The Australian Journal of Agricultural and Resource Economics, 46:1, pp. 1–20

Stated values and reminders of substitute goods: Testing for framing effects with choice modelling

John Rolfe, Jeff Bennett and Jordan Louviere*

Choice modelling, a non-market valuation technique, is used to explore framing issues in the context of environmental valuations. Choice modelling appears to have promise in simultaneously valuing a pool of substitute amenities and goods. Describing choices according to component attributes can also help to frame choices according to a number of trade-offs. The statistical information available helps to determine where framing effects have occurred. Three choice modelling experiments were reviewed to show that framing effects may be more widespread in non-market valuation studies than is commonly thought.

1. Introduction

Researchers using stated preference valuation techniques have always been interested in determining the extent to which survey respondents consider substitute and complementary goods when they value the trade-offs presented to them (Boyle 1989; Mitchell and Carson 1989; Carson and Mitchell 1995). There has often been suspicion that under survey conditions, perceptions about budget and substitute constraints vary markedly from what they would in a 'real life' situation. Information about the influence of substitute goods remains largely hidden to the analyst in applications of the contingent valuation method (CVM). In this paper we discuss the benefits of another stated preference technique termed Choice Modelling (CM).

The NOAA panel (1993) recommended that reminders of substitute goods and budget constraints be included within applications of the contingent valuation method (CVM). Since that recommendation there has been increasing interest in the effectiveness and form of information about substitute goods, and the extent to which they help the respondent to frame the trade-off of interest (Loomis *et al.* 1994; Whitehead and Bloomquist 1995; Kotchen and Reiling 1999; Whitehead and Bloomquist 1999). Framing effects occur when the respondent to a survey is unduly sensitive to the

*John Rolfe, Faculty of Business and Law, Central Queensland University, Emerald, Qld; Jeff Bennett, National Centre for Development Studies, Australian National University, Canberra, ACT; Jordan Louviere, Memetrics Pty Ltd, Sydney, NSW, Australia.

context in which a particular trade-off is offered. There are three areas of particular interest to practitioners of non-market valuation.

The first of these areas relate to the issues involving framing lesser known amenities and trade-offs. It is difficult in the performance of a stated preference survey to provide the same amount of information about various substitutes as well as the issue of interest. The fact that an issue has been selected is also likely to create an implicit signal about the importance of the issue relative to substitutes.

The second framing issue relates to the uncertainty surrounding the effectiveness of direct reminders in CVM surveys. There is some evidence that the inclusion of reminders in CVM studies has little effect on the values estimated (Loomis *et al*. 1994; Kotchen and Reiling 1999). It is possible that such reminders do little to increase the awareness of respondents about substitute goods. This may be because the effect of the reminders may be small compared to other sources of variability (Loomis *et al*. 1994), or because respondents need much more detailed knowledge about substitutes than a simple reminder will provide (Whitehead and Bloomquist 1999). These reasons suggest that as the CVM is applied to lesser-known goods, the problems relating to framing choices against substitutes intensify.

It is also possible that simple reminders may not be effective in changing the structure of choice and information transfer. The position of an amenity in a queue of choices will influence values through a type of sequencing effect (Randall and Hoehn 1996), implying that reminders of substitutes that do not change the order in which items are viewed will not have substantial impact on values. Neill (1995) showed that substitution effects vary according to whether respondents are simply reminded of their existence (as according to the NOAA recommendations) or whether they are directly forced to consider them. Hoehn and Randall (1987), Hoehn (1991), Hoehn and Loomis (1993), Cummings *et al*. (1994) and Neil (1995), have suggested that in order to generate unbiased estimates of value for a particular good, respondents must be asked to simultaneously value the good in question together with relevant substitutes and/or complements.

The third broad framing issue relates to the difficulties in distinguishing differences in value estimates between similar trade-offs. Boyle (1989) suggested that small variations in commodity descriptions will not produce valuation effects, whereas large variations will. However, many issues are complex and multifaceted, meaning that there are a variety of ways in which substitutes are perceived. For example, substitutes for rainforests may be perceived to be other vegetation types, other habitats for birds and mammals, other destinations for tourists to visit, or other areas for Indigenous people to live. It is not always easy for researchers to remind potential respondents of the range of different substitutes that they may need to consider, to determine when a variation is likely to cause a framing effect or, indeed, to determine when a failure has occurred.

In this paper, we explore the use of another stated preference technique termed Choice Modelling (CM) to address some of these framing issues. In the next section, a brief overview of the CM technique is presented. In section three, an analysis of framing issues and CM is presented, followed in section four by details of three CM experiments designed specifically to explore framing and substitute amenities. Final conclusions are drawn in section five.

2. The Choice Modelling technique

Choice Modelling is a stated preference technique that has some similarities to the CVM. In general, CM involves the use of a series of choice sets in which respondents indicate their preferred choice from a pool of alternatives. Each offering consists of at least two choice options described by a number of attributes, the levels of which are varied systematically to form the scenarios and provide certain statistical properties necessary for estimating the parameters of families of probabilistic discrete choice models. In comparison, the CVM is a single-shot approach, where respondents are usually only presented with one option and asked if they prefer it to the current situation.

Respondents to a CM survey usually answer a number of designed scenarios, so there are sufficient data to enable the analyst to estimate the contribution of different attributes to the choices made. This then enables predictions to be made about choices and, hence, valuations to be made. The aim of the CM process is to estimate a model to predict choice on the basis of the attributes that describe the amenities of interest.

The Multinomial Logit model (MNL) used to analyse CM data is motivated by the consideration of a conditional indirect utility function of the form:

$$V = \beta + \beta_1 Z_1 + \beta_2 Z_2 + \ldots . \beta_N Z_N + \theta_1 S_1 + \theta_2 S_2 + \ldots . \theta_M S_M \qquad (1)$$

where V is the indirect utility function, β_1 to β_N are elements of the vector of coefficients attached to the vector of attributes (Z) describing the environmental resource, and θ_1 to θ_M are elements of the vector of coefficients attached to the vector of individual characteristics (S), with the latter usually including income. The intercept (β) represents the influence of unobserved attributes. For some models, particularly when the alternatives are labelled, this intercept term can be disaggregated into alternate specific constants (ASC) to generate more accurate models of choice. An important attribute included in the indirect utility function is some willingness to pay (WTP) measure for provision of the described environmental amenity. The coefficient of the monetary attribute used in describing the choice sets is denoted by $\beta_{\$}$.

This vector of utility parameters includes an element of variance, reflecting an unexplainable component of observations (error term). The random utility maximisation (RUM) model allows choice probabilities to be estimated for utilities that contain random elements (McFadden 2001). The RUM model underpins the use of both logit and probit models to analyse choice data. Assuming an extreme value (Gumbel) distribution of the error terms allows the MNL model to be applied to choice data (McFadden 2001). It is possible to express each of the beta coefficients (relating to both alternatives and attributes) in equation 1 as $\lambda\beta$, where λ represents a scale parameter. This scale parameter is inversely proportional to the standard deviation of the random component inherent in a choice experiment (Louviere 2001). The scale parameter cannot be identified in a specific model, but the comparison of different models does allow identification of the differences in error terms.

It is a condition of the MNL model that the error terms associated with each alternative have to be independently and identically distributed (IID), giving rise to an independence from irrelevant alternatives (IIA) condition. This means that the probability of an option being chosen should be unaffected by the inclusion or omission of other alternatives. The condition is normally tested by omitting an alternative in turn from the choice sets and testing to see if there is significant difference in parameter estimates.

Welfare estimates can be obtained by using the following formula to estimate compensating variation (CV) as described by Hanemann (1984):

$$\mathrm{CV} = -1/\alpha\left[\ln \sum \exp V^0 - \ln \sum \exp V^1\right] \tag{2}$$

where α is the marginal utility of income, and V^0 and V^1 represent the utility before and after the change under consideration. Here the welfare estimate is obtained by finding the difference in utility between two options and scaling that utility to a metric measure with the aid of the marginal utility of income. In CM, the monetary coefficient ($\lambda\beta_{\$}$) generated as a model parameter is used as an estimate of the marginal utility of income. Changes in utility can arise from both changes in the attributes of alternatives, or the inclusion or removal of alternatives altogether.

In situations where the choice set includes a single before and after option, the welfare measure described in equation 2 reduces to:

$$\mathrm{CV} = -1/\alpha[\ln(\exp V^0) - \ln(\exp V^1)] = -1/\lambda\beta_{\$}[V^0 - V^1] \tag{3}$$

In some cases the before and after options may differ only because of changes in a single attribute. For attributes representing non-continuous data, the CV will be represented by the difference between the attribute coefficients for the relevant levels, divided by the monetary coefficient, as prescribed by equation 3. For continuous data though, the marginal value of a change within a single attribute can be represented as a ratio of coefficients, where equation 3 reduces further to:

$$\mathrm{W} = \lambda\beta/\lambda\beta_{\$} \tag{4}$$

This formula for the value of a single attribute change (termed a 'part-worth') effectively provides the marginal rate of substitution between WTP and the attribute in question. The formula also demonstrates that the scale parameter is cancelled out in the estimation process. This means that while model coefficients cannot be directly compared between different CM experiments because of differing (but unknown) scale parameters, resulting value estimates are comparable.

3. Choice Modelling and framing issues

Choice Modelling appears to offer several advantages over other non-market valuation techniques for framing purposes. The first, and perhaps most significant advantage, is that it allows the simultaneous presentation of a pool of alternative and substitute goods. This explicitly requires respondents to consider complementary and substitution effects in the choice process. In addition, problems of bias can be minimised because the amenity of interest can be 'hidden' within the pool of available goods used in a CM experiment.

These strengths are demonstrated in relation to the experiments reported in the following section. The issue of interest was the estimation of non-use values held by Australians for rainforest conservation in Vanuatu, one of the Pacific nations. Because Australians are not well informed about Vanuatu or several other countries where rainforest conservation is an important issue, any potential application of the CVM for that purpose would be problematic. The more appropriate way of framing these choices was to present Vanuatu in a pool of other countries (including Australia) where rainforests could be preserved (see figure 1).

A second major advantage of CM is that it provides a more realistic way for respondents to trade-off opportunity costs than CVM allows. This occurs in two important ways: (i) the WTP attribute is only one of several attributes that defines profiles and, hence, is de-emphasized in importance relative to its central role in the CVM; and (ii) CM allows one to introduce a variety of opportunity costs, not just some WTP mechanism.

In the Vanuatu study, choices were framed against the range of attributes that respondents commonly used to make choices about rainforest

Option 1 - Vanuatu
• 1,000 hectares
• Fairly rare
• No visits allowed
• Protection of rainforest means local people will be worse off
• Special landscapes
$5 donation required

Option 2 - Far North Queensland
• 1,000 hectares
• Fairly rare
• Easy to visit with full facilities
• No local people
• Special landscapes
$5 donation required

Option 3 - Papua New Guinea
• 100 hectares
• Fairly rare
• Easy to visit with full facilities
• No local people
• Special landscapes
$50 donation required

Option 4 - South America
• 10,000 hectares
• Extremely rare
• Easy to visit with full facilities
• Protection of rainforest means local people will be worse off
• No special features
$10 donation required

Option 5 - Thailand
• 100 hectares
• Not rare at all
• Visits possible but moderate access and few facilities
• Protection of rainforest means local people will be worse off
• Special landscapes
$5 donation required

Option 6 - Indonesia
• 1,000 hectares
• Fairly rare
• Easy to visit with full facilities
• Protection of rainforest means local people will be better off
• Special landscapes as well as plants and animals
$50 donation required

Please indicate preference: (Tick one)

☐ 1	***Option 1***	☐ 2	***Option 2***
☐ 3	***Option 3***	☐ 4	***Option 4***
☐ 5	***Option 5***	☐ 6	***Option 6***
☐ 7	***I would not support any option***		

Figure 1 A sample choice set from Experiment A

conservation. Focus groups were held in Brisbane and participants were asked to indicate the key issues that they would consider in choices about rainforest conservation (Rolfe and Bennett 1995). From the results of the focus groups, the attributes chosen to describe the rainforest conservation profiles were:

- Location (country)
- Area (of the conservation proposal)
- Rarity
- Potential to visit
- Effect on local populations

- Special features of the area
- Price of the proposal (framed as a donation)

Apart from the location, the attributes describing each profile could be classified by three environmental attributes (area, rarity and special features), and three socioeconomic attributes (visits, locals and price). This approach had a number of advantages. It de-emphasised price as a trade-off, made the scenarios more realistic, and provided some indication about how people viewed trade-offs between social and environmental factors.[1]

A third framing benefit involves the ability to analyse and compare CM experiments. This allows the analyst to test whether differences in framing the choices to respondents cause variations in the parameters of the resulting choice models. For convenience, differences in framing can be categorised into slight variations in the description of essentially the same good, and larger variations that change the structure of the choices involved (Boyle 1989). Both of these possible differences can be tested by examining the internal validity of models and differences in choice model parameters.

The first test that can be performed is to check that violations in model assumptions have not occurred. The internal validity of choice models can be tested by identifying any IIA/IID violations. The presence of these violations would suggest that choices have not been consistent (independent) and, therefore, that respondents have had difficulty in framing choices through the course of the experiments.

The second test that can be performed is to determine whether slight differences in the way that choices are presented to respondents have impacts on model parameters and, hence, on value estimation. If slight differences in framing do not cause value estimates to change, as Cummings, Brookshire and Schulze (1986) and Boyle (1989) hypothesised, then the parameters for the differently framed choice models should be identical.

The third test that can be performed is to determine whether substantial differences in framing between choice experiments cause changes in value only to those attributes that are not shared by the choice frameworks, or whether they also cause changes to the common attributes as well.[2] Substantial differences can be introduced into CM experiments by including different substitute goods. It would be expected that overall values would differ between choice experiments when there are major differences in the underlying

[1]The part-worth formula described in equation 4 can be used to give any marginal rate of substitution between two attributes by using the relevant beta coefficients.

[2]Boyle (1989) concluded that substantial framing differences caused value changes. This is not surprising to economists. What is more difficult to ascertain from CVM experiments is whether substantial framing differences causes systematic value changes, or only changes in the components not common across the different experiments.

components. However, it could be expected that attributes common between choice experiments are valued in the same way. If they are not, it could be because they have been framed by respondents in different ways.

If the influence on choice of particular attributes is independent of other attributes and choice alternatives and, hence, unaffected by the introduction of other substitutes, this implies that no framing effects are present. Significant changes in beta coefficients common between CM experiments that involve different substitutes will therefore indicate that framing effects have occurred, while insignificant changes will indicate the reverse.

These differences in model parameters can be tested in two main ways. First, confidence intervals for part-worths can be compared to isolate any differences that might exist between models. Significant differences suggest framing problems. Second, log-likelihood tests can be used to identify whether model parameters differ by any more than variations in the relevant scale parameters. Here, the relevant tests are described in more detail.

3.1 The part-worth tests

One way to identify framing effects is to compare the part-worths that are available from models estimated from CM surveys. These are directly comparable between models because the scaling (λ) terms are cancelled out of such equations. In order to estimate whether differences between part-worths generated from different experiments are statistically robust, confidence intervals need to be generated.

This can be done using Fieller's Method as proposed by Krinsky and Robb (1986). It involves the simulation of an asymptotic distribution of the coefficients that are generated in a CM experiment, from which confidence intervals can then be computed. The distribution is achieved by taking repeated random draws of 'the coefficient vectors from a multivariate normal distribution with mean and variance equal to the β vector and variance–covariance matrix from the estimated multinomial logit model' (Morrison *et al.* 1998, p. 10). Implicit prices can then be calculated from each of the random draws of coefficients, and confidence intervals estimated by identifying the values at each tail of the distribution of implicit prices.

3.2 The scale parameter (Swait–Louviere) tests

The Swait–Louviere test entails a proportionality restriction on the parameters of one dataset relative to the second, and a test of whether the sum of the log-likelihoods for the two different datasets differs significantly from the log-likelihood for a model estimated from the pooled datasets with the parameter proportion restriction. The pathway for this analysis is through the estimation of the ratio of scale parameters for the different models.

A scale parameter (the constant of proportionality) is inversely proportional to the standard deviation of the error distribution for each dataset. Steps to find the ratio of the scale parameters involve stacking two datasets, X_1 and X_2, and multiplying one of the datasets by a scalar value (i.e. $X_1{}^* = \phi X_1$) (Swait and Louviere 1993). (The alternate specific constants are not rescaled). The purpose is to determine the value of the scalar ϕ that optimizes the log-likelihood of the MNL model when estimated using the pooled datasets. Rescaling and model estimation continues in an iterative process until the log-likelihood values indicate that a 'best fit' has been obtained.

If both datasets have identical parameters, rescaling is unnecessary and the ratio of scale parameters is 1 (Blamey *et al.* 1998). If dataset X_1 has more random noise than dataset X_2 the variance-scale parameter ratio, $\phi^{1/2}$, will be less than 1; if the opposite is true, ratios will be greater than 1. Significant differences between datasets are tested with a form of the likelihood ratio test (Swait and Louviere 1993). This takes the following form:

$$LR = -2(\mathrm{LogL}_{X_{(1/2)}} - (\mathrm{LogL}_{X_1} + \mathrm{LogL}_{X_2})) \tag{5}$$

where $\mathrm{LogL}_{X_{(1/2)}}$ is the log-likelihood value attached to the MNL model of the stacked dataset at the optimum level of ϕ, and LogL_{X_1} and LogL_{X_2} are the log-likelihoods of the MNL models for the individual datasets (Swait and Louviere 1993; Blamey *et al.* 1998). The resulting likelihood ratio statistic follows an asymptotic χ^2 distribution with $(P + 1)$ degrees of freedom, where P is the number of parameters across the three models involved.

4. Framing estimates of value for international rainforest conservation

Surveys involving CM experiments were designed to estimate the preservation values that Australians might hold for rainforest in Vanuatu. An experimental design procedure was used to design the scenario choices presented to survey respondents, based on the attributes noted earlier. Three successive surveys were run in Brisbane, Australia in 1995 and 1996 by a market research firm. These surveys involved 100, 200 and 100 respondents respectively.

The three surveys consisted of some general ranking and choice exercises designed to remind respondents about a range of environmental issues and budget constraints, a CM experiment section, and questions about the characteristics of respondents. The surveys differed only in attributes and levels used in the CM experiment section.[3] These are displayed in table 1.

In each of the surveys, a labelled model format was employed. Respondents were presented with nine choice sets, each with a standard six locations on offer, as well as a no choice option (figure 1). This meant that seven choices were available in each set of profiles. Hence, each location became a labelled alternative in which a number of other unrecognised attributes specific to the locations may have contributed to choices. These effects were then captured in the estimation of ASC in the models.

Holding the number of locations to six, and allowing the other attributes to vary across only three levels, enabled a powerful experimental design process to be generated. This selected 81 choice sets from the full factorial of possible choices. For convenience, these were blocked into groups of nine, so that there were nine versions of the survey. Each respondent was given a random version, which contained nine choice sets each.

4.1 Experiment A: framing rainforest preservation measures

Some of the strengths of CM can be seen in relation to experiment A, where a sample choice set is presented in figure 1. This demonstrates that CM can be successfully used to disguise a particular issue within a pool of substitutes and frame the choices against a range of component attributes.

In experiment A, respondents were offered one Australian and five international locations for rainforest conservation. The results of the survey, including two significant interactions between attributes, are given in table 2.[4]

Significant interactions were detected between attributes area and potential to visit; and area and special features. The effect of those interactions was to make two of the individual attributes (area and potential to visit) insignificant. All other attributes in the model were statistically significant at conventional levels and their signs were as expected a priori. The overall fit of the model, as measured by McFadden's R-squared, was also very good by the conventional standards used to judge probabilistic discrete choice models. The coefficients for location ASC indicate that Indonesia was the least preferred location.

[3]Copies of the surveys are available from the authors on request.

[4]More accurate models can often be generated by including the socioeconomic characteristics of respondents (Rolfe *et al.* 2000), and/or the use of nested logit models. For brevity, these expanded models have not been reported here.

Table 1 Attributes and levels used in the Choice Modelling experiments

Attribute	Levels – Experiment A	Levels – Experiment B	Levels – Experiment C
Location	Vanuatu Far North Queensland Papua New Guinea South America Thailand Indonesia	Vanuatu Far North Queensland South America Indonesia South-east Queensland Northern New South Wales	Vanuatu Far North Queensland South America Indonesia South-east Queensland Northern New South Wales
Area	100 ha 1,000 ha 10,000 ha	100 ha 1,000 ha 10,000 ha	100 ha 1,000 ha 10,000 ha
Rarity	Not rare at all Fairly rare Extremely rare	Not rare at all Fairly rare Extremely rare	Not rare at all Fairly rare Extremely rare
Visits	No visits allowed Visits possible but moderate access and few facilities Easy to visit with full facilities	No visits allowed Visits possible but moderate access and few facilities Easy to visit with full facilities	No visits allowed Visits possible but moderate access and few facilities Easy to visit with full facilities
Local people	Local people will be worse off No local people Local people will be better off	Local people will be worse off No local people Local people will be better off	Local people will be worse off No local people Local people will be better off
Special features	No special features Special plants and animals Special landscapes & plants and animals	No special features Special plants and animals Special landscapes & plants and animals	
Vegetation			Rainforest Wetlands Rangelands
Price	\$ 5 \$10 \$50	\$ 5 \$10 \$50	\$ 5 \$10 \$50

Table 2 Multinomial logit results for experiments with significant interactions

Variable	Experiment A	Experiment B	Experiment C
ASC†			
Vanuatu	−1.67**	−3.91**	−3.13**
Far North Qld	0.09	−2.81**	−1.85**
South America	−1.61**	−3.97**	−3.22**
Indonesia	−1.94**	−4.11**	−3.22**
Papua New Guinea	−1.83**		
Thailand	−1.92**		
Northern NSW		−3.23**	−2.29**
South-east Qld		−2.92**	−1.58**
Area	−2.69E-05	4.07E-05**	3.03E-05**
Rarity	0.6292**	0.6594**	1.0106**
Potential to visit	−0.0233	−0.1455	−0.4232*
Effect on locals	0.3670**	0.6572**	0.1072
Special features	0.1055*	0.4778**	
Vegetation			−0.1850**
Price	−0.0078**	−0.0266**	−0.0163**
Area/visits	3.93E-05**		
Area/local	3.55E-05**		
Visits/locals		0.1131*	0.2615**
Visits/price		0.0049*	
Locals/special		−0.1665**	
Log likelihood	−1365	−3037	−1400
χ^2 statistic	252	591	456
No. variables	8	15	13
Significance of χ^2	0.000000	0.000000	0.000000
R-squared	0.22047	0.12386	0.20012

*Significant at the 5% level.
**Significant at the 1% level.
†ASC, alternate specific constants.

The model results can be used to estimate values for changes in the supply of rainforest conservation.[5] For example, the part-worth of a change in location from Thailand to South America from experiment A is given by:

$$\text{Part-worth (Thailand to South America)}$$
$$= -1/-0.0078(-1.61 + 1.92) = \$39.74$$

If two rainforest conservation proposals in Thailand and South America were similar in every other respect, the South American proposal would provide \$39.74 more value to Brisbane households than the Thailand proposal.

However, the Hausman–McPherson tests to check error distributions indicated that some IIA/IID violations occurred in this model when an overseas location was dropped from the choice set. The lack of violations for

[5]Examples of part-worths and value differences between profiles are given in Rolfe *et al.* (2000).

the no choice, Far North Queensland (FNQ) and Papua New Guinea (PNG) alternatives meant that dropping either of these alternatives did not change the ratio of choice probabilities for the other alternatives. However, when the Vanuatu and South America locations were dropped, IIA/IID violations were present in the model.[6] This implies that the other overseas locations are being viewed by respondents as substitutes, rather then as independent alternatives. In contrast, the Far North Queensland location was viewed as an independent alternative. The choice alternatives were not being framed by respondents consistently, and value estimates, as demonstrated above, may not be accurate.

These IIA/IID tests lead to an important conclusion: framing problems may be involved in valuation experiments that otherwise appear robust. An inconsistency has been identified in the CM experiment that would have remained hidden in a similar CVM approach. It was hypothesised that these framing inconsistencies may have been occurring because the choice sets were unbalanced between Australian and overseas locations. Some respondents may have always wanted choices that involved attractive Australian profiles. From the wide selection of possible locations for rainforest conservation around the world, the six locations that have been chosen have not been particularly suitable in this survey format.

4.2 Experiment B: testing for minor framing effects

Experiment B was conducted with two location changes from experiment A to give three Australian and three overseas locations.[7] All other factors were held constant in the choice sets, including the experimental design, and the results are reported in table 2.

For the model generated from this survey, no IIA/IID violations could be detected. This suggested that offering a more balanced set of locations could alter significantly the way in which respondents framed choices. To test whether differences between the models reported in table 2 were due to other factors apart from scale parameter differences, the confidence intervals for part-worths were estimated.[8]

These estimates were conducted by taking 200 repeated draws of the vector of coefficients, and then omitting the upper and lower tails of the

[6]The use of nested models with *Pay/NoPay* and *Australian/Overseas* decision branches did not appear to remove these violations.

[7]Papua New Guinea was replaced by northern New South Wales, and Thailand was replaced by South-east Queensland.

[8]It was difficult to apply the Swait–Louviere test for this purpose because the alternatives were not consistent between the choice sets.

Table 3 Independent from irrelevant alternatives (IIA)/independent and identically distributed (IID) tests for Experiment A

Alternative omitted	χ^2 (d.f.)	Pr(C > c)	Significant
No choice	35.9283 (6)	0.000003	Yes
FNQ	46.8908 (6)	0.000000	Yes
Vanuatu	8.3572 (6)	0.213083	No
PNG	14.7594 (6)	0.022212	Yes
South America	7.9187 (6)	0.239448	No

d.f., Degrees of freedom; FNQ, Far North Queensland; PNG, Papua New Guinea.

distributions. The tests are reported below in table 4. The results indicate that the only significant difference in part-worths exists for the area attribute and the Far North Queensland (FNQ) location. The latter result confirms the hypothesis that respondents found it difficult to frame choices in experiment A where only one Australian location was offered in the choice sets. Framing the Australian choices in a different way affected the value of conservation sites in FNQ. The results demonstrate that the part-worth tests isolate the framing effects between experiments to the particular components of choice where they are occurring.

In addition, a comparison of the confidence intervals confirmed that the choices expressed in experiment B were more deterministic than the choices expressed in experiment A. The confidence intervals for the part-worths from experiment B are much tighter than the corresponding part-worths from experiment A, indicating that lower levels of variation in choice occurred in that experiment. This meets with a priori expectations, because experiment B did not have the framing issues associated with experiment A.

Table 4 Part-worths for simple Multinomial Logit models of experiments A and B

	Experiment A (A$)		Experiment B (A$)	
Part-worth	Mean	95% Confidence interval	Mean	95% Confidence interval
Area	0.0084	0.0047, 0.0238	0.0025	0.0016, 0.0036
Rare	55.61	32.79, 124.05	39.17	30.4, 48.6
Potential to visit	15.42	0.26, 46.15	9.94	5.7, 14.8
Effect on locals	64.30	37.98, 190.92	35.20	28.6, 44.6
Special features	11.94	−1.30, 37.61	9.38	5.29, 14.31
Vanuatu	−194.66	−521, −112	−240.27	−303, −192
Far North Qld	7.94	−58.6, 70.1	−168.65	−211, −132
South America	−227.72	−609, −121	−247.44	−306, −201
Indonesia	−272.25	−794, −158	−251.7	−316, −202

4.3 The effect of framing different choices

Experiment C was designed to test for framing effects when a wider group of substitutes were presented to respondents. Instead of concentrating on rainforests, three types of vegetation (i.e. rainforests, wetlands and rangelands) were included in the profiles. This essentially expanded the set of substitutes for respondents to consider. Respondents were told that possible conservation sites for each type of vegetation had been identified in the locations used in the profiles.

To enable comparisons with experiment B, the same experimental design was used and the new vegetation attribute was substituted in place of the special features attribute. All the locations and other attributes used in experiment B were maintained for experiment C. No IIA/IID violations were detected in the results, indicating that respondents were able to frame choices consistently. Results are reported in table 2.

To test for framing effects, we wanted to consider whether the influence on choice of attributes common between experiments B and C had changed. Because the coefficients are not directly comparable, the Swait–Louviere test was performed as a way of identifying whether framing effects could be isolated between the models. The data for the surveys were stacked by stripping out the attributes that were not common to each survey and then combining the sets. The survey codes for the dataset from experiment B were varied according to a scalar factor ϕ, while the codes for the other dataset were maintained.

The dataset from experiment B had twice the number of observations as the dataset from experiment C, so the latter was stacked twice to maintain

Table 5 Results for single and stacked datasets*

Variable	Experiment B	Experiment C	Joint B&C
Vanuatu	−3.5064	−4.5717	−3.9929
Far North Qld	−2.3741	−3.2511	−2.7754
New South Wales	−2.8238	−3.7363	−3.2401
South America	−3.5312	−4.6137	−4.0243
South-east Qld	−2.5126	−3.0229	−2.7300
Indonesia	−3.6600	−4.6014	−4.0884
Rare	0.61269	0.93578	0.79047
Visit	0.15722	0.12009	0.13817
Local	0.55670	0.65376	0.61996
Area	0.00004	0.00004	0.00004
Price	−0.01469	−0.01388	−0.01468
Log-likelihood	−3056.756	−2827.778	−5911.073
R-squared	0.12730	0.19357	0.15667
χ^2 (5)	549.9	863.9	1328.6

*All parameters were significant at the 1% level.

consistency. Repeated MNL models for the stacked dataset were calculated with varying levels of ϕ. The maximum log-likelihood value of the MNL model was achieved when ϕ assumed a value of 0.94. The MNL models for the individual datasets were also calculated, and the results for the three models are reported in table 5.

This enabled the likelihood ratio test to be performed as follows:

$$\begin{aligned} \mathrm{LR} &= -2(-5911.073 - (-3056.756 - 2827.778)) \\ &= 53.078 \end{aligned} \tag{6}$$

There are (11 + 1) degrees of freedom associated with the test, implying that the χ^2 statistic at a 5 per cent significance level is 21.026. This is smaller than the calculated statistic, and means that the hypothesis; that is, that the vector of parameters are equivalent across the two datasets, should be rejected. The differences in the scale parameter are not enough to account for the variations in the coefficients.

The conclusion from this test is that framing effects have occurred between experiments B and C. After differences in the scale parameters had been accounted for, the variations in the coefficients were still significant. This means that the introduction of a wider choice set in experiment C compared with experiment B has impacted on the relative values of the different coefficients.

The size of the scalar factor ratio ϕ identified in the analysis also gives some indication about how respondents framed their choices. This is because the scalar factor ratio is essentially the inverse of the standard deviation of the error distributions for the different models. Because the scalar factor ratio identified is less than 1, the dataset from experiment B has more random noise than the set from experiment C. This indicates that respondents were slightly more comfortable with the wider choice than with the narrowly defined experiment that focused only on rainforests.[9]

More precise evidence about where framing effects have occurred can be gained by comparing the 95% confidence intervals for the part-worths that are common to both experiments. These are set out in table 6.

The results showed that while the means for the part-worths for the attributes appeared to be higher for experiment C than for experiment B, the corresponding means for the part-worths for the ASC (locations) between the same two experiments appeared to be lower. However, there was no significant difference in part-worths between the experiments, apart from the rarity attribute. For that attribute, the value derived from experiment C was significantly different to the value derived from experiment B.

[9]Because the scale parameter identified was very close to a unity value, the framing differences between the two CM experiments were slight.

Table 6 Part-worths for simple multinomial logit models of experiments B and C

	Experiment B (A$)		Experiment C (A$)	
Part-worth	Mean	95% Confidence interval	Mean	95% Confidence interval
Area	0.0025	0.0016–0.0036	0.0029	0.0016–0.0045
Rare	39.17	30.4–48.6	65.59	49.8–92.5
Potential to visit	9.94	5.7–14.8	10.69	3.3–20.0
Effect on locals	35.20	28.6–44.6	44.39	32.7–69.0
Vanuatu	−240.27	−303–−192	−298.65	−347–−222
Far North Qld	−168.65	−211–−132	−202.88	−306–−147
New South Wales	−192.89	−242–−153	−237.65	−364–−179
South America	−247.44	−306–−201	−303.72	−445–−222
South-east Qld	−175.37	−224–−139	−185.74	292–−133
Indonesia	−251.7	−316–−202	−300.21	−444–−221

This indicates that framing effects involved in moving from the rainforest scenarios to broader vegetation conservation scenarios are centred on the rarity attribute.

5. Conclusion

The CM technique has significant strengths for estimating values for environmental goods in ways that minimise potential problems of framing. On one level, the technique offers significant advantages over the CVM in its ability to offer respondents choices from a wide pool of potentially substitutable goods. The ability to disguise an amenity of interest within a pool of potential trade-offs is an important way of minimising information transfer and other potential biases, and modelling realistic choices. This appears to be important where the amenity of interest may not be particularly familiar to respondents.

On another level, the technique has significant advantages in that it can frame choices according to a number of attributes, including offsetting socio-economic and environmental ones. This frames choices in more realistic contexts, as well as providing analysts with a rich information set about value trade-offs.

The CM technique allows for a more rigorous testing of framing effects than does the CVM. The evidence from three experiments reported in this paper suggests that framing effects in relation to substitutes are more widespread in stated preference valuation experiments than may be commonly thought. First, the results indicate that respondents may have difficulties in framing some choices, depending on the pool of substitutes and choice options offered. In contrast to the conclusions of Boyle (1989), it appears that small differences in experiment presentation can influence choice consistency and also lead to significant changes in model parameters. It is notable though that framing effects are often concentrated on a subset of attributes involved, implying that they may not be distinguishable in a similar CVM study.

Second, changes in the range of substitutes that respondents have to consider also may cause framing effects. In the experiments reported in this paper, a substantial expansion of the pool of substitutes on offer led to variation in model parameters. It is notable though that when the part-worths were considered, this variation concentrated on the coefficients for one attribute (rarity), and that no significant differences could be found for the locations and other attributes.

These results give a very different interpretation to framing effects offered by Boyle (1989). The CM experiments reveal that small differences in trade-offs and consideration of trade-offs can lead to framing effects, but that large changes in the range of substitutes does not automatically lead to larger framing effects. Indeed, the experiments presented showed that the framing effects were very limited in the latter case. One conclusion that can be drawn from this result is that framing effects may not automatically be associated with substantial changes in the range of substitutes that respondents are asked to consider. There appears to be significant potential for the CM technique to be used to research this issue further.

The key conclusion that can be drawn about good design practice in CM is that choices should be framed in ways that survey respondents feel comfortable with. When respondents view the number and types of choices as being realistic, then the evidence suggests that common attributes between similar studies are valued in much the same way. When respondents do not view choices as being realistic, then small changes in presentation appear to drive value changes. These results suggest that the time spent in focus groups and pretests and on presentation issues when developing CM applications may be crucial in achieving robust outcomes.

Acknowledgements

The research reported in this paper has been supported by the Australian Centre for International Agricultural Research (ACIAR) Project 9020 (The socioeconomic assessment of native forest preservation proposals in Vanuatu: implications for forest management).

References

Blamey, R.K., Bennett, J.W., Louviere, J.J., Morrison, M.D. and Rolfe, J.C. 1998, *Attribute selection in environmental choice modelling studies: the effect of causually prior attributes,* Choice Modelling Research Report No. 7, University College, University of New South Wales, Canberra.

Boyle, K.J. 1989, 'Commodity specification and the framing of contingent-valuation questions' , *Land Economics*, vol. 65, pp. 57–63.

Carson, R.T. and Mitchell, R.C. 1995, 'Sequencing and nesting in contingent valuation surveys' , *Journal of Environmental Economics and Management*, vol. 28, pp. 155–73.

Cummings, R.G., Brookshire, D.S. and Schultze, W.D. (eds), 1986, *Valuing Environmental Goods: A State of the Art Assessment of the Contingent Valuation Method*, Rowman and Allanheld, Totowa.

Cummings, R.G., Ganderton, P.T. and McGuckin, T. 1994, 'Substitution effects in CVM values' , *American Journal of Agricultural Economics*, vol. 76, pp. 205–14.

Hanemann, W.M. 1984, *Applied Welfare Analysis with Quantitative Response Models*, Working Paper no. 241. University of California, Berkeley.

Hoehn, J.P. 1991, 'Valuing the multidimensional impacts of environmental policy: theory and methods' , *American Journal of Agricultural Economics*, vol. 73, pp. 289–99.

Hoehn, J.P. and Loomis, J. 1993, 'Substitution effects in the contingent valuation of multiple environmental programs: a maximum liklihood estimator and empirical tests' , *Journal of Environmental Economics and Management*, vol. 25, pp. 56–75.

Hoehn, J.P. and Randall, A. 1987, 'A satisfactory Benefit–Cost indicator from contingent valuation' , *Journal of Environmental Economics and Management*, vol. 14, pp. 226–47.

Kotchen, M.J. and Reiling, S.D. 1999, 'Do reminders of substitutes and budget constraints influence contingent valuation estimates? Another comment' , *Land Economics*, vol. 75, pp. 478–82.

Krinsky, I. and Robb, A. 1986, 'Approximating the statistical properties of elasticities' , *Review of Economics and Statistics*, vol. 68, pp. 715–19.

Loomis, J.B., Gonzalez-Caban, A. and Gregory, R. 1994, 'Substitutes and budget constraints in contingent valuation' , *Land Economics*, vol. 70, pp. 499–506.

Louviere, J.J. 2001, 'Choice experiments: an overview of the concepts and issues' , in Bennett, J.W. and Blamey, R.K. (eds), *The Choice Modelling Approach to Environmental Valuation*, Edward Elgar, Cheltenham.

McFadden, D. 2001, 'Economic choices' , *American Economic Review*, vol. 91, pp. 351–78.

Mitchell, R.C. and Carson, R.T. 1989, *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Resources for the Future, Washington.

Morrison, M.D., Bennett, J.W. and Blamey, R.K. 1998, *Valuing Improved Wetland Quality Using Choice Modelling*, Choice Modelling Research Report No. 6, University College, University of New South Wales, Canberra.

Neill, H.R. 1995, 'The context for substitutes in CVM studies: some empirical observations' , *Journal of Environmental Economics and Management*, vol. 29, pp. 393–7.

NOAA 1993, 'Report of the NOAA panel on contingent valuation' , *Federal Register*, vol. 58, pp. 4602–14.

Randall, A. and Hoehn, J.P. 1996, 'Embedding in market demand systems' , *Journal of Environmental Economics and Management*, vol. 30, pp. 369–80.

Rolfe, J. and Bennett, J.W. 1995, *Using Focus Groups to Establish Valuation Frameworks for International Rainforests*, Vanuatu Forest Conservation Research Report No. 9, University College, University of New South Wales, Canberra.

Rolfe, J., Bennett, J. and Louviere, J. 2000, 'Choice modelling and its potential application to tropical rainforest preservation.' , *Ecological Economics*, vol. 35, pp. 289–302.

Swait, J. and Louviere, J.J. 1993, 'The role of the scale parameter in the estimation and comparison of multinomial logit models' , *Journal of Marketing Research*, vol. 30, pp. 305–14.

Whitehead, J.C. and Bloomquist, G.C. 1995, 'Do reminders of substitutes and budget constraints influence contingent valuation estimates: comment?' , *Land Economics*, vol. 71, pp. 541–3.

Whitehead, J.C. and Bloomquist, G.C. 1999, 'Do reminders of substitutes and budget constraints influence contingent valuation estimates: reply to another comment? ', *Land Economics*, vol. 75, pp. 483–4.