The Impacts of Attribute Level Framing and Changing Cost Levels on Choice Experiments Value Estimates

M.E. Kragt[*,a,b], **J.W. Bennett**[a]

[*] *Presenting author. PhD student*

[a] *The Crawford School of Economic and Government, The Australian National University, Canberra*

[b] *Integrated Catchment Assessment and Management Centre, The Australian National University, Canberra*

Paper to be presented at the

54th Annual Conference

Australian Agricultural and Resource Economics Society

Adelaide, 8-12 February 2010

Abstract:

Choice Experiments (CE) are increasingly used to estimate the values of environmental goods and services. CE questionnaires represent the environmental good under valuation by varying levels of non-market attributes. Inclusion of a cost attribute enables the estimation of monetary values for changes in the non-market attributes presented. The ways in which the levels of the attributes are described in the survey - the 'attribute frame' - may affect respondents' choices. Furthermore, varying levels of the cost attribute may impact CE value estimates. The challenge for CE practitioners is to identify the 'appropriate' attribute frames and cost levels.

In this paper, the impacts of changing cost levels and the impacts of describing non-market attributes as absolute levels or in relative terms are assessed. These tests were performed using data from a CE on catchment management in Tasmania, Australia. Contrary to *a priori* expectations, including explicit information cues about relative attribute levels in the choice sets is found not to affect stated preferences. However, comparisons between different split samples provide evidence that respondents' preferences are impacted by changing the range in cost attribute levels, with higher levels leading to significantly higher estimates of WTP for one of the three environmental attributes.

Keywords: Choice Experiments, Environmental Valuation, Bias, Tasmania

1. Introduction

There is an increasing interest in using discrete Choice Experiments (CEs), otherwise known as Choice Modelling (CM), as a stated-preference (SP) technique to estimate values for environmental goods and services. Fundamental to CEs is the use of surveys in which alternative (hypothetical) policy scenarios are described by varying levels of non-market attributes and costs. Respondents are asked to choose their preferred option across the range of alternatives. CE studies have been conducted in fields ranging from health (e.g. Ryan and Wordsworth, 2000) and environmental management (e.g. Hanley et al., 2006) to transportation and infrastructure services (e.g. Hensher and Rose, 2007). The methodology and the survey used to estimate non-market values in a SP study can influence the outcomes and therefore affect both the validity and reliability of value estimates. Validation of methods and results (should) therefore play an important role in SP studies. Many studies have investigated the validity of different SP techniques (see, for example, Bennett et al., 1998; Carlsson and Martinsson, 2001; Grijalva et al., 2002; Johnston, 2006; and Boyle and Özdemir, in press). Some authors have found that CEs can avoid bias from strategic behaviour and reduce embedding effects (Morrison et al., 1996; Hanley et al., 2001) and that CEs are associated with less hypothetical bias than another popular SP technique; the contingent valuation method (CVM) (Murphy et al., 2005). However, recent comparisons between CVM and CE in a health valuation context indicate that the welfare estimates from CE data are significantly higher than estimates from CVM data (Ryan and Watson, In Press; and van der Pol et al., In Press). If CE results are to be used as an input into environmental decision making, research is warranted into what impacts the welfare estimates from CEs and how.

The design of a survey questionnaire forms a vital part of any CE. CE studies are context-specific, that is, the results are specific to the study's circumstances. The context of the survey should match the context of the study setting. Setting the appropriate survey context is critical, in order to estimate the true values respondents hold for the resources under consideration. In this paper, two topics related to survey design context are investigated: the impacts of attribute framing and the impacts of varying the cost vector, that is, varying the range and magnitude of the levels of the monetary attribute. The study uses CE data from a split sample survey of natural resource management changes in the George catchment, Tasmania. The next two sections provide a discussion of attribute level and cost framing issues in previous CE studies. Section two gives an introduction to the modelling framework used to analyse the CE data. This is followed by a description of the case study area and the survey in Section three. In Sections four to six, the results of the data analyses are presented, followed by a discussion of these results in the final Section seven.

Attribute level framing

Framing refers to the context in which choices are made (Rolfe et al., 2002). There is considerable evidence that the framing of questions and the information provided in a survey affects respondents' answers (Ajzen et al., 1996). When using CEs to value non-market goods, it is important to know how respondents' choices are sensitive to the survey context. Not all respondents may have pre-existing preferences for the non-market goods presented in a CE survey. Instead, preferences may be constructed based on the information provided in the survey[1]. In that case, preferences are likely to change with the information provided and with the wording of the questionnaire (i.e. the survey frame), rather than with the nature of the good. It can be argued that framing effects are inherent to SP techniques as these are *contingent* on the information supplied in the survey. Defining the appropriate survey frame forms a vital part of all SP surveys and depends on the purpose of the survey, the context of the issue and the requirements of respondents.

Attribute framing occurs when choices are influenced by the way attributes are described to respondents. The particular focus of the study reported here is the framing of attribute *levels*. Different ways of describing attribute levels may impact on respondents' choices, even when attribute levels are identical. Bateman et al. (2009) stress that respondents should be able to 'evaluate' the information presented in a non-market valuation survey to avoid anomalies in stated preferences. Survey comprehension may be increased when respondents are given information cues to help them to make choices about unfamiliar goods (Schlapfer, 2008). Such cues can be provided by describing attribute levels using information about absolute quantities of the attributes (such as the total number of bird species; Bennett et al., 2008) as well as their relative levels (such as the proportion of floodplain with healthy vegetation; Rolfe and Windle, 2005). It is plausible that respondents will be able to easily evaluate information about one absolute level being higher than another, but that comparisons to relative quantities will allow respondents to more readily assess the relative scarcity of a good. The ways in which attribute quantities are described will vary with the context of each CE study. Decisions about how to define attribute levels are typically made in consultation with scientists, policy stakeholders and focus group discussions but ultimately remain at the discretion of the analyst. In this study, the impacts of defining attribute levels only as absolute values versus including relative quantities are assessed. To the authors' best knowledge, no CE studies have, investigated the impacts on respondents' choices of formulating attribute level in absolute versus relative terms. We hypothesise that absolute attribute levels are more difficult for respondents to interpret. We therefore expect that the variability in responses will be larger in the absence of information about relative quantities, leading to larger variance in value estimates.

Cost framing

Another framing effect occurs when respondents' choices are influenced by varying levels of the cost attribute in a SP survey. In the contingent valuation (CV) literature, this effect is typically observed as a starting point bias. Starting point bias is said to occur when respondents perceive the initial bid levels included in CV questions as a suggestion of 'acceptable' answers and use the proposed bit to develop and/or revise their own 'true' WTP (Mitchell and Carson, 1989). When respondents base their choice on this revised WTP, they are said to anchor their answers on the proposed bid[2]. Ignoring such effects will lead to biased estimation of the mean and the standard deviation of the WTP (see, for

[1] See, for example Braga and Starmer (2005), Bateman et al. (2004) or Tversky and Simonson (1993) on context-dependent preferences.

[2] Specifically, an *anchoring effect* occurs when respondents "fasten upon elements of the scenario that are not intended by the researcher to convey information about the value of the good and use them as cues to the good's approximate 'correct value'". *Starting point bias* is said to occur when "the respondent regards an initial value proposed in the survey as conveying an approximate value of the amenity's true value and anchors his WTP around the proposed amount" (Mitchell & Carsson, 1989, pp 240).

example, Silverman and Klock, 1989; Herriges and Shogren, 1996; Green et al., 1998; Frykblom and Shogren, 2000; and Flachaire and Hollard, 2007).

CEs may also suffer from anchoring effects if different cost-attribute levels, or different ranges in those levels, affect the estimates of implicit prices. Economic theory suggests that models with varying ranges of the cost attribute should produce similar parameter estimates if respondents have stable and well-formed preferences. As long as the cost range used in the survey reflects the distribution of respondents' preferences, a wider or narrower range or a low versus high range in cost levels should not influence the population averaged value estimates if the marginal utility of money is constant (a common assumption in CE) (Stevens et al., 1997). However, given the observed sensitivity to bid levels in CV studies, there is a risk that respondents interpret the proposed levels of the cost attribute in a CE survey as an indication of the "appropriate" value. In such a case, CEs could suffer from similar anchoring effects as CV studies.

Notwithstanding evidence of anchoring bias in the CV literature (Bateman et al., 1999), there are very few studies that have investigated the effects of varying the levels of the monetary attribute in CEs, particularly in an environmental valuation context. In a study of river health improvements, Hanley et al. (2005) investigated whether WTP estimates in a CE are sensitive to the presented levels of the monetary attribute. A split sample survey was used where only the monetary attribute varied between questionnaire designs. [3] In line with *a priori* expectations, the proportion of respondents choosing the status quo option (no payment, no change in environmental attributes) was significantly higher for the questionnaire design with higher costs compared to the lower cost design. Results indicated that the implicit prices estimates in the low-cost split were lower than the WTP estimates in the high-cost split sample, but these differences were not statistically significant because of the high variability of the WTP estimates in the low-cost split sample. Contrary to Hanley et al. (2005), research by Carlsson and Martinsson (2008) showed significantly higher marginal WTP estimates in a CE questionnaire with higher cost levels, compared to a low cost level questionnaire.[4] These results indicate that CE value estimates are impacted by the range of cost levels, but it should be noted that no status quo or 'opt out' alternative was offered to respondents in this study. More in line with the 'traditional' definition of starting point anchoring, Ladenburg and Olsen (2006) tested the impacts of the costs proposed in an "Instruction Choice Set" (ICS) on respondents' answers. The ICS was an example choice set presented to CE survey respondents before the actual choice questions in the survey. To test for starting point bias, the level of the monetary attribute in the ICS was different between two split samples, but the attributes levels in the subsequent choice sets were identical. The authors found that a significantly higher proportion of respondents in the high cost split sample chose the 'more expensive' options in the subsequent choice sets, indicating that respondents may anchor their preferences in the payment levels presented in the ICS. Furthermore, the WTP estimates in the high cost split were significantly higher than in the split sample with the low cost ICS sample.

The available studies provide evidence that varying levels of the monetary attribute can impact WTP estimates but provide no conclusive explanation for these effects. This study contributes to the valuation literature by testing whether respondents' answers are impacted by the proposed levels of

[3] The cost ranges used were £ 2, 5, 11, 15, 24 and £ 0.67, 1.67, 3.67, 5, 8.
[4] The cost ranges used were SEK 125, 200, 225, 275, 375 and SEK 325, 400, 425, 475, 575.

the cost-attribute. Contrary to the work by Ladenburg and Olsen (2006), we use a split sample survey approach in which the range in cost levels differs between all choice questions in the designs. Our work contrasts with Carlsson and Martinsson (2008) in providing respondents with a no-cost alternative. We also use a more pronounced difference in cost range between split samples than Hanley et al. (2005) which is expected to lead to significant differences in value estimates.

2. Modelling framework

Different econometric models can be used to estimate the probability that a particular alternative is chosen from a set of alternatives presented in each choice question (see, for example, Louviere et al., 2000; Alpízar et al., 2001; Bennett and Adamowicz, 2001; and Hensher et al., 2005). In this study, a mixed logit (ML) model specification was used to account for unobserved individual heterogeneity (Hensher and Greene, 2003). In a ML model, the unobserved component of utility U_{ijt} that individual i derives from alternative j in choice situation t is divided into a part that is correlated across individuals and alternatives η_{ij} and a stochastic part that is independently and identically distributed (iid) over alternative and individuals ε_{ijt} : $U_{ijt} = \beta_i \mathbf{X}_{ijt} + [\eta_{ij} + \varepsilon_{ijt}]$ $\quad j=0,1,\ldots,J;\ t=1,2,\ldots,T$

where β_i is a vector of individual specific parameters and X_{ijt} is a vector of observed, explanatory variables; η_{ij} is a random term with zero mean whose distribution varies across individuals and alternatives (Hensher et al., 2005). In a ML model, the analyst needs to define the expected distribution of η_{ij}, such as a normal, lognormal, uniform or triangular distribution (Hensher and Greene, 2003; Hensher et al., 2005). The distributional function of η_{ij} is given by $f(\eta_{ij}|\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is a vector of the unconditional parameters in the distribution. The conditional probability that alternative j will be chosen by individual i in choice situation t is given by:

$$P(j_{it}|\mathbf{X}_{ijt}, \beta_i, \theta) = \frac{\exp(\mu \cdot [\beta_i^{'}\mathbf{X}_{ijt} + \eta_{ij}])}{\sum_{j=1}^{J} \exp(\mu \cdot [\beta_i^{'}\mathbf{X}_{ijt} + \eta_{ij}])}$$

Where μ is a scale parameter that is inversely related to the variance of the error distribution (Swait and Louviere, 1993). Since all parameter estimates within one estimated model have the same scale, μ is typically normalised to one. Note, however, that comparison of estimated coefficients between different experiments is confounded by the different scale parameters in each model.

The estimated model was specified as:

$$U(j = base)_{it} = \beta_i^{'}\mathbf{X}_{ijt} + [\eta_{ij} + \varepsilon_{ijt}]$$
$$U(j = change_1)_{it} = \beta_i^{'}\mathbf{X}_{ijt} + \omega_{change} + [\eta_{ij} + \varepsilon_{ijt}]$$
$$U(j = change_2)_{it} = \beta_i^{'}\mathbf{X}_{ijt} + \omega_{change} + [\eta_{ij} + \varepsilon_{ijt}]$$

Where ω is an error component term, included to allow for unobserved differences in error structures between the different patterns of error correlation between the two “new management” alternatives and the no-cost base alternative (Campbell, 2007). In this study, the ML model was estimated in a panel data format, to control for unobserved heterogeneity across the choices made by the same individual. In a ML-panel model, an individual specific error term is included that is correlated across the sequence of choices made by individual *i*. In this model, the conditional probability of observing a *sequence* of individual choices S_i from the choice sets is the product of the conditional probabilities (Carlsson et al., 2003):

$$S_i(\beta_i) = \prod_t P(j_{it} \left| \mathbf{X}_{ij}, \beta_i, \theta) \cdot t \right.$$

In a typical CE, this sequence of choices is the number of choice questions answered by each respondent. The *unconditional* choice probability now is estimated by the integral over all possible values of β_i, weighed by the density of β_i:

$$P_i(\beta_i \left| \mathbf{X}_{ij}, \theta) = \int S_i(\beta_i) \cdot f(\eta_{ij} \right| \theta) \cdot d\beta_i$$

This model accounts for systematic, but unobserved correlations in an individuals' unobserved utility over repeated choices (Revelt and Train, 1998). An added advantage of using a panel data model is to control for omitted and unobserved variables (Campbell, 2007). Because the ML model does not have a closed form solution, the ML model is estimated using simulated maximum likelihood methods (Train, 2003).

3. The Choice Experiment

The effects of varying attribute level descriptions and cost vectors were tested using data from a CE that was aimed at determining community preferences for alternative catchment management strategies in the George catchment, Tasmania. The George catchment is a coastal catchment in north-east Tasmania, with several small communities, of which St Helens is the largest town (with a population of approximately 2,000; ABS, 2006). Land use in the catchment includes National Parks, agriculture, forestry plantations and State Forests. The rivers in the catchment and the Georges Bay estuary are intensively used for recreational activities. The catchment environment is generally in good condition (Davies et al., 2005; DPIW, 2007) but increased clearing of riparian vegetation, stock access to rivers and streams as well as inputs from forestry operations and other human activities have been identified as threats to catchment water quality and estuary health (DPIWE, 2005; NRM North, 2008). Natural resource management in the George catchment is aimed at preventing water quality decline and maintaining the ecosystem health of the rivers and estuary (Lliff, 2002; BOD, 2007).

Developing the CE survey

The CE survey development involved several rounds of consultations with local decision makers and natural scientists, as well as focus group discussions with community members. Choice attributes and their levels used in the CE survey were identified based on these consultations and results from environmental modelling studies (Kragt and Bennett, 2008). In the George catchment CE survey, three ecosystem attributes were used to describe George catchment environmental conditions: length of native riverside vegetation, number of rare native animals and plant species and are of healthy seagrass beds in Georges Bay. A cost attribute was defined as a one-off levy on rates, to be paid by all Tasmanian households during the year 2009.

The final survey material consisted of an introduction letter, a questionnaire booklet and an information poster. The information poster provided information about the George catchment using maps, photos and charts (see Appendix). Natural resource management in the George catchment, environmental attributes and attribute levels were also described on the poster. The questionnaire comprised four sections. An introductory section contained questions on visitation and activities in the George catchment, plus a question on respondents' perception of current river and estuary quality. The next section explained the choice task at hand, followed by the choice questions. A third section contained questions that aimed to elicit respondents' choice strategies and understanding of the survey. The final section consisted of various socio-economic questions.

The levels of the attributes included in the choice sets reflected the different situations that could occur in the George catchment under different combinations of catchment management actions. Each choice set consisted of a no-cost, no new catchment management base alternative, presented as a likely degradation in catchment conditions in the next twenty years. Two alternative options in each choice set described implementations of new management actions and resulting protection of the environmental attributes (compared to the base alternative). An example choice set is shown in Fig.1.

Split sample versions

To enable testing of attribute level and cost framing effects, three different survey versions were developed. A 'standard' (ST) version provided the base for comparing results between versions. In the ST questionnaire, the levels of native riverside vegetation were measured in km. The choice sets also included 'relative' levels of native riverside vegetation by explicitly stating the proportion of rivers with native vegetation along both sides of the river (Table 1). The area of healthy seagrass beds was measured in hectares, with the 'relative' levels of the proportion of the estuary with healthy seagrass beds (Table 1). The rare species attribute was described as the number of species present in the catchment. The levels of the payment ranged from $0 to $400 (Table 1).

Table 1 Attribute levels used in the standard version of the Geroge catchment CE

Attribute	Description of base level	Alternative levels*
Native riverside vegetation	40km - Healthy native vegetation along 40 km on both sides of the rivers (=35% of total river length)	56, **74**, 81 (km) (50, 65, 70 %)
Seagrass area	420ha – Seagrass growing in 420 ha of Georges Bay (=19% of total bay area)	560, **690**, 815 (ha) (25, 31, 37 %)
Rare native animal and plant species	35 species present – Of the current 80, 35 rare species remain (45 rare species no longer live in the George catchment)	50, 65, 80 (number of species present)
Your one-off payment	0	30, 60, 200, 400 (AU$)

A second version varied from the standard version only in the description of the seagrass and riverside vegetation attribute levels. Although all questionnaire versions described the total river length and total estuary area on the survey poster (see Appendix), the 'absolute levels' survey (AL) version did not include the percentages of river and estuary area explicitly in the attribute description or choice sets (see Figure 2). This sub-sample was used to test whether respondent's choices are impacted by excluding the relative quantities of the attributes.

Figure 1 Choice set in the ST version of the George catchment CE

Question 6

Consider each of the following three options for managing the George catchment.
Suppose options A, F and G are the **only ones** available. Which of these options would you choose?

Features	Your one-off payment	Seagrass area	Native riverside vegetation	Rare native animal and plant species	YOUR CHOICE
Condition now		690 ha (31% of total bay area)	74 km (65% of total river length)	80 rare species live in the George catchment	
Condition in 20 years					Please tick **one** box
OPTION A	$0	420 ha (19%)	40 km (35%)	35 rare species present (45 no longer live in the catchment)	☐
OPTION F	$400	560 ha (25%)	81 km (70%)	65 rare species present (15 no longer live in the catchment)	☐
OPTION G	$400	690 ha (31%)	56 km (50%)	65 rare species present (15 no longer live in the catchment)	☐

Figure 2 Choice set in the AL questionnaire design of the George catchment CE

Question 6

Consider each of the following three options for managing the George catchment.
Suppose options A, F and G are the only ones available. Which of these options would you choose?

Features	Your one-off payment	Seagrass area	Native riverside vegetation	Rare native animal and plant species	YOUR CHOICE
Condition now		690 ha	74 km	80 rare species live in the George catchment	
Condition in 20 years					Please tick one box
OPTION A	$0	420 ha	40 km	35 rare species present (45 no longer live in the catchment)	☐
OPTION F	$400	560 ha	81 km	65 rare species present (15 no longer live in the catchment)	☐
OPTION G	$400	690 ha	56 km	65 rare species present (15 no longer live in the catchment)	☐

A third 'cost range' (CR) version was developed to test whether respondents' answers are affected by the cost levels proposed in the survey. This version varied from the standard version only in the levels of the monetary attribute presented. The cost levels were based on cost used in previous CE studies in Australia and on feedback from the focus groups. During the focus group discussions, $600 had been identified as the "absolute maximum" WTP for natural resource management in the George catchment. This was used as the maximum level of the cost attribute in the CR survey version (Table 2). To avoid a high rate of protest responses from payment levels that would push respondents beyond their maximum cost, the levels in the ST and AL survey versions were scaled by a factor of about 2/3[5] (Table 2). If respondents indeed use the presented cost levels as a 'suggestion' of appropriate costs, rather than basing their choices on their own 'true' WTP (Mitchell and Carson, 1989), the implicit price estimates from the CR survey version will be higher than the estimates from the ST version.

Table 2 Cost levels used in the ST and CR versions of the questionnaire

Split sample version	Levels of the monetary attribute
Standard survey	0, 30, 60, 200, 400 (AU$)
Cost range	0, 50, 100, 300, 600 (AU$)

[5] Using rounded number in the cost levels was considered appropriate to reduce survey complexity and negative reactions from respondents.

Survey experimental design and administration

A total of 24 choice sets were created using a Bayesian *D*-efficient design (Scarpa and Rose, 2008). Prior information on the expected values of the coefficients was elicited from the results of the questionnaire pretested during focus groups in August 2008. Some combinations in the design were not feasible, for example because one alternative completely dominated the others in the levels of the environmental attributes but not in costs. These combinations were removed from the choice design, leaving a total of 20 choice sets to be included in the questionnaire. The total number of choice sets was divided into four blocks, so that each respondent was presented with five choice questions.

In order to achieve a representative sample of Tasmanian households, but within the practical limits of this study, the survey sample was restricted to the two largest population centres in Tasmania (Hobart and Launceston) and the local community around the town of St Helens. Each location was divided into multiple smaller local sampling units, stratified to cover the complete sample location and a range of community types. A random sample was taken from these areas, using a 'drop off/pick up' method[6] with the assistance of local service clubs. Surveyors received a training session and detailed instructions on the sampling locations and procedures. The questionnaires were collected between November 2008 and March 2009.

4. Descriptive statistics

A total of 1,117 surveys was distributed, of which a total of 722 (64.6%) were returned. A series of χ^2-test were conducted to compare the sample characteristics across locations and questionnaire versions. These indicated significant differences in the population characteristics between the urban respondents in Hobart and Launceston and the local population in St Helens. Because of low response rates and to avoid confounding the results from different underlying population characteristics, only the urban samples are included in the analysis reported here. The interested reader is referred to Kragt and Bennett (2009) for more information about the local sample characteristics.

Respondents who consistently chose the no-cost base alternative because they protested against paying a government levy were not included in the analysis. This resulted in a total of 674 useable surveys (Table 3). Because not all respondents answered all the choice questions, the total number of choice observations available for analysis was 2,880.

Table 3 Number of respondents and available choice observations by survey design

Design	Respondents (#)	Choice observations (#)
Standard version	321	1,344
Absolute levels version	151	693
Cost range version	202	843
Total	**674**	**2,880**

[6] This method involved surveyors to visit randomly selected households within each stratified sampling unit with the request for survey participation. When the householder agreed to participate, a copy of the questionnaire was left behind and arrangements were made to pick up the completed survey booklet at a convenient time

Testing the equivalence between the sample and the Tasmanian population statistics (ABS, 2007) revealed no significant differences in average age or education, but showed that the *distribution* of income, education, gender and age in the sample was significantly different from the State average. The main difference with the average Tasmanian population is the larger proportion of respondents with high incomes and/or a university education and the over-representation of women in the sample. The sample is therefore not representative of Tasmanian households and care should be taken when interpreting the results in light of the wider population. The mean descriptive statistics of the sample are presented in Table 4. The number of visits to the George catchment was included in the analysis of the CR data. Respondents had, on average, visited the region 2.6 times in the 5 years before filling out the survey. An attitudinal variable that captures the level of agreement with the survey information was also included. This variable was measured as respondent's agreement with the information presented on the poster on a 5-point Likert scale where 1=strongly disagree and 5=strongly agree.

Table 4 Descriptive statistics of George catchment CE survey sample

Variable	Unit	Mean	Std.	Min	Max
Income	Annual household income ('000 $, before taxes)	76.78	44.52	7.5	210
Education	Respondent education (yrs)	13.50	2.20	8	18
Gender	=1 if respondent is male	0.38	0.49	0	1
Age	Respondent age (yrs)	45.93	14.59	18	91
Visit	Visits to the George catchment (# in past 5 yrs)	2.59	3.53	0	25
Agree*	Agreement with survey information	3.63	0.70	1	5

* Measured on a 5-point Likert scale where 1 = strongly disagree and 5 = strongly agree.

5. Attribute level framing results

ML models were estimated in LIMDEP 9.0 (Econometric Software, 2007) using Halton draws with 500 replications (Train, 2000). In this section, the model results of the ST and AL survey split samples are reported. The results of the CR sample analysis are reported in Section 6.

Model specifications investigated several distributional assumptions for the choice attributes (for example, fixed or log-normally distributed coefficients), the inclusion of a range of socio-demographic variables, various specifications of heteroskedastic or correlated random parameters as well as heteroskedastic latent error components. Since the coefficients of interest in this analysis are the population averaged parameter estimates on the choice attributes, a parsimonious model was specified. The final model (Table 5) includes university degree as a dummy variable to correct for possible bias originating from the relatively highly educated sample. Other socio-economic or behavioural variables were not significant in the split sample models and were not included in the models reported.[7] The four choice attributes were included as random parameters to account for

[7] All models are available upon request from the authors.

variation in respondents' preferences towards the attributes. Following Greene et al. (2006), a constrained triangular distribution was used for the random cost parameter, to ensure a negative sign on each individual's cost parameter. It was not desirable to so constrain the distributions on the environmental attributes, as respondents may have positive or negative preferences towards the attributes. Normal distributions were therefore defined for the environmental attributes. Other distributional forms, or specifying one or more of the environmental attributes as fixed attributes, did not lead to significantly better models.

Table 5 Mixed logit panel model results for the 'standard' (ST) and 'absolute levels' (AL) split samples in the George catchment CE survey

	ST questionnaire		AL questionnaire	
Variable	Parameter	t-value	Parameter	t-value
Random parameter means				
Costs ($)	-0.011***	-13.47	-0.017***	-9.30
Seagrass (ha)	0.001	1.60	0.004***	3.88
Vegetation (km)	0.044***	5.16	0.063***	4.23
Rare species (#)	0.070***	7.31	0.095***	6.58
Random parameter standard deviations				
Cost	0.012***	13.73	0.017***	9.30
Seagrass	0.004***	4.15	0.001	0.23
Vegetation	0.052***	4.76	0.078***	4.85
Rare species	0.100***	8.39	0.076***	4.55
Non-random parameters				
ASC (=1 for change alternatives)	1.395***	2.95	3.388***	2.79
University educ (0/1)	1.393***	2.74	-0.729	-0.57
Stdev of latent error component	2.225***	4.73	4.279***	4.07
Choice observations n	1419		693	
Log-likelihood	-1069.15		-474.11	
Adjusted - $\rho^{2\,(a)}$	0.252		0.278	
AIC	2158.3		968.2	
BIC	2210.9		1013.6	

Note: ***, **, * = significance at 1%, 5% and 10% level. (a) Against a constant-only model; *AIC = -2·(LL-#par) ; BIC = -2·LL + #par·ln(N)*

As shown in Table 5, all attribute parameters have the expected signs. The cost-coefficient is negative and significant in both models, indicating a disutility from higher levels in the cost attribute, *ceteris paribus*. The parameters of the environmental attributes are positive, indicating that respondents derive positive utility from higher levels in vegetation, rare species and seagrass. The parameter estimate on seagrass is only significant in the AL questionnaire version, in which the percentage of the estuary covered by seagrass beds was not explicitly described in the choice sets. Note, however, that the insignificance of the seagrass estimate in the ST sub-sample is irrelevant given the significant heterogeneity towards seagrass in the ST sample - as indicated by the standard deviation on the seagrass random parameter (Hynes, 2008). The positive and significant standard deviations for the random parameters cost, vegetation and species indicate individual heterogeneity in preferences for these attributes.

An alternative specific constant (ASC) for the 'new-management' alternatives was positive and significant in both models, indicating a preference of respondents towards protecting the George catchment that is not captured by the covariates in the models. The coefficient on education was positive in the ST sample, indicating that respondents with some university education were more likely to choose new environmental management actions, *ceteris paribus*. The latent error component is positive and significant, revealing significant unobserved error correlation between the two new-management alternatives. The significant sign on the error component also means that there are significant differences in respondents' perception of, and substitutability between, the new-management alternatives, compared to the base option (Scarpa, 2007).

Attribute framing effects

The first set of hypotheses to be tested are the null hypotheses of equal parameters estimates between the ST and AL versions of the survey: H_0: $\beta_{ST} = \beta_{AL.}$ Because of the confounding effect of the scale parameter μ, the estimated parameters from Table 6 cannot be compared directly. In order to enable a comparison of the parameters, a grid search was conducted to estimate the ratio of the scale parameter (Swait and Louviere, 1993) where the scale parameter for the ST version was constrained to one. The null hypothesis of equal parameter estimates can then be tested using regular likelihood ratio tests: $LR = -2\ [LL_{pooled} - (LL_{ST} + LL_{AL})]$, where LL_{pooled} is the log-likelihood of the pooled model in which one sample has been rescaled by the estimated ratio of scale parameter. LL_{ST} and LL_{AL} are the log-likelihoods of the separately estimated models. The *LR*-statistic is χ^2-distributed with $(k+1)$ degrees of freedom, with k the number of restrictions in the models. The relative scale parameter that maximised the log-likelihood in the pooled ST-AL model was 0.88. This supports our hypothesis that the error variance in the AL version of the questionnaire is larger than the error variance in the ST version.

The χ^2-test value for the AL model against the ST model is 13.97. As this is lower than the χ^2-critical value, we cannot reject the null hypothesis of equal parameter estimates between the standard and 'absolute levels' versions of the questionnaire. To ensure that this conclusion was not a result of differences in scale, a second likelihood ratio test was conducted to test the pooled model with scaling, against the pooled model without rescaling the AL data. The χ^2 value of this test is -0.91, not providing evidence to reject the null hypothesis of parameter equivalence. Hence, it cannot be concluded that the estimated coefficients are significantly different between the ST and AL data.

Implicit price estimates

An alternative way to test whether respondents' answers are influenced by the frame of the attribute levels is to compare the implicit price estimates across models. The marginal willingness to pay (WTP) for each environmental attribute was calculated using parametric bootstrapping from the unconditional parameter estimates with 1,000 replications (Krinsky and Robb, 1986). The marginal WTP estimates are positive and significant for all attributes in both split samples (

Table **6**).

Table 6 Mean marginal willingness to pay (WTP) estimates for the 'standard' (ST) and 'absolute levels' (AL) split samples (95% confidence interval in parentheses) and results of Poe et al. (1994) test for WTP equivalence

Attributes	ST version			AL version			p-value for equivalence†
Seagrass (ha)	0.104*	(-0.02	0.23)	0.239***	(0.12	0.36)	0.070*
Riverside vegetation (km)	3.969***	(2.45	5.48)	3.708***	(1.98	5.51)	0.396
Rare species (#)	6.310***	(4.61	8.08)	5.591***	(3.88	7.34)	0.269

Note: ***, **, * = significance at 1%, 5% and 10% level. 95% confidence intervals based on the 5th and 95th percentile of the simulated WTP distribution. † p-values for a one-sided t-test of statistical insignificant differences between the WTP estimates from the AL sub-sample and the base ST sub-sample

As shown in

Table **6**, the confidence intervals between the implicit price estimates overlap for all attributes. A formal test[8] for statistical differences in WTP estimates was conducted, based on the convolution approach proposed by Poe et al. (2005, 1994). As indicated by the p-values reported in

Table **6**, there are no significant differences in WTP estimates for the riverside vegetation and rare species attributes between the ST and AL samples. The difference in seagrass WTP estimates is significant at the 10% level. There is therefore not enough evidence to conclude that excluding explicit information about changes in relative quantities of the attributes impacts welfare estimates.

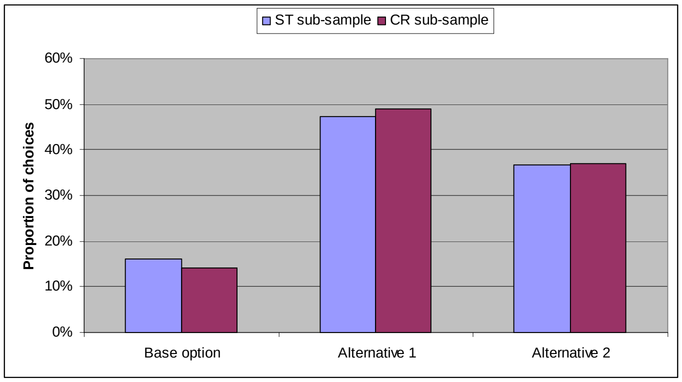
6. Cost framing results

Various tests were used to assess the impacts of different ranges in cost levels on respondents' choices. A first test of differences between the ST and CR questionnaire versions is an analysis of protest responses. It was expected that the higher cost range in the CR questionnaire would lead to a higher rate of protests. The proportion of respondents protesting against the payment was 10.6 percent in the ST survey sample and 12.9 percent in the CR sample. This is not a significant difference across the split samples ($p = 0.512$).

In each choice set, a no-cost base option and two 'new management' alternatives were included. It was expected that a higher proportion of respondents would choose the base-option in the higher cost range version as an opt-out to avoid paying the higher levy. However, the choice data revealed no significant differences in the proportion of choices for the no-cost base option between the ST and CR questionnaire versions (Figure 4; $p = 0.18$).

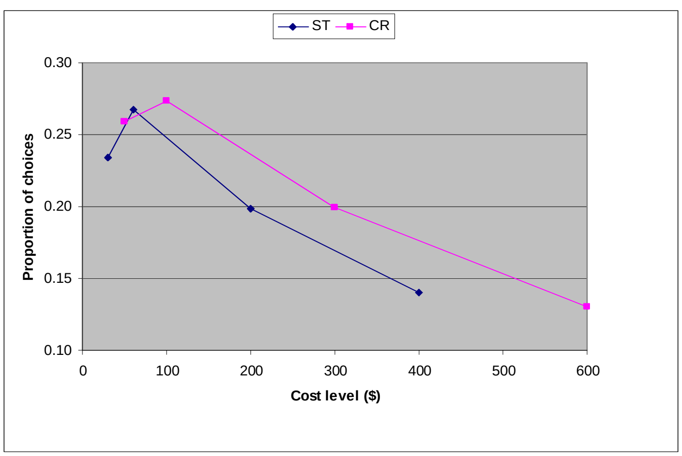
[8] As shown in Poe et al. (1994), comparing confidence intervals between groups is not an appropriate test because it relies on distributional assumptions about WTP that may not be satisfied.

Figure 3 Proportion of choices for the no-cost base option and the 'new management' alternatives



The choice data were further inspected based on the choices by the levels of the cost attribute. Bid-acceptance curves for both survey versions are shown in Figure 4. The figure shows choice sensitivity to the relative cost levels within each sub-sample, with acceptance rates declining with increasing cost levels. However, no statistical significant difference is present between the proportions of respondents who chose the $600 option in the CR sub-sample compared to the proportion of respondents choosing the $400 option in the ST sub-sample. This indicates some insensitivity to the absolute cost levels.

Figure 4 Bid-acceptance for ST and CR questionnaire versions at different levels of the cost attribute



Model results

Mixed logit (ML) model specifications were estimated in Limdep 9.0 using Halton draws with 500 replications (Train, 2000). Similar model estimation procedures as described in Section 5 were followed, with the final model specification reported in Table 7. All attribute parameters have the expected signs. The cost-coefficient is negative and significant for both sub-samples, indicating a disutility from higher levels in the cost attribute, *ceteris paribus*. The parameters on the environmental attributes are positive, indicating that respondents derive positive utility from higher levels in riverside vegetation and rare species. Note that the parameter estimate on seagrass is not significantly different from zero in either of the models. As noted in the previous section, this insignificance is irrelevant if the random parameter has an associated standard deviation estimate that is significant (Hynes et al., 2008). The positive and significant standard deviation for all random parameters reveals considerable unobserved heterogeneity in preferences towards the choice attributes. The standard deviation on the seagrass attribute is not significant in the high cost questionnaire, indicating that seagrass may be better specified as a fixed parameter. Additional models were therefore tested where the parameter on seagrass was modelled as a non-random parameter in the utility function. These specifications did not lead to better model fit ($\chi^2_{\text{LR-test}}$= 6.0 for ST and $\chi^2_{\text{LR-test}}$= 8.5 for CR model) therefore the final reported models include seagrass as a random parameter.

Table 7 Mixed logit panel model results for the 'standard' (ST) and 'high cost range' (CR) split samples in the George catchment CE survey

	ST questionnaire		CR questionnaire	
Variable	Parameter	S.E.	Parameter	S.E.
Random parameter means				
Costs ($)	-0.011***	0.001	-0.007***	0.001
Seagrass (ha)	0.001	0.001	0.001	0.001
Vegetation (km)	0.041***	0.009	0.029**	0.011
Rare species (#)	0.072***	0.010	0.084***	0.012
Random parameter standard deviations				
Cost	0.011***	0.001	0.007***	0.001
Seagrass	0.003***	0.001	0.003	0.002
Vegetation	0.051***	0.010	0.044***	0.012
Rare species	0.094***	0.012	0.067***	0.013
Non-random parameters				
ASC (=1 for change alternatives)	-9.781***	2.444	-13.10***	3.629
Education (yr)	0.435***	0.135	0.502***	0.186
Visitation (# visits)	-0.041	0.081	0.276**	0.134
Agree (1-5) $^{(a)}$	1.686***	0.411	2.473***	0.608
Stdev of latent error component	2.034***	0.383	3.186***	0.536
Log-likelihood	-1006.57		--599.66	
n $^{(b)}$	1,344		843	
Adjusted - $\rho^{2\,(c)}$	0.254		0.271	
AIC	2037.14		1223.32	

Note: ***, **, *= significant at 1%, 5% and 10% level. $^{(a)}$ measured on a 5-point Likert scale where 1 = strongly disagree and 5 = strongly agree. $^{(b)}$ Note that the number of ST choice observations is lower in the ST-CR comparative analysis than in the ST-AL comparison because not all respondents answered the visitation and agreement questions. $^{(c)}$ Against a constant only model of LL_{ST}= -1364.8, LL_{CR}= -839.03.

An alternative specific constant (ASC) for the change alternatives was negative and significant, capturing a mean tendency for respondents to select the no-cost base alternative over the new-management alternatives. However, the significance of the latent random error component indicates that there is considerable heterogeneity across the utilities respondents derive from the new-management alternatives in both the ST and CR models. Similar to the models reported in Section 5, education was positive and significant, indicating that respondents with higher education were more likely to choose new management options. The number of visits to the George catchment was also included in the analysis, to allow for differences in preferences between respondents who visit the region and those who do not.[9] The coefficient for visitation was positive and significant in the CR model, indicating that respondents who visit the region more often are more likely to choose for environmental protection measures. Agreement with the poster information is highly significant in explaining choice probabilities in both the ST and CR survey samples. These results show that respondents who agree with the survey information are more likely to support new environmental management in the George catchment.

Cost range effects

One of the hypotheses to test is whether the parameter estimates across the ST and CR models are equal. To enable a comparison of parameters, a grid search was conducted to estimate the ratio of the scale parameter (Swait and Louviere, 1993) with the scale parameter for the ST version constrained to one. The relative scaling parameter was estimated to be 0.846, which implies that the error variance in the CR version of the questionnaire is larger than the error variance in the ST version (since μ is inversely related to the variance of the error term). The data from both survey versions was pooled and two additional models were estimated: one 'naively' pooled model where all parameters have the same scale, and a 'scaled' model in which potential differences in the variance of responses were controlled for by rescaling the CR data and estimating an additional term on the relative scale parameter. Based on the results of these models and test for equivalence, we cannot reject the hypothesis of equal parameter estimates between the two versions (χ^2-test value of 8.07). To ensure that this result is not a consequence of equal scale parameters, a second test was performed for the 'scaled' pooled model against the 'naively' pooled model. The null hypothesis of equal scale parameters is rejected with $\chi^2_{test} = 6.58$. This implies that the error variance in the CR version is significantly larger than the error variance in the standard survey version: $\sigma_{CR}^2 > \sigma_{ST}^2$. Hence, there is more variability in respondents' choices in the CR sub-sample than in the ST sub-sample. These results contrast with findings by Hanley et al. (2005), who concluded that the error variance in respondents' choices is smaller in a split sample with higher cost levels.

Implicit price estimates

The next hypothesis test involves a comparison of the implicit price estimates across the ST and CR models. The marginal willingness to pay for each environmental attribute was estimated from the unconditional parameter estimates using the WALD procedure in Limdep. 95% confidence intervals

[9] Contrary to the ST and AL sub-samples, visitation and agreement were significant in the CR survey version, which is why these variables are included in the comparative analysis here. Note that no statistical differences were found in visitation rates between split-samples.

were calculated using parametric bootstrapping with 1,000 replications (Krinsky and Robb, 1986). The results are shown in Table 8.

Table 8 Mean marginal willingness to pay (WTP) estimates for the 'standard' (ST) and 'high cost range' (CR) split samples (95% confidence interval in parentheses) and results of Poe et al. (1994) test for WTP equivalence

Attributes	ST version			CR version			p-value for equivalence[†]
Seagrass (ha)	0.09[*]	(-0.03	- 0.21)	0.12	(-0.16	- 0.40)	0.39
Riverside vegetation (km)	3.71[***]	(2.19	- 5.21)	4.22[***]	(0.94	- 7.48)	0.39
Rare species (#)	6.48[***]	(4.77	- 8.26)	12.25[***]	(8.59	- 15.8)	0.00

Note: [***], [**], [*] = significance at 1%, 5% and 10% level. 95% confidence intervals based on the 5th and 95th percentile of the simulated WTP distribution. [†] p-values for a one-sided t-test of statistical insignificant differences between the WTP estimates from the AL sub-sample and the base ST sub-sample

The marginal WTP estimates are positive and significant at the 1% level for the riverside vegetation and the rare species attributes in both split samples. Seagrass is significant at the 10% level of significance in the ST sample only. The confidence intervals around the WTP estimates are wider in the CR sample. This shows larger variance in WTP estimates in the CR sample compared to the ST survey sample.

Conform to *a priori* expectations, the implicit prices estimated in the CR version are higher than the ST version for all environmental attributes (Table 8). A test for statistical differences in WTP estimates was conducted, based on the convolution approach proposed by Poe et al. (2005, 1994). Results from this test show no significant differences in marginal WTP estimates for seagrass and riverside vegetation between the two sub-samples (Table 8). Only the estimated WTP for rare species is significantly higher in the CR sub-sample compared to the ST sub-sample. These results provide only partial support that an upward shift in cost attribute levels provides respondents with a value anchor.

7. *Discussion*

The way in which respondents' make their choices in CE surveys will be affected by the context of the survey. Whereas several studies have investigated the impacts of varying the choice set context on respondents' choices (see, for example, Breffle and Rowe, 2002; DeShazo and Fermo, 2002; Caussade et al., 2005; and Hensher, 2006), there are few studies that have explored alternative ways to frame attribute levels in a CE, or that varied the range in cost levels and the possible impacts on value estimates.

Attribute level descriptions

In this study, the effects of including both absolute and relative descriptions of attribute levels were explored, using results from a CE survey developed to assess community preference for natural resource management in the George catchment, Tasmania. A standard (ST) version of the questionnaire included the absolute quantities of the attributes, and compared these relatively to the total estuary area and total length of rivers. Another, 'absolute levels' (AL), questionnaire version described only the absolute quantities of the seagrass and riverside vegetation attributes. Previous studies have found that survey respondents need information cues to help them make choices about unfamiliar goods (Schlapfer, 2008). It was therefore expected that the exclusion of relative attribute levels would make the information less instructive to respondents. However, results from mixed logit models do not provide enough evidence to show conclusively that preferences are significantly affected when information cues in the form of relative quantities are excluded. Although respondent's variation in choices is higher in the sample without relative attribute level descriptions (as indicated by a scale parameter that is less than one), it cannot be concluded that welfare estimates are different between sub-samples.

Note, however, that the absolute and relative levels of the native riverside vegetation attribute were similar (56, 74, 81km and 50, 65, 70%), while the differences in the absolute and relative levels of the seagrass attribute were more pronounced (560, 690, 815ha and 25, 31, 37%). This could be a reason why we find evidence that respondents evaluated the seagrass attribute differently when information about relative attribute levels was excluded. We speculate that including information about the relative scarcity of an attribute in the form of relative attribute levels will typically be useful to help respondents in evaluating the information presented in the survey.

The attribute frame should be absolutely clear to enable a correct interpretation of the units in which marginal implicit prices are estimated. The 'appropriate' way to describe attribute levels to respondents will depend on the policy and scientific context of the study. The description of attributes and attribute levels presented in a CE questionnaire must match the policy and scientific contexts, needs to be unambiguous and need to be meaningful to respondents. CE practitioners need to be aware that particular attribute frames may influence respondents' choices and that alternative descriptions of attribute levels may affect how respondents comprehend the survey information. Focus group discussions and careful pretesting of CE surveys is essential to assess respondents' reactions to different ways of presenting attribute levels.

Cost levels

Of particular importance to environmental valuation studies is the impact of changing the levels of the *cost* attribute on respondents' preferences. Previous work by Ladenburg and Olsen (2006) and Carlsson and Martinsson (2008) found significant differences between subsamples that were presented with different cost-levels. In contrast, Hanley *et al* (2005) concluded that varying the levels of the monetary attribute did *not* impact WTP estimates between subsamples. In this study, a high cost range (CR) split sample survey version was administered, in which the cost attribute levels were higher than the levels used in the standard (ST) version of the survey. It was expected that a higher proportion of respondents would choose the no-cost 'opt-out' alternative in the high cost split sample. Furthermore, we expected that the levels of the cost attribute might serve as an 'anchor' to respondents about the 'correct' payment for management changes, leading to higher implicit price estimates in the high cost split sample.

Contrary to Hanley et al. (2005), no evidence was found of differences in the proportion of respondents who chose the no-cost base option over costly environmental management alternatives between the ST and CR questionnaire versions. Further analysis of the choice data revealed that the probability of choosing a certain option decreases with increasing costs, indicating choice sensitivity to the cost levels in a CE survey. However, there were no significant differences in the proportion of highest bid acceptance between the ST and CR survey. This may indicate that respondents are more sensitive to relative, rather than absolute cost levels.

The main hypothesis tested in this study is that respondents might 'anchor' their choice on the proposed levels of the cost attribute by interpreting the costs as a hint for a "reasonable" payment for management changes (Frykblom and Shogren, 2000). The higher levels of the cost attribute in the CR survey sub-sample would then have indicated a higher value for the George catchment environment. However, results showed that the estimated taste parameters were not significantly different between the ST and CR survey versions. The scale parameters varied significantly between survey versions. Although it was expected that higher cost levels would invoke 'stronger' (more decisive) reactions in respondents, the error variance was in fact larger in the CR sample. The data thus show that a larger variation in respondent's heterogeneity associated with the expected utility of an alternative in the CR version of the questionnaire. The implicit price estimates are higher in the CR sub-sample for one of the attributes. Therefore, only partial support is provided for the hypothesis that respondents anchor their choices on the levels of the cost attribute.

Given the inconclusive results in the CE literature about the impact of cost levels on respondents' choices and subsequent estimates of WTP, it is important to deliberate on why and how cost levels may affect respondents' choices. Anchoring provides a partial explanation for the findings in this study. Other explanations could be choke price bias, yea-saying or because respondents have unstable preference structures.

In the present study, careful pretesting and focus-group discussions were used to determine respondents' maximum WTP for changes in George catchment natural resource management. The maximum price was set at a level that was considered high enough to reach respondents' choke prices for the management changes proposed, but not so high that the cost levels would seem implausible to respondents. To avoid a high rate of protest responses or hypothetical bias in survey responses, the cost levels were chosen to reflect the relevant (policy) context of the study. However, around 14 percent of respondents choose the highest cost option in both the ST and CR survey versions (Figure 4), indicating that the maximum WTP ('choke price') was not reached for these respondents. Cost levels should be high enough to ensure that respondents consider the monetary attribute in making their choices, but an increase of the maximum cost level presented in the survey should be weight against the plausibility of those costs.

Insensitivity to the absolute price levels could also be due to 'yea-saying' effects, in which respondents always agree to support environmental management options, regardless of their true preferences. Yea-saying may be socially motivated, when the respondent aims to please the interviewer by expressing an opinion considered desirable, or internally motivated, when respondents seek to express their held values (a form of strategic behaviour) (Blamey et al., 1999). However, given that respondents filled out the CE survey in confidence, at their leisure and in the comfort of their own home, no incentive to please an interviewer should have been present in this survey's setting. Furthermore, an increase in the cost vector should have no impact on respondents' choices if yea-saying effects are present, meaning that all WTP estimates will increase when higher cost levels are used. Since significant differences were only found for the WTP estimates for one out of three attributes, yea-saying is unlikely to be the main driver of the findings in this study.

Finally, it is possible that respondents have unstable preference structures for unfamiliar products like environmental goods and services. Again, setting the 'right' survey context is crucial, especially if preferences are (partly) formed by the survey frame, or 'discovered' (Braga and Starmer, 2005) during the surveying process. Different descriptions of attribute levels may influence that 'preference discovery' process. When valuing non-market goods, it is particularly difficult to determine what range in costs levels will be wide enough to cover the possible preferences of all respondents. Consideration also needs to be given to setting a maximum cost level that is high enough to reach respondent's choke price for the management changes proposed.

The design and execution of future CE studies should be aimed at minimising the biases discussed above. Further research is required to investigate effects of attribute level framing and varying cost levels on respondents' choices. There is scope for future research that is aimed at analysing the reasons for respondent's choice behaviour and their reactions to different attribute frames in various choice settings. Studies that compare different types of goods and additional model specifications that incorporate respondents' choice behaviour may provide further insights into the impacts of varying cost vectors on value estimates.

Acknowledgements: This research is supported by the Environmental Economics Research Hub and Landscape Logic, both of which are funded through the Commonwealth Environmental Research Facility of the Australian Department of Environment, Water Heritage and the Arts.

8. References

ABS (2006) 2006 Census. Canberra, Australian Bureau of Statistics.

ABS (2007) Statistics - Tasmania, 2007. *Australian Bureau of Statistics.* Canberra, Australian Bureau of Statistics.

AJZEN, I., BROWN, T. C. & ROSENTHAL, L. H. (1996) Information Bias in Contingent Valuation: Effects of Personal Relevance, Quality of Information, and Motivational Orientation. *Journal of Environmental Economics and Management,* 30**,** 43-57.

ALPÍZAR, F., CARLSSON, F. & MARTINSSON, P. (2001) Using Choice Experiments for Non-Market Valuation. *Economic Issues,* 8**,** 83-110.

BATEMAN, I., LANGFORD, I. & RABASH, J. (1999) Willingness to pay question format effects in contingent valuation. IN BATEMAN, I. J. & WILLIS, K. G. (Eds.) *Valuing environmental preferences: theory and practice of the contingent valuation method in the US, EU and Developing countries.* Oxford, UK, Oxford University Press.

BATEMAN, I. J., DAY, B. H., JONES, A. P. & JUDE, S. (2009) Reducing gain-loss asymmetry: A virtual reality choice experiment valuing land use change. *Journal of Environmental Economics and Management,* 58**,** 106-118.

BENNETT, J. & ADAMOWICZ, W. (2001) Some Fundamentals of Environmental Choice Modelling. IN BENNETT, J. & BLAMEY, R. (Eds.) *The Choice Modelling Approach to Environmental Valuation.* Cheltenham, UK, Edward Elgar.

BENNETT, J., DUMSDAY, R., HOWELL, G., LLOYD, C., STURGESS, N. & VAN RAALTE, L. (2008) The economic value of improved environmental health in Victorian rivers. *Australasian Journal of Environmental Management,* 15**,** 138-148.

BENNETT, J., MORRISON, M. & BLAMEY, R. (1998) Testing the validity of responses to contingent valuation questioning. *The Australian Journal of Agricultural and Resource Economics,* 42**,** 131-148.

BLAMEY, R. K., BENNETT, J. W. & MORRISON, M. D. (1999) Yea-Saying in contingent valuation surveys. *Land Economics,* 75**,** 126-141.

BOD (2007) Georges Catchment and Estuary Project Overview. St Helens, Break O'Day Council.

BOYLE, K. & ÖZDEMIR, S. (in press) Convergent Validity of Attribute-Based, Choice Questions in Stated-Preference Studies. *Environmental and Resource Economics,* In Press, Corrected Proof.

BRAGA, J. & STARMER, C. (2005) Preference Anomalies, Preference Elicitation and the Discovered Preference Hypothesis. *Environmental and Resource Economics,* 32**,** 55-89.

BREFFLE, W., S & ROWE, R., D (2002) Comparing choice question formats for evaluating natural resource tradeoffs. *Land Economics,* 78**,** 298-314.

CAMPBELL, D. (2007) Willingness to Pay for Rural Landscape Improvements: Combining Mixed Logit and Random-Effects Models. *Journal of Agricultural Economics,* 58**,** 467-483.

CARLSSON, F., FRYKBLOM, P. & LILJENSTOLPE, C. (2003) Valuing wetland attributes: an application of choice experiments. *Ecological Economics,* 47**,** 95-103.

CARLSSON, F. & MARTINSSON, P. (2001) Do Hypothetical and Actual Marginal Willingness to Pay Differ in Choice Experiments?: Application to the Valuation of the Environment. *Journal of Environmental Economics and Management,* 41**,** 179-192.

CARLSSON, F. & MARTINSSON, P. (2008) How Much is Too Much? An Investigation of the Effect of the Number of Choice Sets, Context Dependence and the Choice of Bid Vectors in Choice Experiments *Environmental and Resource Economics,* 40**,** 165–176.

CAUSSADE, S., ORTUZAR, J. D. D., RIZZI, L. I. & HENSHER, D. A. (2005) Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B: Methodological,* 39**,** 621-640.

DAVIES, P. E., LONG, J., BROWN, M., DUNN, H., HEFFNER, D. & KNIGHT, R. (2005) The Tasmanian Conservation of Freshwater Ecosystem Values (CFEV) framework: developing a conservation and management system for rivers. *The freshwater protected areas conference 2004.* Sydney, IRN and WWF-Australia.

DESHAZO, J. R. & FERMO, G. (2002) Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency. *Journal of Environmental Economics and Management,* 44**,** 123-143.

DPIW (2007) Annual Waterways Monitoring Reports 2006: George Catchment. Department of Primary Industries and Water.

DPIWE (2005) Environmental Management Goals for Tasmanian Surface Waters. Dorset & Break O'Day municipal areas. Hobart, Department of Primary Industries, Water and Environment.

ECONOMETRIC SOFTWARE (2007) LIMDEP 9.0. Castle Hill, Econometric Software Inc.

FLACHAIRE, E. & HOLLARD, G. (2007) Starting point bias and respondent uncertainty in dichotomous choice contingent valuation surveys. *Resource and Energy Economics,* 29**,** 183-194.

FRYKBLOM, P. & SHOGREN, J. F. (2000) An Experimental Testing of Anchoring Effects in Discrete Choice Questions. *Environmental and Resource Economics,* 16**,** 329-341.

GREEN, D., JACOWITZ, K. E., KAHNEMAN, D. & MCFADDEN, D. (1998) Referendum contingent valuation, anchoring, and willingness to pay for public goods. *Resources and Energy Economics,* 20**,** 85-116.

GREENE, W. H., HENSHER, D. A. & ROSE, J. (2006) Accounting for heterogeneity in the variance of unobserved effects in mixed logit models. *Transportation Research Part B: Methodological,* 40**,** 75-92.

GRIJALVA, T. C., BERRENS, R. P., BOHARA, A. K. & SHAW, W. D. (2002) Testing the Validity of Contingent Behavior Trip Responses. *American Journal of Agricultural Economics,* 84**,** 401-414.

HANLEY, N., ADAMOWICZ, W. & WRIGHT, R. E. (2005) Price vector effects in choice experiments: an empirical test. *Resource and Energy Economics,* 27**,** 227-234.

HANLEY, N., MOURATO, S. & WRIGHT, R. E. (2001) Choice Modelling Approaches: A Superior Alternative for Environmental Valuation? *Journal of Economic Surveys,* 15**,** 435-462.

HANLEY, N., WRIGHT, R. E. & ALVAREZ-FARIZO, B. (2006) Estimating the economic value of improvements in river ecology using choice experiments: an application to the Water Framework Directive. *Journal of Environmental Management,* 78**,** 183-193.

HENSHER, D. A. (2006) How do respondents process stated choice experiments? Attribute consideration under varying information load. *Journal of Applied Econometrics,* 21**,** 861-878.

HENSHER, D. A. & GREENE, W. H. (2003) The Mixed Logit model: The state of practice. *Transportation,* 30**,** 133-176.

HENSHER, D. A. & ROSE, J. M. (2007) Development of commuter and non-commuter mode choice models for the assessment of new public transport infrastructure projects: A case study. *Transportation Research Part A: Policy and Practice,* 41**,** 428-443.

HENSHER, D. A., ROSE, J. M. & GREENE, W. H. (2005) *Applied Choice Analysis: A Primer,* Cambridge, Cambridge University Press.

HERRIGES, J. A. & SHOGREN, J. F. (1996) Starting Point Bias in Dichotomous Choice Valuation with Follow-Up Questioning. *Journal of Environmental Economics and Management,* 30**,** 112-131.

HYNES, S., HANLEY, N. & SCARPA, R. (2008) Effects on Welfare Measures of Alternative Means of Accounting for Preference Heterogeneity in Recreational Demand Models. *American Journal of Agricultural Economics,* 90**,** 1011-1027.

JOHNSTON, R. J. (2006) Is hypothetical bias universal? Validating contingent valuation responses using a binding public referendum. *Journal of Environmental Economics and Management,* 52**,** 469-481.

KRAGT, M. E. & BENNETT, J. (2008) Developing a Questionnaire for Valuing Changes in Natural Resource Management in the George Catchment, Tasmania. *Environmental Economics Research Hub Research Report No.8.* Canberra, Crawford School of Economics and Government, Australian National University.

KRAGT, M. E. & BENNETT, J. (2009) Using Choice Experiments to value River and Estuary Health in Tasmania with Individual Preference Heterogeneity. *Environmental Economics Research Hub Research Report No.16.* Canberra, Crawford School of Economics and Government, Australian National University.

KRINSKY, I. & ROBB, A. L. (1986) On Approximating the Statistical Properties of Elasticities. *The Review of Economics and Statistics,* 68**,** 715-719.

LADENBURG, J. & OLSEN, S. B. (2006) Starting Point Anchoring Effects in Choice Experiments. *Working paper 06/06.* Copenhagen, Danmark, Institute of Food and Resource Economics, University of Copenhagen.

LLIFF, G. (2002) George River Catchment: Plan for Rivercare Works for the Upper Catchment, North George and South George Rivers. St Helens, George River Catchment Coordinator, 2002.

LOUVIERE, J. J., HENSHER, D. A. & SWAIT, J. D. (2000) *Stated choice methods : analysis and applications,* Cambridge, UK ; New York, NY, USA, Cambridge University Press.

MITCHELL, R. C. & CARSON, R. T. (1989) *Using surveys to value public goods : the contingent valuation method* Washington, D.C, Resources for the Future.

MORRISON, M. D., BLAMEY, R. K., BENNETT, J. W. & LOUVIERE, J. J. (1996) A comparison of stated preference techniques for estimating environmental values. *Choice Modelling Research Reports, Report No.1.* Canberra, School of Economics and Management, University College, The University of New South Wales.

MURPHY, J. J., ALLEN, P. G., STEVENS, T. H. & WEATHERHEAD, D. (2005) A Meta-analysis of Hypothetical Bias in Stated Preference Valuation. *Environmental and Resource Economics,* 30**,** 313-325.

NRM NORTH (2008) State of the Region: Water Quality and Stream Condition in Northern Tasmania 2006. Launceston, North Water Monitoring Team.

POE, G. L., GIRAUD, K. L. & LOOMIS, J. B. (2005) Computational Methods for Measuring the Difference of Empirical Distributions. *American Journal of Agricultural Economics,* 87**,** 353-365.

REVELT, D. & TRAIN, K. (1998) Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level. *Review of Economics and Statistics,* 80**,** 647-657.

ROLFE, J., BENNETT, J. & LOUVIERE, J. (2002) Stated values and reminders of substitute goods: Testing for framing effects with choice modelling. *The Australian Journal of Agricultural and Resource Economics,* 46**,** 1-20.

ROLFE, J. & WINDLE, J. (2005) Valuing options for reserve water in the Fitzroy Basin. *The Australian Journal of Agricultural and Resource Economics,* 49**,** 91-114.

RYAN, M. & WATSON, V. (In Press) Comparing welfare estimates from payment card contingent valuation and discrete choice experiments. *Health Economics,* In Press.

RYAN, M. & WORDSWORTH, S. (2000) Sensitivity of Willingness to Pay Estimates to the Level of Attributes in Discrete Choice Experiments. *Scottish Journal of Political Economy,* 47**,** 504-524.

SCARPA, R. & ROSE, J. M. (2008) Designs efficiency for nonmarket valuation with choice modelling: how to measure it, what to report and why. *The Australian Journal of Agricultural and Resource Economics,* 52**,** 253-282.

SCARPA, R., WILLIS, K. G. & ACUTT, M. (2007) Valuing Externalities from Water Supply: Status Quo, Choice Complexity and Individual Random Effects in Panel Kernel Logit Analysis of Choice Experiments. *Journal of Environmental Planning and Management,* 50**,** 449-466.

SCHLAPFER, F. (2008) Contingent valuation: A new perspective. *Ecological Economics,* 64**,** 729-740.

SILVERMAN, J. & KLOCK, M. (1989) The behavior of respondents in contingent valuation: Evidence on starting bids. *Journal of Behavioral Economics,* 18**,** 51-60.

STEVENS, T. H., DECOTEAU, N. E. & WILLIS, C. E. (1997) Sensitivity of contingent valuation to alternative payment schedules. *Land Economics,* 73**,** 140.

SWAIT, J. & LOUVIERE, J. (1993) The role of the scale parameter in the estimation and comparison of multinational logit models. *Journal of Marketing Research,* 30**,** 305.

TRAIN, K., E (2000) Halton Sequences for Mixed Logit. *Paper E00-278.* University of California, Berkeley, Institute of Business and Economics.

TRAIN, K. E. (2003) *Discrete Choice Methods with Simulation*, Cambridge University Press.

VAN DER POL, M., SHIELL, A., AU, F., JOHNSTON, D. & TOUGH, S. (In Press) Convergent validity between a discrete choice experiment and a direct, open-ended method: Comparison of preferred attribute levels and willingness to pay estimates. *Social Science & Medicine,* In Press, Corrected Proof.

Appendix Information poster included in the George catchment CE