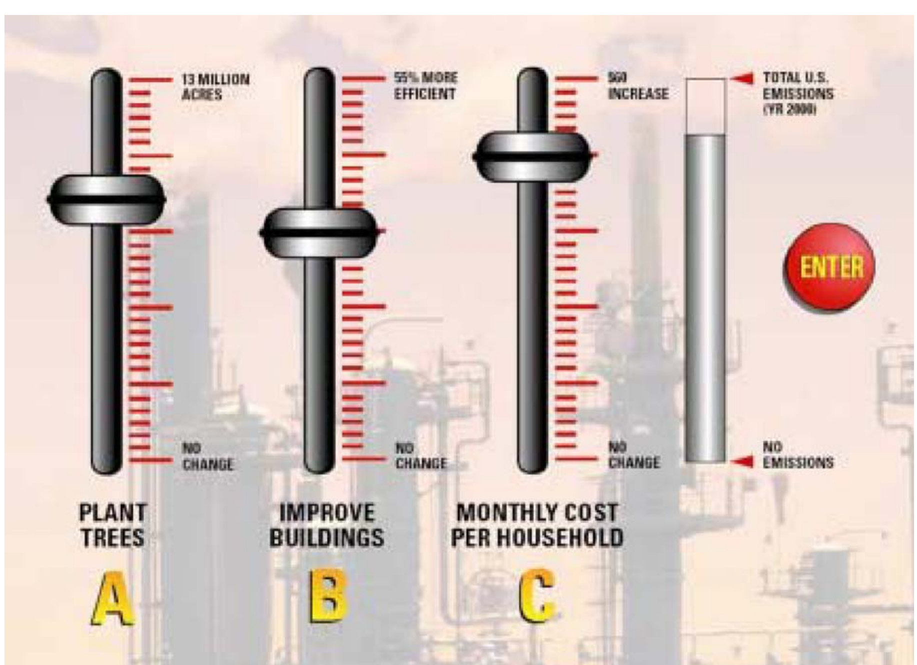

Fig. 1 – CABSCM choice screen.

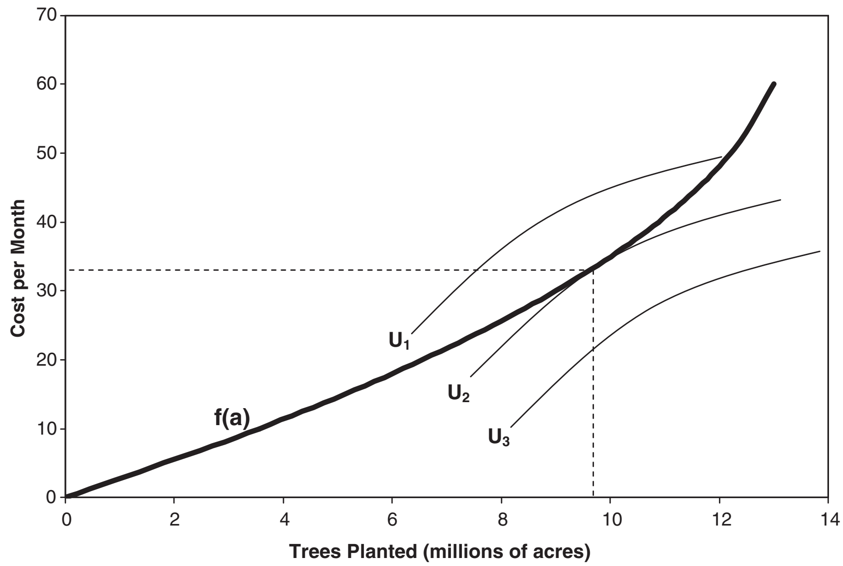

Fig. 2 – Optimal choice in a CABSCM survey.

the attribute slider is set at the top or the bottom of the range, provide truncated observations on marginal WTP. Because individual estimates of marginal WTP are generated, it is possible to explore how marginal WTP varies across respondents. A marginal value function that predicts marginal WTP as a function of characteristics of the respondent can be estimated using regression techniques suitable for truncated data, for example Tobit regression.

3. Methods

3.1. *Museum installation*

The interactive computer survey was designed to be included as an installation in an exhibit on global climate change at the Marian Koshland Museum of Science, which opened at the National Academy of Sciences in Washington, DC, in April 2004. The installation was designed as an interactive kiosk, with a large flat panel screen, a trackball, and a button for entering choices.[1] Part of the purpose of the installation was to teach museum visitors about the role of tradeoffs when dealing with global climate change.

Working in consultation with the museum designers, three separate topics were developed. Visitors could choose to view any or all of the topics. In each topic, a tradeoff between two attributes is motivated with a short film. The respondent then is presented with a choice screen such as that shown in Fig. 1. The attributes include a mix of actions to mitigate (avoid) global climate change and actions to adapt to its consequences. These are briefly described in Table 1.

[1] A web version of the survey can be viewed at http://www.koshland-science-museum.org. Click on “Global Warming Facts & Our Future” and then on “Consider the Alternatives” to start the program.

For each attribute, a slider is provided that is labeled at the bottom and top with the lower and upper bound values shown in Table 1. For drinking water quality, the slider is calibrated using electrical conductivity units, but these units are not presented to the respondents. Instead, three labels are presented: “unfit for human consumption” at the bottom of the slider, “tastes slightly salty” at a level corresponding to 700 conductivity units, and “tastes fresh” at the top of the slider. For Topic 1 only, a fourth bar is included that shows the total impact on net CO_2 emissions from the choices made.

Because the survey would be part of a museum installation, the cost figures had to reflect the best available estimates of actual costs. For each attribute, upper and lower bounds on the levels were chosen so that the cost per US household to achieve the upper bound was about \$60, based on published estimates (National Academy of Sciences, 1992; National Assessment Synthesis Team, 2000; U.S. Department of State, 2002). The respondent then could commit to spend anywhere from \$0 to \$60 on each of the two attributes. While the cost function endpoints were based on published estimates, there was little guidance over the shape of the cost function in between the two endpoints. It is important that the cost function be nonlinear, in order to identify marginal WTP for each respondent. Quadratic cost curves were specified for each attribute that passed through the two endpoints, with increasing marginal cost throughout the range.

3.2. *Statistical analysis of CABSCM data*

Because the sliders have upper and lower limits, the data generated by the CABSCM is censored. The Tobit regression model is used to calculate population averages for both the level chosen and the implied marginal WTP for each attribute. The Tobit regression model assumes that each of these variables follows a normal distribution, but that some observations are censored. If the variable of interest

Table 1 – Topics and attributes

Tradeoff	Units	Lower bound	Upper bound	Inverse cost function
Topic 1—CO_2 emissions				
Plant trees to sequester CO_2	Millions of acres per year	0	13	$A=0.383C-0.00278C^2$
Improve building energy efficiency	Percent increase	0	+55	$A=1.75C-0.0139C^2$
Topic 2—Increased frequency and severity of storms				
Replace wetlands lost to increased storms	Millions of acres lost	−5.8	+1.0	$A=-5.8+0.21C-0.00167C^2$
Avoid property damage from increased storms	Percent increase in damage	+30	−10	$A=30-1.33C+0.0111C^2$
Topic 3—Sea level rise				
Avoid property damage from sea-level rise	Percent increase in damage	+30	−10	$A=30-1.33C+0.0111C^2$
Protect drinking water from saltwater intrusion from sea-level rise	Electrical Conductivity (1000 units)	2.5	0	$A=2.5-0.0783C+0.000611C^2$

is denoted by X, then X is normally distributed with mean, μ, and variance, σ^2. Instead of observing X directly, we observe Y, given by

$$\begin{aligned} Y &= X \quad \text{if} \quad L_l < X < L_u \\ Y &= L_l \quad \text{if} \quad X \le L_l \\ Y &= L_u \quad \text{if} \quad X \ge L_u \end{aligned} \tag{5}$$

where L_l is the lower limit on the slider and L_u is the upper limit on the slider. Maximum likelihood methods are used to estimate μ and σ.

3.3. Internal validity tests

Two tests of the validity of the CABSCM method were designed as part of the museum installation. Three versions of the installation software were developed. Table 1 reflects the design of Version 1. The three versions rotated in the museum on a 12-day cycle, 4 days for each version.

The first validity test explored whether responses were sensitive to the curvature of the cost function. Version 2 of the installation software used a linear cost function instead of the quadratic function shown in Table 1. The slope of the linear cost function was set so that the cost of reaching the upper bound level of the attribute was still \$60. Fig. 3 shows how a respondent would be expected to react to the two different cost functions for the tree planting choice. If preferences are homothetic, a typical respondent would be expected to choose a higher level of the attribute but spend less money on it when faced with the quadratic cost function than when faced with the linear cost function.

The second validity test explored possible range effects. Version 3 of the installation software used the same quadratic cost functions as Version 1, but the upper and lower bounds were set so that the maximum amount that could be “spent” on each attribute was \$30, instead of \$60. If range effects exist, then we would expect more responses between \$0 and \$30 in Version 3 than in Version 1, and a downward shift of those responses, in both the attribute levels and the amount of money committed. If range effects do not exist, then the distribution of responses between \$0 and \$30 should be the same in the two versions.

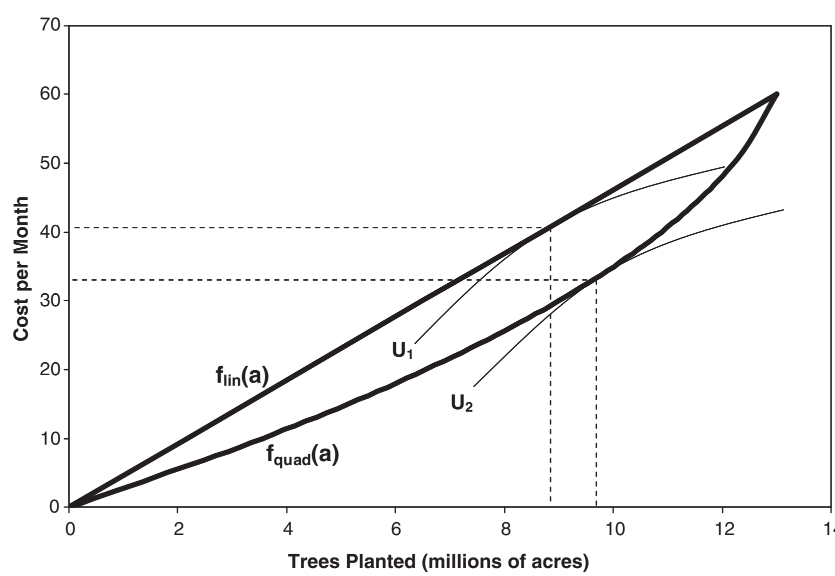

Fig. 3 – Cost function validity test.

3.4. *Comparison between CABSCM and DABSCM*

In addition to the data collected from museum visitors, a second study was conducted that compared results obtained from the CABSCM survey to those from a more typical DABSCM survey. A paper survey was designed that presented respondents with four DABSCM choices; for each choice, respondents chose between two combinations of trees planted and building efficiency. The cost of each option in the DABSCM survey was calculated using the same quadratic cost function, so that each option represented a feasible combination in the quadratic cost CABSCM survey.

Interviews were conducted inside the Koshland museum with museum visitors, and on the Pennsylvania State University campus, with randomly selected faculty and staff. After viewing the introductory video included in the museum installation, each respondent completed both the CABSCM and the DABSCM survey. The order of the two surveys was randomized. The first analysis was a within-respondent comparison of the two survey methods. For each respondent, the marginal WTP for each attribute was determined from the CABSCM response. These were then used to determine which DABSCM response should give higher utility. A preference reversal occurs if the DABSCM choice differed from the choice that was predicted based on the CABSCM response.

A second comparison was made between marginal WTP for the attributes estimated from the DABSCM responses and marginal WTP estimated from the CABSCM survey. The former is estimated using a standard multinomial logit regression. Aggregate measures of the latter are estimated using a Tobit regression on marginal WTP. To simulate what would happen if each method was administered by itself, only responses to the first survey administered are included in this comparison.

4. Results

4.1. *Sensitivity to curvature in the cost function and range effects*

Data was collected from the Koshland Museum installation from April 2004 through January 2005. A total of 14,100 CABSCM choices were recorded in the museum installation. Because the installation was part of a museum exhibit, only limited demographic data was collected. Structured observation of museum patrons using the installation showed that the three topic areas were chosen by roughly equal proportions of users, and that about half of users chose more than one topic. The average dwell time at the installation was 6 min and 14 s, which was over twice as long as any other installation in the global warming exhibit. Of users who approached the installation, 85% completed a choice screen.

For each attribute/cost function combination, the population mean attribute level was estimated using Tobit regression, with upper and lower limits as shown in Table 1. Population mean cost commitment for each attribute is also estimated using Tobit regression, with upper and lower limits of $0 and $60 ($30 for Version 3). These are presented in Table 2. For each topic/cost function combination, the mean level of the attribute or cost is presented as well as the population standard deviation (sigma). Standard errors for each estimated parameter are presented in parentheses.

The first validity test compares responses to the linear CABSCM survey to those from the quadratic CABSCM survey. Theory predicts that the average respondent should choose a higher level of each attribute but commit to a lower cost when facing a quadratic cost function (columns 1 and 2 in Table 2) than when facing the linear cost function (columns 3 and 4). For each of the six attributes, this holds true (note that for property damage and salinity, lower numbers represent higher levels of the attribute). The average respondent chose attribute levels that were 7–24% farther up the slider bar, and 7–21% cheaper when faced with a quadratic cost function than when faced with a linear cost function, consistent with

Table 2 – CABSCM responses from museum visitors

	Quadratic Costs—$60 Cost Range		Linear Costs—$60 Cost Range		Quadratic Costs—$30 Cost Range	
	Attribute	Cost	Attribute	Cost	Attribute	Cost
Trees planted						
Mean	7.87	30.07	7.04	32.49	5.45	17.10
	(0.16)	(0.75)	(0.16)	(0.72)	(0.12)	(0.42)
Sigma	5.96	27.93	6.06	27.95	4.57	15.47
	(0.14)	(0.66)	(0.14)	(0.64)	(0.11)	(0.38)
N	1468		1577		1492	
Building efficiency						
Mean	39.96	36.19	35. 90	39.16	29.23	20.78
	(0.65)	(0.77)	(0.65)	(0.71)	(0.51)	(0.40)
Sigma	23.92	28.44	24.92	27.19	18.72	14.60
	(0.58)	(0.70)	(0.58)	(0.64)	(0.47)	(0.37)
N	1468		1577		1492	
Wetlands lost						
Mean	−1.470	30.81	−1.766	36.67	−2.532	19.25
	(0.065)	(0.61)	(0.0705)	(0.64)	(0.0591)	(0.38)
Sigma	2.564	24.21	2.865	26.05	2.325	14.93
	(0.0553)	(0.52)	(0.0620)	(0.56)	(0.0546)	(0.35)
N	1644		1757		1689	
Property damage from storms						
Mean	16.55	13.29	18.76	16.86	20.40	8.55
	(0.49)	(0.63)	(0.42)	(0.63)	(0.35)	(0.35)
Sigma	19.05	24.32	16.86	25.29	13.40	13.68
	(0.45)	(0.57)	(0.37)	(0.56)	(0.36)	(0.32)
N	1644		1757		1689	
Property damage from sea level rise						
Mean	12.49	17.94	15.90	21.16	17.74	11.16
	(0.48)	(0.63)	(0.49)	(0.73)	(0.34)	(0.34)
Sigma	17.74	23.22	18.20	27.31	12.46	13.02
	(0.42)	(0.55)	(0.43)	(0.65)	(0.34)	(0.30)
N	1434		1506		1533	
Drinking water salinity						
Mean	0.719	34.37	0.8358	39.94	1.274	19.15
	(0.0274)	(0.68)	(0.0267)	(0.64)	(0.0206)	(0.35)
Sigma	1.004	24.96	1.010	24.25	0.776	13.30
	(0.0240)	(0.60)	(0.0229)	(0.55)	(0.0184)	(0.32)
N	1434		1506		1533	

theoretical expectations. Log-likelihood tests confirm that the differences in each case are significant at least in the 5% level.

The second validity test was for range effects. Comparison between the results for the quadratic cost function with the full range available (columns 1 and 2) to the quadratic cost function with a narrower range of opportunities (columns 5 and 6) show that range effects do occur. Theory predicts that respondents facing the narrow cost range who would have chosen an attribute level with a cost greater than $30 will set the slider for that attribute at the upper limit of its range, and that the population average estimated by the tobit regression should not be affected. However, the average respondent chose attribute levels that were 25–31% lower and “spent” 36–44% less money when faced with the narrower range than when faced with the full range of opportunities. Log-likelihood tests confirmed that these differences are all statistically significant at the 1% level. Respondents who faced the narrow range of opportunities did choose the upper limit more frequently than those who faced the full range (18.6% vs. 15.2%), but not as frequently as would be required for the distributions to be similar.

4.2. *Comparison between DABSCM and CABSCM*

A total of 109 in-person interviews were conducted where respondents completed both the DABSCM survey and the CABSCM survey. Of these, 48 were conducted at Pennsylvania State University (PSU), and 61 were conducted at the Koshland Museum of Science. In 47 interviews, the DABSCM survey was completed first. In the other 62 interviews, the CABSCM survey was completed first. In one of the interviews, the DABSCM was only partially completed. The partial results are included in this analysis.

Respondents participating in this part of the study experienced exactly the same CABSCM survey as general visitors to the Koshland museum. The CABSCM responses for this sample can therefore be directly compared to the responses for the general population of museum visitors facing the same cost function (the quadratic cost function with the full set of opportunities). We have no expectations about whether the study participants will exhibit different preferences than general museum visitors, but it is possible that participation in an in-person interview might focus the respondents more, and motivate them to consider their choices more carefully. Tobit estimates of the average attribute levels and cost commitments from the CABSCM responses for this sample were nearly identical to the average for all museum visitors for the same cost function (the differences were less than 4%, and not statistically significant). However, the museum visitor responses exhibited more variability than the responses of participants in the in-person interviews. The Tobit estimate of the standard deviation in attribute levels for museum visitors was 29% higher for both trees and building efficiency than for participants in the in-person interviews, and this difference was statistically significant at the 1% level in both cases. Thus, it does appear that participating in the in-person interview decreased variability in responses, probably due to greater respondent care and effort.

The most striking result from the sample of in-person interviews came from comments made by participants who experienced both survey formats. Of those who volunteered an opinion on the two methods, there was a clear preference for the CABSCM format. Respondents found it easier to understand and found the task easier to complete.

The 109 respondents made a total of 516 DABSCM choices. Of these, 27.6% showed a preference reversal between the two methods. Of these reversals, the CABSCM response implied lower marginal WTP values for the attributes than did the DABSCM response in 57% of cases. The rate of preference reversals was higher among respondents who answered the CABSCM survey first (32.8%) than those who answered the DABSCM survey first (21.8%). This difference in proportions was statistically significant at the 1% level.

From each CABSCM response, marginal WTP for trees and for building efficiency was calculated. Corner solutions provided truncated estimates of marginal WTP. To estimate the population average marginal WTP, Tobit regression is used with an intercept only. Inspection of the marginal WTP values showed that the distribution for each attribute is clearly skewed, with a longer tail to the right. The marginal WTP values were therefore log-transformed before the Tobit regression was estimated, and the upper and lower bounds were adjusted accordingly. First, tests were conducted to determine whether survey order affected the CABSCM responses. Neither a dummy variable for survey order nor a log-likelihood test showed any evidence of a difference in responses due to survey order. Still, to avoid any potential for sequencing effects, only CABSCM responses where that method was administered first are considered in subsequent analyses.

The Tobit regression parameters (the average and standard deviation of the log-transformed marginal WTP values) are presented for trees and building efficiency in Table 3. The estimated median marginal WTP for trees was $5.73 per million acres. The estimated median marginal WTP for building efficiency was $1.71 per 1% improvement. Table 3 also gives 95% confidence intervals for these estimates. Because the distribution of marginal WTP is skewed, population means are larger than the medians ($7.68 for trees and $3.12 for building efficiency).

The DABSCM responses also showed no statistical evidence of a survey order effect. Still, to assure a clean comparison, only DABSCM responses where that method was implemented first are considered here. A multinomial logit regression for these responses is shown in Table 4. All three attributes in the choice (trees, building efficiency and money)

Table 3 – Tobit regression results for CABSCM surveys

	Trees	Building Efficiency
Tobit results for ln(MWTP)		
Mean	1.7458	0.5390
(s.e.)	(0.0993)	(0.1405)
Sigma	0.765	1.0934
(s.e.)	(0.0789)	(0.1013)
N	61	61
Marginal WTP		
Median	5.73	1.71
(95% CI)	(4.72, 6.96)	(1.30, 2.26)

Table 4 – Multinomial logit regression results for DABSCM

Multinomial logit regression		
		Parameter estimate
MU Trees		0.5371
(s.e.)		(0.1383)
U Efficiency		0.2152
(s.e.)		(0.0424)
MU Money		−0.1312
(s.e.)		(0.0291)
N		188
Marginal WTP		
	Trees	Building efficiency
Median	4.09	1.64
(95% CI)	(2.62, 5.90)	(1.3, 2.14)

are statistically significant at the 1% level and of the correct sign. Marginal WTP for trees and building efficiency are calculated and presented in Table 4. Krinsky and Robb's (1986) Monte Carlo technique was used to simulate 95% confidence intervals for these. Here, the multinomial logit regression provides no estimation of the variability in preferences among respondents, and the estimated marginal WTP should be viewed as both the population median and the population mean.

Comparing the median MWTP from the CABSCM to that from the DABSCM, the CABSCM estimate is 40% higher for trees and 5% higher for building efficiency, but the differences are not statistically significant at even the 10% level.

5. Discussion

The internal validity tests of the CABSCM showed mixed results. Responses were sensitive to curvature in the cost function in exactly the ways predicted by economic theory. This is reassuring, since it shows that respondents are reacting to the cost function, which is a key assumption behind the method. However, responses were sensitive to the range of opportunities presented. This is not particularly surprising. Similar range or anchoring effects have been shown for other value elicitation methods, and even in contexts other than valuation (O'Conor et al., 1999). However, this raises questions about whether respondents are reacting to absolute levels of attributes and cost or to relative levels.

The comparison between the CABSCM survey method and the more traditional DABSCM method showed that the two methods generate similar estimates of marginal WTP. This result is somewhat surprising. Several studies have shown that discrete contingent valuation methods consistently generate higher estimates of WTP than continuous methods (Ready et al., 2001). There does not appear to be a similar systematic difference between how respondents answer CABSCM questions and how they answer DABSCM questions.

Comparing each respondent's choices in the two methods, we saw that respondents exhibited contradictions in their responses 27.6% of the time. These reversals appear to be the result of random effects; there was not a systematic direction to the reversals. The rate of reversals seen between CABSCM and DABSCM does not imply that CABSCM responses are invalid. Both the CABSCM and DABSCM generate data that have random components, even within a single respondent. Some reversals would be expected due to these random components. The reversal rate between CABSCM and DABSCM is similar to that seen across questions within methods. Johnson et al. (2000), in a DABSCM survey valuing life expectancy, found that 39% of respondents showed at least one reversal in preferences between questions. Foster and Mourato (2002), in a contingent ranking experiment, found that nearly half of respondents failed at least one internal consistency test.

This study benefited greatly from the collaboration with the Koshland Science Museum. First, the museum had access to professional designers and videographers that resulted in a product that was more polished, professional and user friendly than the typical survey designed by academic researchers. Second, location of the installation in the museum allowed for a very large sample size of respondents at low cost. Third, the installation was located within a larger exhibit where visitors were exposed to high quality, interactive educational material on global climate change. This reduced the need for lengthy explanations in the survey itself.

The museum context imposed some constraints on the research, however. Most importantly, the survey experience had to be short. This limited the amount of information that could be presented about the attributes being valued. Second, for the general population of museum visitors, the social contract between the researcher and the respondent was less salient. The survey did mention that responses would be used by researchers, but respondents likely saw their responses as less consequential than they would have if they had been approached individually by the researcher, either in person or through a mail or telephone contact. The museum context might also include more distractions (friends, other exhibits) than the in-person interviews.

A comparison between the responses of museum visitors and responses of participants in in-person interviews showed that the average responses were quite similar, but that the variability was lower among respondents to the in-person interviews. We conclude that while the museum context is not ideal for data collection, it did provide responses that, on average, were similar to those that we would have obtained had we used a more traditional survey mode. The increased variability is a tradeoff against the ease of data collection and the potential for very large sample sizes.

The goods valued in this study are likely to be particularly susceptible to framing effects. Most respondents have little prior personal experience with global climate change, tree plantings, building energy efficiency or carbon dioxide reductions. Further, of necessity, the descriptions of these goods were brief. A more familiar good might be less susceptible to range effects and reversals than the goods valued in this study.

We conclude that the CABSCM is a promising method for preference elicitation. Respondents prefer it to the DABSCM, finding the task easier to complete. Each CABSCM response generates more information about preferences than a single

DABSCM response. The method does need to be administered on a computer, though it can be done through a web site.

Extensions to the CABSCM are possible. By varying the cost function, it is possible to identify a marginal WTP function that varies with the attribute level (i.e. to measure the curvature in indifference curves). Interactions among attributes in the utility function also can be identified. Future research on this method should investigate the potential for increasing the number of attributes included in the design. While the method is best suited to attributes that are continuously scalable, a combination of scalable and discrete attributes is also possible. The range effects seen in this study require more investigation, but those effects do not necessarily imply that the method is not useful, particularly for goods that are more familiar to respondents.

Acknowledgements

We are grateful to Erika Shugart, Peter Schulz and Patrice Legro of the Marian Koshland Museum of Science for their cooperation, Shayne Gelo and Joan Severson of Digital Artefacts who wrote the interactive software, and Tom Bowman of Bowman Design Group who designed the physical installation. Financial support was provided by the Penn State Institutes of the Environment.

REFERENCES

Arrow, K., Solow, R., Portney, P.R., Leamer, E.E., Radner, R., Schuman, H., January 15 1993. Report of the NOAA Panel on contingent valuation. Federal Register 58 (10), 4601–4614.

Bishop, R.C., Heberlein, T.A., 1979. Measuring values of extramarket goods: are indirect measures biased? American Journal of Agricultural Economics 61 (5), 926–930.

Boyle, K.J., Bishop, R.C., Welsh, M., 1985. Starting point bias in contingent valuation bidding games. Land Economics 61, 188–194.

Boyle, K.J., MacDonald, H.F., Cheng, H.T., McCollum, D.W., 1998. Bid design and yea saying in single-bounded, dichotomous-choice questions. Land Economics 74 (1), 49–64.

Carson, R., Groves, T., Machina, M.J., 1999. Incentive and Informational Properties of Preferences Questions. Plenary Address, European Association of Environmental and Resource Economists, Oslo Norway.

Davis, R.K., 1963. The value of outdoor recreation: An economic study of the Maine woods. PhD dissertation, Harvard University.

Dubourg, W.B., Jones-Lee, M.W., Loomes, G., 1997. Imprecise preferences and survey design in contingent valuation. Economica 64, 681–702.

Foster, V., Mourato, S., 2002. Testing for consistency in contingent ranking experiments. Journal of Environmental Economics and Management 44, 309–328.

Greene, W.H., Hensher, D.A., 2003. A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research. Part B: Methodological 37 (8), 681–698.

Hoehn, J.P., Randall, A., 1987. A satisfactory benefit cost indicator from contingent valuation. Journal of Environmental Economic Management 14, 226–247.

Holmes, T.P., Adamowicz, W.L., 2003. Attribute-based methods. In: Champ, P.A., Boyle, K.J., Brown, T.C. (Eds.), A Primer on Nonmarket Valuation. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Holmes, T.P., Kramer, R.A., 1995. An independent sample test of yea-saying and starting point bias in dichotomous-choice contingent valuation. Journal of Environmental Economic Management 29, 121–132.

Johnson, F.R., Mathews, K.E., Bingham, M.F., 2000. Evaluating welfare-theoretic consistency in multiple-response stated-preference surveys. TER Working Paper T-0003. Triangle Economic Research, Durham, NC.

Krinsky, I., Robb, A., 1986. Approximating the statistical properties of elasticities. Review of Economics and Statistics 68, 715–719.

National Academy of Sciences, 1992. Policy Implications of Greenhouse Warming: Mitigation, Adaptation and the Science Base. National Academy Press, Washington, D.C.

National Assessment Synthesis Team, 2000. Climate Impacts on the United States: The Potential Consequences of Climate Variability and Change. U.S. Global Change Research Program, Washington, D.C.

O'Conor, R.M., Johannesson, M., Johansson, P.O., 1999. Stated preferences, real behavior, and anchoring: some empirical evidence. Environmental and Resource Economics 13, 235–248.

Ready, R.C., Buzby, J.C., Hu, D., 1996. Differences between continuous and discrete contingent value estimates. Land Economics 72, 297–411.

Ready, R.C., Navrud, S., Dubourg, W.R., 2001. How do respondents with uncertain willingness to pay answer contingent valuation questions? Land Economics 77, 315–326.

Rowe, R.D., Schulze, W.D., Breffle, W.S., 1996. A test for payment card biases. Journal of Environmental Economics and Management 31, 178–185.

U.S. Department of State, 2002. U.S. Climate Action Report. U.S. Government Printing Office, Washington, D.C.