*The World's Largest Open Access Agricultural & Applied Economics Digital Library*

**This document is discoverable and free to researchers across the globe due to the work of AgEcon Search.**

**Help ensure our sustainability.**

Give to AgEcon Search

AgEcon Search
http://ageconsearch.umn.edu
aesearch@umn.edu

*Papers downloaded from **AgEcon Search** may be used for non-commercial purposes and personal study only. No other use, including posting to another Internet site, is permitted without permission from the copyright owner (not AgEcon Search), or as allowed under the provisions of Fair Use, U.S. Copyright Act, Title 17 U.S.C.*

New Zealand Agricultural & Resource Economics Society (Inc.)

# Difficult Choices: What Influences the Error Variance in a Choice Experiment?

**Dr Dan Marsh[1] & Yvonne Phillips[2]**
[1]Economics Department, University of Waikato
dmarsh@waikato.ac.nz
[2]PhD student, Department of Economics,
University of Waikato
ysm1nz@gmail.com

**Paper presented at the 2012 NZARES Conference**
Tahuna Conference Centre – Nelson, New Zealand. August 30-31, 2012

*Copyright by author(s). Readers may make copies of this document for non-commercial purposes only, provided that this copyright notice appears on all such copies*

# Difficult Choices: What Influences the Error Variance in a Choice Experiment[1]?

Dan Marsh[2] Yvonne Phillips[3]

## Abstract

In models of choice probability there can be heterogeneity both in individual preferences and in the error in the unobserved portion of utility. The error variance, or its inverse, the scale factor is often assumed to be identically distributed for all individuals and alternatives but this can be an unrealistic assumption. For this study we explicitly model the effect of observed variables on choice reliability through parameterization of the scale factor.

We analyse Canterbury region residents' preferences for water quality in New Zealand's Hurunui River using a fully-ranked choice experiment with two treatment groups for elicitation format: best-worst and repeated-best ranking. We find that error variance decreases with each level of ranking. The best-worst sequential ranking technique is recommended in the literature but we find in practice it is associated with a higher error variance than an alternative, repeated-best technique.

Choices which included one or more alternatives with a negative price (a reduction in local taxes) had a higher error variance and this has implications for estimates of gain/loss asymmetry. Conversely, people who had seen the river, or spent longer on the choice task, or rated their own level of understanding highly had a lower error variance. People who spent more time on a choice task also made more reliable choices, up to a point.

We also find that parameterizing the scale factor reduces the standard deviation of random parameters in a mixed logit model. Scale variation confounds the identification of preference heterogeneity and care should therefore be taken to control for expected sources of this variation.

**Keywords**: Choice experiment, scale factor, error variance, water quality, New Zealand.

[1] We are indebted to Riccardo Scarpa for original ideas and advice on the methodological issue explored in this preliminary paper.
[2] Department of Economics, University of Waikato, Hamilton, New Zealand, Email: dmarsh@waikato.ac.nz
[3] PhD Student, Department of Economics, University of Waikato, New Zealand, Email: ysm1nz@gmail.com

# 1. Introduction

The core purpose of discrete choice experiments is to predict how people would behave when faced with real-world situations, whether buying a particular product or supporting an environmental initiative. The ability of a model to explain and predict choices is therefore of considerable practical, commercial and academic interest.

Choice analysis is based on random utility theory (RUT), which separates the utility gained from a specific choice, into a deterministic and random component. In a multinomial logit (MNL) model the random component, or error term, is assumed to be independently and identically distributed following a type I extreme value (Gumbel) distribution (McFadden 1974). The choice probabilities are defined by:

(1) $P_{jK} = exp\big(\mu\beta' X_{jK}\big) / \sum_{j=1}^{J} exp\big(\mu\beta' X_{jK}\big)$

where $P_{jK}$ is the probability of choosing alternative j in choice set K, $X_{jK}$ is a vector of attribute levels corresponding to alternative j, β is the vector of indirect marginal utilities, and μ (also known as the scale factor) is inversely proportional to the Gumbel error. Louviere et al., (2008) defined choice consistency as "variability in choice outcomes not explained by attributes and associated preference weights". This can be associated with preference heterogeneity, choice uncertainty, variance of the Gumbel error, or heteroscedasticity in the Gumbel error. Generally it is impossible to know for sure whether inconsistency is behavioural or due to model misspecification or both (Louviere, Islam et al. 2008).

In the design phase of a discrete choice experiment, care should be taken to minimise the choice error that can be attributed to design factors. In this study we contribute to the literature on experimental design by investigating whether the type of ranking instruction given to participants affects choice error. Half the participants were directed to sequentially select best and worst alternatives, while the other half were directed to sequentially select the best until all alternatives were ranked. We are indebted to Riccardo Scarpa for suggesting this line of investigation. As far as we are aware there are no published studies that compare best-worst with some other sequential ranking elicitation method although work in this area by Scarpa, Rose, Marley and Collins was presented at the 2011 International Choice Modelling Conference in Leeds, UK.

In the ex-post analysis of our choice data we attempt to control for potential sources of variation in the scale factor or Gumbel error. If unexplained scale variation exists between individuals or choice situations then observations with systematically larger random variance will receive less "weight" in estimating the β coefficients for the observed variables (Bradley and Daly 1994). Welfare estimates may be biased (Adamowicz, Bunch et al. 2008) and parameter distributions may be confounded with unobserved scale factor distributions (Louviere 2004). The direction of the bias is situation specific depending on whether the individuals with larger variances have higher or lower willingness to pay.

Early studies of scale variation used heteroscedastic MNL models to allow the variances of unobserved components from two or more sources of data to differ, for example, scale variance in ranking data (Ben-Akiva, Morikawa et al. 1992) and choice set order (Bradley and Daly 1994). Subsequent research utilized parameterized heteroskedastic MNL models to test the effect of various design factors on scale; for example Swait and Adamowicz (2001); DeShazo and Fermo (2002) and Caussade, Ortúzar et al. (2005). More recent research has shown that it is possible to specify models to accommodate both preference and scale factor heterogeneity using heteroscedastic mixed logit models Louviere, Islam et al. (2008) and Scarpa, Notaro et al. (2011).

Using data on people's preferences for alternative future development scenarios for the Hurunui river catchment in New Zealand we explicitly model the effect of observed variables on choice reliability. We compare the estimation results from MNL, heteroskedastic MNL, mixed logit and heteroskedastic mixed logit models to illustrate the effect of controlling for scale variation and best-worst versus repeated best treatment.

## 2. Empirical context

The Hurunui River is widely regarded as being the most scenic and unspoilt of the seven major alpine rivers in the Canterbury Region of New Zealand's South Island. From its headwaters in the Southern Alps, the Hurunui Rivers flows through alpine lakes and foot hills before crossing the Amuri Plains and flowing through a gorge on its way to the Hurunui Estuary about 200 kilometres from its source. The river is highly significant to Ngai Tahu and nationally important for fishing and Kayaking. It also provides an important habitat for a number of endangered fish and bird species (Environment Canterbury, 2010)[i].

The future of the Hurunui River and its catchment has been hotly contested, between those who seek to store and/or divert water from the river in order to increase agricultural production and those who would like to see the river undeveloped and the quality of natural resources in the river and catchment improved. The Canterbury Regional Council, being concerned about the cumulative effects of intensive land use on surface and ground water quality set out to develop an approach to manage catchment nutrient loads across the region in order to achieve the objectives of its Natural Resources Regional Plan (NRRP) for water quality and aquatic habitats. One approach, drawing on deliberative and systems methods was developed and tested in the Hurunui catchment in 2010/11. The process was carried out both through workshops of stakeholder groups addressing the problem at a regional level and a series of catchment level stakeholder workshops held in the Hurunui District (Wedderburn et al., 2011). A key outcome of this process was the drafting of a preferred approach for the management of the cumulative impacts of land use on water quality in the catchment.

The policy objective of the choice experiment outlined in this paper was to describe and quantify the preferences of Canterbury Region residents with respect to existing conditions (the status quo) and potential future land use and water quality scenarios for the catchment. It was envisaged that this quantitative information on preferences across the region would be used by policy makers at the same time as they considered the outcomes of the stakeholder deliberative process.

## 3. Method

### 3.1 Choice experiment structure

Discrete choice experiments have been widely used in environmental valuation since the earliest application by Boxall, Adamowicz et al. (1996) and are well-suited to situations where policy alternatives have multiple impacts and the objective is to estimate the value of these impacts. Rank-ordered choices have the advantage of providing richer preference information than methods which elicit only the favourite (Hausman and Ruud 1987) particularly since the marginal benefit of asking repeated questions about alternatives within a choice situation is generally greater than the marginal cost.

For this study we obtain full rankings of the five alternatives in each choice situation and use an exploded logit specification to take into account the sequential way in which the ranks are obtained (Lancsar and Louviere 2008). The complete ranking of *J* alternatives in a choice set is a sequence of *J*-1 discrete choices drawn without replacement from the starting set of five alternatives. The utility structure for each choice task is:

$$(2)\quad U_{njk} = \left(\lambda V_{njk} + \varepsilon_{njk}\right) \times \delta_j$$

where n are individual respondents, j are the alternatives, k is the number of alternatives remaining for each choice and $\delta_j$ denotes whether alternative *j* is available or was previously selected. The scale parameter or inverse Gumbel error is denoted by λ.

Participants in ranking tasks may be left to decide how to achieve full ranking or given specific instructions on the order in which to select the rankings, as in Louviere (2004). One elicitation technique known as "best-worst" ranking is to ask respondents to sequentially choose the best and worst alternatives until all are ranked, as in Louviere, Street et al. (2008). However, researchers could theoretically instruct participants to rank the alternatives in any order.

We divided the sample into two groups who were given different instructions. Half the respondents were directed to use a "best-worst" ranking technique and the other half were directed to repeatedly select their favourite from the alternatives remaining ("repeated best"). In both treatments the first choice involved selecting the favourite alternative from a set of five. The favourite alternative was then hidden. Respondents in the best-worst treatment were then directed to select their least preferred option, while the other group was directed to select their next favourite. The process was repeated until the five alternatives were all ranked. Table 1 shows the relationship between selection order and rank for both treatments.

**Table 1- Mapping between treatment, selection order and rank**

| | **Selection order** 1st | 2nd | 3rd | 4th |
|---|---|---|---|---|
| **Number of alternatives** | 5 | 4 | 3 | 2 |
| **Rank under best-worst treatment** | 1st | 5th | 2nd | 4th |
| **Rank under repeated-best treatment** | 1st | 2nd | 3rd | 4th |

In rank-ordered choices the Gumbel error and scale parameter vary across ranks, an issue first addressed by Hausman and Ruud (1987). Errors in welfare estimates may result if rank-order data is pooled without controlling for this scale heterogeneity (DeShazo and Fermo 2002). However, parameter estimates derived from preferred choice models are consistent with those obtained from first rankings once the scale differences are accounted for (Caparros, Oviedo et al. 2008). Using a parameterized heteroskedastic model we test whether the elicitation (best-worst versus repeated-est) method also has implications for the scale parameter.

## 3.2 Heteroscedastic MNL and RPL models

The parameterized scale factor is required to be positive, and so is specified as an exponential function of a sum of explanatory variables to increase or decrease the scale. We also control for other sources of scale heterogeneity, or heteroscedasticity in the Gumbel error.

Scale heterogeneity is essentially caused by variation in the gap between choice task complexity and individual cognitive ability (Heiner 1983). As the gap grows, individuals use more simplifying rules and choices become less certain. One measure of complexity is the number of alternatives in a choice task. The Gumbel error variance increases systematically with the number of alternatives

(Caussade, Ortúzar et al. 2005). In an iterative ranking exercise, however, the error may decrease with the number of alternatives because respondents find it easier to identify favourite or least favourite options and more difficult to rank the remainder (Ben-Akiva, Morikawa et al. 1992). DeShazo and Fermo (2002) and Louviere, Islam et al. (2008) define other measures of complexity based on the number of attributes and levels and the variation of attribute levels between and across alternatives. In our study there is insufficient variation in these complexity measures to identify any effect on the scale parameter. All of the choice cards had the same number of attributes and levels and a similar variance in levels.

The other dimension of the ability-complexity gap is individual cognitive ability. We test several factors expected to be related to individual cognitive ability including education, self-rated understanding, and familiarity with the Hurunui river as defined by seeing or visiting it. Individual cognitive ability may also vary during the course of the survey due to learning effects or fatigue. Authors such as Caussade, Ortúzar et al. (2005) and Scarpa, Notaro et al. (2011) report that Gumbel error tends to decrease as respondents complete more choice tasks. This effect peaks at some point and then declines for subsequent tasks as fatigue sets in. This issue is often addressed in choice experiments by randomising the choice card order. Our approach is to both randomise and include a set of parameters denoting choice task sequence to explicitly model for the scale effect. Finally, we also include a parameter for the amount of time spent on each choice task, as an indicator of relative cognitive effort.

### 3.3 Heteroscedasticity and WTP/WTA Asymmetry

We include an additional measure of choice complexity which is the presence of both positive and negative costs (an increase or decrease in total household rates) in the available alternatives.

Asymmetry between willingness to pay (WTP) and willingness to accept (WTA) is a widely reported phenomenon in discrete choice experiments (e.g. Bateman, Day et al. (2009); Hess, Rose et al. (2008); Lanz, Provins et al. (2009)). Losses of a good tend to be valued larger than gains. Welfare attributed to the status quo alternative may be artificially inflated if the asymmetry is not accounted for (Scarpa, Ferrini et al. 2005).

This asymmetry is a typical manifestation of loss aversion known as the “endowment effect”(Kahneman, Knetsch et al. 1991). Explanations for loss aversion include the role of substitution and income effects (Hanemann 1991). Disposable income constrains demand in terms of WTP, but not demand for compensation. Low substitutability implies that it is not possible to compensate an individual for the loss of a good, resulting in extreme WTA values. However, many investigations report discrepancies which are larger than those predicted by theory.

If choices with both positive and negative prices are more difficult this would cause variation in the scale factor and could exaggerate the WTP/WTA discrepancy. We employ a piecewise linear specification of the price coefficient similar to Hess, Rose et al. (2008) and Lanz, Provins et al. (2009) and test whether the inclusion of a scale parameter affects the magnitude of the asymmetry.

The specification of the scale factor is:

$$(3)\ \lambda_{bkt} = exp\left(\sum_{bk} \theta_{bk} + \sum_{t} p_t + V_n + \omega_{kt} + \propto m_{kt} + \gamma m_{kt}^2\right)$$

Where $\theta_{bk}$ is the set of parameters denoting rank *k* in treatment type *b*, $p_t$ is the set of parameters denoting the order of the choice task, $V_n$ is the vector of individual cognitive ability parameters, $\omega_{kt}$ is a dummy variable indicating whether the available set includes an alternative with a negative price and $m_{kt}$ is the time taken, in minutes, to perform the ranking task.

# 4. Survey design

## 4.1 Attribute selection

Attributes selected for inclusion in the choice experiment were informed by catchment level stakeholder workshops where qualitative methods were used to identify the most important attributes for different stakeholders. Survey design was also informed by discussion with environmental economists familiar with local water quality issues and with the technical experts who were assisting with development of the ‘preferred approach’. Advice from experts developing the preferred approach was also used to define attributes and levels for a range of future scenarios. The final set of 6 attributes were suitability for swimming and recreation, ecological health, salmon and trout populations, tributary water quality and change in number of jobs in the region.

Since some scenarios would result in reduction in environmental quality, the payment variable (local taxes) could either increase, indicating willingness to pay for improved environmental quality or could fall, indicating willingness to accept compensation for reduced quality. A specific attribute describing water quality in tributaries was included in order to better understand the relative importance of water quality in the main river (currently satisfactory) versus the lowland tributaries (currently not satisfactory).

Attribute levels are categorical and defined using minimum standards set by Canterbury Regional Council. An attribute that meets the minimum standard is defined as “satisfactory”. If it does not it is “unsatisfactory”. Exceeding the minimum standard is defined as “good”. Tributary water quality is currently unsatisfactory and expected to decline under some scenarios so an extra level “poor” was added to represent this decline. The levels for changes in jobs were based on potential effects on the agricultural sector and the wider economy resulting from different water management scenarios. The levels are: 250 fewer jobs, no change, 250 more jobs or 500 more jobs.

Early versions of the questionnaire were piloted with selected workshop participants, Canterbury residents and technical experts. At this stage interviewees were debriefed on their experience in filling in the questionnaire with several questions being improved and clarified as a result. An on-line version of the questionnaire was then pretested using Canterbury region residents. Respondents for the final version of the survey were recruited from an online market research panel in June 2011 and invited to fill in the survey online. There were quotas on age, gender and education level in order to help achieve a representative sample. People who resided outside the Canterbury region were excluded, as were people who completed the survey in less than five minutes

Figure 1 shows an example of a choice card as it was presented to participants. When participants selected an alternative it was hidden and they were then instructed to select another alternative.

Choice card 1 of 6

Please think about the options presented below and select the option you think is **best**.

| | | Current Situation | Scenario A | Scenario B | Scenario C | Scenario D |
|---|---|---|---|---|---|---|
| Main river | Suitability for Swimming and Recreation | Satisfactory √ | Good √√ | Not Satisfactory x | Satisfactory √ | Good √√ |
| | Ecological Health | Satisfactory √ | Not Satisfactory x | Not Satisfactory x | Not Satisfactory x | Good √√ |
| | Salmon and Trout | Satisfactory √ | Satisfactory √ | Good √√ | Satisfactory √ | Not Satisfactory x |
| Tributaries | Tributary water quality | Not Satisfactory x | Good √√ | Satisfactory √ | Not Satisfactory x | Poor xx |
| Economy | Number of Jobs | Stay about the same | Stay about the same | 250 *more jobs in region* | Stay about the same | 500 *more jobs in region* |
| | Cost to you | $0 *increase* | $75 *increase* | $-100 *decrease* | $25 *increase* | $200 *increase* |

**Figure 1 – Example of a choice card**

### 4.2 Efficient design

We generated a D-efficient design in six blocks using the Ngene software package (Institute of Transport and Logistics Studies 2007). Efficient designs require a smaller number of respondents to achieve a given level of statistical significance of the parameters (Scarpa and Rose 2008).

We used information from other water quality non-market valuation studies in New Zealand such as Marsh, Mkwara et al. (2011); Tait and Baskaran (2011) and incorporated this information into the initial Bayesian priors. Bayesian priors make the design efficiency more robust to misspecification than optimising with fixed priors (Ferrini and Scarpa 2007). We then updated the prior distributions with values obtained from pilot tests of the survey. The design mean D-error was 0.21 with a standard deviation of 0.008.

Rather than specifying cost as a continuous attribute we specified a large number of levels at $25 increments between -$100 and $200. A constraint was imposed so that each level appeared at least once in a block. This meant that participants saw a variety of costs without imposing too much of a penalty on design efficiency. A negative cost represents a decrease in the household's annual rates bill. Negative costs were required because water quality attributes are expected to decline under some scenarios of agricultural intensification. If cost was constrained to be positive it would be difficult to avoid dominated choice situations and design efficiency would be much lower.

## 5. Results

### 5.1 Sample statistics and model estimation

Sample statistics for the final sample of 505 completed surveys are presented in Table 2. Comparison with data from the 2006 census suggests that the sample is broadly representative of the region although it should be noted that certain groups are over or under represented. In particular, our sample somewhat over represents females, those with a post-school qualification and those in the 18-30 age bracket and under represents low income households (less than $30,000).

**Table 2 - Sample statistics**

| Treatment group | | Best-worst | Repeated-best | 2006 Census |
|---|---|---|---|---|
| Count of participants | | 250 | 255 | 521,832 |
| **Per cent** | | | | |
| Gender | Female | 62% | 55% | 51% |
| | Male | 38% | 45% | 49% |
| Age | 18-30 | 23% | 27% | 20% |
| | 30-44 | 34% | 27% | 29% |
| | 45-59 | 24% | 24% | 26% |
| | over60 | 19% | 22% | 25% |
| Post-school qualification | | 56% | 62% | 51%[4] |
| Annual household income | Less than $30,000 | 20% | 16% | 22% |
| | $30,000 to $50,000 | 16% | 21% | 20% |
| | $50,000 to $70,000 | 20% | 20% | 21% |
| | $70,000 to $100,000 | 17% | 18% | 19% |
| | Greater than $100,000 | 12% | 15% | 18% |
| | Declined | 14% | 10% | |
| Location of residence | Christchurch city | 70% | 75% | 67% |
| | Other Canterbury | 30% | 25% | 23% |
| Involved in farming | | 10% | 4% | |
| Seen the Hurunui or a tributary in last 12 months | | 38% | 44% | |
| Visited the Hurunui or a tributary in last 12 months | | 16% | 15% | |
| **Average** | | | | |
| Concern about water pollution from farming | | 4.16 | 4.25 | |
| Self-rated understanding of choices (1 to 10 - understood ) | | 6.04 | 5.88 | |
| Self-rated ease of making choices (1 to 10 – easy) | | 5.60 | 5.24 | |
| Time taken per choice card (seconds) | | 63 | 54 | |

The fully-ranked choice sets are decomposed into a series of choices as per the exploded logit specification detailed by Lancsar and Louviere (2008). The best-worst method results in different comparisons to the repeated-best method, which means the selection order needs to be taken into account in the decomposition. The sign of the utility parameters were reversed in situations where respondents were selecting the “worst” alternative.

We estimated four different models using the maximum (simulated) likelihood estimate in BIOGEME (Bierlaire 2003) and present the attribute coefficients in Table 3. Model 1 is a fixed parameter MNL model; model 2 is a fixed parameter MNL with scale parameterization; model 3 is a panel random parameters logit (RPL) model and model 4 a panel RPL with scale parameterization. Results excluding the scale parameters are presented in Table 3. The scale parameters for models 2 and 4 are presented in Table 4.

[4] Statistics New Zealand Household Labour Force Survey 2011

**Table 3 – Results - attribute coefficients and model fit**

| | | MNL | | MNL + scale params | | RPL | | RPL + scale params | |
|---|---|---|---|---|---|---|---|---|---|
| Attribute | Measure | Coef. | P-value | Coef. | P-value | Coef. | P-value | Coef. | P-value |
| Normalised Cost | $\hat{\mu}$ | -0.652 | <0.000 | -0.340 | <0.000 | -1.070 | <0.000 | -0.242 | <0.000 |
| | $\hat{\sigma}$ | | | | | 1.950 | <0.000 | 0.412 | <0.000 |
| Dummy for neg cost | $\hat{\mu}$ | 0.222 | 0.04 | 0.035 | 0.600 | 0.527 | <0.000 | 0.084 | 0.02 |
| | $\hat{\sigma}$ | | | | | 2.440 | <0.000 | 0.453 | <0.000 |
| "Good" ecology | $\hat{\mu}$ | 0.134 | <0.000 | 0.112 | <0.000 | 0.264 | <0.000 | 0.062 | <0.000 |
| | $\hat{\sigma}$ | | | | | 1.060 | <0.000 | 0.194 | <0.000 |
| "Unsatisfactory" ecology | $\hat{\mu}$ | -0.713 | <0.000 | -0.440 | <0.000 | -0.999 | <0.000 | -0.213 | <0.000 |
| | $\hat{\sigma}$ | | | | | 1.310 | <0.000 | 0.210 | <0.000 |
| "Good" fishing | $\hat{\mu}$ | 0.064 | 0.03 | 0.067 | <0.000 | 0.142 | <0.000 | 0.041 | <0.000 |
| | $\hat{\sigma}$ | | | | | 0.694 | <0.000 | 0.098 | <0.000 |
| "Unsatisfactory" fishing | $\hat{\mu}$ | -0.671 | <0.000 | -0.409 | <0.000 | -0.955 | <0.000 | -0.202 | <0.000 |
| | $\hat{\sigma}$ | | | | | 1.180 | <0.000 | 0.203 | <0.000 |
| 500 more jobs | $\hat{\mu}$ | 0.071 | 0.05 | 0.053 | 0.010 | 0.158 | <0.000 | 0.030 | <0.000 |
| | $\hat{\sigma}$ | | | | | 0.650 | <0.000 | 0.107 | <0.000 |
| 250 more jobs | $\hat{\mu}$ | 0.048 | 0.17 | 0.051 | 0.010 | 0.095 | 0.04 | 0.014 | 0.14 |
| | $\hat{\sigma}$ | | | | | 0.172 | 0.81 | 0.050 | 0.09 |
| 250 less jobs | $\hat{\mu}$ | -0.519 | <0.000 | -0.325 | <0.000 | -0.781 | <0.000 | -0.162 | <0.000 |
| | $\hat{\sigma}$ | | | | | 1.250 | <0.000 | 0.206 | <0.000 |
| "Good" recreation | $\hat{\mu}$ | 0.216 | <0.000 | 0.098 | <0.000 | 0.269 | <0.000 | 0.039 | <0.000 |
| | $\hat{\sigma}$ | | | | | 0.886 | <0.000 | 0.084 | 0.01 |
| "Unsatisfactory" recreation | $\hat{\mu}$ | -0.784 | <0.000 | -0.513 | <0.000 | -1.130 | <0.000 | -0.246 | <0.000 |
| | $\hat{\sigma}$ | | | | | 1.510 | <0.000 | 0.268 | <0.000 |
| "Good" tributaries | $\hat{\mu}$ | 0.552 | <0.000 | 0.358 | <0.000 | 0.858 | <0.000 | 0.182 | <0.000 |
| | $\hat{\sigma}$ | | | | | 0.866 | <0.000 | 0.065 | 0.03 |
| "Satisfactory" tributaries | $\hat{\mu}$ | 0.362 | <0.000 | 0.264 | <0.000 | 0.582 | <0.000 | 0.125 | <0.000 |
| | $\hat{\sigma}$ | | | | | 0.928 | <0.000 | 0.063 | 0.09 |
| "Poor" tributaries | $\hat{\mu}$ | -0.568 | <0.000 | -0.352 | <0.000 | -0.823 | <0.000 | -0.185 | <0.000 |
| | $\hat{\sigma}$ | | | | | 1.410 | <0.000 | 0.262 | <0.000 |
| Cost * post-school education | $\hat{\mu}$ | 0.244 | <0.000 | 0.135 | <0.000 | 0.237 | <0.000 | 0.101 | <0.000 |
| Cost * seen the site | $\hat{\mu}$ | 0.269 | <0.000 | 0.173 | <0.000 | 0.541 | <0.000 | 0.101 | <0.000 |
| Status quo | $\hat{\mu}$ | 0.266 | <0.000 | 0.0981 | <0.000 | 0.364 | <0.000 | 0.040 | <0.000 |
| Log-likelihood | | -12219 | | -11916 | | -14506 | | -11316 | |
| AIC | | 2.02 | | 1.97 | | 1.90 | | 1.87 | |
| BIC | | 2.03 | | 1.99 | | 1.91 | | 1.90 | |
| Adjusted rho-square | | 0.16 | | 0.18 | | 0.21 | | 0.22 | |

## 5.2 Homogenous tastes: models 1 & 2

In the first MNL model all the parameters except for "250 more jobs" are significant at least at the 5% level. The coefficients are all of the expected sign, with levels representing a decline in quality being negative. The attributes which have two improvement levels, jobs and tributary water quality, have a larger coefficient for the best level, thus conforming with the weak axiom of revealed preference. The parameters are not directly comparable between the two models but are similar in relative magnitude with some important exceptions that are detailed below.

The cost levels include both negative and positive values so we use a piecewise linear specification similar to Hess, Rose et al. (2008) to account for potential asymmetry. Cost is normalised to be a similar range to the other parameters by dividing by $200. The cost coefficient is negative and the negative cost dummy parameter is positive in all models. This is consistent with the endowment effect and means people are more willing to forgo a reduction in rates than spend their existing monetary endowment. In model 1 the WTP values are 152% higher when the overall package cost is negative. However, in model 2 it is only 110% higher. This finding suggests that scale variation may have artificially inflated the degree of asymmetry in model 1 results.

Among the other parameters the absolute value of the negative coefficients are larger than the improvement parameters, a common finding in studies which compare WTP and WTA (Lanz, Provins et al. 2009). The WTA to avoid 250 jobs lost is much higher than the WTP to gain 500 jobs. The other parameters have categorical levels so the degree of asymmetry cannot be determined.

The status quo parameter is significant and positive in both models, indicating that respondents slightly preferred the no change scenario, all else being equal. The status quo bias is another manifestation of loss aversion (Kahneman, Knetsch et al. 1991). In model 2 the status quo parameter is relatively lower compared with all other attributes. This may be another effect of controlling for scale variation caused by negative cost alternatives.

A large number of interaction terms were tested but the two that were consistently significant were the cost x post-school education interaction and cost x seen where “seen” is a dummy variable indicating the individual has personally seen the Hurunui River or its tributaries. Both of these interactions are positive, which means that people who have more education or have seen the site tend to be willing to pay more for environmental quality. Income and education are highly correlated so the education interaction effect is probably a combination of income effect and environmental awareness.

### 5.3 Heterogeneous tastes: models 3 & 4

Models 3 and 4 are panel mixed logit models with random parameters for cost, jobs, and environmental attributes. The unconditional mean parameter estimates are very similar to those in the fixed parameter models. Five hundred Halton draws were used to estimate the random parameters. Uniform distributions were used because this carries a lower risk of misspecification than less flexible distributions (Hess and Axhausen 2005). The RPL models have improved model fit, with adjusted McFadden r-squared values 0.21 and 0.22 versus 0.16 and 0.18 for models 1 and 2.

The negative cost dummy parameter is relatively larger in the RPL models. In model 3 WTP is 197% higher when the overall cost is negative. In model 4 it is 154% higher. Similar to the MNL models, the inclusion of the scale parameters has the effect of reducing the relative magnitude of the negative cost and status quo parameters. There is a small decrease in variance of the asymmetry as well.

Most of the random parameters standard deviations are significant at the 1% per cent level. In model 3 the standard deviation for “250 more jobs” is not significantly different to zero. In model 4 “250 more jobs” and “satisfactory tributaries” are significant at the 10% level only, while “good tribuaries” is significant at 5%. Almost all of the random parameter standard deviations are smaller in model 4 than model 3. The exception is “250 more jobs” but neither the means nor standard deviations were statistically significant for this parameter in either model. It appears that failing to control for scale variation in model 3 magnified the estimated variance in individual preferences, as predicted by Louviere (2004)

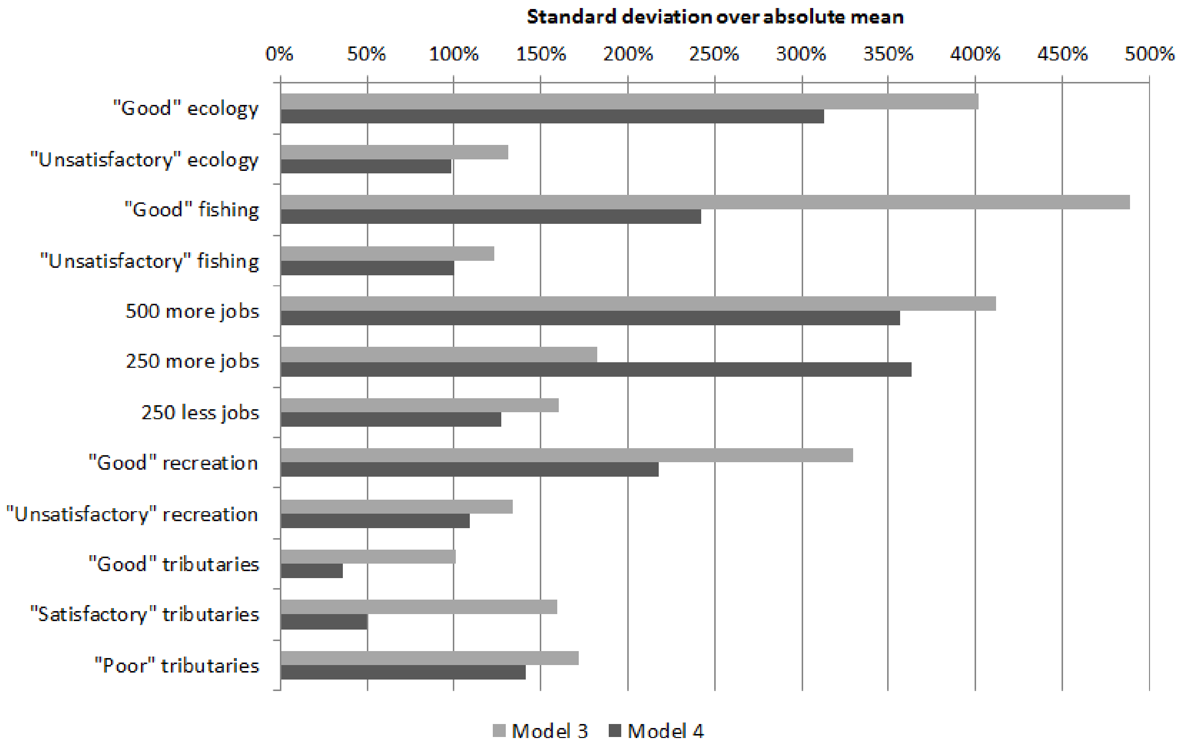


**Figure 2 - Magnitudes of standard deviations relative to parameter means**

**Table 4 – Scale-shifting parameter coefficients**

| | | Model 2 | | Model 4 | |
|---|---|---|---|---|---|
| Variable | | Coef. | P-value | Coef. | P-value |
| $p_2$ | Choice card 2 | 0.165 | 0.010 | 0.190 | 0.010 |
| $p_3$ | Choice card 3 | 0.259 | <0.000 | 0.343 | <0.000 |
| $p_4$ | Choice card 4 | 0.229 | <0.000 | 0.391 | <0.000 |
| $p_5$ | Choice card 5 | 0.331 | <0.000 | 0.547 | <0.000 |
| $p_6$ | Choice card 6 | 0.412 | <0.000 | 0.560 | <0.000 |
| $\theta_{1,2}$ | Best-worst 2nd selection | 0.563 | <0.000 | 0.870 | <0.000 |
| $\theta_{1,3}$ | Best-worst 3rd selection | 0.441 | <0.000 | 0.535 | <0.000 |
| $\theta_{1,4}$ | Best-worst 4th selection | 0.975 | <0.000 | 1.250 | <0.000 |
| $\theta_{2,1}$ | Repeated-best 1st selection | 0.042 | 0.430 | 0.006 | 0.930 |
| $\theta_{2,2}$ | Repeated-best 2nd selection | 0.315 | <0.000 | 0.494 | <0.000 |
| $\theta_{2,3}$ | Repeated-best 3rd selection | 0.476 | <0.000 | 0.574 | <0.000 |
| $\theta_{2,4}$ | Repeated-best 4th selection | 0.784 | <0.000 | 0.923 | <0.000 |
| $\propto$ | Minutes spent on choice set | 0.883 | <0.000 | 0.881 | <0.000 |
| $\gamma$ | Minutes squared | 0.186 | <0.000 | 0.179 | <0.000 |
| $V_u$ | Understanding >5 | 0.294 | <0.000 | 0.342 | <0.000 |
| $\omega$ | Available alts included a neg cost | 0.122 | 0.010 | 0.164 | <0.000 |

## 5.4 Scale parameter estimation results

In Model 2 and 4 the Gumbel error scale factor is parameterized as per equation (3) and the scale parameter coefficients are reported in Table 4 above. There is little difference in the scale parameter coefficients between the two models. Because model 4 has the best overall fit, with the highest adjusted rho-square of 0.215 and the minimum Akaike and Bayesian information criterion, we focus on model 4 for the remainder of this discussion.

## 5.5 The effect of choice task order on scale

The first five coefficients control for the effect of choice sequence on scale. The dummy coefficients for choice task 2-6 are all positive and significant at the 5% level, if not 1%. Figure 3 illustrates the coefficients for task order and 95% confidence intervals from model 4. The scale factor increases after the first task, which means the Gumbel error is lower for subsequent tasks. This result is similar to that reported by Caussade, Ortúzar et al. (2005) and Scarpa, Notaro et al. (2011) and is an indication of learning effects. There is no evidence of a fatigue effect, or decline in scale factor by the end of the survey but six choice tasks was probably not enough to cause fatigue. Scarpa, Notaro et al. (2011) report a decline in scale only after the eleventh choice task.

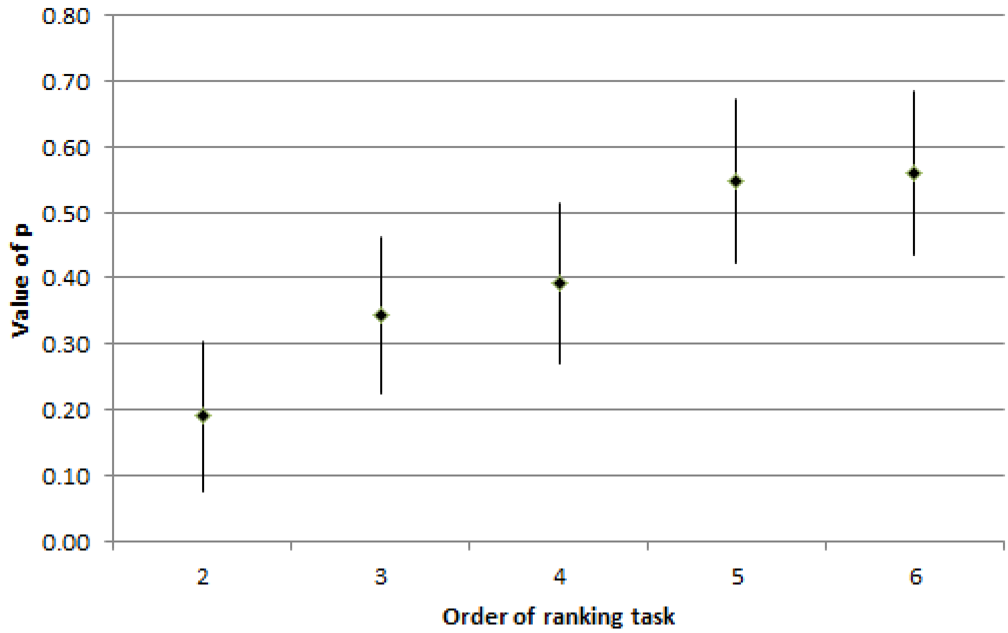


**Figure 3 - Scale effects in choice task order**

## 5.6 The effect of treatment and rank on scale

The explanatory variables for scale (scale shifters) include a set of dummy parameters for selection order in both treatment types. The base case is the first selection under the “best-worst” elicitation treatment. The coefficient estimate for the dummy denoting first selection in the repeated best treatment is statistically no different to zero. In both treatment groups the respondent is first asked to select his or her favourite alternative so there is no reason to expect any difference.

The coefficient estimates for the scale shifters for the second selection are negative for both treatments. Although fewer alternatives should make the choice comparatively easier, all else being equal, the amount of random noise is expected to increase as rank decreases because the most preferable alternative has already been removed (Ben-Akiva, Morikawa et al. 1992). The coefficient estimate for the dummy variable denoting second selection is much more negative under the best-worst treatment than it is in the repeated best case, which means this treatment group had a higher Gumbel error. This is an unexpected finding because Marley and Louviere (2005) suggest that the extremes (best and worst alternatives) are easier for people to identify than middle ranks. Our finding is not necessarily inconsistent with this theory. It is possible the choices appear more random because some people failed to follow instructions and actually selected their second favourite alternative instead of the least favourite. The words “best” and “worst” in the instructions were highlighted, enlarged, different colours, included in mouse over text and the completed rankings were shown to respondents to check before they moved on to the next card. This still does not guarantee that people properly followed the ranking instructions.

The estimated scale shifters for the third selection are of similar magnitude for both treatments. In both cases the respondent is asked to select his or her preferred alternative from the three remaining. For the fourth selection the best-worst treatment group have to select the least preferred of 2 alternatives to order ranks 3 and 4. The repeated-best group have to select their preferred alternative to order ranks 4 and 5. The best-worst treatment again has a lower scale factor thus underlining our finding that this treatment was less effective in eliciting preferences.

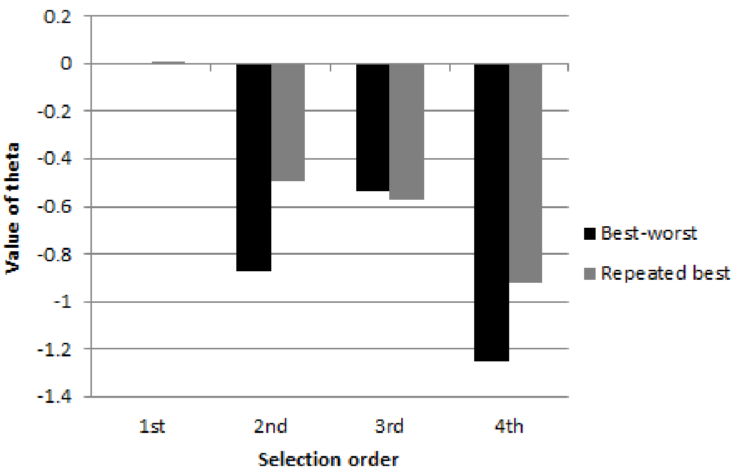


**Figure 4 - Scale effects rank order**

## 5.7 The effects of cognitive ability and cognitive effort on scale

We found that age, education level and site familiarity each had some explanatory power but self-rated understanding of the choice task proved to be a better explanatory variable and is highly correlated with all of these. Individuals rated their understanding from 1 (did not understand at all) to 10 (understood completely), and a dummy variable for greater than 5 captured most of the effect of individuals with better understanding on the scale factor. The coefficient of the variable for self-rated understanding is positive and highly significant but not as large in absolute terms as the parameters for task or rank order.

Cognitive ability or even understanding of a task does not perfectly predict individual performance on that task. Effort and concentration are also important. In this study we use response time as a simple measure of individual effort. We automatically recorded timestamps along with the choice answers so we know how long people took to fully rank each choice card (but not the time taken for each selection on that choice card). Rose and Black (2006) investigated the effect of response time on parameter variance using mixed logit models but we are not aware of any other studies which have included response as scale-shifting parameter. Response time appears to explain the scale of Gumbel error better than sequence, rank, or any other parameter we tested.

The average time taken to fully rank a choice card was 61 seconds, with a standard deviation of 45 seconds. The time parameter is positive and the time squared parameter is negative, indicating a quadratic relationship between time taken and scale (see Figure 5). Choice accuracy improves up to 140 seconds and declines thereafter, perhaps because people eventually give up and use a simplifying heuristic, or they may have been doing something else for at least part of the time.

We note that the direction of causality of time taken and size of scale factor is not clear-cut. Tyebjee (1979) reported that people took longer to decide when a choice situation was not strongly dominated. As well as being a measure of individual effort, a time variable may therefore capture residual choice complexity that is not controlled for by the other scale parameters. We tested additional measures of complexity reported by DeShazo and Fermo (2002) but these were not significant for our data set and did not affect (nor were they affected by) the time parameter. In any case, time appears to be an important parameter to include if the goal is to control for sources of scale variation.

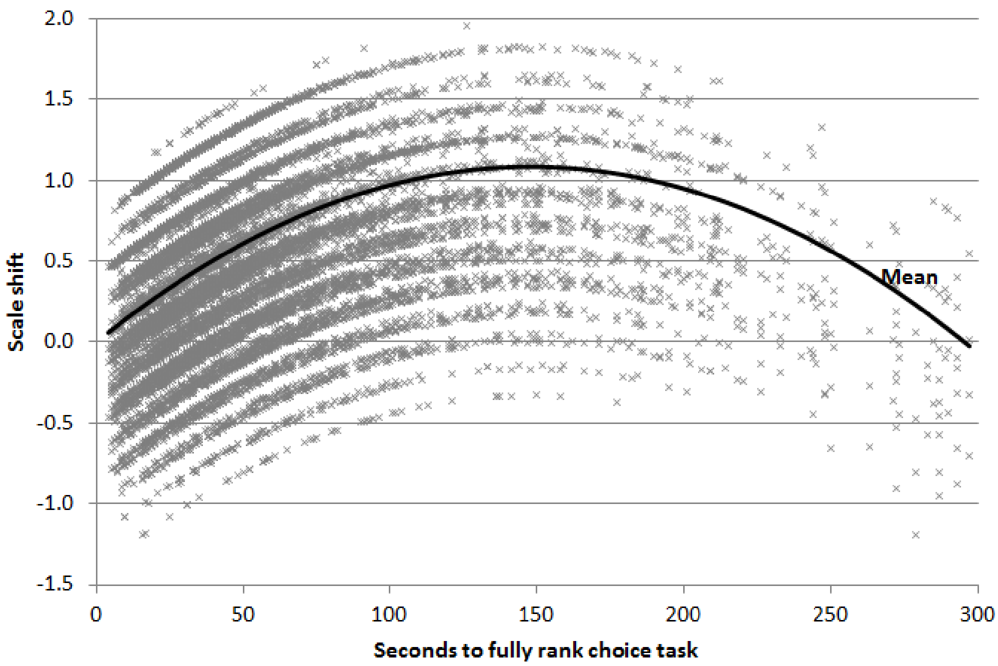


**Figure 5 – Fitted scale shift versus time**

## 5.8 Negative cost and scale variation

Forty percent of alternatives in the experimental design had a negative cost (a decline in rates), forty per cent had a positive cost and the remainder were the zero cost status quo alternatives. We included a dummy variable to indicate whether the available alternatives at each selection included at least one with a negative cost. This parameter is negative and significant at the 1% level. This implies that a choice situation which includes a negative cost is associated with larger Gumbel error.

The negative-cost scale shifter is not as large as the other scale factors but it affects the magnitude of the asymmetry between positive and negative cost coefficients. If we omit this scale parameter the asymmetry between positive and negative cost is only slightly less than the model with no scale parameters. This is true for both the fixed-taste and random parameter version. The other determinants of scale only inadequately control for the effect of scale variation on positive/negative cost asymmetry.

## 5.9 Willingness-to-pay results

Willingness-to-pay for an improvement in environmental quality, or willingness-to-avoid a decline in quality is calculated by dividing the attribute coefficient by the cost coefficient. Due to asymmetry in our cost parameter, we report two sets of unconditional mean WTP/WTA values for each model in Table 5**Error! Reference source not found.**. The first column for each model is WTP/WTA under a situation where the household faces an overall increase in rates. The second column is WTP/WTA in a situation where household faces an overall decrease in rates. We also include the effect of cost*education and cost*seen interactions by using the population mean for education and sample mean for number of people who have seen the river (due to the lack of population data).

We report the results for three models in the interest of brevity. Model 3 is the RPL model, model 4 is RPL with scale parameters, and model 5 is the same as model 4 except a latent class is used to exclude people who did not attend to (ignored) the cost attribute. Non-attendance to cost is a form of protest behaviour in which people select the scenario which gives their preferred environmental outcome regardless of the cost. No relative implicit price can be calculated for these individuals, and pooling the data will lead to upward biased welfare estimates. See Scarpa, Gilbride et al. (2009) for an in-depth explanation of attribute non-attendance and the latent class method. We find that only 36 per cent of individuals attend to the cost parameter under a latent class framework with different attribute coefficients constrained to zero.

In model 4 the difference between the two columns of WTP/WTA values is smaller than in model 3 due to the smaller estimated asymmetry effect. In model 5 all the values are smaller in magnitude, as expected.

**Table 5 – Marginal WTP/WTA under tax increase/decrease scenarios**

| | Model 3 (random parameters) | | Model 4 (incl scale parameters) | | Model 5 (excl non-attenders to cost) | |
|---|---|---|---|---|---|---|
| | Rates increase | Rates decrease | Rates increase | Rates decrease | Rates increase | Rates decrease |
| "Good" ecology | $74 | $290 | $88 | $219 | $44 | $67 |
| "Unsatisfactory" ecology | -$282 | -$1,098 | -$302 | -$753 | -$166 | -$254 |
| "Good" fishing | $40 | $156 | $58 | $143 | $25 | $39 |
| "Unsatisfactory" fishing | -$269 | -$1,049 | -$287 | -$714 | -$160 | -$244 |
| "Good" recreation | $76 | $296 | $55 | $136 | $33 | $50 |
| "Unsatisfactory" recreation | -$319 | -$1,242 | -$349 | -$869 | -$206 | -$315 |
| "Good" tribs | $242 | $943 | $258 | $643 | $147 | $225 |
| "Satisfactory" tribs | $164 | $640 | $177 | $442 | $87 | $133 |
| "Poor" tribs | -$232 | -$904 | -$262 | -$654 | -$147 | -$224 |
| 500 more jobs | $45 | $174 | $43 | $106 | $29 | $44 |
| 250 more jobs | $27 | $104 | $19 | $48 | $23 | $35 |
| 250 less jobs | -$220 | -$858 | -$230 | -$572 | -$135 | -$205 |

# 6. Discussion and conclusion

This study contributes to the choice modelling literature in several ways. First, it is one of only a small number of non-market valuation studies concerned with willingness-to-pay for environmental quality in New Zealand. We provide quantitative information on the preferences of Canterbury Region residents that can be used by policy makers as they make decisions about the future development of the Hurunui catchment. We also contribute by investigating methodological issues relating to scale variation and manners to rank alternatives in choice experiments.

Under our specific set of modelling assumptions, we find evidence that the iterative best-worst elicitation format, while praised in the literature, may have some shortcomings in practice. We hypothesize this may be because some people don't pay enough attention to notice the instructions change from "best" to "worst". Future research could investigate whether this is in fact the case. Our finding may only apply to web or mail-based surveys where there is no interviewer to ensure participants follow instructions, and it may be possible to correct with improved survey design. However, we recommend the alternative repeated-best elicitation format which shares the advantage of being sequential (which reduces the cognitive burden of full ranking) but has consistent instructions.

Similar to other studies which parameterize the scale factor, we find that rank order and choice card sequence have significant effects on the scale factor. As a point of difference we also included response time (and time squared) as a scale parameter to account for individual effort and/or residual unexplained choice complexity. We find that the effect of response time on scale is larger than that of rank or position in the choice sequence. So it appears to be a determinant of scale variation.

Discrete choice experiments commonly include attribute levels which represent a decline compared with the current situation. Our study is relatively unusual since we also include a cost parameter which may be positive or negative, representing an increase or decrease in the household rates bill. We find that choices which include a negative cost are associated with a lower scale factor, suggesting they are more difficult for respondents to process. It could be useful to know whether it is the negative cost alone that is the source of the choice error, or whether it is the combination of positive and negative costs in the set of available alternatives. An additional, or alternative measure of complexity to those developed by DeShazo and Fermo (2002) could be the number of attributes with both positive and negative levels in each choice card. This could be tested in future research with a more systematic arrangement of positive and negative levels.

The asymmetry between positive and negative cost, as modelled with a piecewise linear function, was lowest when using a heteroskedastic model with the negative cost scale parameter. This suggests that the magnitude of reported endowment effects may be biased if scale variation is not accounted for. As far as we are aware, this effect has not been reported before and could be an area for further investigation in other studies which model gain/loss asymmetry.

The heteroskedastic model offered improved fit and had relatively smaller random parameter standard deviations, which is consistent with the notion that determinants of systematic preference variation and scale variation are usefully separable in discrete choice analysis. This highlights the importance of controlling for scale variation when the researcher is interested in predictive power.

**Acknowledgements**

This research was funded under FRST Programme C10X0603: Delivering tools for improved environmental performance funded by The Foundation for Research Science and Technology (FRST) and the Pastoral 21 partners Dairy New Zealand, Meat and Wool New Zealand and Fonterra.

## References


Adamowicz, W., D. Bunch, et al. (2008). "Behavioral frontiers in choice modeling." Marketing Letters **19**(3): 215-228.

Allison, P. D. and N. A. Christakis (1994). "Logit models for sets of ranked items." Sociological methodology **24**(1994): 199-228.

Bateman, I. J., B. H. Day, et al. (2009). "Reducing gain-loss asymmetry: A virtual reality choice experiment valuing land use change." JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT **58**(1): 106-118.

Ben-Akiva, M., T. Morikawa, et al. (1992). "Analysis of the reliability of preference ranking data." Journal of Business Research **24**(2): 149-164.

Bierlaire, M. (2003). "BIOGEME: a free package for the estimation of discrete choice models."

Boxall, P. C., W. L. Adamowicz, et al. (1996). "A comparison of stated preference methods for environmental valuation." ECOLOGICAL ECONOMICS **18**(3): 243-253.

Bradley, M. and A. Daly (1994). "Use of the logit scaling approach to test for rank-order and fatigue effects in stated preference data." Transportation (1986-1998) **21**(2): 167.

Caparros, A., J. L. Oviedo, et al. (2008). "Would You Choose Your Preferred Option Comparing Choice and Recoded Ranking Experiments." American Journal of Agricultural Economics **90**(3): 843-843.

Caussade, S., J. d. D. Ortúzar, et al. (2005). "Assessing the influence of design dimensions on stated choice experiment estimates." Transportation Research Part B: Methodological **39**(7): 621-640.

DeShazo, J. R. and G. Fermo (2002). "Designing Choice Sets for Stated Preference Methods: The Effects of Complexity on Choice Consistency." JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT **44**(1): 123-143.

Ferrini, S. and R. Scarpa (2007). "Designs with a priori information for nonmarket valuation with choice experiments: A Monte Carlo study." JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT **53**(3): 342-363.

Hanemann, W. M. (1991). "Willingness to pay and willingness to accept: how much can they differ?" The American Economic Review **81**(3): 635-647.

Hausman, J. A. and P. A. Ruud (1987). "Specifying and testing econometric models for rank-ordered data." JOURNAL OF ECONOMETRICS **34**(1–2): 83-104.

Heiner, R. A. (1983). "The Origin of Predictable Behavior." The American Economic Review **73**(4): 560-595.

Hess, S. and K. W. Axhausen (2005). Distributional assumptions in the representation of random taste heterogeneity.

Hess, S., J. M. Rose, et al. (2008). "Asymmetric preference formation in willingness to pay estimates in discrete choice models." Transportation Research Part E: Logistics and Transportation Review **44**(5): 847-863.

Institute of Transport and Logistics Studies (2007). Ngene – A Software Capability to Design and Generate Choice Experiments, The University of Sydney.

Kahneman, D., J. L. Knetsch, et al. (1991). "Anomalies: The endowment effect, loss aversion, and status quo bias." The Journal of Economic Perspectives **5**(1): 193-206.

Lancsar, E. and J. Louviere (2008). "Estimating individual level discrete choice models and welfare measures using best-worst choice experiments and sequential best-worst MNL." University of Technology, Centre for the Study of Choice (Censoc): 08-004.

Lanz, B., A. Provins, et al. (2009). Investigating willingness to pay—willingness to accept asymmetry in choice experiments.

Louviere, J. J. (2004). "Complex statistical choice models: Are the assumptions true, and if not, what are the consequences?" Advancing the Methodology of Discrete Choice Experiments in Health Economics, Oxford University, September.

Louviere, J. J., T. Islam, et al. (2008). "Designing Discrete Choice Experiments: Do Optimal Designs Come at a Price?" Journal of Consumer Research **35**(2): 360-375.

Louviere, J. J., D. Street, et al. (2008). "Modeling the choices of individual decision-makers by combining efficient choice experiment designs with extra preference information." Journal of choice modelling **1**(1): 128-163.

Marley, A. A. J. and J. J. Louviere (2005). "Some probabilistic models of best, worst, and best-worst choices." Journal of Mathematical Psychology **49**(6): 464-480.

Marsh, D., L. Mkwara, et al. (2011). "Do Respondents' Perceptions of the Status Quo Matter in Non-Market Valuation with Choice Experiments? An Application to New Zealand Freshwater Streams." Sustainability **3**(9): 1593-1615.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. Frontiers in Econometrics. P. Zarembka. New York, Academic Press.

Rose, J. M. and I. R. Black (2006). "Means matter, but variance matter too: Decomposing response latency influences on variance heterogeneity in stated preference experiments." Marketing Letters **17**(4): 295-310.

Scarpa, R., S. Ferrini, et al. (2005). Performance of error component models for status-quo effects in choice experiments. Applications of simulation methods in environmental and resource economics. R. Scarpa and A. Alberini, Springer Publisher, Dordrecht, The Netherlands**:** 247-273.

Scarpa, R., T. J. Gilbride, et al. (2009). "Modelling attribute non-attendance in choice experiments for rural landscape valuation." European Review of Agricultural Economics **36**(2): 151-174.

Scarpa, R., S. Notaro, et al. (2011). "Exploring Scale Effects of Best/Worst Rank Ordered Choice Data to Estimate Benefits of Tourism in Alpine Grazing Commons." American Journal of Agricultural Economics **93**(3): 809-824.

Scarpa, R. and J. M. Rose (2008). "Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why*." Australian Journal of Agricultural and Resource Economics **52**(3): 253-282.

Swait, J. and W. Adamowicz (2001). "The influence of task complexity on consumer choice: a latent class model of decision strategy switching." Journal of Consumer Research **28**(1): 135-148.

Tait, P. and R. Baskaran (2011). "Incorporating Local Water Quality in Welfare Measures of Agri-environmental Policy: A Choice Modelling Approach Employing GIS." Proceedings of EAAE 2011 Congress "Change and Uncertainty", Zurich, Switzerland.

Tyebjee, T. T. (1979). "Response time, conflict, and involvement in brand choice." Journal of Consumer Research: 295-304.

Wedderburn, L, D. Bewsell, P. Blackett, M. Brown, S. Kelly, M. Mackay, K. Maani, O. Montes and T. Payne (2011) "Developing a preferred approach for managing the cumulative effects of land uses on freshwater quality". Report prepared for P21 environment programme and Environment Canterbury

[i] Environment Canterbury Regional Council (2010, 19 July) Potential moratorium on the Hurunui River and its tributaries [Letter to the Minister for the Environment]. Downloaded from http://beehive.govt.nz/sites/all/files/nz_19_July_letter_to_Minister_for_the_Environment.pdf