The Non-Market Value of Biodiversity Enhancement in New Zealand's Planted Forests

A thesis

submitted in fulfilment

of the requirements for the degree

of

Doctor of Philosophy in Economics

at

The University of Waikato

by

RICHARD YAO

THE UNIVERSITY OF

WAIKATO

Te Whare Wānanga o Waikato

2012

Abstract

This study investigates the non-market value of biodiversity enhancement in New Zealand's planted forests using the stated choice experiments (CE) approach. This study focuses on two issues. One issue is policy orientated where we estimate the non-market value of biodiversity enhancement and the determinants of this value. The other issue is about the neutrality of major experimental design criteria used in CE. Specifically, we examine the impact of using different criteria on attribute non-attendance, choice variability, choice determinism and learning.

To estimate the non-market value of biodiversity enhancement, a random parameters logit model with error components is used to analyse choice data collected from 209 respondents across New Zealand. The panel nature of the choice data set is exploited to calculate the marginal willingness-to-pay (WTP) for environmental attributes of each respondent. Panel random-effects regression models are subsequently employed to determine the factors that influence individual-specific WTP values. Results suggest that New Zealand taxpayers would be willing to pay $26.5 million per year for five years for a proposed biodiversity enhancement programme. Random effects regression analysis suggest that respondents living close to large planted forests (i.e., less than 10 kilometres away) would pay more for the programme.

To study whether the selection of experimental design criterion affects attribute non-attendance and choice variability, we analyse a balanced sample with split designs. The balanced sample is composed of 1509 choice observations equally distributed across three experimental designs, namely: orthogonal, Bayesian D-efficient and optimal orthogonal. Results from latent class logit

analysis suggest that tasks derived from the Bayesian D-efficient design (BDD) criterion are more attended than those derived from orthogonal and optimal orthogonal designs. Heteroskedastic logit analysis indicates that, unlike the two other designs, higher choice task complexity (as measured by entropy proxies) in the BDD does not increase choice variability of respondents. This is indicated by the absence of a significant increase in the variance of the Gumbel error in the choice data collected using BDD unlike the data collected using the two other criteria.

To study whether the three experimental designs vary in terms of choice determinism and task order effects, a separate analysis of the balanced data set using heteroskedastic logit models is undertaken. Results show that higher levels of choice task complexity (as measured by attribute dispersion proxies) in BDD contribute to increasing choice determinism of respondents but not in the orthogonal design. Choice data collected using BDD choice tasks exhibit a steady learning effect, unlike the other designs which do not exhibit any form of continuous learning.

We conclude that the BDD criterion provides choice tasks that are superior compared to the other two design criteria. Choice data collected using this criterion has a higher quality as indicated by more attended choice tasks, lower choice variability and a pattern of continuous learning. These results point to a higher behavioural efficiency of respondents in evaluating complex choice tasks. However, these results might be specific to the choice data collected in this current study. We suggest that future studies should further investigate the impacts of different experimental designs to verify the findings of this study.

Acknowledgements

The completion of my PhD program was surely a long, tough, and challenging journey but at the same time a very productive and pleasant experience.

I would like to express my profound gratitude to my chief supervisor, Professor Riccardo Scarpa, for his excellent mentoring and valuable comments throughout the conduct of my research. I thank Dr. Pam Kaval for bringing me to New Zealand and for guiding me during the early stage of my PhD programme. I appreciate the support and encouragement provided by Professor Mark Holmes, Dr. Dan Marsh, Mrs. Bridget Daldy and Dr. Sayeeda Bano during my PhD years in the Department of Economics. I thank Ms. Maria Fitzgerald and Mrs. Leonie Pope for providing valuable assistance during the time I have been in the department.

I acknowledge Scion and Future Forest Research for providing for my research expenses and tuition fees. I thank my economic survey assistants for recruiting my survey respondents across New Zealand. I also thank my survey respondents and focus group participants for their time and valuable inputs to my research.

To my managers and fellow staff members at Scion, namely: Tim$_1$, Tim$_2$, Peter, Brian, Trevor, James, Lisa, Andy, Brian, Luke, Duncan, Anne, Brenda, Greg, Dave, Dean, Marie, Alyson, Dina, Paulyn, thank you for your never-ending support and encouragement for me to finish my thesis.

I thank Dr. João Palma and Duncan Harrison, for helping me construct the geo-spatial data set that I analysed for Chapter 3 of this thesis.

I also thank several Scion staff members as well as Rotorua-based forest managers and conservationists, especially Dr. Lyndsay Bulman, Dr. Lisa Berndt, Dr. Brenda Baillie, Greg Steward, Colin Maunder, Debbie Stewart, Noel Hyde and Dr. Richard Seaton, for all those very useful pointers about New Zealand's forest biodiversity.

To my fellow students, especially Tinh, Ramelan, Lena, Rey, Ariel, Manny, Joy, thanks for being a part of my campus life at the University of Waikato. I consider you all as my family and the University as my second home.

To Papa, Mama, Ate Lilian, Kuya Roland, Kuya Glenn – my dear parents, sister and brothers – thank you for always being there for me.

To Leah, Maui and Reign, my beloved wife and daughters, you always serve as my great inspirations.

And to our Lord God for being the source of my strength, seeing me and my family through all life's struggles, and blessing us always with moments of joys and victories to carry us through each day.

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Abbreviations

AD: Advance design

ANA: Attribute non-attendance

ASD: Average standard deviation of attribute levels across alternatives in a choice task

AVC matrix: Asymptotic variance-covariance matrix

BDD: Bayesian D-efficient design

CE: Choice experiments

CL: Conditional logit

DOC: Department of Conservation

DSD: Dispersion of standard deviation (of attribute levels across alternatives in a choice task)

DSF: DeShazo and Fermo (2002)

ED: Experimental design

FIM: Fisher information matrix

NZ: New Zealand

OOD: Optimal orthogonal design

ORD: Orthogonal design

PWLC: Piecewise linear coding

RPL: Random parameters logit model

RPPanel: Random parameters logit model with panel specification

RPECPanel: Random parameters logit model with error components with panel specification

RUM: Random utility maximization

PLCM: Panel latent class model

UB: Utility-balance

WTP: Willingness-to-pay

YK: Yao and Kaval (2010)

Chapter 1: Introduction

1.1 Overview

Biodiversity conservation is essential to human existence (Pimentel, et al., 1992; Folke, et al., 1996). Biodiversity services provided by planted forests (e.g. habitat provision to threatened native species) have gained increasing ecological importance (Carnus et al., 2006; FAO, 2010; Humprey et al., 2003) and have been found to have a high economic value (Scarpa, 2003; Hanley et al., 2002; Bienabe and Hearne, 2006). In the case of New Zealand, although policy makers are keen to include biodiversity values into forest management, identifying the management options that would provide the greatest biodiversity benefit to society remains a big challenge (Maunder, 2008). This is because the value that New Zealanders place on biodiversity enhancement in planted forests remains unclear and therefore unlikely to be included in policy decision making. However, with the development of economic valuation techniques (i.e. Choice Experiment), complex biodiversity values can be estimated. Choice Experiment (CE) can be used to investigate the preferences and willingness-to-pay values of an individual on the changes in biodiversity outcomes in planted forests. A crucial component in CE is the selection of a criterion to construct the Experimental Design (ED) of the choice questions in a survey instrument. ED refers to the systematic arrangement of the changes in the levels of attributes of an environmental good presented to respondents in a series of choice situations called choice sets (or choice tasks). In generating the ED, the analyst must select what statistical properties the design should exhibit, such as orthogonality or statistical efficiency.

As there are a number of ED criteria (e.g., orthogonal design criterion, D-efficient design criterion, utility-balanced design criterion) that have been developed for CE, the choice analyst is faced with the decision to select a design criterion that would support his/her objective. For example, an analyst who would like to minimize standard errors of coefficients estimates would likely choose a D-efficient design, whilst an analyst who aims to generate a design with alternatives that have equally likely chance of being selected would elect to have a utility-balanced design.

Despite the fact that different design criteria are now being used in CE, many CE studies to date assume that the selection of the ED criterion is neutral to the estimated parameters (e.g., coefficients and scale parameters). In other words, many recent CE studies continue to assume that different design criteria would likely have the same effect on the parameter estimates. This may be because many choice analysts assume that the parameter estimates are mainly influenced by the preferences of respondents. However, given that different ED criteria have different objectives (e.g., higher statistical efficiency, utility balance) they could also have a systematic effect on the choice tasks generated that could influence the mean and variance of the estimated parameters. If this systematic effect existed and was not taken into account, parameter estimates would likely be biased.

As described above, it is important to identify the best options for managing biodiversity in planted forests so as to benefit society. This is done here by examining the preference and values of a sample of potential biodiversity enhancement beneficiaries using CE. In using CE, it is important to investigate whether the selection of the experimental design criterion would be likely to influence the estimation of the parameters. This thesis aims to answer two general questions:

(1) Would New Zealanders be willing to pay for a biodiversity enhancement programme in the country's planted forests? If so, how much would this amount be?

(2) Are the most commonly employed selection criteria for experimental design neutral to parameter estimates? If not, how does one criterion differ from the other in terms of behavioural efficiency?

The questions above are addressed by studying the preferences and willingness-to-pay (WTP) values of New Zealanders using the CE exercise that employed different experimental design criteria. Between November 2009 and August 2010, 209 respondents sampled across New Zealand completed a choice experiment questionnaire through phone-mail survey and phone-internet survey. The choice questions collected data on stated preferences for a proposed biodiversity enhancement programme in New Zealand's planted forests. WTP estimates were then derived from random utility models estimated on such data.

To estimate the WTP for a biodiversity programme that aims to enhance the habitat for threatened native species in planted forests, random parameters logit models are used to analyse the *full sample* of data. Within it five different experimental designs are used. The full sample has 209 survey respondents each of whom provided preferred options from nine choice tasks. The random parameters logit model with error components was used to provide estimates that were used to simulate the median and variance WTP for each choice attribute.

Aggregated WTP values under different scenarios, accounting for different sources of bias (e.g., hypothetical, aggregation biases), are also estimated to show the national value of the proposed biodiversity enhancement programme.

To study whether the selection of an experimental design criterion is neutral, we analyse a *balanced sample* with split designs. The balanced sample is a subset of the full sample mentioned above. The data is composed of 1509 choice observations equally distributed across three experimental designs, obtained from the full factorial using three separate criteria: orthogonality, Bayesian D-efficiency and optimised orthogonality. Heteroskedastic logit models are used to examine the effect of each design on the choice behaviour of respondents. We explore whether choice variability of respondents are affected differently by higher complexity levels of choice tasks derived from the three criteria. We employ two methods in examining the effect of choice task complexity. The first method is based on *entropy* as a measure of choice task complexity which is proposed by Swait and Adamowicz (2001a). The second method is based on another choice task specific measure that we call *attribute dispersion* which is described in DeShazo and Fermo (2002). By also using the Heteroskedastic logit approach, we examine whether the ED criteria have different learning effects along the sequence of choices made by each respondent. In addition to choice complexity and learning effects, we also compare the three design criteria in terms of attribute non-attendance as described in Scarpa et al. (2009, 2010) where we employ latent class logit models in the analysis.

Results in Chapter 3 suggest that the proposed biodiversity enhancement programme is highly valued by New Zealand taxpayers. Our calculations show that taxpayers had an aggregate WTP value of NZ$26 million per year for the proposed five-year programme. The calculation of the aggregate WTP value accounts for different sources of biases which include aggregation and hypothetical biases.

Findings on the comparison of different ED criteria are presented in Chapters 4 and 5. In Chapter 4, we report that attribute non-attendance is found to be different across designs when using an approach based on latent classes, with choice tasks derived from the Bayesian D-efficient design (BDD) being more attended to relative to the other two designs. Similarly, using the heteroskedastic logit models, we find that higher levels of choice task complexity affect variance heterogeneity differently across designs. The null of design entropy affecting error variance fails to be rejected only by the BDD, which is hence deemed to be comparatively superior. In Chapter 5, we report our findings that different ED criteria vary in terms of the effect of attribute dispersion and learning on the scale factor. We conclude that in our case BDD provided behaviourally more efficient choice tasks. We also recommend that further investigation should be done on the impacts of different ED criteria on behavioural efficiency.

1.2 Background and research questions

New Zealand's 1.8 million ha of planted forests account for approximately 7% of the country's land area. These forests consist mainly of exotic trees such as radiata pine, Douglas-fir and eucalyptus. They provide habitat for at least 118 threatened native animals and plants (Pawson, et al., 2010; Brockerhoff, et al., 2008). Studies also suggest that habitat for threatened native species can be enhanced through forest management (Humpreys et al., 2003; Carnus, et al., 2006; Seaton, et al., 2010; Maunder et al., 2005). While there is evidence that a typical New Zealander would be willing to pay to support a biodiversity enhancement programme on private land by planting more native trees (Yao and Kaval, 2010), it is not known if this also holds for biodiversity in privately owned exotic planted forests. This motivates us to ask the first research question or RQ1:

RQ1. Would New Zealanders be willing to pay for biodiversity enhancement in planted forests, and if so, approximately how much?

The answer to this question is the subject in Chapter 3 of this thesis, where we investigate whether a typical New Zealand taxpayer would be willing to financially support a proposed biodiversity enhancement programme.

There are several criteria for generating a choice experimental design. Three design criteria are examined in this thesis. The first is the orthogonality criterion (the most frequently employed design criterion used in the beginning of the CE literature) which constrains the correlation between attribute levels to zero. This criterion is also efficient when the model to be used in the data analysis is the multivariate linear regression model. The second criterion is the optimised orthogonality criterion, which selects the design from the various orthogonal designs available assuming that a non-linear regression model (e.g., logit) will be used in analysing the data collected, and hence based on the properties of this model's asymptotic variance covariance matrix. While the orthogonal design criterion does not make any assumption on the coefficient estimates, the optimal orthogonal criteria generate designs with the assumption that the model coefficient estimates are all zero, according to the approach by Street and Burgess (2005). Thus, the two design criteria basically assume that the contribution to the indirect utility of the choice attributes is zero. We therefore classify orthogonal and optimal orthogonal designs to belong to the group called *utility neutral designs.* The third criterion is the Bayesian D-efficiency criterion, which assumes that the coefficient estimates are non-zeroes. To generate an experimental design, the Bayesian D-efficiency criterion requires some prior information about the mean and distribution of coefficient values to be estimated, which may come either from a pilot survey, from related previous study, or from some experts' opinions.

In a statistical sense, using a variety of indicators, experimental designs derived by means of the Bayesian efficiency criterion have been proven to outperform choice experimental designs that did not take into account prior information (Kessels et al., 2006; Bliemer and Rose, 2011; Bliemer and Rose, 2009; Vermuelen et al., 2011; Scarpa et al., 2009). However, in a behavioural sense, we still have very limited empirical studies to show how Bayesian efficient designs compare with utility neutral designs.

A recent study by Louviere et al. (2008) suggests that EDs with "higher statistical efficiency" resulted in less consistent choice responses. Louviere et al. compared 44 different designs (40 optimal orthogonal and 4 adaptive) with varying levels of efficiency following the "D-efficiency" measure described in Street and Burgess (2007). Their results suggest that responses to choice tasks were systematically less consistent as statistical efficiency increased. However, in the study by Louviere, designs that were compared were not derived from using *a priori* information on coefficient values.

Swait and Adamowicz (2001a, 2001b) and DeShazo and Fermo (2002) modelled the effect of choice task complexity on choice variability by examining the effect on the variance of the error component, but these studies both used choice observations derived from choice tasks with varying numbers of alternatives and attribute levels that were all generated from designs derived using the orthogonal criterion.

Hess et al. (2008) compared three different designs: orthogonal with random blocking, orthogonal with blocking and a D-efficient design. Their findings indicate that the D-efficient design performed only slightly better (in terms of behavioural efficiency) than the orthogonal design with blocking. Although the D-efficient design used *a priori* information, the design treated the *a priori* estimates with perfect certainty. In contrast, the Bayesian D-efficiency criterion accounts for the uncertainty of the given prior distribution of parameters (see Ferrini and Scarpa, 2007).

Huber and Zwerina (1996) suggest that the utility-balanced design increases statistical efficiency of the design that could lead to a reduction in the theoretically minimum number of respondents needed to estimate a basic conditional logit model. To investigate the impact of utility-balanced design on choice behaviour, Viney et al. (2005) empirically investigated three different designs – utility-balanced, orthogonal, and random designs. Their results suggest that choice tasks derived from utility-balanced designs yielded data with greater inconsistency or random variability in responses compared to the other two designs.

While the abovementioned studies have compared the behavioural impacts of different experimental design criteria, none of them compared the behavioural effects of two particular experimental design criteria (i.e., *with Bayesian a priori* versus *without a priori* information) on choice behaviour. This is probably the first study to compare the impacts of these two experimental design criteria on different aspects of choice behaviour.

Going back to the study of Viney et al. (2005) as mentioned above, where they found greater random variability of utility balanced design, a possible cause of greater variability is greater choice complexity. Swait and Adamowicz (2001a) suggest that the complexity of a choice task can be represented by entropy. Entropy is a choice task specific measure of complexity where the theoretically maximum entropy is achieved when each of the three alternatives in a choice task had an equally likely chance of being selected. A choice task with alternatives that have an equally likely chance of being selected has alternatives that are identical in utility terms. Given that different experimental design criteria would lead to different systematic arrangements of attribute levels in a choice task, this might contribute to differences in entropy levels between different designs. This study aims to answer the second set of research questions:

RQ2. Do different experimental designs differ in entropy levels? If so, would a higher entropy level have the same effect on choice variability across designs?

Scarpa et al. (2009, 2010) provide evidence that CE respondents may tend not to process all the attributes that are used to illustrate the choice alternatives. Their results suggest that accounting for attribute non-attendance in choice analysis results in significant improvement in model goodness of fit and higher efficiency of coefficient estimates. As different designs would likely have different choice tasks generated, we ask the third set of research questions:

RQ3. Does the selection of experimental design criterion influence attribute non-attendance? If so, what are the effects on the parameter estimates and WTP values?

DeShazo and Fermo (2002) have shown two characteristics of choice tasks that contribute to choice complexity giving a respondent greater cognitive burden in selecting the preferred alternative in a choice task. These characteristics are the average standard deviation between attribute levels across alternatives in a choice task and the dispersion of standard deviation of attribute levels across alternatives in a choice task. We collectively call these two choice task specific characteristics *attribute dispersion*. The higher the attribute dispersion, the greater the complexity; this corresponds to providing greater cognitive burden to respondents. It is important to note that attribute dispersion is another measure of choice task complexity and the calculation of this measure is different to the entropy measure described in Swait and Adamowicz (2001a). This leads to the fourth research question:

RQ4. Is there a relationship between the variance of the attributes and the Gumbel error variance?

Several CE studies have shown that the ordering of choice tasks influences the estimation of indirect marginal utility and the the Gumbel error variance (Caussade et al., 2005; Holmes and Boyle, 2005; van der Waerden et al., 2006; Kjær et al., 2006; Day and Pinto-Prades, 2010; Day et al., 2010). To illustrate this, as a respondent answers a sequence of choice sets (e.g., 1^{st}, 2^{nd},… …, 9^{th}), the first choice set would likely involve the highest degree of lack of familiarity with the choice context. As a respondent selects the preferred alternatives from the second, third and fourth choice sets, he/she would likely find them easier to select than earlier ones because of the learning effects. Selecting from the latter choice sets (7^{th}, 8^{th} and 9^{th}) could tend to make a respondent tired or start to experience fatigue. This leads us to the fifth research question:

RQ5. Do different choice experimental design criteria generate sets of choice tasks with different learning/fatigue effects?

1.3 Structure of the thesis

This thesis has five chapters. Chapter 1 (this chapter) motivates the study and provides an overview of the thesis and its overall objectives. It presents the five sets of research questions and discusses some things that can be expected from this thesis. Chapter 2 provides an overview of choice experiments, the choice models used for the analyses, the design efficiency measures and the choice data sets analysed in Chapters 3, 4 and 5.

Chapter 3 is the first main chapter which is policy orientated. We present the estimates of WTP values using the collected choice data set that we analysed using logit models. This chapter also describes how individual willingness-to-pay values are aggregated to represent a national level value for a proposed biodiversity enhancement programme in planted forests.

To answer research questions 2 and 3 (i.e., RQ2 and RQ3), Chapter 4 focuses on studying impacts of experimental design selection in terms of entropy as a complexity measure and attribute non-attendance as a behavioural response to complexity. We calculate the entropy measure for each design and compare how entropy levels affect choice variability across designs using heteroskedastic logit models. For analysing attribute non-attendance, we use latent class logit models where we identified the different groups of respondents following different attribute processing patterns based on reported non-attending behaviour. We also calculate the WTP value for each attribute and compare the differences in WTP values between the three design treatments.

To further investigate if the selection of criteria for choice experiment design matters, Chapter 5 presents results from an examination of the effects of higher attribute dispersion and learning on the variance of the Gumbel error terms across designs. Heteroskedastic logit models are used to examine the differences in attribute dispersion and learning/fatigue effects across designs. Chapter 6 summarises the results. Policy recommendations are formulated and future research directions are suggested.

Chapter 2: Methods, designs and data

This chapter starts by providing a description of choice experiments and the econometric models used in Chapters 3, 4 and 5 of this thesis. This is followed by the experimental designs used in this study and their evaluations in terms of statistical measures of design efficiency. We also describe here the sampling strategies and how well these sampling strategies were achieved. Finally, the two types of data sets (i.e., full sample and balanced sub-sample) used in this study are described. This chapter provides an overview of the key elements of the thesis and these will be referred to in the next three chapters.

2.1 Choice experiments for biodiversity valuation in planted forest

2.1.1 Overview of stated choice experiments

Stated choice experiments (CEs) are conducted in the field of environmental economics to obtain data on the hypothetical behaviour of individuals in regard to the changes in the provision of environmental goods and calculate measures of values on the changes of attribute levels such as willingness-to-pay.[1] In a CE survey, a respondent is provided with a series of choice tasks that leads to the collection of a panel of choice responses. Each choice task contains a set of alternatives that may include a status quo alternative (with attributes at their current levels of provision) and hypothetical alternatives (including current and improved attribute levels) constructed from an experimental design (ED). Each alternative is described by several attributes of relevance to the respondent that

---

[1] Choice experiments in environmental economics are also called as Attribute-based methods. This is described in Holmes and Adamowicz (2003). For further details about choice experimental design criteria, one may read Ferrini and Scarpa (2007) and Scarpa and Rose (2008).

include environmental attributes and a cost for each option in the choice task. When a respondent selects the preferred alternative (from among the three alternatives), he or she implicitly makes trade-offs between the levels of attributes in all the alternatives shown in a choice task.

2.1.2 The choice model

The Random Utility Maximization (RUM) model proposed by McFadden (1974) provides the standard framework for modelling the choice behaviour of an individual. Under the RUM framework, an individual evaluates different alternatives in a choice task and selects the one that provides the highest expected utility level. To illustrate this, we first describe the structure of the utility function that has deterministic and stochastic components as modelled by the basic conditional logit model. The analyst aims to estimate a $1 \times K$ row of utility weights or utility coefficients β for a column of vector X of $K \times 1$ attributes for respondent n's indirect utility function V_i. The estimation is based on data showing respondents' chosen alternative among the set of J competing alternatives presented in choice task s. For this exercise, each respondent was presented with nine choice tasks (S = 9). As shown in Figure 3.2 (on page 68), each choice task has three alternatives (J = 3); with one representing the status quo (*sq*) or the current condition identified based on expert opinion and facts from the environmental literature, while the other two are changed alternatives (a_1, a_2) composed of different combinations of attribute levels generated using an experimental design.

We represent the utility perceived by respondent n from selecting alternative j in choice task s as U_{njs}. Based on random utility theory, utility has two components: the observed indirect utility V_{njs} and the unobserved error component ε_{njs}. V_{njs} is associated with the satisfaction derived by respondent *n* from the changes in attribute levels, ε_{njs} represents the stochastic component of utility that is independently and identically Extreme Value Type I (or Gumbel) distributed across the alternatives. The utility function can be shown as

$$U_{njs} = V_{njs} + \varepsilon_{njs} \tag{2.1}$$

The deterministic component V_{njs} is specified to be linear in parameters (i.e., $V_{njs} = \beta' X_{njs}$) where X_{njs} is a vector of observed variables relating to alternative *j*. The conditional logit probabilities can be specified with Gumbel error scale $\lambda > 0$:

$$P_{nls} = \frac{\exp\left(\lambda\left(\beta' X_{nls}\right)\right)}{\sum_{j=1}^{J} \exp\left(\lambda\left(\beta' X_{njs}\right)\right)},\ J\text{=1,2,3} \tag{2.2}$$

where P_{nls} represents the probability that alternative *l* will be selected by respondent *n* from the set of *J* alternatives shown on choice task *s*. The values of X_{njs} are defined by the experimental design. An efficient design is expected to maximise the amount of information the design conveys to identify the estimates for the vector of marginal utilities, β. The information matrix for the design assuming a conditional logit model is defined by the matrix of second derivatives of the log-likelihood function presented as

$$I\left(\beta, X_{njs}\right) = \frac{\partial^2 \ln L}{\partial\beta\partial\beta'} = \sum_{n=1}^{N}\sum_{j=1}^{J}\sum_{s=1}^{S} P_{njs}\left(X_{njs} - \bar{X}_{njs}\right)\left(X_{njs} - \bar{X}_{njs}\right)' \tag{2.3}$$

where $\overline{X}_{njs} \equiv \sum_{j=1}^{J} P_{njs} X_{njs}$

The $I(\beta, X_{njs})$ matrix above that has a dimension of $K \times K$ represents the Fisher Information Matrix (FIM). FIM is a measure of the amount of information that observable sources of utility X_{njs} carry about β where choice probabilities depend upon.

The popularity of the conditional logit model can be attributed to the fact that its set of choice probabilities takes a closed form (Train, 2009). This refers to the simple mathematical formulation of the Jacobian (vector of first derivatives of the Log-likelihood function) and the Hessian (matrix of second derivatives of the Log-likelihood function). As these two matrices are functions of utility coefficients β and the experimental design X_{njs}, an experimental design that increases the magnitude of the elements in $I(\beta, X_{njs})$ with respect to a baseline design is therefore a more informative design. It is important to note that the negative of the inverse of the expected FIM is the maximum likelihood estimator of the asymptotic variance-covariance (AVC) matrix that can be shown as

$$AVC = \Omega(\beta, X_{njs}) = \left[E\left[I(\beta, X_{njs})\right]\right]^{-1} = \left[\frac{\partial^2 \ln L}{\partial\beta\partial\beta'}\right]^{-1} \tag{2.4}$$

where $\ln L$ is the log-likelihood of design X_{njs}:

$$\ln L = \sum_{n=1}^{N} \sum_{j=1}^{J} \sum_{s=1}^{S} Y_{njs} \ln P_{njs}\left(X_{njs}, \beta\right) \tag{2.5}$$

and Y_{njs} represents the indicator of choice that takes the value of 1 if chosen or 0 otherwise. The diagonal and off-diagonal elements of *AVC* represent, respectively, the variance and covariance of estimated coefficients β. The smaller the elements of *AVC* of the design, the more efficient the design is. A good criterion for choosing an efficient design is the one that minimises the determinant of the *AVC* matrix.[2] An appropriate algorithm to generate and search for an efficient design would need to generate new designs from an existing coded design matrix, evaluate each design based on efficiency as a function of the arrangement of attribute levels, and identify the generated design that has *AVC* with the lowest determinant.

2.1.3 Measures of choice experimental design efficiency

As the *AVC* matrix of a design contains many elements, it first needs to be transformed into a single number for a straightforward comparison of efficiency between different designs. One single measure of matrix size is the determinant. The determinant of a matrix refers to the summation of the terms, each term representing a product of systematically selected elements of a square matrix. For a square matrix to have a non-zero determinant (or non-singular), it should be full rank which implies that matrix columns are independent or not collinear. Therefore, the determinant of the *AVC* matrix provides a valid measure of design efficiency. However, as the number of matrix column (K) increases, the determinant also becomes larger. Thus, the determinant should be normalised by K. This determinant of the *AVC* matrix is the D_p*–error* presented formally as

---

[2] As AVC and FIM are inversely related, minimising the determinant of the AVC corresponds to maximising the determinant of FIM.

$$D_p - error = \det\left(AVC\left(\beta, x_{njs}\right)\right)^{1/K} \quad (2.6)$$

Another measure of design efficiency that has been occasionally reported in experimental design literature is the *A-error*.

$$A - error = \frac{trace\left(AVC\left(\beta, x_{njs}\right)\right)}{K} \quad (2.7)$$

A–error is similar to the D_p *–error* in that it uses the *AVC* matrix and also account for the number of columns. However, instead of a determinant, it employs the trace of the *AVC* matrix that only accounts for the diagonal elements (variance) and not off-diagonals (covariance). Not fully accounting for all the elements of the *AVC* matrix might be the reason for the relatively lower acceptance of the *A–error* in the literature.

The D_p*–error* (also called *point D-error* or *local D-optimal*) is not without drawbacks. Under this measure, the values of β are treated with certainty at the experimental design stage. This is not plausible because if the analyst already has good estimates of β, there is no need to generate an efficient design for a choice experiments survey. Usually, the estimates of β are derived from relevant previous studies, expert opinion or from a pilot survey. In this case, estimates of β would likely have a degree of uncertainty. This uncertainty can be accounted for by providing adequate *a priori* distributions (Sandor and Wedel, 2001, 2002, 2005; Ferrini and Scarpa, 2007). This makes the *Bayesian D-error* or D_b*-error* a more attractive design measure because it accounts for the uncertainty where an expectation is taken over the assumed *a priori* distributions of β. Formally, D_b *-error* can be presented as

$$D_b - error = \int \left[\det\left(AVC\left(\beta, x_{njs}\right)\right)\right]^{1/K} N\left(\mu, \Sigma\right) d\beta \qquad (2.8)$$

where the term $N(\mu,\sigma)$ tells that the values of β are *a priori* distributed normal with mean μ and variance-covariance Σ. Imposing a normal distribution for prior β may indicate that reliable *a priori* information has been used (e.g., coming from a pilot survey). However, if priors came from less reliable source due to limited availability of relevant studies or lack of resources to do a pilot survey, a less reliable prior assuming a uniform distribution may be employed leading to $D_{b_u} - error$ shown below

$$D_{b_u} - error = \int \left[\det\left(\Omega\left(\beta, x_{sj}\right)\right)\right]^{1/k} U\left(\mu, \Sigma\right) d\beta \qquad (2.9)$$

As *D*-error and *A*-error are measures of design efficiency, Huber and Zwerina (1996) have shown evidence that utility balance contributes to improving design efficiency particularly if *a priori* parameter estimates are accounted for in the construction of the design. The utility balanced criterion is centred on choice probabilities of the alternatives in a choice task. Given a choice task with three alternatives, if these alternatives are equally attractive to a respondent, then each alternative gets a choice probability of 0.33, resulting to a perfectly utility balanced choice task. If a choice task has an alternative with all attributes being more attractive (e.g., low cost, greater numbers of threatened species sighted) than the other two alternatives, that dominating alternative would get a choice probability of 1.00. This is based on the assumption that respondents' utility levels would increase monotonically with improvements in environmental attributes. A choice task with a dominating alternative would be very easy to respond to, however, the choice data collected would be uninformative to model parameters resulting in less precise estimates of parameters. Huber and Zwerina (1996) indicate that if one out of three alternatives in a choice task is chosen almost all of the time, then this leads to extreme values of the cumulative probability function. In general, choice tasks that generate extreme probabilities (e.g., 1.0 or 0.0) are less effective at constraining the choice model parameters than those generating moderate ones. To ensure that good information from an exercise in choice experiments is obtained, one should minimise (and if possible eliminate) the occurrence of choice sets with dominant alternatives (Krieger and Green, 1991; Huber and Zwerina, 1996).

Dominance in an experimental design can be detected manually by examining each choice task. Alternatively, the Utility-Balanced (*UB*) measure proposed by Kessels et al (2004) can be used to measure the degree of dominance in a design. In contrast to the design efficiency measures mentioned above, that account for the AVC matrix, the *UB* measure focuses on the choice probabilities. *UB* can be presented as

$$UB = \left(\frac{\sum_{s=1}^{S} \left(\prod_{j=1}^{J} P_{nls} \right)}{S\left(1/J \right)^{J}} \right) \times 100\% \qquad (2.10)$$

where l represents the chosen alternative and S is the number of choice tasks in a design. This *UB* measure is expressed in percentage form with 0% indicating that each choice task in the design has a dominant alternative and with 100% suggesting that every alternative in each choice task has an equal chance of being selected. ChoiceMetrics (2011 p. 95) suggests that observed utility balance measure of efficient designs lies between 70 to 90 percent. Viney et al. (2005) compared empirically three different experimental designs: orthogonal main effects, utility-balanced, and random. Their results indicate that although the three experimental designs did not impact the underlying parameter estimates, the utility-balanced design yielded greater random variability in the responses. A possible reason for the increase in variability could be as a result of a higher level of choice complexity because choosing an alternative from a choice task with equally attractive alternatives would require more effort that could lead respondents to make different choices (Swait and Adamowicz, 2001a, 2001b). The variability of responses can be modelled by parameterising the scale of the coefficients of the indirect utility function under the heteroskedastic logit framework that is discussed in the next section.

2.1.4 Heteroskedastic logit model

The cumulative distribution function (cdf) of an individual error component of the Conditional Logit (CL) model in Equation (2.1) can be presented as

$$F\left(\varepsilon_{nj}\right) = \exp\left(-\exp\left(-\varepsilon_{nj}\lambda\right)\right) \quad , \quad -\infty < \varepsilon_{nj} < \infty \quad , \quad \lambda > 0 \qquad (2.11)$$

where λ represents the scale parameter. The above cdf suggests that the variance of ε_{nj} is $\sigma^2 = \pi^2 / 6\lambda^2$ (Ben Akiva and Lerman, 1995). This follows a set of assumptions that gives rise to CL as presented in Equation (2.2), where λ is a scalar constant that allows scale to vary based on factors that would likely influence the variance of the error component σ (or the unobserved component of utility). The conditional logit model assumes that the error variance is constant across individuals. However, the assumption of constant error variance has been questioned in several papers (Hensher et al. 1999; Louviere, 2001; Swait and Adamowicz, 2001a; Swait and Adamowicz, 2001b DeShazo and Fermo, 2002; Louviere et al., 2002). Accounting for scale differences offers advantages which include getting superior model fits from choice analysis where the analyst can pool and rescale the data set (Ben Akiva and Morikawa, 1990). Since λ cannot be identified, the analyst should estimate the product $\lambda(\beta X)$ based on a reference point (Swait and Louviere, 1993). Given that $\lambda = \sqrt{\pi^2 / 6\sigma^2}$, on one hand as σ approaches infinity, λ approaches zero which makes the CL model allocate equal choice probabilities for all three alternatives. On the other hand, as σ approaches zero, λ approaches infinity leading to a CL model that predicts a probability of one to the alternative that provides the highest systematic utility (Ben-Akiva and Lerman, 1985).

To parameterise the scale parameter, we follow Swait and Adamowicz (2001a) where we use a heteroskedastic logit model to account for *q* factors that would likely influence the scale parameter. This we formally present as

$$\lambda_s = \exp\left(\sum_{q=1}^{Q} \gamma_q C_q\right) \tag{2.12}$$

where *q*-factors may include experimental design, choice complexity, order (or learning) effects and interaction between these factors. γ_q represents an estimate of slope shifter of the scale for the q^{th} factor where a negative sign implies that the factor contributes to a decrease in scale (higher error variance), while a positive sign indicates an increase in scale (lower error variance). As we relaxed the assumption of a constant scale parameter, we also relaxed the assumption that the Gumbel error is independent and identically distributed (i.i.d.).[3] This is because the error variance is no longer constant as we allow it to vary based on *q*-factors that would likely influence λ. This set up allows the simultaneous estimation of the utility coefficients and error variance as a function of *q*-factors.

2.1.5 Latent class model with attribute non-attendance

The success in making CE methods more realistic contributed to an increase in its use for economic valuation (Louviere et al., 2000; Hess and Rose, 2009). A desirable feature of CEs is their ability to place respondents into situations in which they must make trade-offs among multiple attributes of alternatives. However, as an analyst tries to make choice tasks as realistic as possible—e.g. by including the most relevant attributes that were carefully identified from literature reviews, focus group meetings and pre-testing—some respondents may attend only to attributes that they are most interested in and ignore the others. One reason for attribute non-attendance (ANA) is that some respondents may tend to reduce cognitive effort in the evaluation of alternatives by attending only to a subset of attributes. The issue of ANA has been corroborated from empirical evidence drawn from many CE studies in the field of transport, marketing and health, environmental economics and food choice (Swait, 2001; Hensher et al., 2005; Hensher, 2006, 2008, 2010; Swait and Adamowicz, 2001a, 2001b; Fasolo et al., 2007; Islam et al., 2007; McIntosh and Ryan, 2002; Lancsar and Louviere, 2006;

[3] The term i.i.d. implies that the variances associated with a component of random utility expression describing each alternative (capturing all the unobserved influences on choice) are identical, and that the unobserved effects are not correlated between all pairs of alternatives.

Bruschi et al., 2010; Scarpa et al., 2011a; Hensher and Greene, in press). The presence of attribute ANA leads to the violation of the continuity axiom that assumes fully compensatory choice behaviour, which implies that respondents attended to all attributes in a choice task (see Hensher 2006 for details of this axiom). In essence, non-attendance to one or some attributes results in non-compensatory choice behaviour because despite any improvements in the levels of unattended attributes, they will fail to compensate for the worsening in the levels of other attributes (Lockwood, 1996; Spash, 2000; Sælensmine, 2002; Rekola, 2003). Scarpa et al (2009) present some empirical evidence showing the different types of ANA behaviour where some respondents ignored one attribute, others ignored more than one, while a few ignored all attributes (hence made random choices). Their results suggest that accounting for different non-attending behaviour of respondents in choice analysis contributes to a significant improvement in model goodness of fit and more accurate estimates of parameter values. Scarpa et al. (2009) suggest a modelling technique that allows the grouping of respondents (up to a probability) into different latent classes that could represent groupings based on non-attendance to certain subsets of attributes.

We model ANA following the Panel Latent Class Logit Model (PLCM) described in Scarpa, et al. (2009). PLCM can be presented as

$$P_n(Y_n|\boldsymbol{\beta}_c) = P_s(i_1, i_2, \dots, i_S|\boldsymbol{\beta}_c) = \prod_{s=1}^{S} \frac{exp(X_{is}\boldsymbol{\beta}_c)}{\sum_k exp\,(X_{ks}\boldsymbol{\beta}_c)} \qquad (2.13)$$

where c represents non-attendance latent classes, P_n represents the probability of respondent n observing a set of S choices $Y_{n=\{y_1, y_2, \dots, y_S,\}}$ is a product of logits

$\prod_{s=1}^{S} \frac{exp(X_{is}\boldsymbol{\beta}_c)}{\sum_k exp(X_{ks}\boldsymbol{\beta}_c)}$. To obtain the unconditional probability of the panel of choices of respondent *n*, the law of total probability is used. This is by summing up the conditional probabilities over the finite set of membership probabilities, $P(c)$, of the specified ANA classes. The unconditional probability can be expressed as:

$$P_n(Y_n) = \sum_c P(c)\, P_n(i_s|\boldsymbol{\beta}_c) = \sum_c \frac{exp(\alpha_h)}{\sum_c exp(\alpha_c)} \prod_{s=1}^{S} \frac{exp(X_{is}\boldsymbol{\beta}_c)}{\sum_k exp\,(X_{ks}\boldsymbol{\beta}_c)} \qquad (2.14)$$

where α represents class-specific constants indentified by imposing that they sum to zero.

In the PLCM above, ANA is operationalised by allowing individuals to be classified to latent classes with utility coefficients restricted to zero for selected attributes, while unrestricted (non-zero) attributes are assumed to have exactly the same value across classes. For the current study, an example of a latent class would be a group of respondents who attended only to the bird attributes (i.e. falcon and brown kiwi) while ignoring the non-bird attributes (i.e. kokopu, kakabeak and gecko). For this latent class, we constrain the utility coefficients of the non-bird attributes to zero while allowing the bird utility coefficients to vary. We can also include other latent classes such as a class that attended to all attributes and a class that ignored the status quo option. For the class that attended all attributes, all utility coefficients are allowed to vary; while for the class that ignored the status quo option, we restrict the utility coefficient for SQ option to be zero. Suppose the three ANA latent classes above represent the most applicable specification for our sample data, then the statistical fit of the model should significantly increase (relative to the conditional logit model) indicating the presence of non-attendance (suggesting that discontinuous preference exists). For this exercise, to identify the most applicable number of latent classes and types of latent classes (e.g., class ignoring the cost attribute, class ignoring bird attributes) we use the minimum Akaike Information Criterion (AIC) approach (Swait, 1994; Boxall and Adamowicz, 2002). AIC is one of the alternative measures of goodness of fit to pseudo R^2 in non-linear regression models (e.g. conditional logit). Under the conditional logit model, AIC minimizes $-2\ln L + 2k$ where lnL represents the log-likelihood value and k is the number of parameters (Kennedy, 2008). However, as AIC does not account for the number of choice observations N, we elected to use the normalized AIC criterion which can be expressed as AIC/N. Normalised AIC is a relative measure allowing for the comparison of two or more models or model specifications. The smaller the normalised AIC value the better the model fit while accounting for the number of parameters estimated.

2.2 Overview of experimental design criteria used in the study

Experimental design in CE provides a means to construct choice tasks in an efficient way as it can influence the accuracy of WTP estimates (Lusk and Norwood, 2005; Campbell, 2007). The literature on experimental design for CE has progressed significantly over the last two decades. Several experimental design strategies have been developed (Kuhfeld et al., 1994; Huber and Zwerina, 1996; Carlsson and Martinsson, 2003; Street et al., 2005; Johnson et al., 2007; Scarpa and Rose, 2008). Although there are several experimental design criteria, this study focuses on empirically examining three criteria: (1) orthogonal design

(ORD), (2) optimal orthogonal design (OOD), and (3) Bayesian D-efficient design (BDD). We describe each in turn in the following.

2.2.1 Orthogonal design

The first experimental design criterion used for CE was ORD (Louviere and Hensher, 1982; Louviere and Woodworth, 1983). The rationale for constructing ORD is derived from linear multivariate models originally used for analysing treatment effects in biological experiments using linear regression models (e.g., Ordinary Least Squares) (Ferrini and Scarpa, 2007). However, choice data is analysed using non-linear regression models (e.g., logit) to examine changes in utilities, hence orthogonality is not a criterion for statistical efficiency for discrete choice experiments (Train, 2009; Bliemer and Rose, 2006). Kessels, et al (2006) and Bliemer and Rose (2009 p. 21) demonstrate that the statistical efficiency of designs following the ORD criterion are relatively lower compared to the more recently developed class of efficient designs (e.g., BDD) where the estimated parameters are more precise as indicated by lower covariances. Lower covariances correspond to a smaller *D*-error which is the determinant of the asymptotic variance-covariance (AVC) matrix as shown in Equation (2.6).

There are two main approaches to generate orthogonal designs: *sequential* and *simultaneous* (Rose et al. 2008). The sequential approach generates designs with attributes that are uncorrelated within, but not between, alternatives. To construct a sequential ORD, one initially creates an ORD for the first alternative then generates subsequent alternatives by re-arranging the rows of the first alternative. The sequential approach allows the analyst to construct designs with a lesser number of choice tasks compared to a simultaneously generated ORD. One drawback of a sequentially generated ORD is that it may not be appropriate for choice tasks with labelled alternatives that, under the orthogonality criterion, would also require orthogonality between alternatives.

The simultaneous orthogonal approach generates designs with a set of attribute levels that are independent both within and between alternatives. Under the simultaneous approach, all alternatives are constructed at the same time. The advantage of simultaneous orthogonal designs is that they can be used for both labelled and unlabelled choice experiments. Given this advantage, this study employed the simultaneous approach in generating the orthogonal design that was used to generate the choice tasks for collecting the choice data for the ORD sample. The simultaneous ORD was generated using the experimental design software NGENE version 1.02.

2.2.2 Optimal orthogonal design

Street et al. (2001, 2005) propose a design criterion called *Optimal Orthogonal Design* (OOD). Following the OOD criterion, one can generate choice tasks (mainly applicable for unlabelled or generic alternatives) with improved statistical properties compared to the traditional orthogonal design (Street and Burgess, 2005; Rose and Bliemer, 2008). OOD offers a two-fold improvement over ORD. First, respondents are forced to make trade-offs on all attributes of a choice task as all pairs of attributes take different values, an improvement compared to ORD that would likely include two attributes in a choice set having the same level (Rose et al., 2011). Second, OOD takes into account that the model used for the analysis is a non-linear regression model (e.g., conditional logit model) where we analyse utility differences (see Train, 2009). The measure of statistical efficiency for OOD is called *D-efficiency* and is expressed in percentage form where the closer the D-efficiency to 100%, the more efficient the design. The calculation for the D-efficiency is explained in detail in Burgess and Street (2003, 2005), Street and Burgess (2004, 2007), Street et al. (2005) and Rose and Bliemer (2008a). It is important to note that this measure of efficiency is different from the "D-error" in Equation 2.6. For this study, we generated an optimal orthogonal design with D-efficiency rating of 100% using NGENE 1.02.

Similar to ORD, OOD does not use prior information of parameter estimates. This design employs an algorithm that searches through different experimental designs generated, assuming that all parameter estimates from a multinomial logit model are equal to zero (Street and Burgess, 2005; Sandor and Wedel, 2005). Assuming a set of prior parameter estimates to be all equal to zero can be too naïve because an analyst could easily access information about some approximation of parameters from related studies (Huber and Zwerina, 1996; Chaloner and Verdinelli, 1995; Ferrini and Scarpa, 2007; Scarpa and Rose, 2008). One could readily assume the sign of the parameter estimate for the cost attribute to be negative. In addition, assuming that all parameters are equal to zero may be unrealistic because the contribution of attributes to the utility of an individual can be large as attributes were carefully identified by the analyst as those that would likely influence an individual's utility level (Kessels et al., 2006).

2.2.3 Efficient design

Huber and Zwerina (1996) suggest the importance of constructing experimental designs based on prior information that could lead to higher design efficiency (or lower D-error).[4] Their work focused on using priors estimated from pre-test interviews that were treated with exact certainty. This approach was extended by Sandor and Wedel (2001, 2002, 2005) who proposed experimental designs that account for uncertainty in prior information used for design construction in a Bayesian fashion.[5] Several types of Bayesian experimental designs have been developed to suit the needs of analysts. These include Bayesian D-efficient, Bayesian C-efficient and Bayesian S-efficient designs. The Bayesian D-efficient design strategy aims to minimise the standard errors of parameter estimates by using *a priori* information and generating (or updating) a design based on this prior information (Chaloner and Verdinelli, 1995; Ferrini and Scarpa, 2007). The Bayesian C-efficient design aims to reduce the variance of the ratio of the parameters i.e., Willingness-to-Pay (WTP) which is a scale free measure of value (Scarpa and Rose, 2008; Vermuelen et al., 2011; Kerr and Sharp, 2010). This design strategy favours analysts who prefer to have narrower confidence intervals of WTP. The Bayesian S-efficient design criterion minimises the required sample size of the experiment without compromising the accuracy of parameter estimates (Bliemer and Rose, 2005; 2009a; 2009b). This favours analysts who face a limited budget for conducting surveys, by reducing the theoretically minimum required number of respondents. Thus, the Bayesian S-efficient design criterion helps to reduce the cost of conducting choice surveys that traditionally require a large sample of respondents to produce quality model estimates (i.e., significantly high t-ratios).

[4] Gain in statistical efficiency reduces the theoretically minimum sample size and, to a certain extent, allows the reduction in the number of choice tasks, which can be considered advantageous for both analysts and survey respondents.

[5] Construction of Bayesian experimental designs is described in Chaloner and Verdinelli (1995).

The *Bayesian D-efficient design* (BDD) generated and examined in this study assumes a conditional logit model. The design was generated following the criterion of minimising the determinant of the AVC matrix or the D-error. Reduction in D-error could also lead to a decrease in the theoretically minimum required sample size that can be translated into savings in time and money from the perspective of the analyst. Prior information used to generate BDD came from a pilot survey of 35 respondents (randomly drawn from the New Zealand population) who each completed nine choice tasks generated using simultaneous orthogonal design. According to Ferrini and Scarpa (2007), prior information from pilot surveys can be considered reliable and this can be used for improving the efficiency of an existing design. One way to improve the efficiency of an existing design is to employ a sequential survey method where one first collects an initial wave of choice survey data, estimates the model parameters and uses these estimated parameters to update the existing design (Scarpa, Campbell and Hutchinson, 2007). This technique was implemented using NGENE 1.02 to generate the Bayesian D-efficient design to construct the BDD choice tasks. These choice tasks were used to collect the choice data for the BDD sample that we analysed in this present study.

As an aside, there are many other design criteria in addition to orthogonal, optimal orthogonal and efficient designs. These other criteria include adaptive (Toubia et al., 2007; Tilahun et al., 2007), random (Train and Wilson, 2008), choice percentage (Toner et al., 1999; Kanninen, 2002; Johnson et al., 2006), Bayesian A-optimal, G-optimal and V-optimal designs (Kessels et al., 2006).

2.3 Generation and evaluation of the three experimental designs

The three main experimental designs (ORD, BDD and OOD) discussed above were all generated using NGENE version 1.02. ORD was generated using the simultaneous approach. This simultaneous orthogonal design was used in the first batch of survey where we collected 35 completed surveys (9 choice sets × 24 respondents = 216 choice observations). From these choice responses, we constructed a choice data set following the dummy coding procedure described in Hensher et al. (2005).[6] This data set was used to estimate the coefficients and standard errors of a conditional logit model. These estimates were used as *a priori* information for the generation of the three Bayesian efficient designs (D-, S- and C-efficient designs). For the generation of OOD, we did not use those coefficients because this criterion assumes that beta coefficients are all zeroes.

Equations 2.6 to 2.10 show the different experimental design measures which include D-error and Utility-balanced measures. We know that the lower the D-error, the more statistically efficient the design becomes. Utility balance was considered important by Huber and Zwerina (1996) which suggested that the more utility balanced the design is, the higher the quality of information we collect from respondents. We use these design measures to evaluate the three designs that we compare in Chapters 4 and 5 and to assess empirically the claim made by the proponents.

OOD is an experimental design criterion that minimises D-efficiency measure assuming that we do not have prior information about the parameter estimates. In this case, Table 2.1 shows that the D_z-and A_z errors for OOD are

[6] We also describe how we implemented dummy coding on Page 70 of this thesis.

lower (or better) than those of BDD and ORD. If we evaluate the designs assuming that we have fixed prior information, BDD is the most efficient design with a D_p-error of 0.213 whilst OOD becomes the worst design with a D_p-error twice as that of BDD. The Bayesian D-efficiency measure also accounts for prior information but it also accounts for a degree of uncertainty around the prior information used for updating the design. The D_b-error is slightly higher for BDD (0.22) but four times as much for OOD (0.94). This is not surprising because the OOD criterion maximises the differences between alternatives in choice tasks and this could lead to a higher determinant of the AVC matrix, thus leading to a higher D_b-error. Table 2.1 also shows the effects of optimising the three updated designs (ORD, BDD and OOD) following the BDD criterion using *a priori* information from larger sample sizes.

This three EDs that we are comparing here were all generated using NGENE 1.02. Before using the BDD generated from NGENE, we have checked first for the presence of dominant alternatives. With the assumption that the utility of an individual increases monotonically with the improvement in attribute levels (i.e., Level 2 is strictly preferred to Level 1 which is strictly preferred to the current condition), we found two choice tasks with dominant alternatives in one of the three blocks. To eliminate the presence of dominance, we relabelled and swapped attribute levels across choice tasks within a block. We are aware that this procedure has implications on the efficiency of the design. In addition, the BDD design generated from NGENE was not dummy coded so we then converted the designs to dummy coding. As we have relabelled, swapped and converted into dummy coding, the efficiency measures previously calculated for the initial BDD design have therefore been altered. After constructing the new dummy coded

BDD design, we recalculated the efficiency ratings using the design evaluation feature on NGENE 1.02 (e.g., ";eval = recoded_design_BDD.ngd").

The evaluation of the three designs in terms of statistical efficiency is presented on Table 2.1. Columns 2, 3 and 4 of Table 2.1 show the three sets of efficiency measures. In the first set, we assumed that parameter values are zeroes ($\boldsymbol{\beta}_s = \mathbf{0}$). OOD has the lowest D_z-error and A_z-error hence it is the most efficient design under this measure. This is followed by BDD and ORD, respectively. In the second set, where we assumed that parameter values are not equal to zero ($\boldsymbol{\beta}_s \neq \mathbf{0}$), BDD is the most efficient design based on having the lowest D_p and A_p errors while OOD has the lowest efficiency. This is not surprising because BDD follows the Bayesian D-efficiency criterion, where we used parameters values estimated using conditional logit model from the initial set of choice data (shown on Table 2.1). These sets of priors can be considered "reliable" because they came from actual survey respondents. The term *reliable* is mentioned in Ferrini and Scarpa (2007) where a sample from a pilot survey can be considered *reliable* if the difference in the marginal rates of substitutions (e.g. marginal WTP) between pilot and final sample is small. To check for the reliability of priors that we derived from the pilot survey of 35 respondents, we compare the calculated marginal Willingness to Pay (WTP) from this pilot WTP_P with the WTP from the full sample (WTP_F) of 209 respondents. Table 2.2 shows the percentage difference between WTPs for the increase in abundance of brown kiwi which is approximately 10% between the pilot and the full sample. For the highest feasible increase in Falcon abundance (or attribute level two), the WTP from the full sample is lower by 20%. This relatively small difference of WTPs between the pilot and full samples suggests that our set of priors can be considered reliable.

The WTP for the non-bird attributes may be difficult to compare because the utility coefficients from the pilot sample are not statistically significant.

Despite BDD being the most efficient, its theoretically minimum required sample size or S_p estimate (4,157) is more than ten times the S_p estimate for ORD (375). As the calculation for S_p estimate is based on having all parameters being statistically significant at 5% level, this might have been influenced by the presence of non-bird attributes in the choice task which have been found to have a comparatively low contribution to individual utility and hence have utility coefficient values very close to zero. These coefficient estimates might have required a significantly large number of choice observations to become significant at the 5% level. One may argue that those attributes should not have been included in the investigation at all, but they were included because of their importance for wildlife management.

In the second set of efficiency measures above, although we accounted for the effect of parameter values not being equal to zero ($\boldsymbol{\beta} \neq \mathbf{0}$), we have assumed those parameters to be fixed and therefore to be known with certainty. However, there typically exists a considerable amount of uncertainty about parameter values $\boldsymbol{\beta}$ and such uncertainty should be accounted for. We accounted for this uncertainty by following the sequential Bayesian framework suggested in Ferrini and Scarpa (2007) and applied in Scarpa et al. (2007). The third set of design efficiency measures is based on the sequential Bayesian approach. As expected, the BDD is the most efficient design based on Bayesian D-error (D_b-error). The Bayesian measure for theoretically minimum sample size, or S_b estimate, for OOD is 6 million choice observations while BDD and ORD respectively got S_b estimates of 1.3 million and 0.2 million.

After evaluating the three EDs that we used in the survey, we also optimised each ED to derive updated Bayesian D-efficient designs using conditional logit model estimated parameter values from respective design sample of 503 choice observations in Table 4.1. Using NGENE, the ORD design used in the survey was optimised for Bayesian D-efficient design using the utility coefficient estimates of logit model from the ORD sample in Table 4.1. Similarly, we generated updated Bayesian D-efficient designs for the OOD and BDD using the same procedure. We then evaluated the design efficiency of the three updated Bayesian D-efficient designs. These new design efficiency measures are presented in columns 5, 6 and 7 of Table 2.1. The three new designs all demonstrate considerable improvement in terms of *D*-error, *A*-error and *S*-estimate. Although BDD was previously the most efficient design, it still improved in efficiency from D_b-error of 0.223 to 0.150. The OOD, which was previously the least efficient, not surprisingly had the most remarkable improvement in design efficiency from D_b-error of 0.937 to 0.170. The OOD design that was optimised for Bayesian D-efficiency has outperformed the efficiency of ORD which got a D_b-error = 0.185. These results corroborate the notion put forward in other studies (Scarpa et al., 2007, Kerr and Sharp, 2010) that EDs can be updated and be made more statistically efficient using more reliable prior information from a bigger sample of respondents. However, if an analyst had used a utility neutral design for the first wave of survey, that ED could still be optimised following the BDD criterion and significantly gain statistical efficiency (i.e., lower D_b-error). As described in Scarpa et al. (2007) an improvement in the statistical efficiency of ED leads to more accurate parameter estimates in addition to a reduction in the theoretically minimum sample size. As a result of optimising for the BDD criterion, all efficiency measures (i.e., D-error, A-error, S_b estimates) for the three designs significantly improved, with OOD exhibiting the highest degree of improvement. In terms of the percentage of improvement, more weight should be given to the improvement (or reduction) in D-error as it is a measure of statistical efficiency of the overall experimental design. Lesser weight may be given to the S_b estimate (or S-efficiency score) which represents the maximum of the individual scores for parameters. Although OOD achieved an impressive improvement of 79,517%, this may be considered irrelevant because one cannot get an economically feasible sample size to retrieve significant parameters for the relevant attributes however efficient the design process is.

Table 2.1: Evaluation of the three experimental designs used in the falcon survey

	Evaluation of efficiency of existing designs			Optimised for Bayesian D-efficiency of existing design using parameter estimates from each design treatment		
	ORD	**BDD**	**OOD**	**ORD**	**BDD**	**OOD**
Assuming βs = 0						
D_z-error	0.205	0.178	0.091	--	--	--
A_z-error	0.542	0.478	0.308	--	--	--
Assuming βs ≠ 0 but fixed						
D_p-error	0.290	0.213	0.589	0.173	0.143	0.161
A_p-error	0.801	0.595	3.417	0.345	0.274	0.309
S_p estimate	375	4,157	4,114	174	1,237	478
Assuming βs ≠ 0 and accounting for uncertainty						
D_b-error	0.307	0.223	0.937	0.185	0.150	0.170
A_b-error	0.850	0.622	18.886	0.369	0.289	0.327
S_b estimate	212,740	1,265,695	6,091,078	562	6,432	7,660
% of improvement from optimisation for BDD				After 6,990 evaluations on NGENE	After 6,659 evaluations on NGENE	After 6,657 evaluations on NGENE
D_p-error	--	--	--	167%	149%	365%
A_p-error	--	--	--	232%	217%	1,105%
S_p estimate	--	--	--	215%	336%	862%
D_b-error	--	--	--	166%	148%	550%
A_b-error	--	--	--	230%	215%	5,770%
S_b estimate	--	--	--	37,861%	19,678%	79,517%

Note: Conditional logit model estimates of *βs* from pilot survey data used as priors are presented on Table 2.3.

Table 2.2: Willingness to Pay estimates from pilot and full survey samples

	Pilot Sample (n=35)				Full Sample (n=209)				
	Coefficient	Standard Error	P-value	Marginal WTP_P	Coefficient	Standard Error	P-value	Marginal WTP_F	% diff in WTP[a]
Brown kiwi 1	**0.462**	**0.252**	**0.07**	**$ 22.00**	**0.504**	**0.098**	**<0.01**	**$ 20.16**	**9.1%**
Brown kiwi 2	**0.591**	**0.251**	**0.02**	**$ 28.14**	**0.622**	**0.095**	**<0.01**	**$ 24.88**	**13.1%**
Native fish 1	0.242	0.241	0.32	NS	**0.287**	**0.093**	**<0.01**	**$ 11.48**	**--**
Native fish 2	0.286	0.248	0.25	NS	0.143	0.095	0.13	NS	**--**
Native plant 1	0.335	0.233	0.15	NS	0.145	0.094	0.13	NS	**--**
Native plant 2	0.112	0.251	0.66	NS	**0.210**	**0.094**	**0.03**	**$ 8.40**	**--**
Green gecko 1	0.190	0.246	0.44	NS	0.017	0.093	0.86	NS	**--**
Green gecko 2	**0.549**	**0.241**	**0.02**	**$ 26.14**	0.092	0.093	0.32	NS	**--**
Bush falcon 1	**0.550**	**0.253**	**0.03**	**$ 26.19**	**0.453**	**0.098**	**<0.01**	**$ 18.12**	**44.5%**
Bush falcon 2	**0.706**	**0.246**	**<0.01**	**$ 33.62**	**0.700**	**0.094**	**<0.01**	**$ 28.00**	**20.1%**
Cost to respondent	**-0.021**	**0.004**	**<0.01**	--	**-0.025**	**0.002**	**<0.01**	--	--
Indicator for status quo	**0.876**	**0.413**	**0.03**		0.177	0.158	0.26		
Pseudo R2	0.060				0.245				
Number of choice obs	314				1850				

[a] To calculate for the percentage difference in marginal WTP, we used the formula: $\%diff = [(WTP_P – WTP_F)/ WTP_F] \times 100\%$

Note1: NS means *not significant* at the 90% confidence level.
Note2: Values in boldface font represent statistical significance of utility coefficients at the 90% confidence level.

Table 2.3: Conditional logit model estimates using the data set with the first 35 respondents (first wave)

	Coefficient	Standard Error	T-ratio	P-value
Brown kiwi 1	**0.462**	**0.252**	**1.832**	**0.067**
Brown kiwi 2	**0.591**	**0.251**	**2.354**	**0.019**
Native fish 1	0.242	0.241	1.002	0.316
Native fish 2	0.286	0.248	1.155	0.248
Native plant 1	0.335	0.233	1.441	0.150
Native plant 2	0.112	0.251	0.446	0.655
Green gecko 1	0.190	0.246	0.771	0.441
Green gecko 2	**0.549**	**0.241**	**2.278**	**0.023**
Bush falcon 1	**0.550**	**0.253**	**2.174**	**0.030**
Bush falcon 2	**0.706**	**0.246**	**2.865**	**0.004**
Cost to respondent	**-0.021**	**0.004**	**-5.136**	**<0.001**
Indicator for status quo	**0.876**	**0.413**	**2.122**	**0.034**
Log-likelihood value				-324.473
Pseudo Rho2				0.078
Adj Pseudo R2				0.060
Number of choice observations				314
Number of respondents				35
Number of iterations				5

Note: Text in boldface font indicates statistical significance at the 90% confidence level

2.4 Sampling procedure

We planned to have four different levels of sample stratification. The first type was about getting representative samples from regions with a large proportion of planted forests and those regions with relatively smaller proportion of these forests. The second was to get representative samples of respondents living in rural and urban areas of the country. The third focus was to have 50-50 split samples of respondents who completed mail survey and online survey. The fourth was to split the full survey sample into groups of respondents who completed choice tasks generated from different experimental designs. This section of the thesis presents the sample stratification that we tried to achieve. At the end of each sub-section below we compare the sampling stratification we aimed for with the actual split of the sample data we collected.

2.4.1 Regional groupings

In order to get a representative survey sample of respondents across New Zealand, we employed a stratified sampling approach based on the distribution of the population. The strata employed include the location of residence across the 18 regions of the country. To get a balanced representation of respondents living in regions with large planted forests and those with smaller ones, the 18 regions were grouped into two categories based on the proportion of the area planted forests to the total area of the region. To do this, a digital map of New Zealand called Land Cover Database version 2 (LCDB2) was used. A spatial software called ArcGIS was used to extract the total planted forest area by region and intersect these with the total area by region. Six regions have been found to have at least 12% of planted forest area (Group 1) while the remaining 12 regions have less than 12%

of planted forests (Group 2). Group 1 represents the regions where one can find the country's largest plated forests (e.g., Kaingaroa, Matariki). Group 2 regions generally have smaller and sparse planted forests. From each regional group, we tried to sample 50% of the respondents. Table 2.4 presents the two regional groups with Group 1 having six regions and Group 2 with 10 regions. Based on the 2006 population census data of Statistics New Zealand, 75% of the country's population lived in Group 2 regions while only 25% lived in Group 1 regions. We aimed to over-sample in Group 1 regions to come up with a 50-50 split (50% of the sample from Group 1 and 50% from Group 2).

Table 2.4: Groupings of New Zealand regions by proportion of planted forests

Region	Land Area (in 1000 ha)	Planted Forest Area (in 1000 ha)	Percentage
Group 1 (Large planted forests)			
Nelson	422,397	131,049	31.0%
Bay of Plenty	12,160,133	3,036,148	25.0%
Gisborne	8,360,456	1,590,658	19.0%
Waikato	24,442,870	3,640,194	14.9%
Northland	12,508,417	1,819,227	14.5%
Hawkes Bay	14,173,620	1,730,250	12.2%
Group 2 (Smaller planted forests)			
Auckland	4,517,341	519,078	11.5%
Tasman	9,636,113	1,034,553	10.7%
Wellington	8,103,633	687,051	8.5%
Marlborough	10,222,844	732,016	7.2%
Manawatu-Wanganui	22,210,568	1,458,517	6.6%
New Plymouth	7,258,142	287,316	4.0%
Otago	31,873,471	1,253,361	3.9%
Canterbury	45,226,480	1,207,128	2.7%
Southland	31,379,319	810,510	2.6%
West Coast	23,356,245	473,449	2.0%
Overall total	**265,852,049**	**20,410,506**	**7.7%**

Source: Data adapted from http://koordinates.com/#/search/?q=lcdb2

Table 2.5 shows the planned stratification and actual distribution of respondents by regional grouping and by region. The planned stratification was based on the population distribution of New Zealand across regions (Statistics New Zealand 2012). As the Waikato region has the highest proportion of population in Group 1, we allocated the highest proportion of respondents (36%) for that group. As Auckland has the highest proportion of population of the regions in Group 2 (and in the country as well), we tried to allocate to it the highest proportion of respondents (40%) in that group. However, since we got more respondents in Group 1 than in Group 2, some targeted proportions of respondents per region were not achieved. In Group 1, the actual proportion of respondents in Waikato was 45% instead of the targeted 36%. While in Group 2, instead of getting 40% for Auckland, we only got 26%. Overall, as we suffered from a low response rate, this sampling strategy was not fully achieved. We got 66% of respondents from Group 1 regions and 34% from Group 2 regions.

Table 2.5: Planned stratification following the population distribution in New Zealand and actual distribution of respondents

Grouping	*Planned*		*Actual*		*Difference*	
	Number	*%*	*Number*	*%*	*Number*	*%*
Group 1 (Large planted forests)						
Northland Region	15	14%	17	12%	2	2%
Waikato Region	38	36%	63	45%	25	9%
Bay of Plenty Region	25	24%	34	24%	9	0%
Gisborne Region	6	6%	9	6%	3	0%
Hawke's Bay Region	15	14%	8	6%	-7	8%
Nelson Region	6	6%	8	6%	2	0%
Group 1 Total	***105***	**100%**	***139***	**100%**	**34**	**18%**
Group 2 (Smaller planted forests)						
Auckland Region	42	40%	18	26%	-24	14%
Taranaki Region	4	4%	2	3%	-2	1%
Manawatu-Wanganui Region	8	8%	12	17%	4	9%
Wellington Region	15	14%	17	24%	2	10%
Tasman Region	4	4%	0	0%	-4	4%
Marlborough Region	4	4%	4	6%	0	2%
West Coast Region	2	2%	1	1%	-1	1%
Canterbury Region	17	16%	10	14%	-7	2%
Otago Region	6	6%	3	4%	-3	2%
Southland Region	2	2%	3	4%	1	2%
Group 2 Total	**104**	**100%**	**70**	**100%**	**-34**	**46%**

Note: For the target number of respondents, we have assumed here that 209 was the target to provide comparison with the actual number of respondents.

2.4.2 Urban-rural split

We also attempted to use another stratum which is the urban-rural split. Statistics New Zealand (2010) reports that in 2006, 72% of the households lived in urban areas while 28% lived in rural communities. As we have drawn our survey sample from the Whitepages, we were able to compile phone numbers of households residing in the 14 urban centres of the country. The 14 urban centres were composed of the 14 key cities in the country which are: Whangarei, New Plymouth, Wanganui, Nelson, Auckland, Gisborne, Palmerston North, Christchurch, Hamilton, Napier-Hastings, Kapiti Coast, Tauranga, Rotorua, Wellington and Invercargill. To draw rural respondents, we selected people living outside these urban centres. However, due to low response rate, we ended up with a slightly different urban-rural distribution of 60-40 instead of the 72-28 split we originally aimed for.

2.4.3 Mail-online split

In 2006, 92 percent of New Zealand households had land-based telephone units while 66% had internet connection (Statistics New Zealand, 2010). We therefore employed a phone-mail and phone-internet survey. We initially aimed to get 60% of the respondents from phone-mail while 40% of the respondents from phone-internet. In doing this two-stage survey technique, we first called people listed in the White Pages and asked if they were interested in participating in a survey. Three economic survey assistants, all New Zealand-born native English speakers, so as to minimize interviewer bias, were hired and trained to call people on the phone list. A total of 2,996 phone calls were made between December 2009 and August 2010. The calling exercise, suggests a pattern that for every four numbers dialled, two ended up getting in contact with a New Zealand household member while the remaining two ended up with either answering machine or continuous ringing. For every two persons contacted, one would likely agree to participate in the survey while the other one would likely be uninterested or too busy to participate. People who agreed to participate were asked to choose whether they preferred to participate in the survey by mail or online. Those who preferred online were asked to provide their email addresses. Each online participant was sent an email containing a link to an online questionnaire that corresponded to a particular version of the questionnaire. For each survey link that was emailed, we included an identification number to track the completion of the survey and to facilitate sending a follow up email message in case the online survey had not been completed within two to three weeks. If the person talked to preferred to participate in the survey through a mail questionnaire, the home address listed in the White Pages was verified for accuracy. Each questionnaire sent by mail had an ID number to facilitate tracking for follow up phone calls if the survey was not returned within three weeks. During the follow up phone call, we emphasized that we valued their participation in the survey and if the questionnaire was not received or got misplaced, a new copy of the questionnaire would be sent off.

A total of 781 mail surveys or online invitations were sent out. A larger proportion of people expressed interest in completing the survey by mail so, it was later decided to focus solely on collecting survey data using the phone-mail approach. We have also seen that the online questionnaire had some formatting issues when two other internet browsers Google Chrome and Mozilla Firefox compared with Internet Explorer where the online survey could be better viewed. We finished with 261 filled-out surveys, 84% of which were mail and 16% online. Mail surveys had a relatively higher valid survey rate of 81% compared to online with 74%.

2.4.4 Experimental design split

We employed the sequential survey method described in Scarpa, et al. (2007) where we sent surveys in two waves. The experimental design technique used for the first wave followed the orthogonal design criterion. The structure of the orthogonal design, and so as the subsequent designs in this study, is composed of 27 choice tasks divided into three blocks. Each respondent was provided with nine choice tasks (therefore potentially nine choice observations per respondent). Each choice task had three alternatives. The first alternative represents the current situation with cost = $0 (not included in the design). The other two alternatives represent changed alternatives with combination of levels derived from a particular experimental design with the cost attribute getting $30, $60 or $90. For the first wave, we allocated 108 respondents who agreed on the phone to participate in the survey. Unfortunately we were able to get back only 35 completed questionnaires out of these 108 people. Choice observations compiled from the first wave served as out pilot choice data.

The second wave of survey involved 432 respondents who agreed to participate. This wave consisted of four experimental designs with each design targeted to have 108 respondents. The four designs were: (1) Optimal orthogonal, (2) Bayesian D-efficient, (3) Bayesian S-efficient, and (4) Bayesian C-efficient. The second, third and fourth designs belong to the class of Bayesian efficient designs that assume a conditional logit model will be used to analyse the collected choice data. These four designs were generated using NGENE 1.02 using the conditional logit coefficient estimates from the initial set of surveys completed by 35 respondents, as *a priori* distribution of the parameters of the indirect utility function. Table 2.3 presents the conditional logit model estimates from the first 35 respondents.

We had originally planned to conduct a third wave survey that would have 432 respondents distributed over four advanced experimental designs with each design allocated with 108 respondents. The four designs were: (AD1) Random parameter logit with panel implementation S-efficient design; (AD2) Random parameter logit with error components panel S-efficient design; (AD3) Model Averaging 1 with equal weights for CL, RPPanel and RPECPanel; and (AD4) Model Averaging 2 with more weights to RPPanel and RPECPanel than CL.

Advance designs AD1, AD3 and AD4 are described in detail in Scarpa and Rose (2008).

We used parameter estimates from the completed first and second wave of surveys as *a priori* information to create these more advanced designs. However, we were not able to generate any of the more advanced designs with the property that we desired which had a realistic sample size requirement. When we use the panel and cross sectional specifications, we arrived at a minimum required sample size of 1.5 million choice observations which seemed impossible to satisfy. A possible reason for the extremely high sample size requirement is that we included environmental attributes that represent the less charismatic native species (i.e., native fish, native plant, green gecko) which seem to be less attractive to many respondents. This resulted in having parameter estimates with low *t*-ratios (hence not significant) which consequently contributed to the requirement of larger sample sizes for the four advance designs. We therefore decided not to pursue the empirical examination of the more advanced designs. We leave this task for future research and different operational conditions.

2.5 Choice data

From the choice experiments survey, two types of data sets were constructed: the *full sample* which includes all the completed questionnaires; and the *balanced sample* which excluded several observations from the full sample to facilitate comparison of the three choice experimental designs. The full sample was used in the analysis reported in Chapter 3 whilst the balanced sample was used in the analysis presented in Chapters 4 and 5.

2.5.1 The full sample

A total of 821 people agreed on the phone to join the survey. Each person was sent a survey package which included the questionnaire, return stamped envelope, a pen and a note pad. 261 respondents filled out and returned the survey which corresponds to a second stage survey response rate of 32%. Out of these 261 respondents, 209 provided valid entries to the choice data set. 201 respondents evaluated all nine choice tasks provided to them while eight respondents completed only part of the nine choice tasks. Those eight respondents ended completing 1, 3, 6, 7 or 8 choice tasks. As we have sent out self administered questionnaire, it is difficult to determine the reasons why these eight respondents did not complete the nine choice tasks. We speculate that some accidentally missed a couple of choice tasks while some might have found the last choice tasks to be too tiring. Some must have preferred to skip some choice tasks rather than provide random answers. Because of the non-completion of some choice tasks, instead of collecting 1881 choice observations, we ended up with 1850 observations for the final full data set.

We used five choice experimental designs in collecting the full data set, namely: orthogonal (ORD), optimal orthogonal (OOD), Bayesian D-efficient (BDD), Bayesian C-efficient (BCD) and Bayesian S-efficient (BSD). Table 2.6 shows the distribution of choice data across designs. As we focused on comparing the first three experimental designs, a large majority (83%) of the choice observations of the full sample came from those designs.

Table 2.6: Sample distribution by choice set order and experimental design of the full sample

Choice set order	ORD	OOD	BDD	BCD	BSD	Pooled
1st	57	60	56	17	18	208
2nd	57	59	56	16	18	206
3rd	57	59	55	17	18	206
4th	57	57	56	17	18	205
5th	57	58	56	17	18	206
6th	57	58	56	17	18	206
7th	56	57	56	17	18	204
8th	56	57	56	17	18	204
9th	57	57	56	17	18	205
Total choice Observations	**511 (28%)**	**522 (28%)**	**503 (27%)**	**152 (8%)**	**162 (9%)**	**1850 (100%)**
Total number of respondents	**58 (28%)**	**60 (29%)**	**56 (27%)**	**18 (8%)**	**17 (8%)**	**209 (100%)**

We originally planned to compare all five different designs. However, due to very low response rates and limited resources (i.e., time and money), we decided to focus on comparing three experimental designs and there are ORD, OOD and BDD. To address low response rates and unevenly distributed design blocks for these three designs, we recruited more respondents on the telephone and sent them the questionnaires that were not completed and returned from previous mail outs to fill in the gaps and to increase the existing sample size. In August 2010, we recruited 204 additional respondents over the phone and sent them the questionnaires with blocks from the three experimental designs that were not completed in the previous mail outs. Forty-two respondents fully completed and returned the surveys. The additional 42 respondents increased the number of choice observations for ORD, BDD and OOD while also contributing to a more balanced distribution across three blocks (Table 2.7). As we did not send out surveys to fill in the gaps for BCD and BSD, the distribution across blocks for these designs was relatively unbalanced compared to the first three designs. The observation in the BSD design treatment was concentrated on block 3 (50%) while BCD on block 2 (46%).

Table 2.7: Sample distribution by block and experimental design of the full sample

Block number	ORD (%)	BDD (%)	OOD (%)	BCD (%)	BSD (%)	Pooled (%)
1	162	152	125	45	27	511
	(32%)	(30%)	(24%)	(30%)	(17%)	(27%)
2	171	198	187	70	54	680
	(33%)	(39%)	(36%)	(46%)	(33%)	(37%)
3	178	153	210	37	81	659
	(35%)	(31%)	(40%)	(24%)	(50%)	(36%)
Total	511	503	522	152	162	1850

Note: “(%)” above indicates the proportion of the choice observations per design treatment.

2.5.2 The balanced sample

As we decided to focus on comparing the three experimental designs which are the ORD, OOD and BDD, we constructed a balanced data set with split design. Table 2.8 shows the distribution of the three design samples (each with 414 choice observations) before we added the 42 additional respondents. Table 2.9 shows the more balanced distribution across blocks for the three designs after adding choice observation from the additional sample.

Table 2.8: Sample distribution by block and experimental design of the earlier sample

Block	ORD	%	OOD	%	BDD	%	Pooled	%
1	117	28%	98	24%	134	32%	349	28%
2	153	37%	127	31%	189	46%	469	38%
3	144	35%	189	46%	91	22%	424	34%
Total	414	100%	414	100%	414	100%	1242	100%

Table 2.9: Sample distribution by block and experimental design of the final balanced sample

Block	ORD	%	OOD	%	BDD	%	Pooled	%
1	162	32%	125	25%	152	30%	439	29%
2	171	34%	187	37%	198	39%	556	37%
3	170	34%	191	38%	153	30%	514	34%
Total	503	100%	503	100%	503	100%	1509	100%

To construct the final balanced data set in Table 2.9, as we got higher response rates for ORD and OOD compared with the BDD design (Table 2.7), we excluded 8 and 19 observations from the orthogonal and optimal orthogonal samples, respectively. This is to make the distribution of choice samples on a per block and a per order basis exactly the same across the three designs. Table 2.10 shows the number of observations per choice task order. As each respondent was provided with nine choice tasks, choice tasks were ordered as 1^{st}, 2^{nd}, 4^{th},… …, until the 9^{th} choice task. The number of observations for each choice task order was 56 for each design treatment with the exception of the 3^{rd} choice task that had 55 observations for each sample. The reason for this imbalance was because some respondents were not able to complete the evaluation of the nine choice tasks.

The criteria used for excluding choice observations to construct the balanced sample were: (1) choice observations from respondents who did not complete the nine choice tasks; (2) choice observations from respondents who sent back the questionnaire very late because we had already completed the planned design when we received those; and (3) for convenience, other choice observations at the bottom of the worksheet were removed.

After excluding the above choice observations, we constructed a balanced data set from 172 respondents. The 503 choice observations from the ORD came from 57 respondents, while choice observations for OOD and BDD samples came from 59 and 56 respondents, respectively (Table 2.10). The reason for the differences in the number of respondents (despite the same number of choice observations per design sample) is that not all respondents completed the nine choice tasks assigned to them. For the ORD sample, three respondents did not complete the nine choice tasks, while for OOD and BDD samples, seven and one respondents did not complete, respectively.

Table 2.10: Sample distribution by choice task number and experimental design of the final balanced sample

Choice set order	No. of Observed Choice Sets			
	ORD	OOD	BDD	Pooled
1st	56	56	56	168
2nd	56	56	56	168
3rd	55	55	55	165
4th	56	56	56	168
5th	56	56	56	168
6th	56	56	56	168
7th	56	56	56	168
8th	56	56	56	168
9th	56	56	56	168
Total Choice Observations	**503**	**503**	**503**	**1509**
Total number of respondents	**57**	**59**	**56**	**172**

2.6 Summary

In this chapter, we set up the formal theory of choice models that will be estimated in Chapters 3, 4 and 5 of the thesis. These three types of logit models include conditional, heteroskedastic and latent class models. We also described the different measures of design efficiency such as the Bayesian D-error, D-efficiency and Utility Balanced measures. In the next three chapters, we will be referring to the above description of the models and data.

Chapter 3: Valuing biodiversity enhancement in planted forests: socio-economic and spatial determinants of willingness-to-pay

3.1 Introduction

The world's planted forests cover approximately 264 million hectares accounting for about seven percent of the global forest area (FAO, 2010). A planted forest, which can be composed of a single exotic forest species, is generally considered as a legitimate land use to address the global demand for roundwood, pulp, non-wood products and other forest goods (Bauhus et al., 2010). Planted forests contribute to the conservation of natural forests by off-setting pressure on primary and old growth forests (UNCED, 1992; Dyck, 2003). In addition, they provide important ecosystem services that include habitat provision for native species, including those threatened with extinction (Jukes and Peace, 2003; Brockerhoff et al., 2008; Pawson et al., 2010). Planted forests can be managed to enhance the provision of habitats for rare and protected native species (Jactel et al., 2006; Pawson et al., 2005; Hartly 2002; Maunder et al., 2008; Bauhus and Schmerbeck, 2010). However, enhancing the provision of habitats for threatened species comes at a price (Seaton et al., 2006; Maunder et al., 2008; Weir, 2010). It is therefore important to examine whether the general public would benefit from a biodiversity enhancement initiative and if they did would they be willing to pay to support such initiative.

New Zealand (NZ) has a total of 1.8 million hectares of planted forests accounting for 22% of the country's total forest area (MAF, 2010). In 2009, total revenue derived from sale of planted forests products was the country's third largest export earner contributing NZ$3.7 billion (2.8% of GDP) to the economy.

NZ planted forests are composed of exotic trees, with Radiata pine (*Pinus radiata*) as the dominant species accounting for 90% of the forest area. The remaining species include Douglas fir (*Pseudotsuga menziesii*), Cypress (*Cupresus sp.*) and Eucalypts (*Eucalyptus sp.*) (MAF, 2010). Exotic forests provide habitats for at least 118 threatened native species that include the brown kiwi (the country's national symbol) and the bush falcon (Pawson et al., 2010). Areas in between clear cut and remaining forest stands of the 135,000-hectare Kaingaroa forest in the Central North Island region provide habitats for the bush falcon which are better than any other habitat areas with stands of native forests in isolated hilly areas of the country (Seaton, 2006; 2010). The Kaingaroa forest area has the highest concentration of bush falcon in the country (Stewart and Hyde, 2004). The bush falcon is the country's fastest bird and it preys mainly on exotic bird species and insects (Seaton, 2006).

The Department of Conservation (2000) reports that New Zealanders place a high value on native plants and animals, as they form a basis to the culture and a sense of national identity. Native birds and plants can be seen all over the country both in public conservation lands (e.g., national parks, forest parks) and private lands (e.g., residential lands, planted forests). Using a dichotomous choice contingent valuation method, Yao and Kaval (2010) (referred to as YK) have shown that a typical New Zealand individual would be willing to pay about $82 per year in additional local taxes (or local rates) to support the planting of more native trees and shrubs on public land and $42 per year for more natives on private land.[7] Private land in YK mainly referred to private properties large

[7] It was mentioned in the survey that additional native trees and shrubs would provide additional habitat to native birds, fish and geckos.

enough to accommodate the planting of native trees. Private land in YK did not include planted forests which could be as large as 135,000 hectares as the Kaingaroa Forest or the 16,000-hectare City Forests in Dunedin. Although YK has shown that additional native trees are valued on private land, it remains unclear if increasing the number of threatened species in exotic planted forests by improving the habitat would be valued by New Zealanders.

This chapter is motivated by the general question *Is a proposed biodiversity enhancement programme in planted forests valued by New Zealanders?* Answers to this question would provide some insights for the country's existing biodiversity programme on private land that is part of New Zealand's 20-year Biodiversity Action Plan (2000 to 2020). We also envision that those answers would provide some pointers in the formulation of future policies for the management of planted forests in countries where similar conditions exist. Specifically, this thesis chapter aims to answer two research questions:

(1) Would New Zealanders be willing to pay for biodiversity enhancement in planted forests, and if so, approximately how much?

(2) What are the factors that would likely influence the willingness-to-pay (WTP) of an individual for biodiversity enhancement? And by how much would these factors affect the median WTP?

The first question is addressed by analysing a survey data collected using the stated choice experiments (CE) approach (please refer to Chapter 2 for details of CE). WTP values (or WTPs) are calculated by taking the ratios of the coefficient of the attribute level over the marginal utility of income. By using the Monte Carlo simulation, we have accounted for the distribution of WTP in these ratios.

Simulated median WTPs are subsequently aggregated (accounting for potential biases) to represent a national value of biodiversity enhancement in New Zealand's planted forests. This method is described in detail in Section 3.3 of this chapter. The second question is answered by determining the factors that would likely influence individual specific WTPs using panel random effects regression models with a panel of 10 WTP values per respondent as dependent variable. This follows the panel random effects ordinary least squares (OLS) regression method described in Campbell (2008a) and Scarpa et al. (2011b). We made an innovation here as we have combined socio-economic and attitudinal covariates with geospatial distance of respondents from large planted forests as additional explanatory variables of individual specific means of marginal WTP.

Marginal WTPs calculated from utility coefficient estimates of logit models suggest that New Zealand taxpayers, accounting for potential sources of bias, would pay an aggregate value of NZ$26 million per year for five years to support a proposed government coordinated programme on enhancing the provision of habitat for threatened native species found in planted forests. Results from panel regression analyses indicate that the factors that influence individual specific WTP include higher education, attitude toward conservation and proximity to large planted forests.

The next section of this chapter (section 3.2) provides a brief overview of the importance of biodiversity around the world and in New Zealand. Section 3.3 describes how this study contributes to, or extends previous studies. In 3.4, we describe the econometric models and other methods used in the analyses. Section 3.5 illustrates how choice data were collected and constructed for the choice analysis. Section 3.6 presents the results of econometric analyses and discusses the answers to the two main questions above. This chapter ends with conclusions and policy implications in section 3.7.

3.2 Biodiversity and planted forests

The term "biodiversity" is defined by the Convention on Biological Diversity as "the variability among living organisms from all sources including, terrestrial, marine and other aquatic ecosystems and the ecological complexes of which they are part; this includes diversity within species, between species and of ecosystems" (CBD 1992). This suggests that biodiversity includes diversity within species populations (genetic variation); the number of species, and the diversity of ecosystems. For this study we focus on some of New Zealand's threatened native species whose numbers considerably declined in the past decades.

Biodiversity decline is considered as the most important global environmental issue (FAO 1992; World Bank 2002). In 2009, 191 countries ratified the Convention on Biological Diversity (CBD) indicating that world leaders recognise the importance of putting a halt to biodiversity decline (CBD, 2010). Countries like the United Kingdom, United States, Australia and New Zealand developed long-term Biodiversity Action Plans to provide a platform for the incorporation of biodiversity conservation in policy decision making to address biodiversity decline. Despite the development and implementation of these action plans, biodiversity levels in those countries continue to decline (e.g., the population of threatened species continue to decrease). A possible reason is that although many governments aim to incorporate biodiversity in their environmental policies, many types of data, such as robust non-market values that can be used for cost benefit analysis, are not available. Very limited studies have attempted to estimate the value that people would be willing to pay to support biodiversity conservation in planted forests (de Groot and van der Meer, 2010). This restricts the inclusion of biodiversity values into the decision making framework and limits the value of planted forests only in terms of timber products (or only in terms of market values). This situation persists around the world despite the increasing evidence showing that planted forests provide ecosystems services such as carbon sequestration, soil erosion control, water regulation and biodiversity conservation (Norton, 1998; Scott et al., 2005; Byrne and Milne, 2006). Over the last decade, the capacity of planted forests to provide habitats for plants and animals has gained increasing attention (Schroth and da Mota, 2004; Carnus et al., 2006), especially in areas in where native forests have become rare (Humpreys et al., 2006; Berndt et al., 2008).

New Zealand's planted forests consist mainly of radiata pine (Pinus radiata), an introduced species typically managed as monoculture stands and harvested between 26 and 32 years after planting (Carnus et al., 2006; Dyck, 2003). Although these productive planted forests are mainly managed for timber production, they also provide an excellent habitat for indigenous plants and animals (Pawson et al., 2010; Brockerhoff et al., 2003; Norton, 1998; Spellerberg and Sawyer, 1995). Radiata pine forests, especially large ones with areas greater than 5000 hectares, provide habitats for threatened native animals (e.g., NZ bush falcon, brown kiwi, giant kokopu fish, green gecko) and plants (e.g., kakabeak shrub, native orchids) (Seaton, 2006; Pawson et al., 2006; Maunder et al., 2005; Brockerhoff et al., 2003; Pierce et al., 2002; Borkin, 2009). They also provide connectivity between areas of native ecosystems such as corridors that enable native fauna to gain better access for food from neighbouring native ecosystems (Norton, 1998; Carnus, et al., 2006; Brockerhoff et al., 2008; Pawson et al., 2010).

NZ forest companies recognise the importance of planted forests in providing habitat for native species (Maunder, 2008). Comprehensive ecological studies were undertaken and sponsored by forest companies to examine how planted forests can be managed to better suit the needs of threatened native animal inhabitants such as the bush falcon in Kaingaroa forest and long-tailed bat in Kinleith forest (Seaton, 2006; Borkin, 2008). Management of planted forests for biodiversity is encouraged in voluntary forest certification schemes (e.g., Forest Stewarship Certificate (FSC)). FSC certified companies regularly conduct ecological surveys which include monitoring for the presence of threatened native species and coordination with concerned non-government organisations (e.g., Wingspan, Forest and Bird) for the protection native birds (and other species) from forest harvesting operations (PF Olsen, 2009; Maunder, 2005). The management of planted forests for the provision of habitats for threatened native species indicates that NZ forest companies value and also benefit from biodiversity conservation as certified timber products gain better access to both global and domestic markets. For instance home improvement chains such as Home Depot in the US, B&Q in the UK and Bunnings Warehouse in New Zealand all prefer to buy and sell more FSC certified wood products than non certified ones.

Habitat provision is a type of ecosystem service. There is a rising public and corporate awareness of the importance of well-functioning ecosystem (Fisher et al., 2008; TEEB, 2010). Economic valuation of habitat provision for native wildlife had been undertaken on several forests (e.g., Czajkowski et al., 2008;

Christie et al., 2006). In the case of New Zealand planted forests, to our knowledge, the economic value of habitat provision has not yet been studied. This study is probably the first to estimate the economic value of a proposed habitat enhancement programme in planted forests. This is in line with the New Zealand Department of Conservation's (DOC's) policy to protect and enhance threatened native species which include the brown kiwi (Holzapfel et al., 2008).

With the development of economic valuation techniques (e.g., CE), complex biodiversity values can now be estimated. CE can be used to examine preferences and estimate WTPs of an individual on the changes in biodiversity outcomes in planted forests. However, in using CE there is a need to account for biases (e.g., hypothetical) that may arise in the elicitation of WTPs. Aggregation of the WTP values have inherent biases (e.g, aggregation bias). This study has accounted for both hypothetical and aggregation biases. We describe how we accounted for these biases in Section 3.3.

This thesis chapter presents aggregated WTP values of an important aspect of biodiversity that is the enhancement of iconic (e.g., brown kiwi) and less visible (e.g., giant kokopu fish) threatened native species in planted forests. Previous studies have shown that New Zealanders value indigenous biodiversity in general, however, those estimates of values mainly refer to indigenous biodiversity in national parks or regional forest parks which are composed mainly of native trees and shrubs (Yao and Kaval, 2010). National parks and forest parks are part of DOC's owned and managed conservation lands that usually have native trees of different ages and different types. Those landscapes are extremely different from planted forests. This is because a stand or plot of planted forest usually consists of a single type of exotic pine trees often with same tree age. Yet, despite the highly modified environment in an exotic forests stand, they still contribute to providing habitats for more than 100 threatened native species. In fact some threatened native species, like the bush falcon, has benefited more, living along the forest edges of large stands of Pinus radiata in Kaingaroa forests than in stands of native forests in the country's hilly areas (Seaton, 2006). Although the contribution of planted forests to biodiversity conservation has been recognized to be quite phenomenal, the monetary value of the benefits from this service still remains unclear and needs further examination to be included in future policy decision making. It is therefore the aim of this study to provide an estimate of an aggregate value of biodiversity enhancement in the country's planted forests to recognise that planted forests not only provide timber products that can be sold in the market, but also native biodiversity that could be enhanced to improve the welfare of society.

3.3 Approaches for valuing biodiversity enhancement

This study aims to estimate the WTP of New Zealand taxpayers to improve the habitats of selected threatened species that can be seen (e.g., bush falcon) or potentially sighted (e.g., giant kokopu, a native fish) in New Zealand's planted forests. WTP estimation is undertaken using the CE valuation framework. We describe in this section how the choice attributes were identified and how the questionnaire was developed. The econometric models used that are not described in Chapter 2, which include the Random Parameters Logit Model and Panel Data Random Effects Regression, are also discussed.

3.3.1 Focus groups and identification of attributes

Between June and August 2009, we conducted four focus group meetings. These were held in Rotorua on 6 June, in Whakatane on 8 July, in Taupo on 16 July and in Hamilton on 5 August (Table 3.1). The focus group in Rotorua was held at Scion (Rotorua) and this mainly served as a practice run to provide a feel for the author in moderating a focus group as he never moderated a focus group before. (The author was mentored by a senior staff at Scion who conducted a number of focus groups in the past.) Five volunteer Scion staff members attended this practice session. In the next three focus groups, participants were drawn from the general public through links with local councils and a university. Focus groups in Whakatane and Taupo were done in coordination with key council staff members who helped in disseminating the call for participation. They helped to post flyers in bulletin boards that would attract the attention of potential survey respondents who could be anybody from the general public. The fourth (and last) focus group in Hamilton was done in coordination with the University of Waikato and was attended by five students and a Maori community volunteer. Participants in the three focus groups consisted of labourers, retired people, council staff members, unemployed, students, a Maori community volunteer, and council staff members.

Table 3.1: Location, date and occupation of focus group participants

Location	Date	Number of attendees	Occupations of attendees
Rotorua	6 June 2009	5	1 chemist, 1 spatial analyst, 1 entomologist, 2 communication specialists
Whakatane	8 July	5	1 retired, 1 labourer, 1 clerk, 2 council staff members
Taupo	16 July	8	1 retired, 1 office staff, 1 farm worker, 1 labourer, 3 council staff members, 1 unemployed
Hamilton	5 August	6	4 students, 1 unemployed, 1 Maori ecological restoration volunteer

A range of native species were suggested by focus group participants which include brown kiwi, kereru, frog, worm, mistletoe, snail, eels, bat, pohutukawa, wood rose, falcon, gecko, weta, giant kokopu, inanga, morepork, kokako, fernbird, kakabeak, and spotless crake. This list was discussed with ecologists and was trimmed down to five based on the species conservation status and if the species' visibility could be enhanced through forest management. Consultations with ecologists helped identify the feasible set of five threatened species composed of the New Zealand bush falcon, the Auckland green tree gecko, the giant kokopu fish, the kakabeak plant and the brown kiwi. The frog and bat were excluded because of their low visibility in planted forests.

Figure 3.1 shows the five identified species with corresponding description of different levels of presence in planted forests. The column labelled as "current condition" represents the existing level of abundance in specific planted forests. On page 5 of the survey questionnaire (see Appendix A), we provided a detailed description of the current situation of the five attributes or the five threatened native species in planted forests. This is to introduce or familiarise the respondents with the current condition wherein planted forests provide suitable habitats to indigenous plants and animals. From the current condition, we explored the feasible ranges of increase for Level 1 (intermediate level improvement) and Level 2 (highest level of improvement) in consultation with forest ecologists and forest managers. The range of payment values (i.e., dollar bid values $30, $60 and $90) to increase the abundance of these species were also identified in focus group meetings.

Figure 3.1: The five native species, their current condition and two feasible levels of enhanced conditions

Threatened Animal/Plant		*Current Condition*	*Level 1*	*Level 2*
Brown Kiwi		Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 10 out of 200** planted forests	Kiwi calls heard **in 20 out of 200** planted forests
Giant Kokopu		Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 3 out of 10** suitable streams	Kokopu seen **in 5 out of 10** suitable streams
Kakabeak		At least **3 naturally occurring** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs	At least **20 actively managed** Kakabeak shrubs
Auckland Green Gecko		Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 3 out of 50** walks	Gecko sighted **in 5 out of 50** walks
NZ Bush Falcon		Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 3 out of 8** drives	Bush falcon sighted **in 5 out of 8** drives

Figure 3.1 above shows that the description of levels is framed using words such as "sighted", "heard" and "occurring" in order to collect a combination of use, option, existence and bequest values from respondents. Use values include direct use such as the value a recreationist derives from bird watching and indirect use such as the value derived from knowing a forest provides habitat for wildlife. Option value includes knowing that one would hear a kiwi in a forest in the future. Existence value comes from knowing threatened birds exist in a forest. Bequest value comes from ensuring that a threatened bird will be conserved for future generations.

The abovementioned values represent the different components of economic value. In this case, economic value refers to the degree to which biodiversity enhancement in planted forests satisfies individual preferences. Therefore, economic value can be measured by the amount of money that the individual is willing to pay for supporting biodiversity services in planted forests. Under the choice experiments approach to valuation, each survey respondent is presented with a series of choice tasks (as described in Chapter 2). To systematically populate the choice tasks with bundles of attribute levels, an experimental design is used. Attribute levels in Figure 3.1 are coded to allow each level to be accounted for into an experimental design framework. (We describe the different experimental designs in Chapter 2). Figure 3.2 presents a sample of a choice task used in the study. This choice task is part of a choice set series following an orthogonal design. Column 1 of Figure 3.2 shows the five threatened species and their corresponding locations. This is to emphasize that the five species can be found in planted forests in different parts of the country. For instance, the native plant kakabeak can be seen mainly in planted forests in the East Coast while the bush falcon can found mainly in the Kaingaroa forest and in Nelson.

Figure 3.2: Example of a choice task showing five environmental attributes, a cost attribute and three alternatives

Threatened Animal/Plant		***Current Condition***	***Option A***	***Option B***
Brown Kiwi (Frequency of hearing calls in planted forests in North Island)		Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 20 out of 200** planted forests	Kiwi calls heard **in 1 out of 200** planted forests
Giant Kokopu (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand)		Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 3 out of 10** suitable streams	Kokopu seen **in 1 out of 10** suitable streams
Kakabeak (Occurrence in 20% of the planted forests on the East Coast and Hawke's Bay)		At least **3 naturally occurring** Kakabeak shrubs	At least **3 naturally occurring** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs
Auckland Green Gecko (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions)		Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 5 out of 50** walks	Gecko sighted **in 1 out of 50** walks
NZ Bush Falcon (Bush falcon sightings while driving through pine forests in Central North Island and Nelson)		Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 3 out of 8** drives	Bush falcon sighted **in 1 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only		**$0**	**$30**	**$60**
I would choose (please tick)		☐	☐	☐

3.3.2 Choice attributes, levels and coding

The sample choice task in Figure 3.2 is composed of six attributes; five environmental attributes and a cost for the given policy alternative. Each environmental attribute is represented by a threatened native species that was identified as important to New Zealanders from a series of focus group meetings shown in Table 3.1. The identification of the five key species in the choice task was undertaken in consultation with forest ecologists and supplemented by the ecological literature. Key species selection was also guided by the conduct of focus group meetings. Focus group participants suggested that increasing the abundance of threatened native species, including non-bird species, is very important for forest wildlife management, and this would likely be valued by the general public (DOC, 2000) .

Brown kiwi and bush falcon are popular native bird species which can be considered as iconic species in the country. These two iconic species inhabit several planted forests and these forests can be managed to increase their abundance (Maunder et al., 2005; Colbourne et al., 2005). The brown kiwi can be found in planted forests in the North Island, particularly in Northland, Coromandel, Tongariro and Hawke's Bay (Pawson et al., 2010). Recognition of the importance of plantations for kiwi conservation is increasing. New forest management guidelines have been developed to minimise the effect of forest operations to kiwi population (http://rarespecies.nzfoa.org.nz/fauna/forest_birds/species/kiwi.htm). Sporle and Bliss (2008) suggest conservation orientated management regimes with the aim of achieving kiwi safe forestry operations in plantations inhabited by brown kiwi. On the other hand, it was mentioned earlier that New Zealand bush falcon is doing well in the Kaingaroa forest (Stewart and Hyde, 2004; Seaton 2006, 2010). The bush falcon is at the top of the food chain making it a good biodiversity indicator (Stewart, 2012).

In addition to the two bird species above, a number of threatened non-bird species (which include the kokopu, gecko and kakabeak), can also be found in planted forests (Pawson et al., 2010). Many planted forests have rivers and streams which provide habitats to native fish such as giant kokopu (Hanchet, 1990). Some planted forests on the East Coast and Hawke's Bay, provide habitat to kakabeak (Shaw and Burns, 1997). Auckland green gecko had been sighted in planted forests in Northland, Bay of Plenty and Waikato (BioWeb, 2009).

The three non-bird (or less iconic) species have been included in the choice task as ecologists that we approached pointed out their high ecological importance and their potential to be seen or heard in planted forests (therefore capturing both use and non-use values). The green gecko is a pollinator and seed disperser of certain species of native plants (Rowlands, 1989). It can be seen in tree branches and open ground. It can bark or chirp by clicking its tongue against the roof of the mouth. Planted forests can have native understoreys which can benefit from the increase in abundance of green geckos.

The native plant Kakabeak has special significance to New Zealanders because it is widely known and commonly used as an image in gifts, tourist souvenirs and the like. The presence of Kakabeak indicates good control of browsing animals (e.g., deer, goats, introduced snails) (Shaw, 1993). The Maungataniwha pine forest in northern Hawke's Bay has some parts with securely fenced enclosures; a Kakabeak plant was found in one of the enclosures (Slui, 2011).

The presence of Giant Kokopu is an indicator of good water quality in the waterways of planted forests. It indicates that a planted forest maintains a good riparian cover and clear running water. The threatened giant kokopu is listed as 'vulnerable' on the IUCN Red List of Threatened Species.[8] The term "giant" comes from the fact that it is the largest of all the 34 Galaxias species worldwide.

---

[8] http://www.iucnredlist.org/

Each attribute in the choice task has three levels representing different levels of population abundance of the species in planted forests. The initial, or base level, represents the current level of population abundance identified in consultation with forest ecologists and forest managers. From the current condition, we explored the feasible ranges of increase for Level 1 (intermediate level improvement) and Level 2 (highest level of improvement) again in consultation with forest ecologists and forest managers. The payment values were proposed to be paid per year for a five year period (i.e., dollar bid values \$30, \$60 and \$90) to increase the population of threatened native species. The range of “realistic” values was identified in a final focus group.

In presenting the choice tasks to respondents, we did not vary the order of the environmental attributes. The brown kiwi was always on top of the other species while the bush falcon was always the species at the bottom. In constructing the choice data set, we employed dummy coding where two dummy variables are assigned to each environmental attribute. The first dummy variable takes the value of 1 if the attribute is on level 1 and 0 otherwise. The second dummy takes the value of 1 if the attribute is on level 2 and 0 otherwise. If the environmental attribute level is on current condition, then both dummy variables take the value of zero. The four-level cost attribute was assigned with one variable that takes the values of \$0, \$30, \$60 and \$90. The cost variable takes the value zero if it was a status quo option and \$30, \$60 and \$90 if it was a changed alternative.

3.3.3 Survey questionnaire and the valuation scenario

The attribute levels and bid amounts were coded to allow them to be entered into an experimental design framework following the orthogonality criterion. This initial design was used to populate the nine choice tasks that were included in the questionnaire that was pre-tested.

The survey questionnaire was first tested with a pilot survey involving five respondents. Some of the pilot respondents indicated that the survey was too long primarily because of the lengthy description. We therefore trimmed down the questionnaire by deleting some of the unnecessary words, as indicated by the pilot respondents. We then conducted a second pilot test with another five respondents. Four out of five respondents mentioned that the questionnaire was long, but they were able to fully understand the questions and recognised the reasons for the presence of descriptions, which was mainly to make the valuation scenario as clear and realistic as possible. Nevertheless, in the following revision we still cut out a few more irrelevant words mentioned by the second set of pilot respondents.

In the CE valuation scenario, we included a “cheap talk” script as first described in Cummings and Taylor (1999). Some of the reasons for including the script are: to draw the respondent’s attention specifically to the cost variable; to remind respondents that they could use their money to buy other things they enjoy or to simply remind respondents of the opportunity cost of their money (Cameron and DeShazo, 2010). Although we did not use deception in the elicitation process, we have carefully designed a cheap talk script to address hypothetical bias inherent in the valuation scenario (Cummings and Taylor, 1999; Bishop and Heberlin 1979). In the aggregation process, we accounted for the remaining hypothetical bias by considering that hypothetical WTP is about twice the actual WTP (Christie, 2007).[9] We have also accounted for the aggregation bias by considering that people who agreed to participate but did not complete the survey would possibly indicate a WTP of zero (Morrison, 2000). Furthermore, as not all of the respondents were taxpayers, we also took account of the fact that non-taxpayers such as those retired or students would also have a WTP of zero.

3.3.4 Determinants of WTP

Campbell (2007) and Scarpa et al. (2011b) have used panel random effects regression models to determine the factors influencing WTP for the improvement of environmental goods. Campbell used a panel of individual specific median WTP estimates (estimated using a mixed logit model with panel specification) as dependent variable and socio economic characteristics and location as explanatory variables. Results suggest that income levels, community type and location significantly influence the variation of individual means of marginal WTPs. Similarly, Scarpa et al. (2011b) explained the variation of individual specific means of marginal WTP estimates (from exploded logit model with panel specification) using socio-economic characteristics such as marital status (e.g., single, married) and education level, and found that these explain reasonably well the observed variability. The above two studies identified the determinants of variation of individual means of marginal WTPs in terms of socio-economic

[9] The disparity between actual and hypothetical WTP has been studied in the 1970s (e.g., Bishop and Heberlin, 1979), 80s (Sinden, 1988), 90s (Foster et al., 1997; Frykblom, 1997; List and Shogren, 1998) and 2000s (Hofler and List, 2000; List and Gallet, 2001). For our current study, we find the Christie (2007) as the most appropriate for accounting for the disparity.

characteristics, attitude and affiliations but did not include the distance of the respondents to the location of those public amenities.

Choice analyses that account for the effects of distance on WTP are a very limited but growing area of research in the stated preference literature (Johnston et al., 2011). Several contingent valuation studies have used global distance decay models and found that WTP is negatively associated with the distance of the individual from the environmental good in question (e.g., Bateman et al. 2000, 2006; Hanley et al., 2003). However, Johnston et al. (2011) find no clear pattern of global distance decay on WTPs from a choice experiments exercise because of the occurrence of non-continuous spatial variation. Johnston has identified the presence of WTP hotspots in a stated discrete choice experiments framework by applying the Getis-Ord statistic (Getis and Ord, 1992). However, both studies examined the distance effects on WTP for a particular environmental good (i.e., a river, a watershed) in one specific location.

Campbell et al. (2009) explored the spatial variation of choice experiments estimates with the application of spatial kriging methods to interpolate information from individual specific WTP estimates for landscape improvements across the Republic of Ireland. These authors found that WTP values for a rural landscape improvement scheme are not evenly distributed as they vary across the country. In a related study, Campbell et al. (2008b) examined the spatial dependence of individual specific WTP values for the landscape improvement scheme and found that the values are not spatially uniform, but rather are globally clustered.

The present study examines the effects of distance of the location of residence of respondents with respect to large planted forests which can be found in many different areas in the country. In addition to collecting data on personal characteristics, an approach was also developed to locate the geo-referenced spatial coordinates of each respondent's place of residence. Respondents' existing addresses in the database were first verified using New Zealand Post's address-postcode-finder. Once confirmed, specific latitude and longitude coordinates for all addresses were found using the web site http://stevemorse.org/jcal/latlon.php which uses GoogleMaps to identify coordinates. The advantage of this technique is that we did not need to use a Global Positioning System (GPS) data logger to locate the coordinates nor ask each respondent to report his/her coordinate. We simply used the physical addresses of respondents on the mailing list to identify the coordinates.[10] Spatial coordinates of several online respondents were not located because of the absence of accurately verified address as the White Pages did not have their complete address.

Given that there are multiple sites with large planted forests, we developed a method in collaboration with geo-spatial analysts, where the geo-spatial coordinate of each respondent was used to create geographical buffers. Using a digital layer of the New Zealand map, geographical buffers with radius of 10, 50 and 100 km were created using ArcInfo © 9.10 and the programming language Python 2.6. Using a second digital layer that contains the New Zealand Land Cover Database version 2 (please see MfE, 2011), each buffer was intersected

---

[10] We initially thought of sending the GeoBatch link to online respondents and request them to report the coordinates on the survey. However, this option was not employed as some online respondents might get suspicious as to why we are tracking their exact locations, which might adversely affect the response rate.

with the sum of the area of planted forests, thus enabling the identification of planted forest areas around each geo-spatial coordinate. A further step was taken to consider that threatened native species could only establish themselves in large forests. To determine those large forest areas, contiguous planted forests of more than 5000 ha were aggregated and all the other scattered forests were ignored and this procedure created the final buffer intersections. We used the area of large planted forests derived from these final buffer intersections to create the spatial variables that we used as spatial covariates in the random effects regression model. In addition to these spatial covariates, we also included other covariates collected from the survey such as socio-economic characteristics, attitudes, affiliation to further explain the variation in calculated individual specific means of marginal WTPs.

Results from this analysis may be useful for policy decision makers involved in the formulation and implementation of afforestation schemes to provide insights as to how WTPs are influenced by distance and socio economic characteristics of people in nearby or faraway communities. One important measure is price elasticity of demand for biodiversity in planted forests. Estimating the price elasticity would answer the question, *if an existing large planted forest area (i.e., at least 5000 hectares in size) is situated less than 10 kilometres away from an individual, would this increase his/her WTP for biodiversity enhancement?* The answer to this question would be useful for the planning of one of the country's proposed afforestation scheme as described in Watt et al. (2011).

3.4 Models

The models we estimate in this chapter are the conditional logit (CL) model, the latent class logit model (LCM), the random parameters logit (RPL) model, the random parameters logit model with error components (RPLEC) and the panel data random effects regression model. To estimate the final utility coefficients that were used to simulate the willingness to pay values, RPLEC has been used. From RPLEC, we simulate individual specific estimates of WTP which allows us to generate a new panel data set. The constructed panel data set has a 10-period panel which is used to estimate a panel random effects regression model where we identified the determinants of willingness to pay of respondents.

3.4.1 Random parameters logit (RPL) model

We have described CL and LCM in Chapter 2 of this thesis. We therefore start by describing the RPL model. The RPL model (also known as mixed logit model) provides computationally practical and flexible econometric approach for discrete choice models derived from random utility maximisation (McFadden and Train, 2000). RPL overcomes major limitations of the basic conditional logit model by (1) taking into account that different individuals have different taste intensities or preferences; (2) allowing unrestricted substitution patterns; and (3) accounting for correlation in unobserved factors (Train, 2003, 2009; Hensher and Greene, 2003). The RPL approach relaxes the strong assumption of independent and identically distributed (i.i.d.) error terms, which corresponds to the behavioural property of independence of irrelevant alternatives (IIA) (Revelt and Train, 1998). The consequence of assuming that error terms are distributed i.i.d. is that it does not allow for the error components of different alternatives to be correlated. To account for this correlation, the unobserved portion of utility (i.e., error components) is partitioned into two additive terms where one term is heteroskedastic and correlated over alternatives (η) while the other is i.i.d. over alternatives (ε) as show in Equation 3.1

$$U_{njs} = \beta X_{\boldsymbol{njs}} + [\eta_{\boldsymbol{njs}} + \varepsilon_{njs}] \tag{3.1}$$

where η is the first random term with zero mean and with distribution over individuals and alternatives depends on underlying parameters and observed data relating to respondent *n* selecting alternative *j* in choice set *s*; ε is the second random term that is i.i.d. extreme value Type I distributed (Hensher and Greene, 2003). The η may be assumed *a priori* to have a particular distribution, which can be assumed to be normal, lognormal, truncated normal, triangular, Weibull and exponential (or any other). Assuming normal and lognormal distributions can be problematic as the former is sensitive to “wrong” signs (e.g., positive cost coefficient) while the latter exhibits fat tails (Train and Weeks, 2005). These properties are relevant to the current study of valuing biodiversity enhancements where taste intensities are expected to be positive for various improvements from the status quo. After evaluating the estimates from a number of specifications and distributional assumptions, we found the bounded triangular distribution as described in Hensher et al. (2005) was the most appropriate approach for this exercise. We employed an RPL model with a panel specification that facilitates the estimation of individual specific WTPs (Train, 2009).

3.4.2 Error components RPL model

Although the RPL model, as mentioned above, accounts for individual heterogeneity, it still does not account for status quo effects. In CE for environmental valuation, a typical choice task consists of three alternatives: a status quo alternative (hereby called SQ) that serves as the reference point (e.g., current condition) and is held fixed across all choice tasks; and two alternatives depicting a scenario different from the status quo with attribute levels generated from an experimental design. Respondents are likely to consider the SQ utility in a systematically different manner from the utility associated with the designed alternatives because SQ is experienced while the designed options are hypothetical (Scarpa et al., 2005). The utilities derived from the two designed options would likely be more correlated between themselves than with the utilities derived from SQ. This correlation structure can be accounted for by specifying a RPL model with additional errors that consider the difference in correlation across utilities (Herriges and Phaneuf, 2002). Specifying this RPL model with the additional error component addresses SQ effects as reported in previous studies (e.g., Samuelson and Zeckhauser, 1988; Haaijer, 1999; Haaijer, et al., 2001). Given the three alternatives for each choice task in this choice experiments exercise, the error component model may be specified as:[11]

$$U(sq) = Asc + \beta X_{sq} + \varepsilon_{sq} \quad (3.2)$$

$$U(c1) = \beta X_{c1} + \gamma_{c1} + \varepsilon_{c1} \quad (3.3)$$

$$U(c2) = \beta X_{c2} + \gamma_{c2} + \varepsilon_{c2} \quad (3.4)$$

---

[11] In equations 3.2 to 3.4, we excluded the subscripts *njs* that are shown in Equation 3.1 for a parsimonious presentation. Subscript *j* represents the j^{th} alternative which can either be *sq*, *c1*, *c2*.

where β represents a vector taste parameters for biodiversity enhancements which can either be random or fixed; *Asc* accounts for the systematic effect on SQ; ε represents the unobserved component of utility with extreme value Type I distribution; and γ is a normally distributed error component with zero mean that applies to changed alternatives $c1$ and $c2$. An important feature of γ is that it allows flexible patterns of substitution through an induced correlation structure across utilities amongst designed alternatives (Scarpa et al., 2005; Scarpa et al., 2007).[12] The RPL error component (RPLEC) logit model may be considered as an analogue of the nested logit model as it allows for correlation of utilities across alternatives in the same nest but different correlation across nests. But unlike nested logit, the RPLEC model relaxes the IIA assumption within alternatives of the same nest. For this exercise, we employed the RPLEC with panel specification which implies that the additional error component is the same across the choices made by the same individual (Scarpa et al., 2005).

3.4.3 Panel random effects regression models

To determine the factors that influence the individual means of the marginal WTP estimates we employed the panel random effects regression models. The formula for simulating individual specific marginal WTPs used here was proposed in Greene et al. (2005) and was applied in Scarpa et al. (2011b). The random effects models can address two important problems of cross-sectional data analysis, namely: unobserved heterogeneity and omitted variable bias.[13] Panel models can

---

[12] Note that η in Equation 3.1 is different from γ in Equations 3.3 and 3.4. The η accounts for the correlation of all three alternatives while γ is an additional error term of the RPL model that induces the correlation amongst changed alternatives *c1* and *c2*.

[13] More information on panel data analysis can be found in Baltagi (2008) and Greene (2008).

control for unobserved heterogeneity by accounting for it as having either a fixed or a random effect. In this study, random effects models are used to account for the correlation of WTPs across attributes for each respondent. Random effects models also allow researchers to control for certain types of omitted and unobserved variables where omitted variables are treated to be different between respondents but constant across different biodiversity enhancement outcomes. The panel random effects approach employed here is described in Campbell (2007).

Panel data models can account for systematic group effects. In this exercise, we create a variable that contains a 10-period panel of individual specific WTP estimates for the 10 biodiversity enhancement attribute levels (e.g., respondent *n*'s WTP for a level one increase in the number of falcons sighted). This represents the dependent variable, W_{na}. The panel model can be specified as:

$$W_{na} = \alpha_n + \gamma' A_{na} + \beta' X_n + \delta' S_n + \varepsilon_{na} \qquad (3.5)$$

where W_{na} represents a panel vector of respondent-specific means of marginal WTPs for attribute level *a* for respondent *n*, α_n represents independent random variables with constant mean and variance, A_{na} is a vector of indicator variables for $k-1$ attribute levels, X_n represents a vector of socio-economic characteristics, attitude and indicator of affiliations reported by respondent *n*, S_n is a vector of the natural log of areas of large planted forests included within a particular unit of radius from respondent *n* (e.g., 10 km radius, between 10 and 50 km radius, between 50 and 100 km radius).

The dependent variable W_{na} is derived from the estimates of the best fitting model where we calculated the respondent-specific means of the marginal WTP distributions for each of the 209 respondents and for each of the 10 attribute levels. We then regress those values on socio-economic, attitudinal, affiliation and spatial covariates in the form of a 10-period panel to account for dependencies of the values from the same respondent. Unlike the W_{na} that varies within a respondent, the four groups of covariates mentioned above are fixed for the 10-period panel for each respondent. For example, the indicator for gender of respondent n remains the same for the 10-period panel. To control for the effect of the 10 different WTPs for particular attribute levels, we included a vector of $k-1$ indicator variables with one attribute level serving as reference to avoid the dummy variable trap.

We start with the fixed effects regression followed by the random effects to explore the determinants of the means of the marginal WTPs. This is to identify patterns of sensitivity of estimated individual means of marginal WTP to both personal and spatial characteristics of respondents. This study therefore attempts to enhance the theoretical validity of the hypothetical survey on the value of improving existing biodiversity services in planted forests.

3.5 Data summary

Data used for discrete choice analysis consists of 209 respondents who provided valid responses to the choice experiments questions. Choice responses from these 209 respondents provided 1850 choice observations which are referred to in Chapter 2 as the *full data set*. Table 3.2 shows that the sample of respondents is slightly biased towards the high income group as demonstrated by the 34% of the respondents having an income above $100,000 while the New Zealand population averaged 22%.

Table 3.2: Income distribution of respondents versus the national proportion

Household income	Proportion of Respondents	National Proportion reported by Statistics New Zealand
$20,000 or less	8%	9%
$20,001–$30,000	11%	13%
$30,001–$50,000	17%	20%
$50,001–$70,000	12%	19%
$70,001–$100,000	17%	19%
$100,001 or more	34%	22%
TOTAL	100%	100%

Table 3.3 presents a summary of the socio economic and attitudinal characteristics of the 209 choice respondents. About 44% of the respondents had tertiary or post graduate education while 64% were female. These proportions are slightly higher compared to the national proportions of 40% for higher education and 51% for female. A small proportion of respondents reported they were volunteers to conservation organizations such as Forest and Bird and New Zealand Government's Department of Conservation (DOC). One out of five of the respondents wanted to include the Tui bird, a popular non-threatened native bird, in the choice tasks. Respondents were asked about their attitude toward supporting the proposed biodiversity programme and we found that one out of five had a "Government-should-pay" attitude. As respondents were provided with a description of the proposed programme and a walk-through of how to select the preferred alternative in each choice task, we asked each respondent to rate his/her level of understanding of the choice questions after completing the nine choice tasks. Twenty-one percent of the respondents gave a rating of 10 indicating that

they completely understood the choice questions. Less than half (42%) gave a rating between five and seven while 11% gave a rating of four and below.

Table 3.3: Summary statistics of socio-economic and attitudinal covariates

Item	Percentage of respondents	SNZ National Proportion
Completed higher education	44%	40%
Female	64%	51%
Forest and Bird member	8%	--
DOC volunteer	3%	--
Tui should be in the choice set	21%	--
Government should pay	18%	--
Self-rated understanding of CE questions ("10" represents "completely understood" and "1" represents "did not understand at all")		
– 8 to 10	47%	--
– 5 to 7	42%	--
– 1 to 4	11%	--

Table 3.4 presents a summary of the spatial variables used as covariates in the random effects panel regression analysis. We located the geo-spatial referenced coordinates of only 115 choice respondents. We did not find the coordinates for the other 94 choice respondents due to several reasons which include insufficient details of the address provided and respondents lived in a very rural area. Of the 115 spatial choice respondents, 28 (24%) were found to be situated less than 10 spatial kilometres away from large planted forests with an area of at least 5,000 hectares.[14] Using ESRI ArcMap 10.0, we intersected a 10-kilometre radius for each respondent's spatial coordinate with the digitally mapped areas of large planted forests in New Zealand. The intersection resulted in the identification of forest areas included in the 10-kilometre radius of each of the 28 respondents with areas ranging between 17 and 14,000 hectares. About half of the respondents lived

[14] We assumed that 5,000 hectares of planted forests would be sufficiently large to provide habitat to threatened native species like the falcon and other native species (e.g., bush robins, native fish). In addition, large planted forests also form a landscape view where people would be able to recognise their presence. Furthermore, a large forest area would likely provide connectivity between areas with native forests therefore contributing to biodiversity service.

in areas with large planted forests situated between 10- and 50-kilometre radii. While 25% of respondents lived in properties where one could find large planted forests between 50 and 100 km of spatial distance.

Table 3.4: Summary statistics for the three spatial covariates

Spatial Covariate	Area of planted forests within the radius (hectares)	Number of respondents (% of 115 respondents with spatial coordinates)
10-km radius		28 (24.4%)
– Average	3,936	
– Minimum	17	
– Maximum	14,000	
Between 10- and 50-km radius		58 (50.4%)
– Average	40,175	
– Minimum	1,900	
– Maximum	220,000	
Between 50- and 100-km radius		29 (25.2%)
– Average	62,334	
– Minimum	6,200	
– Maximum	770,000	

The 10, 50 and 100 km buffers in Table 3.4 above were chosen to represent different types of visits that would have an implication on the use value of planted forests. Respondents living within the 10 km buffer would be able to visit the planted forest either by bicycle or a short drive. Those living within the 50 km buffer but beyond 10 km would likely be able to make a day trip. Those beyond the 50 km buffer but within the 100 km buffer would be at the border of a one day trip and might require a place a spend the night.

While we accounted for the distance of planted forests to respondents, we elected not to study in detail the impacts of some recreational attributes as we mentioned at the outset that this study focuses on valuing biodiversity enhancement. As some planted forests like Whakarewarewa in Rotorua and

Woodhill in Auckland offer recreational facilities such as walking trails, mountain biking tracks, tree adventures and horse riding tracks (Dhakal et al., 2012; Auckland Council, 2011), we opted not to collect those data for this study. However, I plan to examine those attributes spatially in my future research work at Scion which include valuing avoided erosion and the economics of tangibles and intangibles.

3.6 Results

3.6.1 Logit models

We analysed the full choice data set with 1850 choice observations using four logit models. Model 1 is the basic conditional logit (CL) model where estimated coefficients for both environmental and cost attributes demonstrate the expected signs (i.e. negative sign for cost, positive sign for marginal utilities) (Table 3.5). The coefficient for the indicator for SQ is positive but not significant. This might indicate that there is no additional utility associated with the status quo over and above that associated with its attribute levels. The conditional logit model imposes the restrictive IIA assumption which assumes that all respondents have the same preference (Greene and Hensher, 2003). To relax the assumption of preference homogeneity, we estimate Model 2 which is a Latent Class logit model with panel specification (PLCM) that assumes that a sample of respondents would have different types of preferences that can be grouped into latent classes or class memberships (Heckman and Singer 1984; Greene and Hensher, 2003). Following the panel latent class framework in Scarpa et al. (2009) we estimate the class memberships of different groups of respondents based on the attributes that they likely did not attend to. Under this approach, the coefficients of those particular attribute levels are constrained to zero following a latent class logit model framework. Initially, 20 latent class logit models with different latent classes were estimated. The PLCM estimates reported in Model 2 of Table 3.5 has the best model fit among the 20 specifications tested as indicated by the lowest normalised AIC of -1.154. Appendix Table 1 shows the AICs of the 20 latent class model specifications distributed into two groups: 10 cross-section and 10 panel specifications. The panel type specification exhibited lower normalised AIC values not exceeding 1.28 compared to the higher AICs of cross sectional models (always above 1.93). The cross section specification of LC Model 5 has the highest normalised AIC (worst model fit) while the panel LC Model 5 has the lowest AIC (best model fit). The improvement in model fit in the panel specification (versus cross-section) suggests that it is important to take into account that the process of evaluating choice task *s* by respondent *n* is correlated to the way one evaluates other choice tasks in the series assigned to him/her.

Going back to the coefficient estimates in Model 2, the coefficient for the SQ indicator is negative and significant, indicating that a typical respondent derives more utility by choosing the changed alternative or enhanced biodiversity level than the current condition. In terms of non-attendance to choice attributes, Model 2 results suggest that about one-third of the respondents ignored the cost attribute. This is consistent with Scarpa et al. (2009) which report that cost is one of the most non-attended attribute in a choice task. Model 2 estimates also suggest that 37% of the sample ignored the SQ option, 23% did not attend to less iconic species (i.e., kokopu (native fish), kakabeak (native plant) and gecko), while 6% ignored all attributes. Non-attendance to SQ is consistent with the negative coefficient for the indicator of SQ, which means that respondents on average would be more satisfied with the creation of a new programme that aims to enhance the level of biodiversity in planted forests. While more than half of the respondents have still attended to the SQ option in selecting the preferred alternative in a choice task, about 37% might have simply ignored it because they elected to focus more on the changed options that presented combinations of mostly higher attribute levels while attribute levels in SQ are fixed and represent the lowest attribute level.

Although Model 2 provides interesting results regarding attribute non-attendance, the coefficients for the changed levels of the less iconic species are no longer significant while in Model 1 the coefficients for Native Fish 1 and Native Plant 2 are positive and significant. To address this issue, we employed the Random Parameters Logit (RPL) model with panel specification that accounts for individual heterogeneity. Compared to LCM, RPL takes into account that each respondent has a unique set of preferences for the environmental good in question.

Before finally settling with the RPL specification reported in Model 3, we first determine the superior RPL model specification from among a series of preliminary RPL models. First we ran an RPL model where we assume that all utility coefficients are random. From there we identify which random coefficients would likely have significant effects to utility by running different model specifications. Out of the more than 20 different specifications tested, we identified four random parameters to have significant effects. We also identified the suitable distributional assumption of these four random parameters by testing on normal, log normal, uniform and triangular distributions. The triangular distribution resulted to the best model goodness of fit (i.e., highest log-likelihood value) with the log-likelihood value of -1034.95 compared to -1035.93 for normal,

-1089.23 for uniform, and -1811.48 for log normal. The selection of the triangular distribution is also based on Train (2009 p. 138) which states that “The triangular distribution has positive density.. … taking the form of a tent. … These densities have the advantage of being bounded on both sides, thereby avoiding the problem that can rise with normals and log normals having unreasonably large coefficients for some share of decision makers.” The triangular distribution is also known for avoiding the allocation of shares to extreme values of coefficient, which is a drawback in other distributional assumptions, such as in the normal or log-normal.

From the above series of tests for RPL, we settled with a preferred RPL model with four random parameters that are assumed to have triangular distributions. The cost parameter is assumed to have a constrained triangular distribution wherein it has been constrained to be negative with an upper limit of zero. The three other random parameters were not constrained. The standard deviations of the four random parameters are significant at the 5% level indicating taste heterogeneity. The coefficients for changed attributes for Native plant 1, Native plant 2 and Native fish 1 are positive and significant while the coefficient for Gecko remains not significant similar to Model 1. However, the RPL panel model in Model 3 does not account for SQ effects. Thus we employed Model 4 to induce the correlation amongst SQ and the changed alternatives as described in Scarpa et al. (2006). Estimates for Model 4 indicate a strong correlation between the two changed alternatives as indicated by the coefficient for the error components being positive and significant. Model fit significantly improves with the addition of error components as exhibited by the significant increase in the log-likelihood value from -1,035 in Model 3 to -991 in Model 4. Although the *p*-values in Model 4 remained virtually the same as in Model 3, the magnitude of the coefficient for cost in Model 4 is significantly lower, while the coefficients for the environmental attributes remain virtually the same. This translates to higher WTP for biodiversity enhancement. This implies that the RPL panel with error components model better allows the accounting of changes in attribute levels by inducing the correlation between the changed alternatives. We therefore used Model 4 estimates to simulate the median WTP values that we subsequently used to compute for the aggregated WTP as described in the next section. We also used Model 4 to estimate the individual specific WTP values that we included in the construction of the panel data to identify the determinants of WTP.

As an aside, it is also important to mention that estimates of utility coefficients for native fish levels 1 and 2 in Table 3.5 demonstrate pattern of lack of insensitivity to scope. This is indicated by complete insensitivity to higher levels of fish protection. Lack of sensitivity to scope (sometimes called as *lack of sensitivity to scale*) has been identified as a potential issue in contingent valuation and in choice experiments (Ryan and Wordsworth, 2000; Foster and Mourato, 2003; Goldberg and Roosen, 2007; Rolfe and Windle 2010). Although we are aware that this chapter is more policy orientated, in that the issue of insensitivity to scope may not be important (Ryan and Wordsworth, 2000), we still attempted to address this issue by using a non-linear coding approach called *piecewise linear coding* (PWLC). PWLC captures the sensitivity within the intervals as well as enforces continuity and weak monotonicity of the utility function (Bierlaire, 2008).

PWLC is different from dummy coding. In dummy coding, we can assign two dummy variables (e.g., Kiwi1, Kiwi2) per attribute with three levels. For status quo these two variables can be assigned respective values (0,0), for level 1 increase in abundance (1,0), for level 2 increase (0,1). From dummy coding, one can create PWLC variables by using the Kiwi1 and Kiwi2 to generate a new variable which is defined as Kiwi1a = Kiwi1 + Kiwi2. From here the analyst would now need to use Kiwi1b and Kiwi2 for each three level attribute. However, the interpretation of estimates from PWLC is different because Kiwi1a is now also shared by Kiwi2.

Appendix Table 2 shows the estimates for Models 1 to 4 where we used PWLC. Comparing the estimates for Model 4 with PWLC to Model 4 with dummy coding in Table 3.5, the coefficients for Brown kiwi 2 and Bush falcon 2 (of the former) are no longer significant. We interpret the PWLC coefficients for level 1a (e.g., Brown kiwi 1a) differently compared to dummy coding because Brown kiwi 1a now relates to both levels (Brown kiwi 1 and Brown kiwi 2). Results from PWLC show a pattern of a big jump with level 1a of improvement (e.g., Brown kiwi 1a) while the coefficient for the second level (e.g., Brown kiwi 2) are never significant meaning that they do not produce further benefits. These results are actually more consistent with utility theory in economics. Although model estimates in Table 3.5 using dummy coding provide better statistical estimates, estimates in Appendix Table 2 provide better results from the stand point of economic theory of utility.

Table 3.5: Estimates of logit models (n = 1850 choice observations)

Item	Model 1 Conditional Logit			Model 2 Latent Class Logit Panel			Model 3 Random Parameters Logit Panel			Model 4 Random Parameters Logit Panel with Error Components		
<u>*Attributes and SQ*</u>	Coef	Std err	*p*-value	Coef	Std err	*p*-value	Coef	Std err	*p*-value	Coef	Std err	*p*-value
Brown kiwi 1	**0.504**	**0.098**	**<0.01**	**0.669**	**0.121**	**<0.01**	**0.921**	**0.141**	**<0.01**	**0.898**	**0.137**	**<0.01**
Brown kiwi 2	**0.622**	**0.095**	**<0.01**	**0.818**	**0.123**	**<0.01**	**1.094**	**0.138**	**<0.01**	**1.048**	**0.128**	**<0.01**
Native fish 1	**0.287**	**0.093**	**<0.01**	0.163	0.131	0.21	**0.330**	**0.134**	**0.01**	**0.307**	**0.153**	**0.04**
Native fish 2	0.143	0.095	0.13	0.024	0.138	0.86	0.201	0.138	0.15	0.138	0.145	0.34
Native plant 1	0.145	0.094	0.13	0.181	0.136	0.18	**0.348**	**0.138**	**0.01**	**0.343**	**0.163**	**0.04**
Native plant 2	**0.210**	**0.094**	**0.03**	0.129	0.130	0.32	***0.299***	***0.143***	***0.04***	***0.329***	***0.161***	***0.04***
Green gecko 1	0.017	0.093	0.86	-0.115	0.135	0.40	-0.047	0.139	0.74	-0.053	0.135	0.70
Green gecko 2	0.092	0.093	0.32	-0.061	0.139	0.66	*0.003*	*0.164*	*0.99*	*0.124*	*0.159*	*0.43*
Bush falcon 1	**0.453**	**0.098**	**<0.01**	**0.476**	**0.120**	**<0.01**	**0.860**	**0.145**	**<0.01**	**0.909**	**0.147**	**<0.01**
Bush falcon 2	**0.700**	**0.094**	**<0.01**	**0.914**	**0.122**	**<0.01**	***1.178***	***0.153***	***<0.01***	***1.188***	***0.147***	***<0.01***
Status Quo Indicator	0.177	0.158	0.26	**-5.864**	**0.504**	**<0.01**	**-3.767**	**0.318**	**<0.01**	**-1.594**	**0.637**	**0.01**
Cost	-0.025	0.002	<0.01	**-0.123**	**0.011**	**<0.01**	***-0.169***	***0.013***	**<0.01**	***-0.063***	***0.004***	***<0.01***
<u>*Attribute non-attendance*</u>												
Ignoring cost				**0.347**	**0.133**	**0.01**						
Ignoring status quo				**0.369**	**0.158**	**0.02**						
Ignoring non-iconics				**0.227**	**0.060**	**0.00**						
Ignoring all attributes				**0.057**	**0.020**	**0.00**						
<u>*Random Parameters*</u>												
Bush falcon 2							**1.755**	**0.500**	**<0.01**	**1.606**	**0.658**	**0.01**
Native plant 2							**1.244**	**0.538**	**0.02**	**1.446**	**0.557**	**0.01**
Cost							**0.320**	**0.023**	**<0.01**	**0.063**	**0.004**	**<0.01**
Green gecko 2							**2.197**	**0.493**	**<0.01**	**1.369**	**0.520**	**0.01**
Error Component										**7.652**	**1.005**	**<0.01**
Log-likelihood	-1785.14			-1052.57			-1034.95			-990.68		
Normalised AIC	1.943			1.154			1.136			1.088		
McFadden Pseudo R^2	0.116			0.482			0.491			0.512		

Note1: Values in *italics* represent coefficient estimates for random parameters.
Note2: Values in **boldface font** represent estimates statistically significant at 5% level.

Another issue worth mentioning is that in Chapter 2 (on Pages 41 and 42 of this thesis), we can divide the full sample into two regional groupings. We got 66% of the full sample from Group 1 regions (regions with a large proportion of planted forests) and 34% from Group 2 regions (regions with a small proportion of planted forests). We therefore estimate Model 4 in Table 3.5 using split samples of Groups 1 and 2 respondents. We present these model estimates in Appendix Table 3.

Both Groups 1 and 2 sub-samples appear to value levels 1 and 2 increases in abundance of brown kiwi and bush falcon in planted forests. However, it appears that only the group who lived in regions with a larger proportion of planted forests, controlling for other factors, would have a significant increase in utility relative to status quo if a biodiversity programme would be implemented. This might indicate that the presence of more forests or greater accessibility to planted forests could be an important factor for consideration if there is a need to prioritise regions for biodiversity enhancement.

However, despite Appendix Tables 2 and 3 provide estimates that address insensitivity to scale as well as heterogeneity across regional groupings, we lost the statistical significance for non-bird attributes. Compared to Model 4 in Table 3.5 with nine utility coefficients being statistically significant, in the Appendix Table 2 only four utility coefficients are statistically significant, while for the split samples in Appendix Table 3, only five or six coefficients are significant. As Model 4 results using the full sample in Table 3.5 demonstrate superior statistical properties we elect to use these estimates in the simulation of marginal WTPs and aggregation.

3.6.2 Median and aggregate WTPs

As Model 4 in Table 3.5 provides the best model fit as indicated by the lowest normalised AIC and highest Pseudo R^2 among the models, we simulated the median WTP (or the 50^{th} percentile) for each attribute and the corresponding confidence intervals around the median using Monte Carlo simulation with 10,000 random draws following the R-code described in Thiene and Scarpa (2009). We simulated median WTPs instead of mean WTPs because the latter cannot be simulated when the cost coefficient is distributed log-normal, triangular or constrained triangular as suggested in Daly, et al. (2011). In that recent paper, it was mentioned that "the moments of the WTP distribution might not exist for a given distribution of the cost coefficient" which include the constrained triangular distribution that we used in this present study. For this present study, simulated median WTPs suggest that the two most valued attribute levels are level 2 increases in Falcon ($19/year) and Brown Kiwi ($17/year) (Table 3.6). We also report the 95% confidence interval around the median WTP for more falcons ($17 to $21) and kiwis ($15 to $18) that we derived from the simulation exercise. A level 1 increase in the number of endangered native plant kakabeak and the native fish kokopu were also valued at around $5/year for five years through an addition in the amount payable to income tax.

Table 3.6 also shows the total WTP of a typical respondent for increasing the abundance of threatened native species is approximately $65 per year for five years. This total WTP also includes the willingness to pay for the development of a new forest biodiversity programme of approximately $25 per year. If the sample of respondents represented the New Zealand population of taxpayers, we could simply multiply the total WTP per year by the total number of people who pay their annual income tax. However, that would be biased. In addition, the elicitation method used was based on a hypothetical market. Therefore, in aggregating the WTP values to the national level, we accounted for two of the major sources of biases as mentioned in the stated preference literature. These are hypothetical (List and Gallet, 2001) and aggregation (Morrison, 2000) biases. List and Gallet (2001) suggest that hypothetical WTP would likely be two to three times as the actual WTP. To account for this bias we divided the hypothetical WTP by two.[15] Table 3.6 shows the discounted WTP values for each attribute where the WTP for the level 2 option decreased to about $33 per year for five years. Morrison (2000) suggests that one way to address aggregation bias is to consider survey non-responses to have WTP of zero. For every 100 surveys we sent out, 43 completed surveys returned, thus a response rate of 43%. We therefore assumed that 57% of the New Zealand taxpayers have a WTP of zero. As of 2006, New Zealand had a total of 3 million taxpayers. As we have a response rate of 43%, we multiply 3 million by 43% of the total taxpayers which is 1.29 million. We have also taken into account that in our sample of 209 respondents, only 64% of them were taxpayers as other people were retired, students or homemakers. From this sample proportion, we multiply 1.29 million taxpayers by 64% which results to 825,600 willing taxpayers.

Overall, the aggregated WTP values for a Level 2 increase in threatened species amounted to a national value of about $26.5 million per year for five years

[15] As we have included a cheap talk script in the description of the choice experiment scenario, we assumed that the script addresses the upward WTP bias. We therefore opted to multiply the hypothetical WTP by 0.50 instead of 0.33.

(with present value of this five-year annuity at more than $100 million)[16]. It is envisioned that the amount will be used to fund the proposed five-year biodiversity enhancement programme only and will not cover administration fees. If the biodiversity enhancement programme managers opted to focus only on the more visible and more popular bird species (i.e., brown kiwi and bush falcon), the national WTP value for a level 2 increase corresponds to about $22.3 million per year for five years.

---

[16] Present Value (PV_A) of a five year annuity was calculated using the formula $PV_A = A \cdot \left[\left(1 - \left((1+i)^n\right)^{-1}\right) \cdot i^{-1}\right]$ where A represents the aggregated annual WTP, i is the annual interest rate of 8 percent, and n is the number of years (n = 5 years).

Table 3.6: Simulated WTP estimates from Model 4 and aggregated to the national level

Threatened Species	**All Attributes (in NZ$/year for five years)**		**Bird Species Only (in NZ$/year for five years)**		**Non-bird species (in NZ$/year for five years)**	
	Level 1	*Level 2*	*Level 1*	*Level 2*	*Level 1*	*Level 2*
Brown kiwi	14.26 (12.46 – 16.10)	16.64 (14.89-18.45)	14.26 (12.46 – 16.10)	16.64 (14.89-18.45)	-	-
Giant Kokopu	4.87 (3.00-6.78)	NS	-	NS	4.87 (3.00-6.78)	NS
Kakabeak	5.45 (3.45-7.48)	5.22 (3.25-7.23)	-	-	5.45 (3.45-7.48)	5.22 (3.25-7.23)
Green gecko	NS	NS	NS	NS	NS	NS
Bush falcon	14.43 (12.52-16.39)	18.86 (16.85-20.93)	14.43 (12.52-16.39)	18.86 (16.85-20.93)		
Indicator for changed alternative	25.28 (17.36-33.46)	25.28 (17.36-33.46)	25.28 (17.36-33.46)	25.28 (17.36-33.46)	25.28 (17.36-33.46)	25.28 (17.36-33.46)
					-	-
T O T A L	**64.29**	**66.00**	**53.97**	**60.78**	**35.60**	**30.50**
Discounted WTP per taxpayer (50%) - to account for hypothetical bias	32.15	33.00	26.99	30.39	17.80	15.25
Aggregated annual WTP (multiplied by 825,600 willing NZ taxpayers)	**26,538,912**	**27,244,800**	**22,278,816**	**25,089,984**	**14,695,680**	**12,590,400**
Present value of five annual payments (PV_A)	**105,962,180**	**108,780,586**	**88,952,852**	**100,177,031**	**58,675,589**	**50,269,816**

Note 1: “NS” means not significant at the 95% confidence level.
Note 2: Figures in parentheses represent simulated 95% confidence intervals of median WTP.

3.6.3 WTP determinants

Using the specification in Model 4 in Table 3.5, which is the RPLEC panel model, we estimated individual-specific means of the conditional distributions of marginal WTPs. As each of the 10 means of WTP for each respondent would likely be correlated, we used panel random effects regressions to explain patterns of variation. In the set of explanatory variables of the random effects models, we included indicator variables for k–1 changed attribute levels to control for WTP variation within each respondent. We explore the role of socio-economic, attitudes and spatial characteristics of each respondent on individual specific WTP values. Table 3.7 presents the estimates for three panel random effects regression models. Model A includes socio-economic and attitude covariates. It has 161 respondents each with 10-period panel (or observed 10 times). The reduction from the full sample size of 209 to 161 respondents is because not all respondents provided data on these covariates. Because we have mainly employed a mail survey, some respondents did not report their highest educational attainment. Others did not complete page 18 of the questionnaire which could have provided the data as to whether they were DOC volunteers or Forest and Bird member. This created gaps in the socio economic data that resulted to the exclusion of 48 respondents in the panel random effects regression for Model A.

Model B has the same set of covariates as Model A but with a reduced sample size that matches the sample size in Model C. As mentioned earlier, the spatial coordinates for some respondents were not located due to insufficient data on addresses of respondents especially those who completed the survey online whose home addresses were not verified. As a result, the sample size was reduced to 1110 observations for the model with spatial covariates.

Estimates in Models A, B, and C suggest that having tertiary education as highest educational attainment has a significantly positive effect of about $1.25 to $3.10 on median WTP. Similar to the results of the panel random effects regression in Campbell (2007), results from Model A suggests that being female positively influences WTP where a typical female respondent would contribute approximately $2.00 more than male. However, keeping the variable *Female* in Model B, which used a smaller sample, resulted to an insignificant coefficient estimate. Keeping the variable female in Model C resulted to the panel random effects regression model not converging as indicated by a note in NLOGIT saying "Error 249: Random effects. Did not find positive estimated component." We therefore dropped this variable in Model C.

Estimates from Model C indicate that being affiliated with or serving as volunteers for conservation institutions such as DOC and the environmental NGO Forest and Bird has even greater positive significance of approximately $9.16. As expected, having a negative attitude towards contributing a dollar amount for biodiversity (i.e., having a "Government should pay" attitude) significantly reduces WTP by about $3.00. Having a good level of understanding of the choice questions, from a self-rated scale of 1 to 10, with 10 as the highest rating, indicates that an increase in one level of understanding increases WTP by $0.41. Interestingly, respondents who would like to include the non-threatened but popular native bird Tui also demonstrate a positive marginal impact on median WTP of about $1.26 (significant at the 15% level).

The main highlight of the estimates in Table 3.7 is the coefficient estimates for the spatial covariates. Including the three spatial covariates in Model C significantly increases the log-likelihood value from -4233 in Model B to -4222 in Model C. The calculated likelihood ratio test statistic of 23.06 between the two models exceeds the critical chi square value of 16.27 at the 99.9% confidence level. Thus, the null hypothesis that individual specific WTPs for biodiversity enhancement are not a function of geo-spatial distance of respondents from large planted forests is strongly rejected.

Using the group of respondents who lived more than 100 kilometres away from large planted forests as reference, the coefficient for the 10 km radius suggests that a respondent living close to large planted forests would be willing to pay $0.17 more for biodiversity enhancement than a respondent living further away. This is consistent with the results from previous studies suggesting some form of distance decay in environmental use values (e.g. Bateman et al. 2006; Pate and Loomis, 1997). This is because, as mentioned earlier, a person situated within a 10-km radius can have better access to an enhanced forest biodiversity (biking distance or less than a five-minute drive to the site). This somewhat indicates that a respondent situated close to a planted forest has both use and non-use value for the enhancement of the resource.

Estimates in Table 3.7 indicate that the WTP of an individual living in a place with large planted forests beyond the 10-kilometre radius (i.e., 10- to 50-kilometre radius) does not increase with biodiversity enhancement. A possible reason for this is that people perceived that the potential to benefit is low as they live further away. People living within this range seem to be concerned more about option value than direct use value.

In terms of respondents living in places where planted forests had a spatial distance between 50 to 100 kilometre radii, they demonstrate a pattern of increase in WTP of $0.15 per respondent for biodiversity enhancement. This might indicate the presence of non-use value whereby respondents who reside very far away from planted forests would be willing to pay more by simply knowing that the habitat was enhanced to increase the abundance of threatened species even though they might not be able to visit those forest areas. However, the estimate for *Between 50- and 100-km radius* is significant only at the 81% confidence level and therefore statistically weaker than that for the *10-kilometre radius*.

Table 3.7: Panel random effects model parameter estimates

	Model A *Model with Socio-economic Covariates*			***Model B*** *Model with Socio-economic Covariates Reduced sample size*			***Model C*** *Model with Socio-economic and Spatial Covariates – Reduced sample size*		
	Coeff	Std Err	*p*-value	Coeff	Std Err	*p*-value	Coeff	Std Err	*p* -value
Indicator for attribute level									
Brown kiwi 1	**23.234**	**1.167**	**<0.01**	**24.399**	**1.483**	**<0.01**	**24.399**	**1.457**	**<0.01**
Brown kiwi 2	**26.907**	**1.167**	**<0.01**	**28.256**	**1.483**	**<0.01**	**28.256**	**1.457**	**<0.01**
Native fish 1	**8.799**	**1.167**	**<0.01**	**9.241**	**1.483**	**<0.01**	**9.241**	**1.457**	**<0.01**
Native fish 2	**4.657**	**1.167**	**<0.01**	**4.891**	**1.483**	**<0.01**	**4.891**	**1.457**	**<0.01**
Native plant 1	**9.677**	**1.167**	**<0.01**	**10.162**	**1.483**	**<0.01**	**10.162**	**1.457**	**<0.01**
Native plant 2	**8.842**	**1.167**	**<0.01**	**9.426**	**1.483**	**<0.01**	**9.426**	**1.457**	**<0.01**
Green gecko 2	**4.477**	**1.167**	**<0.01**	**4.839**	**1.483**	**<0.01**	**4.839**	**1.457**	**<0.01**
Bush falcon 1	**23.496**	**1.167**	**<0.01**	**24.674**	**1.483**	**<0.01**	**24.674**	**1.457**	**<0.01**
Bush falcon 2	**29.228**	**1.167**	**<0.01**	**30.472**	**1.483**	**<0.01**	**30.472**	**1.457**	**<0.01**
Socio-economic covariate									
Tertiary	**1.264**	**0.595**	**0.03**	**3.098**	**0.827**	**<0.01**	**2.393**	**0.851**	**<0.01**
Female	**1.963**	**0.553**	**<0.01**	0.535	0.728	0.46	-	-	-
Forest and Bird	**6.194**	**1.014**	**<0.01**	**8.324**	**1.192**	**<0.01**	**9.165**	**1.176**	**<0.01**
DOC Volunteer	**10.930**	**1.706**	**<0.01**	**9.012**	**1.841**	**<0.01**	**8.675**	**1.787**	**<0.01**
Understanding of CE questions	**0.310**	**0.114**	**0.01**	**0.344**	**0.148**	**0.02**	**0.410**	**0.149**	**0.01**
Tui should be in the choice set	**3.308**	**0.641**	**<0.01**	**1.838**	**0.799**	**0.02**	*1.257*	*0.798*	*0.12*
Government should pay	**-2.755**	**0.692**	**<0.01**	**-2.968**	**0.866**	**<0.01**	**-2.968**	**0.850**	**<0.01**
Constant	**-3.956**	**1.162**	**<0.01**	**-5.149**	**1.490**	**<0.01**	**-6.477**	**2.723**	**0.02**
Spatial Covariate									
Log of forest area in 10-km radius	-	-	-	-	-	-	**0.168**	**0.051**	**<0.01**
Log of area in 10-50 km radius	-	-	-	-	-	-	-0.049	0.098	0.62
Log of area in 50-100 km radius	-	-	-	-	-	-	*0.147*	*0.110*	*0.18*
Log-likelihood	-6023.50			-4233.07			-4221.54		
Pseudo R^2	0.513			0.517			0.527		
No. of observations	1610			1110			1110		

Note: Values in **boldface font** represent estimates statistically significant at 5% level.

3.7 Conclusions and policy implications

Results from data collected from 209 choice respondents across New Zealand indicate that a typical respondent values biodiversity enhancement in the country's 1.8 million hectares of planted forests, especially those in large planted forests. An aggregate WTP of NZ$26.5 million per year for five years would be paid through income tax to fund a proposed biodiversity enhancement programme that aims to increase the abundance of threatened species seen or heard in planted forests.[17] This study provides empirical evidence that New Zealanders would collectively be willing to financially contribute a considerable amount to biodiversity enhancement in planted forests. This extends previous study by Yao and Kaval (2010) that New Zealanders would be WTP for biodiversity enhancement on private land by demonstrating that even in exotic planted forests they still value habitat enhancement for threatened species. The estimated value may be useful not only for future policy decision making but also to satisfy the growing interest of large corporations in incorporating ecosystem services values in business plans (TEEB, 2010; WBCSD, 2011).

Both socio-economic and spatial factors are found to influence individual specific means of marginal WTP estimates. Those that significantly contribute positively to WTP are affiliation to conservation organisations, higher education, and having an appreciation of native birds. Those that contribute negatively to WTP include an attitude of reliance mainly on the government to fund the

[17] Perhaps a way to check for the robustness of the estimated WTPs is to include an option in the Internal Revenue Department's (IRD's) income tax return form where taxpayers would be provided an option to donate a portion of their tax refund to a briefly described biodiversity enhancement in planted forests.

proposed enhancement programme. These empirical results suggest the characteristics of relevant individuals (or maybe groups) to focus on in winning support for enhancing biodiversity in planted forests. Analytic results also indicate the enhancement of biodiversity provides more benefits those living within a 10-kilometre radius from large planted forests compared to those living further away. The above findings might be useful for a future study that seeks to identify an appropriate funding mechanism for biodiversity enhancement in planted forests.

We also acknowledge that due to the low response rate, we collected a small sample size 209 choice respondents. This sample size might be too small to calculate a national estimate of value for biodiversity enhancement. We suggest that future studies that would aim to come up with a national estimate of value would require more resources to allow the collection of a bigger sample size to estimate national WTP values from a more representative national sample.

Chapter 4: An investigation of experimental design criteria and their behavioural efficiency: entropy and attribute non-attendance

4.1 Introduction

The Choice Experiments (CE) method has been widely used to study behavioural responses in different fields which include transportation, health economics, marketing, energy, political science and environmental economics. Part of its wide acceptance can be attributed to the fact that it provides a theoretically valid framework that allows the examination of individual preferences. The framework allows an individual to reveal the tradeoffs amongst alternatives with different combinations of attribute levels in a choice task. A crucial component of CE is the systematic arrangement of attribute levels on each alternative in a choice task which is addressed by using a fraction of the full factorial, or the Experimental Design (ED).[18] A common approach in constructing an ED is the fractional factorial approach to generate an initial series of single alternatives that are then allocated to choice tasks using various methods which include randomised, cyclical, and Bayesian (Bunch et al., 1996; Sándor and Wedel, 2001, 2002, 2005; Kanninen, 2002; Bliemer and Rose, 2006).

The ED is usually optimised following a certain criterion chosen by the analyst.[19] One of the first ED criteria used for CE was the *orthogonality criterion*

[18] Chapter 2 of this thesis (on pages 26 to 31) provides an overview of the different statistical measures of design efficiency. This current introductory section of Chapter 4 focuses on how different experimental design criteria evolved to suit the needs of choice analysts.

[19] Some of the earlier choice tasks were generated by randomly populating the alternatives with the identified attribute levels. This criterion is referred to as random design criterion.

derived from linear multivariate models originally used for statistical analysis of treatment effects in biological experiments (Louviere and Woodworth, 1983; Louviere and Hensher, 1983). The orthogonality criterion generates fractional factorial designs that exhibit no correlation between each row of design attributes and/or between columns of alternatives.[20] One advantage of this criterion is that the analyst does not need any *a priori* knowledge of the parameter estimates and their distribution (or parameter mean and standard error). The analyst can generate an orthogonal design by simply knowing the number of attributes, alternatives and number of choice tasks per respondent for a CE exercise. However, orthogonality is a design property appropriate mainly for linear regression models (e.g., Ordinary Least Squares). Since CE data are analysed using non-linear regression models (e.g., logit) to examine changes in utilities, orthogonality is not necessarily a criterion for efficiency (Bliemer and Rose, 2006). Bliemer and Rose (2009 p. 21) demonstrate using a logit model that the statistical efficiency of an orthogonal design is relatively low compared with an efficient design. They find that the theoretically minimum required number of CE respondents, following a conditional logit model, with orthogonal design is seven times more compared to a Bayesian D-efficient design. The gain in statistical efficiency enables the analyst to reduce the required sample size and/or reduce the number of choice tasks. The former translates into a reduction in survey costs while the latter leads to lesser time required for respondents to complete the survey.

[20] Orthogonal designs are described in detail in Louviere, et al. (2000) and Hensher, et al. (2005). An electronic library of orthogonal designs is available at http://www2.research.att.com/~njas/oadir/

However, the Bayesian D-efficient Design (BDD) criterion requires reliable prior knowledge of parameter estimates which could come from a pilot survey (Ferrini and Scarpa, 2007). Given that analysts would not have enough time to collect *a priori* information but would still like to generate a design derived from logit model, he/she might opt to employ the Optimal Orthogonal Design (OOD) criterion in generating the ED. OOD criterion employs an algorithm that search through different EDs generated assuming that all parameter estimates (from a conditional logit model) are equal to zero (Street and Burgess, 2005; Sandor and Wedel, 2005). However, assuming a set of prior parameter values of zero might be too naïve because an analyst could easily access information about the priors from related studies. One could also readily assume that the sign of the parameter for the cost attribute to be negative and for obviously positive changes to be positive.

Given that analysts select a particular ED criterion depending on their objectives and specific situation, very limited studies have accounted for the behavioural impact of using different design criteria. In terms of statistical efficiency, previous studies compared different design criteria using simulations and found that designs following the BDD criterion provide more statistically efficient parameter estimates than those from ORD criterion (Bliemer and Rose, 2009; Vermuelen et al., 2011; Scarpa et al., 2009). However, no study to date has empirically investigated the impact of different ED criteria on measures of behavioural efficiency, such as choice complexity and attribute non-attendance, using a real sample and with a specifically controlled design.

Choice task complexity has been linked to the manner in which attribute levels are arranged across alternatives in a choice task. We use the complexity measure called "entropy" described in Swait and Adamowicz (2001a, 2001b). The higher the entropy value the greater the complexity level of a choice task. Such choice task would require respondents to exert more cognitive effort in the selection of the preferred alternative.

Attribute Non-Attendance (ANA) refers to the tendency of some respondents to systematically ignore particular attributes in the evaluation of alternatives in a choice task (Scarpa et al., 2009). (Pages 34 to 37 (Chapter 2) include a description of how to model ANA). There are currently two ways of addressing ANA, namely: *Inferred* and *Stated*. Inferred ANA is derived from observed patterns of choice made by respondents, while Stated ANA is obtained from respondents self-reporting their non-attendance to specific attributes after completing a choice task (or a series of choice tasks) (Scarpa et al., 2011a). In this study, as we collected very limited data on stated ANA, we focus our analysis on inferred ANA. The data analysis reported in this chapter aims to answer the following research questions:

(1) Does the selection of ED criterion affect inferred Attribute Non-Attendance? If so, what are the effects on the parameter estimates and WTP values?

(2) Is there an effect of higher choice task complexity on the variance of the Gumbel error (the unobserved component of utility)?

(3) If so, does this vary across different experimental designs? Which particular experimental design provides the most benefit to a choice experiments exercise?

Answers to the above questions would cast some light on the issue of criterion selection for stated choice experimental design. Identification of an experimental design that provides the highest behavioural efficiency would benefit not only choice analysts, but also the respondents who evaluate a series of complex choice questions. Attribute Non-Attendance (ANA) is described in Chapter 2 and very briefly in Section 4.2. The choice complexity metric is discussed in 4.3 and 4.4. Section 4.5 briefly describes the balanced data set used in the analysis. Section 4.6 presents the results in comparing ANA of different design criteria where we estimated panel latent class logit models to examine this. We also present the estimates of heteroskedastic logit models where we test the null hypothesis that the selection of ED has no effect on choice variability (via the entropy proxies). Conclusions are reported in section 4.7 where we show which design performed the best among the three designs studied here.

4.2 Attribute non-attendance and experimental design

The concept of attribute non-attendance and its modelling is described in Chapter 2 of this thesis. Although it is evident that ANA in CE studies exists and econometric models have been developed to account for its presence, to our knowledge, there has been no study yet that examined the factors that contribute to ANA. This study aims to contribute to answering this research question by empirically investigating the effect of different ED criteria on attribute non-attendance and on the estimated parameter values. We therefore attempt to test the null hypothesis: *Attribute non-attendance is the same for both utility neutral and*

Bayesian efficient designs. Utility neutral designs are EDs derived from design criteria that do not use prior knowledge of the parameters (e.g., ORD and OOD) while BDD falls under the group of Bayesian efficient designs that account for *a priori* information of the parameters.

4.3 Choice complexity

CE approaches have demonstrated to yield greater information than contingent valuation (CV) approaches. However, the higher amount of information collected from CE comes at the cost of requiring respondents to exert additional cognitive effort (Swait and Adamowicz, 2001a, 2001b; DeShazo and Fermo, 2002, 2004). CE respondents are expected to have a full understanding of how to select the preferred alternative on each choice task, process the information provided and then choose the preferred alternative by making tradeoffs. This can be quite complex and the level of complexity in processing choice questions vary between CE studies as complexity level can be affected by the number of attributes and the number of attribute levels (Swait and Adamowicz, 2001a, 2001b; De Shazo and Fermo, 2002; Arentze et al., 2003). One consequence of higher complexity might be that respondents would tend to select the status quo alternative leading to status quo bias, which could seriously affect the welfare measure (Dhar, 1997a; Dhar, 1997b; Boxall et al., 2009). In addition, different complexity levels have been found to significantly influence decision strategy selection (Payne, 1976; Olshavsky, 1979; Payne et al., 1988; Simonson and Tversky, 1992). Despite the importance of studying choice complexity in CE, very few studies have been carried out to examine its impacts or to account for its presence (e.g., Mazzotta and Opaluch, 1995; Boxall et al., 2009; Bliemer and Rose, 2011).

Our literature review indicates that very limited number of studies examined the effects of different ED criteria on choice complexity. A few studies examined choice complexity in CE and employed the orthogonal design criterion (e.g., Arentze et al. 2003; DeShazo and Fermo, 2002; Swait and Adamowicz, 2001a, 2001b). Louviere et al. (2008) compare the effects of different experimental designs on complexity by examining 44 EDs with systematically varying design dimensions as well as statistical efficiency. These authors focused on two ED criteria (i.e., optimal orthogonal and adaptive designs) with neither groups of designs being derived from *a priori* information on coefficient values. Viney et al (2005) focused on examining the effects of three different EDs (i.e., orthogonal, utility-balanced and random) on complexity by looking at the variance of the error term. Similar to Louviere et al., Viney et al. examined different EDs that were generated without any assumptions on parameter priors called "utility neutral designs" (Kessels et al., 2006).

Bliemer and Rose (2011) compare Bayesian D-efficient design with orthogonal design but their empirical analysis focused mainly on the gains in statistical efficiency (i.e., lower standard errors of attribute coefficients) but did not examine their effects on complexity. Many other studies compared utility neutral designs and Bayesian efficient designs and similarly they have shown empirical evidence that CEs based on efficient designs provide more accurate parameter estimates than utility neutral design (e.g., Kessels et al., 2006; Scarpa and Rose, 2008; Kerr and Sharp, 2010).

This is probably the first study to test whether the selection of different ED criteria influences the cognitive effort exerted by respondent as measured by choice complexity in CE. In doing so, we employ the method described by Swait and Adamowicz (2001a) where complexity is measured by entropy that we describe in section 4.4. Following Swait and Adamowicz 2001a, we calculated the entropy value of each choice task. Using the heteroskedastic logit model we evaluate the impact of entropy on the scale parameter for each of the three design treatments. Doing this helped us to answer the question *Does the effect of choice task complexity (via the entropy proxy) on choice variability vary across different experimental designs?*

4.4 Entropy as a measure of complexity and choice variability

Swait and Adamowicz (2001a) suggest that the complexity of a choice task can be represented by entropy. Following their paper, Equations 4.1 and 4.2 shows the formulae to calculate for the entropy value *E* of choice set *s* represented as E_s:

$$E_s = -\sum_{j=1}^{J} Q_{js} \ln Q_{js} \tag{4.1}$$

where

$$Q_{njs} = \frac{\exp(\beta X_{njs})}{\sum_{j \in k_n} \exp(\beta X_{nks})} \tag{4.2}$$

where Q_{njs} represents the choice probability that individual *n* chooses alternative *j* among *k* alternatives in choice set *s*. The betas are the estimated values from a conditional logit model.

We now show how calculated choice task specific entropy value E_s can be incorporated into the heteroskedastic logit model that we described in Chapter 2. Swait and Adamowicz (2001a) show that E_s or the complexity of choice task s affects the error variance (σ_s^2). σ_s is inversely related to the scale factor λ_s which can be presented as $\lambda_s = \frac{\pi}{\sigma_s \cdot \sqrt{6}}$ where π is the constant that is approximately 3.1416. We assume that λ_s is a quadratic function of E_s of the choice situation so as to capture nonlinearities of entropy.

$$\lambda_s(C_s) = \exp\left(\gamma_1 E_s + \gamma_2 E_s^2\right) \tag{4.3}$$

The quadratic form in 4.3 above allows λ_s to account for the reaction of a respondent described by Keller and Staelin (1987) where one may tend to exert greater effort to making decisions (which could enhance preference consistency across respondents) up to a certain level of complexity. After reaching a particular level, respondents may tend to employ simplifying decision heuristics resulting to collection of data with greater preference inconsistencies. If this situation applies to one of the design treatments, then we could expect that 4.3 would have $\gamma_1 < 0$ and $\gamma_2 > 0$. If $\gamma_1 < 0$ and $\gamma_2 = 0$, then we have a situation where an individual reacted to an increase in complexity level mainly by resorting to simplifying heuristics and very limited cognitive effort to cope with higher entropy. Finally, if $\gamma_2 = 0$ and $\gamma_1 = 0$ then we fail to reject the null hypothesis that entropy has no effect on scale. This implies that the increase in complexity would not likely reach a point at which greater preference inconsistencies could occur.

4.5 Data

For examining the effects of different ED criteria on attribute non-attendance and choice complexity, we use the *Balanced sample* described in Section 2.5 of this thesis. This sample consists of 1509 choice observations that are evenly distributed across three different ED samples composed of Orthogonal Design (ORD), Optimal Orthogonal Design (OOD) and Bayesian D-efficient Design (BDD). For an objective comparison of the three EDs, we allocated each design treatment with 503 choice observations. The three samples all have equal number of observed choice set orders (i.e., 56 observations for the 1st, 2nd, 4th, 5th, 6th, 7th, 8th and 9th choice set orders; and 55 observations for the 3rd choice set order).

4.6 Results

4.6.1 Conditional logit model

We estimated the coefficients for the same utility specification of a conditional logit model (Table 4.1) from three different samples of choices each based on different design criterion. In terms of parameter estimates, the coefficients for cost for the three samples are all negative and significant, suggesting that the decision of respondents to choose their desired alternative is negatively influenced by the amount of money that they would pay to enhance biodiversity in planted forests. All significant coefficients for the environmental attributes (e.g., Brown kiwi 1, Brown kiwi 2, Bush falcon 2) have positive signs which implies that the proposed biodiversity enhancement outcomes contribute positively to the utility of an individual. Although some estimated coefficients (e.g., Green gecko 1, Native plant 1) have negative signs, these are not statistically significant. A relatively larger proportion of coefficient estimates in the OOD sample are not statistically significant particularly the non-bird species which maybe considered as less charismatic species. We suspect that this situation might indicate non-attendance to that particular subset of attributes. This might be attributed to the fact that the bird species would likely be more familiar or more readily seen in planted forests while the native plant, gecko and native fish seem less visible in planted forests or other areas of New Zealand.

The coefficient for the indicator for status quo option (SQ) for the ORD sample is positive and significant while those from the two other designs are negative but not significant. We surmise that respondents with choice tasks generated from the ORD criterion would likely choose SQ or they have a higher tendency to opt out compared to the other two designs. We investigate this conjecture by using the latent class panel models that account for attribute non-attendance in the next section.

Table 4.1: Conditional logit model estimates for the three design treatments

Attribute	ORD			BDD			OOD		
	Coeff	Std Err	*p*-value	Coeff	Std Err	*p*-value	Coeff	Std Err	*p*-value
Brown kiwi 1	0.471	0.209	0.02	0.377	0.179	0.04	0.198	0.198	<0.01
Brown kiwi 2	0.702	0.206	<0.01	0.456	0.168	0.01	0.191	0.191	<0.01
Native fish 1	0.349	0.195	0.07	0.378	0.161	0.02	0.180	0.180	0.36
Native fish 2	0.242	0.202	0.23	-0.031	0.169	0.86	0.175	0.175	0.28
Native plant 1	0.259	0.185	0.16	-0.039	0.180	0.83	0.187	0.187	0.25
Native plant 2	-0.092	0.205	0.65	0.436	0.165	0.01	0.184	0.184	0.58
Green gecko 1	0.132	0.200	0.51	-0.053	0.167	0.75	0.190	0.190	0.78
Green gecko 2	0.443	0.197	0.03	-0.179	0.167	0.29	0.180	0.180	0.45
Bush falcon 1	0.499	0.208	0.02	0.567	0.170	<0.01	0.196	0.196	0.14
Bush falcon 2	0.823	0.202	<0.01	0.789	0.172	<0.01	0.186	0.186	<0.01
Cost to respondent	-0.026	0.003	<0.01	-0.020	0.003	<0.01	0.003	0.003	<0.01
Indicator for SQ option	0.734	0.329	0.03	-0.039	0.307	0.90	-0.273	0.273	0.17
Log-likelihood	-459.28			-497.66			-469.62		
Rho-square	0.169			0.099			0.150		
Adjusted rho-square	0.147			0.078			0.128		
Observations	503			503			503		

4.6.2 Panel latent class logit model results

To identify some of the possible latent classes for attribute non-attendance, on page 16 of the questionnaire, we asked each respondent which attributes that he/she did not attend to. This question immediately followed after a respondent completed the evaluation of the nine choice tasks on pages 7 to 15. Table 4.2 presents the self-reported attribute non-attendance of the three design treatments. BDD has the lowest average non-attendance rate of 6.4% while ORD has the highest average. About 5% of the responses in the ORD sample did not attend to brown kiwi while the two other treatments both have zero non-attendance rate. The ORD treatment has the highest proportions of stated ANA in four out of the five environmental attributes. The BDD treatment consistently demonstrates the lowest proportion of non-attendance in all five attributes. The *p*-values in the sixth column of Table 4.2 indicate significant differences in the proportions of stated non-attendance between ORD and BDD.

Table 4.2: Percentage (%) of respondents stating non-attendance and testing the equality of proportions between treatments

	ORD	BDD	OOD	Pooled	Test of equality of proportions between treatments (*p*-value)		
					ORD vs BDD	ORD vs OOD	BDD vs OOD
Brown kiwi	5.4	0.0	0.0	1.8	<0.001	<0.001	1.000
Native fish	17.5	12.5	17.9	16.0	<0.001	0.775	<0.001
Native plant	14.3	10.7	14.3	13.1	0.003	1.000	0.003
Green gecko	17.5	7.2	7.2	10.6	<0.001	<0.001	1.000
Bush falcon	3.6	1.8	1.8	2.4	0.002	0.002	1.000
Average	11.7	6.4	8.2	8.8	--	--	--
Minimum	3.6	0.0	0.0	1.8	--	--	--
Maximum	17.5	12.5	17.9	16.0	--	--	--

Table 4.2 also shows a pattern that the non-bird attributes (fish, plant and gecko) have higher rates of non-attendance relative to the two native bird species. This allows the identification of non-attendance to non-bird species as one latent class. The next latent class identified is non-attendance to status quo option. This is based on the fact that in choice experiments, there could potentially be a status quo bias where a respondent would tend to choose the status quo option or simply opt out from evaluating the change alternatives (Boxall et al., 2009). We assume here that the opposite can also hold true wherein respondents could also tend to ignore the status quo and focus only on the changed alternatives. There is a possibility that the underlying experimental design could affect this potential source of bias. A third possible latent class is full attendance where a class of respondents considered all the six attributes in the evaluation of choice tasks.

Ten latent class model specifications have been tested through a grid search procedure. The grid search was done to identify a group of latent classes across designs that produce the best model fit and, at the same time, latent class models should have converged for the three design treatments. This is to allow the comparison of the three treatments. Table 4.3 shows the normalised AICs of 10 different latent class model specifications of the design treatments. As much as we would like to use the estimates of specifications with the lowest AICs, there were convergence issues in those groups. Some specifications that converged show *p*-values of 1.00 in the latent classes making them unusable. For instance, although specification number 5 has lower AICs, the *p*-value for the cost coefficient in the OOD gets a p-value of 1.00 which can be a sign of misspecification.We settled on using the estimates from LCM specification number 3 because the model for all three treatments converged and all *p*-values of latent classes made sense.

Table 4.3: Estimates of normalised AICs of latent class logit models using the three design samples

Specification Number	Latent classes (LCs) – Attributes ignored	Normalised AIC (AIC/N)		
		ORD	BDD	OOD
1	LC1 - Ignored SQ, LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Ignored all attributes	1.126	Did not converge	1.309
2	LC1 - Ignored SQ, LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Ignored Cost	1.168	1.339	1.243
3	LC1 – Ignored SQ, LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Full attendance	1.135	1.362	1.309
4	LC1 – Ignored SQ LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Full attendance LC4 – Ignored all attributes	1.077	Did not converge	1.332
5	LC1 – Ignored cost LC2 – Ignored SQ LC3 – Ignored Gecko, Kakabeak and Kokopu LC4 – Ignored all attributes	1.147	1.340	1.413
6	LC1 – Ignored cost LC2 – Ignored SQ LC3 – Ignored Gecko, Kakabeak and Kokopu LC4 – Ignored Falcon	1.172	1.342	1.085
7	LC1 – Ignored cost LC2 – Ignored SQ LC3 – Ignored Gecko, Kakabeak and Kokopu LC4 – Ignored Kiwi	1.131	1.335	1.247
8	LC1 – Ignored SQ LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Full attendance LC4 – Ignored Kiwi	1.139	1.365	Did not converge
9	LC1 – Ignored SQ LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Full attendance LC4 – Ignored Falcon	1.139	1.366	1.362
10	LC1 - Ignored SQ LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Ignored Kiwi LC4 – Ignored Falcon	Did not converge	1.366	1.371

Table 4.4 presents the estimates of latent class panel model for the three design treatments for specification 3 from Table 4.3. For each design treatment, the three behaviourally defined latent classes were: (1) a class that attended to all six attributes; (2) a class that did not attend to non-bird species which were Green gecko, Kakabeak plant and Kokopu fish; and (3) a class that ignored the SQ alternative. For the first class, as all attributes were assumed to have been attended, all coefficients were not constrained. For the second class, we defined this latent class by constraining to zero the coefficients of the three non-bird species. For the third class, the coefficient for SQ was constrained to zero.

As expected, the model goodness of fit significantly improved when the latent class panel model is used compared to the conditional logit model (Table 4.4). This is indicated by the statistically significant increases in the log likelihood values (e.g., for ORD, from -459 to -271). This provides evidence of the presence of heterogeneity in attribute attendance in the three choice data sets. Table 4.4 reports results for the three design treatments, the estimated probabilities for the class that ignored the SQ alternative are significant at the 99.9% level. The BDD sample has the highest latent class probability of ignoring the SQ (0.647) while OOD has the lowest probability (0.481) and closely followed by ORD (0.489). This result might indicate that choice tasks generated from BDD criterion would likely have better encouraged respondents to focus more on the designed alternatives compared to the two other design treatments. On the other hand, the latent class probabilities for OOD and ORD are virtually the same. To compare these two design treatments, we look at the signs of the utility coefficients for SQ, which is negative and significant for OOD while positive and significant for ORD. This suggests that in the ORD sample, the expectation of moving away from the current situation would partly affect people’s utility negatively. While in OOD sample, moving away from the current situation would partly positively affect utility. These results indicate that using ORD will lead to more respondents opting out compared to OOD.

Table 4.4 shows that the coefficients of the indicators for the SQ option are both positive and significant for both ORD and BDD treatments. This suggests that, after controlling for non-attendance, respondents in both treatments who tended not to ignore the SQ option, seem to like the current condition and would be willing to support it. This is the opposite of what we found in the OOD sample, where the coefficient for the SQ option is negative and significant, indicating that a typical respondent would not be likely to stay with the current scenario but would gain more utility from the scenario with greater level of biodiversity.

Table 4.4 also shows that the OOD sample has the highest probability value (0.519) for the class that ignored the less charismatic species. A possible reason for greater occurrence of non-attendance is that respondents may have felt that attending to all attributes is more complex or imposing greater cognitive burden in choice tasks designed using the optimal orthogonal criterion. This design may have made them attend to a smaller number of attributes, perhaps as an attempt to reduce their effort in doing tradeoffs so as to still complete all the nine choice tasks they were expected to finish.

The BDD design has the highest probability value (0.647) for the class that ignored SQ. This indicates that this class would be likely to focus more attention on the scenario alternatives proposing change and therefore more likely to place a higher value to the changed alternatives. We speculate that a possible reason for this is that the combination of attribute levels presented in the alternatives different to the SQ have been found to be “more coherent” by respondents and this somewhat reduced the cognitive effort in evaluating those alternatives. What I meant by “more coherent” is that attribute levels from choice tasks derived from BDD are more logically arranged compared to the two other designs. There seem to be a criterion in BDD that imposes moderate overlaps of attribute levels between alternatives. By visually comparing choice tasks across the three design treatments, ORD choice tasks, in general, exhibit the most number of overlaps; OOD does not have any overlap; while BDD choice tasks seem to be in-between or have moderate number of overlaps. (This was based on actual observation of the choice tasks generated from the three different designs.) We speculate here that the overlaps in BDD might be at the right frequency that makes the attribute levels in the choice tasks to somewhat appear more logically arranged to respondents. Appendix Figures 1, 2 and 3 show examples of choice tasks of the three designs with different levels of overlaps. Chrzan and Orme (2000) define minimal overlap as “Within choice sets, attribute levels are duplicated as little as possible.” Chrzan and Orme provide results suggesting that choice based conjoint designs with moderate or balanced overlaps would be a desirable design strategy compared to minimal or higher frequencies of overlaps.

To further examine if the BDD treatment results to greater non-attendance to the SQ option, we used the balanced pooled sample and ran the panel latent class model (PLCM) that we used earlier on each design sub sample. Table 4.5 (columns 5, 6 and 7) presents the PLCM estimates from the pooled sample. All three non-attendance latent class probabilities are significant at the 90% confidence level with the class ignoring the SQ option predicted to have the highest proportion of respondents at 50%, followed by those who ignored the non-bird attributes at 36% and then full attendance at 14%. Columns 8, 9 and 10 of Table 4.5 present the estimates from the panel latent class model that allows each class membership to be a function of design. For the class that ignored the SQ option, the coefficient for BDD is the highest among the three designs and is significant at the 90% level. This corroborates the finding in the split samples that BDD choice tasks have the highest proportion of respondents who ignored the SQ option.

Table 4.4: Panel latent class model estimates for the three design treatments

Attribute	ORD			BDD			OOD		
	Coeff	**Std Err**	***p*-value**	**Coeff**	**Std Err**	***p*-value**	**Coeff**	**Std Err**	***p*-value**
Brown kiwi 1	0.806	0.508	0.11	0.873	0.192	<0.01	0.985	0.229	<0.01
Brown kiwi 2	1.176	0.437	0.01	1.081	0.223	<0.01	1.094	0.252	<0.01
Native fish 1	0.997	0.378	0.01	0.556	0.255	0.03	-0.167	0.386	0.67
Native fish 2	0.867	0.450	0.05	0.267	0.208	0.20	-0.361	0.412	0.38
Native plant 1	0.986	0.366	0.01	0.534	0.233	0.02	-1.336	0.565	0.02
Native plant 2	0.653	0.470	0.16	0.820	0.207	<0.01	-0.603	0.397	0.13
Green gecko 1	0.339	0.264	0.20	0.255	0.227	0.26	-0.861	0.447	0.05
Green gecko 2	1.236	0.411	<0.01	0.189	0.244	0.44	0.054	0.408	0.90
Bush falcon 1	1.307	0.409	<0.01	1.031	0.246	<0.01	0.477	0.217	0.03
Bush falcon 2	1.976	0.280	<0.01	1.348	0.234	<0.01	0.668	0.220	<0.01
Cost to respondent	-0.037	0.003	<0.01	-0.014	0.003	<0.01	-0.045	0.005	<0.01
Indicator for SQ option	4.419	0.843	<0.01	4.277	0.580	<0.01	-4.002	0.508	<0.01
Latent Class (LC)	LC Prob	Std Err	*p*-value	LC Prob	Std Err	*p*-value	LC Prob	Std Err	*p*-value
C1 - Full Attendance	0.114	0.303	0.71	0.083	0.328	0.80	<0.001	<0.001	0.98
C2 - Ignored 3 non-bird attributes	0.397	0.171	0.02	0.270	0.182	0.14	0.519	0.104	<0.01
C3 - Ignored the SQ alternative	0.489	0.161	<0.01	0.647	0.071	<0.01	0.481	0.112	<0.01
Log-likelihood	-271.47			-328.49			-315.27		
Normalised AIC	1.135			1.362			1.309		
Normalised Finite Sample AIC	1.137			1.364			1.311		
Normalised BIC	1.252			1.479			1.427		
Observations	503			503			503		

Table 4.5: Panel latent class model estimates from the pooled balanced sample

Item	Multinomial Logit Model			Panel Latent Class Logit (PLCL) Model on Pooled Balanced Sample			PLCL Model with class membership as a function of design		
	Coeff	Std Err	p-value	Coeff	Std Err	p-value	Coeff	Std Err	p-value
Kiwi 1	0.495	0.109	<0.01	0.667	0.128	<0.01	0.670	0.130	<0.01
Kiwi 2	0.654	0.105	<0.01	1.125	0.132	<0.01	1.156	0.135	<0.01
Kokopu 1	0.318	0.101	<0.01	0.653	0.135	<0.01	0.690	0.134	<0.01
Kokopu 2	0.134	0.103	0.19	0.565	0.145	<0.01	0.598	0.147	<0.01
Kakabeak 1	0.179	0.103	0.08	0.695	0.127	<0.01	0.726	0.130	<0.01
Kakabeak 2	0.228	0.103	0.03	0.637	0.118	<0.01	0.667	0.122	<0.01
Gecko 1	0.019	0.102	0.85	0.227	0.113	0.04	0.246	0.115	0.03
Gecko 2	0.098	0.101	0.33	0.195	0.104	0.06	0.206	0.105	0.05
Falcon 1	0.481	0.106	<0.01	0.538	0.140	<0.01	0.536	0.141	<0.01
Falcon 2	0.720	0.104	<0.01	0.883	0.119	<0.01	0.891	0.123	<0.01
Indicator for status quo	0.159	0.171	0.35	-0.429	0.235	0.07	-0.404	0.253	0.11
Cost	-0.026	0.002	<0.01	-0.095	0.006	<0.01	-0.093	0.006	<0.01
LC1 - Full attendance				0.137	0.076	0.07			
LC2 - Ignored SQ				0.502	0.046	<0.01			
LC3 - Ignored non-bird attributes				0.361	0.083	<0.01			
Full attendance (LC1 as a function of design)									
- Constant (ORD)							-1.021	0.786	0.19
- BDD							-0.632	0.871	0.47
- OOD							0.531	0.685	0.44
Ignored SQ (LC2 as a function of design)									
- Constant (ORD)							-0.123	0.377	0.74
- BDD							0.758	0.445	0.09
- OOD							0.524	0.438	0.23
Log-likelihood	-1460.32			-964.09			-960.51		
Normalised AIC	1.951			1.296			1.297		
Number of choice observations	1509			1509			1509		

4.6.3 Complexity and heteroskedastic logit by design treatment

Swait and Adamowicz (2001a) suggest that entropy is a convenient scalar measure of complexity that summarises the impacts of number of alternatives, number of attributes, number of attribute levels, and preference similarity among alternatives. The three design treatments we examine here are identical in terms of number of alternatives, number of attributes and attribute levels. However, since they were constructed using three different design criteria, they are likely to differ in terms of preference similarity among alternatives. The theoretically maximum entropy is achieved when each alternative in a choice task has an equal chance of being selected compared with other alternatives. Having an equal chance of being selected could have two outcomes: (1) respondents exerted more effort in evaluating choice tasks with higher complexity, or (2) respondents chose the "preferred" alternative randomly without applying any effort.

Following Equation 4.1, we calculate the entropy of each design block of the three designs. Table 4.6 shows a summary of these entropy measures by block and by design. Among the three design criteria, ORD has the lowest mean and median entropy values of 0.894 and 0.942. However, ORD also has the highest entropy standard deviation that is at least twice as large as the other two designs. One reason for this is that the choice tasks in Block 2 of ORD, on average, have low entropy values and these have reduced the overall mean entropy. This resulted in a wide dispersion of entropy values across ORD choice tasks. The above implies that although ORD is the design with the lowest overall entropy, complexity levels across choice tasks vary twice as much as the choice tasks in the other designs. As BDD has the least standard deviation, the variation in complexity across choice task is the lowest among the three designs.

We illustrate the distribution of entropy measures of the choice tasks (E_s) for each design using the kernel density plot shown in Figure 4.1. Each kernel density shows the proportion of the 27 choice tasks (distributed into three blocks) for each design. ORD has the widest entropy range while BDD and OOD have virtually the same entropy range. However, despite the similarity of entropy ranges for BDD and OOD, the entropy for the former is more concentrated towards less than 1.0 with the latter towards greater than 1.0. This illustrates the fact that the BDD design has more entropy values distributed over the lower range than the OOD. The narrower BDD kernel density also illustrates that it has the lowest dispersion of entropy as suggested by the lower standard deviation of 0.078 compared to ORD and OOD with 0.171 and 0.085, respectively (please see Table 4.6).

Figure 4.1: Kernel density of entropy by experimental design

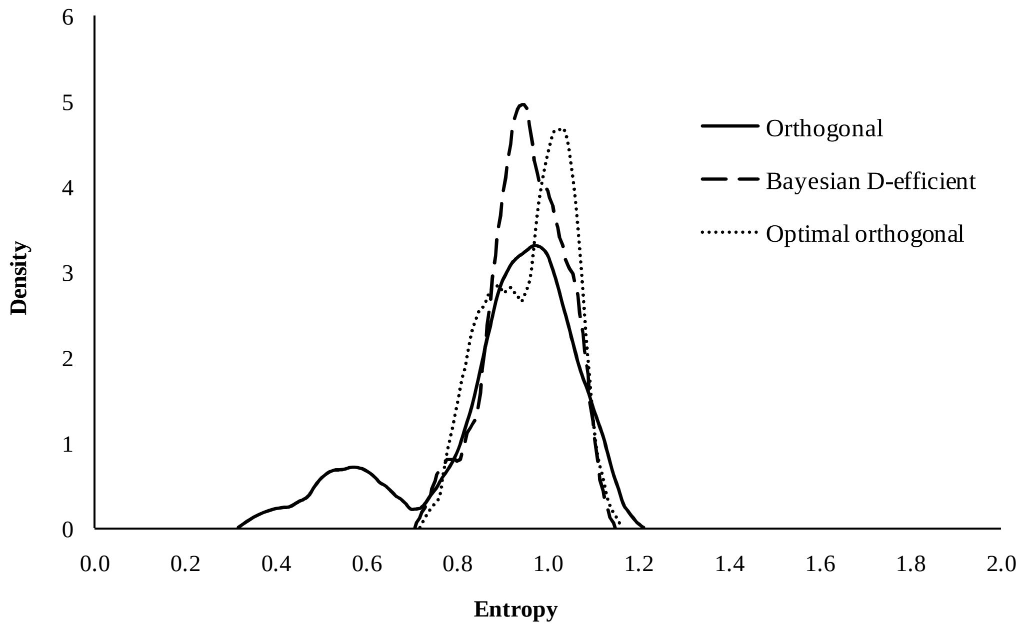


Table 4.6: Distribution of entropy values by design and by block

	Orthogonal				Bayesian D-efficient				Optimal orthogonal				Pooled Sample			
	B1	B2	B3	All	B1	B2	B3	All	B1	B2	B3	All	B1	B2	B3	All
Mean	0.984	0.717	0.982	0.894	0.952	0.959	0.954	0.955	0.927	0.991	0.962	0.960	0.954	0.889	0.966	0.939
Median	0.963	0.744	0.989	0.942	0.951	0.971	0.961	0.962	0.912	1.015	1.015	1.013	0.944	0.936	0.989	0.962
Standard deviation	0.083	0.178	0.053	0.171	0.052	0.112	0.058	0.078	0.082	0.080	0.083	0.085	0.077	0.178	0.066	0.122
Minimum	0.838	0.431	0.906	0.431	0.871	0.772	0.865	0.772	0.830	0.788	0.832	0.788	0.830	0.431	0.832	0.431
Maximum	1.097	0.960	1.052	1.097	1.029	1.085	1.050	1.085	1.092	1.068	1.060	1.092	1.097	1.085	1.060	1.097
Number of choice tasks	27	27	27	81	27	27	27	81	27	27	27	81	81	81	81	243

Table 4.7: Estimates from conditional and heteroskedastic logit models

Design	No. of observed choice sets	Log-likelihood Conditional Logit	Log-likelihood Heteroskedastic Logit	Likelihood ratio test statistic	Estimated Coefficient of Entropy (Robust *p*-value)	Estimated Coefficient of Entropy Squared (Robust *p*-value)
ORD	503	-459.28	-455.09	8.38	-11.1 (<0.01)	9.46 (<0.01)
BDD	503	-497.66	-496.14	3.04	5.67 (0.90)	-4.75 (0.84)
OOD	503	-469.62	-468.92	1.40	-5.86 (0.03)	2.41 (0.21)
Pooled	1509	-1460.32	-1458.86	2.92	-4.82 (0.21)	2.67 (0.28)

Table 4.7 presents the estimates from the conditional and heteroskedastic logit models for the three design treatments. (Conditional and heteroskedastic logit models are earlier introduced in Chapter 2). The conditional logit model imposes the restriction that scale is not a function of entropy or complexity while the heteroskedastic logit model relaxes this assumption by allowing the scale to vary based on complexity level of choice tasks. Of the three designs, only the ORD has a significant improvement in log likelihood value when scale is considered to be a function of entropy. This is indicated by the likelihood ratio (LR) test statistic of 8.38 which exceeds the critical Chi-square value with two degrees of freedom at the 95% confidence level of 5.99.

The pooled sample did not have a significant improvement in fit with scale as indicated by the Chi-square test statistic of only 2.92. This result is not surprising as OOD and BDD samples (or two-thirds of the pooled sample) both have low Chi-square test statistics. Similarly, both BDD and OOD treatments failed to reject the null hypothesis that scale is not a function of complexity as indicated by Chi-square test statistics of 3.04 and 1.40, respectively. This result is consistent with the findings in Swait and Adamowicz (2001a) which suggests that “the lack of variability in entropy across respondents is leading to this non-significant impact of complexity”. Swait and Adamowicz suggest that the design of CE should include some fraction of simpler tradeoffs (e.g., entropy range [0.60,0.80]) and some fraction with an extreme entropy range (e.g., [1.00,1.20]) to allow the separation of variance and taste effects. However, that paper also mentioned that the objective of allowing a wide entropy range would be advantageous for analysing data sets that combine Stated Preference (SP) with Revealed Preference (RP) data. The EDs examined in that paper consisted of 10 data sets (6 SPs and 4 RPs) that were all derived from the orthogonal main effects design criterion which is the criterion used in constructing ORD for this present study. In the present study ORD varies highly in complexity as denoted by the 0.431-1.097 entropy range, while BDD and OOD had ranges of 0.772-1.085 and 0.788-1.092 , respectively. It appears that when the design criterion assumes that the choice data will be analysed using a conditional logit model, the entropy range is minimised and this contributes to a reduction in the statistical significance of the effect of entropy on the scale factor.

The scale coefficients for entropy in column 5 of Table 4.7 show that OOD and ORD are negative and significant which is consistent with respondents in both treatments resorting, to some degree, to simplifying decision heuristics. This might include attribute non-attendance and avoiding cognitive burden by defaulting to the SQ option (Swait and Adamowicz, 2001b; Boxall et al., 2009). Those simplifying heuristics generate greater preference inconsistencies, which can be reflected by an increase in the variance of the unobserved component of utility.

Columns 5 and 6 of Table 4.7 show that, in the ORD treatment, scale is a function of entropy as indicated by very low *p*-values of the scale coefficients for entropy and entropy squared. The *p*-value of the scale coefficient for entropy is less than 0.05 while for entropy squared is nearly significant. Thus we cannot jointly omit the scale for OOD. For the BDD treatment, high *p*-values of both scale coefficients indicate that we can jointly omit scale. The ED derived from the BDD criterion seems to have provided choice tasks that somewhat reduced the effort on the part of the respondents to find the utility maximising choice. This might provide the reason why 65% of the respondents who completed the BDD choice tasks were likely to ignore the SQ option. This shows a pattern that the absence of influence of entropy on scale resulted in more coherent experimentally designed alternatives in choice tasks where respondents paid more attention to all attributes.

It is important to note that the sample size of each design treatment only has 503 choice observations which might not be enough to make the above results generalisable. The above results might be specific to the relatively small choice data set analysed here. Appendix Table 4 presents the utility coefficients of the heteroskedastic logit models, where scale coefficients for entropy are presented in Table 4.7. Unfortunately all utility coefficients (in Appendix Table 4) which are significant in the conditional logit model estimates in Table 4.1 are no longer significant with the addition of entropy scale coefficients. We speculate that a possible reason for this is that we used a small sample size that resulted to a lack of variability. This might suggest that the findings in this study may be specific to the choice data collected here.

4.6.4 Complexity and heteroskedastic logit (pooled sample)

We have pooled the three ED samples and analysed the effect of entropy on scale while controlling for the design effects. Table 4.8 begins with Model 1 which is the basic conditional logit model. Model 1 shows significant coefficient estimates very similar to those of the conditional logit estimates for the BDD sample in Table 4.1. Model 2 is a heteroskedastic conditional logit model which allows for factors that could influence scale. Similar to the results in the split samples, all estimates of utility coefficients, including cost, are no longer significant. However, as we added four scale factors related to entropy and design criteria, these contributed to significantly improving model fit from a log-likelihood of -1460 in Model 1 to -1453 in Model 2. The log-likelihood ratio test statistic of 14.3 exceeds the critical Chi-square value at the 95% confidence level of 9.49 and implies rejection of the restricted form. Model 2 allows the testing of entropy measures on scale where the signs of the two scale coefficients suggest a quadratic effect but not statistically significant. While controlling for the quadratic effects of entropy on scale, we also tested for the effects of EDs. Using ORD as the reference design, the scale coefficient for BDD is negatively significant at the 98% confidence level while OOD is also negative but not significant. This might indicate that relative to ORD, and after controlling for entropy, BDD contributes to increasing the variance of the error term leading to greater choice inconsistency. This contradicts the results presented in sections 4.6.2 and 4.6.3 which show that BDD contributes more to increasing attendance to the designed alternatives and provides greater choice consistency relative to the other two designs. We further examine this result using Model 3, which is a heteroskedastic logit model where we interacted entropy variables with ED indicators to capture the net effect of the designs. Model 3 estimates show positive scale coefficients for BDD and OOD but not statistically significant. Actually it may be difficult to say anything from Model 3 estimates as all utility and scale coefficients are no longer significant. There is also no significant improvement in model fit from moving from Model 2 to Model 3. However, focusing on the signs and magnitudes of the design coefficients, these results are consistent with those in the previous sections suggesting that the net effect of BDD results to lower error variance relative to ORD. Consequently, given that we neither see any clear net effect of design nor any joint effect of the interaction between design and entropy in Table 4.8, we examine further the effect of choice complexity via the attribute dispersion proxy (which is another component of choice task complexity) in the next chapter of this thesis.

Table 4.8: Estimates from Logit models from the pooled sample

	Model 1 Conditional Logit Model			Model 2 Heteroskedastic Conditional Logit Order, Designs and Entropy as Scale			Model 3 Heteroskedastic Conditional Logit Order, Designs & Entropy with Interaction		
	Coeff	**Std Err**	***p*-value**	**Coeff**	**Std Err**	***p*-value**	**Coeff**	**Std Err**	***p*-value**
Brown kiwi 1	0.495	0.111	<0.01	1.960	2.620	0.45	8.460	15.900	0.60
Brown kiwi 2	0.654	0.105	<0.01	2.730	3.820	0.48	11.700	23.200	0.61
Native fish 1	0.318	0.101	<0.01	1.170	1.600	0.47	4.960	9.810	0.61
Native fish 2	0.134	0.104	0.20	0.527	0.768	0.49	2.550	5.090	0.62
Native plant 1	0.179	0.102	0.08	0.728	1.090	0.50	2.970	5.850	0.61
Native plant 2	0.228	0.105	0.03	0.658	0.911	0.47	2.810	5.210	0.59
Green gecko 1	0.020	0.103	0.85	0.040	0.420	0.92	0.083	1.930	0.97
Green gecko 2	0.098	0.101	0.33	0.485	0.680	0.48	2.430	5.320	0.65
Bush falcon 1	0.481	0.107	<0.01	1.780	2.520	0.48	8.220	16.400	0.62
Bush falcon 2	0.720	0.101	<0.01	2.760	3.770	0.46	12.100	23.500	0.61
Cost to respondent	-0.026	0.002	<0.01	-0.103	0.140	0.46	-0.445	0.863	0.61
Indicator for SQ	0.159	0.176	0.37	-0.590	1.090	0.59	-3.500	7.920	0.66
Indicator for Bayesian D-efficient design (BDD)				-0.473	0.170	0.01	10.100	13.100	0.44
Indicator for Optimal orthogonal design (OOD)				-0.030	0.139	0.83	7.190	14.300	0.61
Entropy of a choice task				-2.910	3.600	0.42	-7.640	5.600	0.17
Entropy square				1.670	2.310	0.47	5.070	3.780	0.18
BDD * Entropy							-21.700	29.900	0.47
BDD * Entropy square							11.000	17.000	0.52
OOD * Entropy							-12.900	31.600	0.68
OOD * Entropy square							5.420	17.400	0.76
Log-likelihood	-1460.32			-1453.17			-1451.44		
Number of choice observations	1509			1509			1509		

In Chapter 5, the same pooled balanced sample examined in this chapter will be used to examine whether the effect of higher attribute dispersion on the error variance vary across experimental designs. Attribute dispersion was described and used in DeShazo and Fermo (2002) to show that respondents evaluated choice situations less consistently as task complexity increases. It is a different set of measures of task complexity in that it accounts mainly for the dispersion of attribute levels in a choice task. Compared to entropy, it does not use coefficient estimates to calculate the complexity measure. DeShazo and Fermo (2002) evaluate groups of choice tasks with varying number of alternatives, attributes and attribute levels which were all derived from the orthogonal criterion. In contrast, the analysis in Chapter 5 focuses on the effect of attribute dispersion on the error variance of choice data collected from choice tasks with the same number of alternatives, attributes and attribute levels but derived from three different design criteria.

4.7 Conclusions

We conclude that based on the sample choice observations studied here, using the BDD criterion may offer some advantages. These include more statistically efficient parameter estimates, reduction in the theoretically minimum number of respondents and providing respondents with choice tasks that are relatively more behaviourally efficient. Higher behavioural efficiency is indicated by higher rates of attendance to the designed alternatives and lesser occurrence of choice inconsistencies. The above may be translated to higher data quality, more reliable parameter estimates and reduction in survey time and cost. However, we acknowledge that the above results might be specific to the split sample studied here. It would be good to have a future study that could verify this result with a much larger choice data set with similar split designs.

This chapter provides some evidence of the superiority of the Bayesian D-efficient design criterion in terms of behavioural efficiency relative to optimal orthogonal and orthogonal designs. We found that the selection of ED criterion strongly affects both stated and inferred attribute non-attendance. Results indicate that the sequence of choice tasks derived from the Bayesian D-efficiency criterion tend to minimise stated attribute non-attendance. Design alternatives from this criterion were found to have lower inferred attribute non-attendance relative to the other two designs. The lower attribute non-attendance rate is found to lead to more accurate welfare estimates as respondents tend to be more engaged in the evaluation of designed alternatives and demonstrated a relatively lower incidence of opting out.

In terms of the impact of choice task complexity on choice variability of respondents, we found that the three design criteria have varying impacts. Results suggest that the variation in complexity levels (via the entropy proxy) of choice tasks derived from the Bayesian D-efficiency criterion does not lead to an increase in choice inconsistency (via the error variance proxy) of respondents, whilst the variation in task complexity of the other two criteria does increase choice inconsistency. We found that, unlike in the two other design treatments, the entropy levels in choice tasks derived from the Bayesian D-efficiency criterion do not increase in any way, the variance of the Gumbel error.

We have shown that, in terms of attribute non-attendance and contribution to choice consistency, Bayesian D-efficiency is the superior criterion. However, there may be other factors that could show the difference between the three EDs examined here, such as attribute dispersion of choice tasks and learning/fatigue effects (DeShazo and Fermo, 2002; Plott, 1986; Bateman et al., 2008; Caussade et al., 2005). We present and discuss our examination on the effects of attribute dispersion and learning on choice variability across designs in the next chapter of this thesis.

Chapter 5: Design criteria effects on choice complexity and learning

5.1 Introduction

Stated Choice Experiments (CE) has been widely used in many different fields to study the preferences of individuals. Over the last two decades, several Experimental Design (ED) criteria (e.g., Bayesian D-efficient design (BDD), Optimal orthogonal design (OOD)) have been developed to address the drawbacks of the most widely used criterion Orthogonal Design (ORD). However, despite analysts employ different EDs for conducting CE exercises, most of them assumed that the selection of ED criterion is neutral to the estimated parameter values. This chapter explores this issue and investigates whether the three different EDs have the same effect on the estimates of the coefficients (β) of the indirect utility function and scale coefficients (λ), while controlling for the effects of task complexity and task order.

In Chapter 4, we presented the results of our examination of three ED criteria (i.e., ORD, BDD and OOD) based on choice task specific *entropy* levels. We reported some empirical evidence that higher entropy levels in the choice tasks derived from the different EDs have varying effects on scale. We have shown that different ED criteria can result in different patterns of attribute non-attendance which could lead to different estimates of the utility coefficients and willingness-to-pay (WTP) values. We concluded in Chapter 4 that using the BDD criterion, relative to the two other ED criteria, results in the generation of a superior ED that has the following characteristics: (1) highest design efficiency as indicated by having the lowest D_b-error; (2) greater choice consistency as the entropy level does not reduce the scale factor unlike the two other EDs; and (3) provides more realistic WTP estimates as the designed alternatives have been more attended to compared to the designed alternatives in the other two EDs.

Chapter 4 has examined the effect of entropy on choice consistency. Entropy is one of the proposed measures of choice task complexity. However, there is another measure of complexity that is described in DeShazo and Fermo (2002, 2004). We call this task specific measure *attribute dispersion*. Similar to entropy, attribute dispersion is choice task specific. However, unlike entropy, which is a single overall measure of complexity of a choice task, attribute dispersion is broken down into two subcomponents, namely: *a*verage *s*tandard *d*eviation of attribute levels across alternatives in a choice task (ASD) and *d*ispersion of *s*tandard *d*eviation of attributes levels across alternatives in a choice task (DSD).[21] The calculation of attribute dispersion values does not include any coefficient values (unlike entropy) as we will present soon. In addition, although attribute dispersion is a different component of choice task complexity, it can directly influence entropy (Adamowicz and Swait, 2001b).[22] Given the association between entropy and attribute dispersion, we examine the effect of attribute dispersion on the variance of the unobservable component of utility (or error variance) separately from entropy. This chapter examines how attribute dispersion is associated with the Gumbel error variance in the three EDs employed in our study here. We also investigate how the ordering of choice tasks in the three EDs can influence the error variance in more detail which we very briefly examined in Chapter 4. In the final part of the analysis, we used a pooled sample to

[21] We describe ASD and DSD in detail in section 5.2.

[22] Our experimental design data shows strong and significant relationship between entropy and attribute dispersion. We present the strength and significance of these relationships in Section 5.5.

jointly examine the effects of different EDs, attribute dispersion and choice task order on the error variance.

5.2 Measures of choice complexity

Discrete choice models generally assume that an individual is certain about his/her preferences. However, when an individual deals with complex decisions, he/she may become uncertain about the utility derived from the available alternatives. This may be due to the complexity of the choice environment where an individual may not fully understand the implications of the tradeoffs between alternatives. Many CE studies have shown that varying certain aspects of choice tasks influences the cognitive cost (or choice complexity) of evaluating the choice tasks (e.g., Dellaert, et al., 1999; DeShazo and Fermo, 2002; Ohler, et al., 2000; Hensher, 2003). Several studies have shown some empirical evidence that increasing the complexity levels of the choice tasks is positively associated with greater error variance. Some aspects of choice tasks found to be positively associated with the error variance include the following:

- number of alternatives in a choice task (Widlert 1998; Hensher et al. 2001; De Shazo and Fermo 2002; Arentze et al. 2003; DeShazo and Fermo 2004; Caussade et al. 2005; Hensher 2006; Rose et al. 2009);

- number of attributes per choice alternative (Mazotta and Opaluch 1995; Ohler et al. 2000; Pullman et al. 2000; Wang and Li 2002; DeShazo and Fermo 2002, 2004; Caussade et al. 2005; Hensher 2001, 2006);

- number of attribute levels (Dellaert et al. 1999; Wang and Li 2002; DeShazo and Fermo 2002, 2004; Caussade et al. 2005; Boxall et al. 2009); and

- dispersion of attribute levels across the alternatives in a choice set (that we refer to here as *attribute dispersion*) (DeShazo and Fermo 2002, 2004).

Whilst the number of alternatives, number attributes, and number of attribute levels have been shown to be positively associated with the error variance in many studies (including those enumerated above), few studies have examined the effect of attribute dispersion on error variance. DeShazo and Fermo (2002, 2004) provide some empirical evidence that attribute dispersion is positively associated with error variance suggesting that higher attribute dispersion would likely lead to decreasing choice determinism. This is because under the Random Utility Maximisation (RUM) theory, a higher error variance leads to lower contribution of the deterministic component of the utility function whilst the contribution of the stochastic component in explaining utility increases. We refer to choice determinism as inversely related to choice stochasticity. We define the increase in choice determinism as the decrease in variation in choice outcomes not explained by the underlying utility function. For example, an individual with perfect information given choice task t with three attributes would rank-order those attributes as A3 > A2 > A1. However, as choice task complexity increases (resulting to a decrease in choice determinism), an individual would likely provide inconsistent choices that could possibly make A3 no longer the most important attribute. A number of factors can contribute to lower choice determinism (higher choice stochasticity) and these include: (1) preference ordering may be incomplete such that A3 > A1 while A2 is not included in the rank-order; and (2) a respondent may become indifferent between the two alternatives and may choose randomly.

DeShazo and Fermo (2002) [hereby referred to as DSF] demonstrate that greater attribute dispersion contributes to a greater cognitive burden to respondents that result in lower choice determinism in the evaluation of alternatives in choice tasks. DSF compared different blocks of choice tasks with varying characteristics (e.g., different number of attributes, different number of attribute levels). All those choice tasks were generated using the ORD criterion. In this present study, we explore the effect of higher attribute dispersion of choice tasks to the cognitive cost of respondents. We formulate a null hypothesis that the effect of attribute dispersion on cognitive costs does not vary across different EDs. This study aims to answer the research question: *Do attribute dispersions in different ED criteria demonstrate varying effect on choice determinism? If so, which ED criterion provides the most benefit to a choice analyst based on the effect to choice determinism?*

In this chapter, we show how choice tasks from the three EDs differ in terms of attribute dispersion and how different dispersion levels influence the error variance. As ED criteria have different objective functions (e.g., ORD imposes orthogonality between attribute levels, BDD minimises the Bayesian D-error assuming $\beta \neq 0$, and OOD maximises D-optimality measure assuming $\beta = 0$), it can be expected that EDs generated from those criteria would differ in terms of the two dispersion measures – ASD and DSD. We describe below the formulae that we used to calculate for ASD and DSD of the set of choice tasks of each ED.

We calculate ASD_s based on the standard deviation in attribute levels of alternative j ($\boldsymbol{SD_j}$). SD_j is defined as the standard deviation among the normalized attribute levels of alternative j in choice task s and can be shown as

$$SD_j = \sqrt{\frac{\left(\sum_{i=1}^{I}\left(x_{ij} - \bar{x}_j\right)\right)^2}{J}}$$	(5.1)

where x_{ij} is the normalized i^{th} attribute of alternative j, i is the total number of attributes of alternative j. We calculate SD_j using an ordinal-integer metric where the attribute levels vary along three monotonically increasing levels of 1, 2, 3. These values represent the values we used in generating the three EDs. These levels were translated in the choice task as categorical metrics familiar to respondents. For example, the three levels of the attribute (that we include in the choice task for this study) on the occurrence of a threatened bird species can be translated as "sighted once", "sighted 3 times", and "sighted 5 times". Using the ordinal-integer metric for calculating SD_j ensures that attribute levels for each attribute are equally weighted on this alternative specific measure of dispersion. This measure of dispersion varies based on the similarity of attribute levels in an alternative. If all attribute levels in alternative j are all highly desirable, then the value of SD_j would be lower compared to say alternative h that contains a combination of the most desirable and the least desirable attribute levels. This is because an alternative with very similar attribute levels would tend to be cognitively easier to process since it does not require a respondent to make intra-alternative tradeoffs. An alternative with a highly dispersed set of attribute levels can be expected to be more difficult to process.

SD_j is used to calculate for the choice set specific measure ASD_j, which is the average standard deviation of attribute levels across alternatives in choice task s. Equation 5.2 shows that we simply divide the summation of SD_j by J which represents the total number of alternatives in choice task s:

$$ASD_s = \frac{\sum_{j=1}^{J} SD_j}{J} \tag{5.2}$$

ASD_s represents the <u>average</u> effect of SD_j for choice task s.

The second attribute dispersion measure for choice task s is DSD_s. DSD_s describes the <u>dispersion</u> of average standard deviation of attribute levels across alternatives of choice task s.[23] Higher values of DSD_s suggest greater degree of spread across alternatives in the within-alternative attribute dispersions. The study by DSF suggests that higher values of DSD_s correspond to greater cognitive cost that could contribute to increasing choice task complexity.

$$DSD_s = \frac{\sum_{j=1}^{J} \left(SD_j - ASD_s\right)^2}{J} \tag{5.3}$$

Other measures of complexity described in DSF, which include number of attributes per alternative and number of alternatives per choice task, are not used in this study because all respondents were provided with choice tasks with the same number of alternatives, with each alternative having the same number of attributes. Each respondent was provided with nine choice tasks. Given the nine choice tasks, there would likely be some order effects that could result in a respondent learning how to more efficiently evaluate the alternatives of the initial orders (1st, 2nd, 3rd) of choice tasks. Answering the latter sequence of the choice tasks could later result in

[23] One may also refer to DSD as the *dispersion of the dispersion* of attribute levels while ASD as the *average of the dispersion* of attribute levels.

fatigue or boredom whereby a respondent gets tired of evaluating the alternatives after several replications (e.g., 7^{th}, 8^{th}, 9^{th}). We describe how we examine learning and fatigue effects in the next section.

5.3 Learning and fatigue effects

Similar to the effects of attribute dispersion to choice determinism, several applications of CE have shown that the ordering of choice tasks influences the estimates of indirect marginal utility and the error variance (Bradley and Daly 1994; Stopher and Hensher, 2000; Hensher et al., 2001; Ortúzar et al., 2000; Ortúzar and Rodríguez, 2002; Pérez et al., 2003; Caussade et al., 2005; Holmes and Boyle, 2005; van der Waerden et al., 2006; Kjær et al., 2006; Bateman, et al., 2008; Day and Pinto Prades, 2010; Day et al., 2010). However, unlike attribute dispersion which would likely be positively (negatively) associated with the variance (scale) of the error term, the order effect can vary from initial choice task replications with increasing scale and then in the latter choice tasks with decreasing scale. Holmes and Boyle (2005) show a pattern of increasing scale factor as respondents learn to evaluate the alternatives as they proceed through a series of choice tasks. A number of empirical studies suggest that choice determinism is low for the first choice tasks, increases in the next ones then decreases again at some point (e.g., Caussade et al. 2005; Day et al. 2010, Scarpa et al. 2011b). This trend follows a common sequence where at first the respondent tries to learn the choice task and the effort needed to accomplish it, then he/she applies the learned behaviour in the next choice tasks, and finally the respondent gets tired or bored in the last choice tasks. Possible reasons for this trend include the occurrence of learning, boredom and fatigue (Adamowicz and Swait 2001b; DeSarbo et al. 2004; Holmes and Boyle 2005; DeSarbo et al. 2005). Given that different ED have different complexity levels they might tend to exhibit different patterns of learning/fatigue effects. This study also aims to answer the research questions:

(1) *Does learning/fatigue effects vary in different experimental designs?*

(2) *If so, which ED criterion provides the most benefit to choice analyst based on learning/fatigue effects?*

Task order effects in CE surveys would likely occur because each respondent is provided with several choice questions which can range between 3 and 64 replications. Caussade, et al. (2005) provided each respondent with 16 choice tasks and analytical results suggest that the order effects in the series of choice tasks can be divided into three parts: (1) the first eight choice tasks exhibited a trend of decreasing error variance indicating that a respondent tended to gain a better understanding of how to evaluate a sequence of choice tasks as they went through the first few replications; (2) the 9^{th} to the 11^{th} choice tasks show a decline in the learning pattern as fatigue or boredom overpowered a respondent's evaluating effort; and (3) the third part is from the 12^{th} choice task onwards where a typical respondent exhibited increasing fatigue levels as indicated by a pattern of increasing error variance.

While Caussade et al found fatigue effects, many other CE exercises did not find sufficient evidence of fatigue effects (e.g., Ohler et al., 2000; Savage and Waldman 2008) despite respondents completing relatively large number of choice tasks e.g., up to 64 choice tasks (Brazell and Louviere, 1997). This could possibly be due to the fact that the choice tasks evaluated were either simple or at least not too complex, so as not to engender the occurrence of fatigue (Day et al., 2010). Another reason could be associated with the ED criterion as used to construct the choice tasks, because different ED criteria would result in different choice task complexity levels.

For instance, the ORD criterion creates EDs that have uncorrelated attribute levels within alternatives and does not assume any parameter values. The ORD criterion also assumes that respondents have an equal preference for all attribute levels and therefore assumes that the contribution of an alternative in a choice task to the observed utility is the same as the other alternatives (Grossmann, Holling, and Schwabe, 2002). There is also an ED criterion that imposes utility balance where respondents would likely place equal weight on the set of alternatives in a choice task and this would probably be difficult for respondents to answer (Huber and Zwerina, 1996). However, a choice task derived from a utility balanced design would likely have a high entropy level because of high similarity in utilities across alternatives in a choice task (Swait and Adamowicz, 2001a). Viney, et al. (2006) report evidence of having a set of utility balanced choice tasks which positively correlates with error variance. This suggests that ED criterion selection could influence the way respondents answer a series of choice tasks. Based on the above discussion we formulate the null hypothesis that *the ED criterion does not influence the learning, fatigue or boredom effects*. We test this hypothesis using the data collected from a CE exercise with nine choice tasks where we collected choice data using the three EDs that we described in Chapter 2 and evaluated in Chapter 4. We describe how we model choice determinism as a function of choice complexity and order effects in the next section.

5.4 Measuring choice determinism

In measuring choice determinism, we use an observable proxy which is the standard deviation of the random error in the individual’s utility function represented as “σ”. A lower value of σ indicates an increase in choice determinism. Assuming that the unobserved effects or error terms ε, are Extreme Value Type I distributed,

$\sigma = \frac{\pi}{\lambda \cdot \sqrt{6}}$ where λ represents the scale factor that is inversely related to σ. The assumption that the error terms are independent and identically distributed (*i.i.d.*) allows the scale factor to be fixed which implies that the utility function can be scaled by an arbitrary constant without affecting logit choice probabilities, P_{js}. Under the *i.i.d.* assumption, λ is often assigned a fixed value of approximately 1.28255 corresponding to the usual assumption of $\sigma = 1$.

To relax the *i.i.d.* assumption we allow the systematic component of the error term to be explained by the scale factor λ. To do this we employ the heteroskedastic (also called covariance heterogeneity) logit model to parameterise λ to vary across different measures of choice determinism. This can be shown as

$$P_{iqs} = \frac{\exp(\lambda_q (C_q) \cdot (V_{iqs}))}{\sum_{A_j \in A(q)} \exp(\lambda_q (C_q) \cdot (V_{jqs}))} \tag{5.4}$$

where V_{iqs} represents the observed component of utility; λ_q is a function of a vector of q choice determinism measures, C_q, which includes ASD and DSD as well as measures of choice task order effects described in Section 5.3. This parameterisation of the scale factor follows Swait and Adamowicz (2001a) where λ is specified as an exponential function to preclude negative scale parameters. Although this specification results to a highly non-linear-in-parameters model, it has excellent convergence properties (Swait and Adamowicz, 2001a; DeShazo and Fermo, 2002; DeShazo and Fermo, 2004). To estimate heteroskedastic logit models for this exercise we have used Biogeme 1.8 (Bierlaire, 2009).

$$\lambda_q(C_q) = \exp\left(\sum_{q=1}^{Q} \gamma_q C_q\right) \tag{5.5}$$

The scale factor in Equation 5.5 above is no longer constant, as it is a function of choice determinism, which is in turn approximated by measures of attribute dispersion and task order effects. The sign of the estimated scale coefficient γ_q shows how the scale factor is affected by the q^{th} choice determinism measure (e.g., DSD). A negative sign indicates a reduction in scale which implies a decrease in the level of choice determinism. If the scale coefficient for *ASD* is positive, this would imply that higher *ASD* would contribute to increasing the scale factor or increasing choice determinism. The effects on scale in the initial choice tasks of the sequence can be expected to be positive as these contribute to learning, which in the early stages is strong. The coefficients for the latter choice tasks can be expected to have negative coefficients as fatigue and/or boredom would likely contribute to a decrease in choice determinism. We also expect that the effects of attribute dispersion and order effects on the scale factor would vary across choice data sets collected using different EDs. We present our analytical results in the next section.

5.5 Results

Using the attribute dispersion measures in equations 5.1 to 5.3, we calculate the ASDs and DSDs of the three EDs (i.e., ORD, BDD, OOD) that we used to collect choice data for the *Balanced Data Set*[24]. We present and compare the calculated values of ASD and DSD of the three EDs in section 5.5.1. In section 5.5.2, we discuss the relationships of ASD and DSD with the entropy values (presented in Chapter 4 of this

[24] The *Balanced Data Set* is described in Chapter 2.

thesis) based on pairwise correlation coefficients. Section 5.5.3 presents the effect of ASD and DSD on the scale factor across the three designs based on the analysis of the Balanced Data Set. In section 5.5.4, we analysed again the Balanced Data Set where we specifically evaluate the effects of learning and fatigue on the scale factor for each ED sample (e.g., choice data set collected using ORD). In section 5.5.5, we analyse the pooled sample to investigate how measures of choice determinism (i.e., ASD, DSD, order effects) jointly affect the scale factor while controlling for the effects of different EDs.

5.5.1 Attribute dispersion levels of the three experimental designs

We present the attribute dispersion measures of the three experimental designs (EDs) in Table 5.1. Each ED has a total of 243 choice alternatives. Two-thirds of these alternatives are generated from one of the three ED criteria (and are called *designed alternatives*) whilst one-third are Status Quo (SQ) alternatives that represent the current set of biodiversity levels that are fixed and were not derived from an ED. The *two-third* to *one-third* ratio arises due to the fact that each choice task has three alternatives: two designed alternatives and an SQ alternative. The 243 alternatives of each ED are divided into three blocks (Blocks 1, 2 and 3) with each block distributed into nine choice tasks.

Following Equations 5.1 to 5.3, we calculated the two choice task specific measures of attribute dispersion *ASD* and *DSD*. Table 5.1 shows that ORD has the lowest overall mean *ASD* of 1.10 while BDD and OOD have virtually the same mean *ASD*. As *ASD* measures the average standard deviation in attribute levels in a choice task, this indicates that attribute levels within alternatives in ORD are more similar compared to BDD and OOD. DeShazo and Fermo (2002) suggest that an alternative with similar levels (e.g., all high, all low, all moderate) requires less cognitive effort to process. The minimum and maximum ASD values for ORD are always lower compared to the two other designs which might indicate that ORD has the least complex choice tasks among the three EDs, given the relatively higher similarity in attribute levels of each alternative. However, the spread of *ASD* is highest for ORD as indicated by the standard deviation of 0.19 compared to BDD and OOD with 0.13 and 0.15, respectively. This suggests that although on average ORD choice tasks have relatively lower complexity, the level of complexity between choice tasks vary the most for ORD. This is further demonstrated by the fact that the ranges (*Max less Min*) of *ASD* for all three blocks of ORD, especially Block 1, are all greater than the ranges of the three blocks of the other two EDs (please see Table 5.1). The distributions of *ASD* for the three *EDs* are illustrated using histograms with kernel density graphs in Figures 5.1a, 5.1b and 5.1c. The figures show that ORD has the widest spread of *ASD* followed by OOD and then BDD. *ASD* figures and graphs in Table 5.1 and Figure 5.1b suggest that BDD has the least spread of *ASD*. This implies that the complexity levels of the set of choice tasks generated using BDD criterion have the smallest range of variation of complexity among the three designs. The dispersion of ASD for the three EDs are further illustrated in Figure 5.2 using kernel density graphs demonstrating that the ASD values for ORD have relatively lower densities and exhibit the widest spread.

Both *ASD* and *DSD* are choice task specific measures of complexity. However, unlike *ASD*, which represents a measure of the average complexity of the three alternatives in a choice task, *DSD* is a measure of the spread of complexity across alternatives in a choice task. Table 5.2 shows that ORD and OOD are tied as having the highest overall mean *DSD* of 0.19. ORD demonstrates greater variation in DSD across three blocks while OOD have consistently high level of DSD across blocks. BDD has the lowest overall mean *DSD* of 0.16 and this value is virtually the same across three blocks suggesting that BDD consistently demonstrates low dispersion of complexity across blocks. In addition, the standard deviation of DSD and the maximum of DSD are lowest for BDD which indicate that this design has the lowest spread of dispersion of complexity among the three designs. The histograms with kernel density graphs in Figure 5.3a, Figure 5.3b and Figure 5.3c illustrate the spread of dispersion of the three EDs. Again BDD exhibits the lowest level of complexity among the three designs in terms of overall mean and overall spread of dispersion of complexity of choice tasks. We also present kernel density graphs of DSD in Figure 5.4 to further illustrate that ORD demonstrates the greatest spread in terms of the DSD measure of attribute dispersion. We can see here a pattern that ORD choice tasks have wider spread of ASD and DSD than the two other design criteria. This is probably because BDD and OOD are optimised for particular beta values (e.g., $\beta = 0 \ or \ \beta \neq 0$). We speculate that the optimisation processes in BDD and OOD might have contributed to the narrowing of spread of dispersion. For instance, minimising the D-error would likely contribute to minimising the spread of dispersion. However, this might not be generalisable and further investigation should be done to verify this.

The effects of *ASD* and *DSD* on choice determinism can be tested by using choice survey data and examine how these two measures of attribute dispersion are associated with the scale factor. We conjecture that the wide range of variability of ASD of the choice tasks in ORD would contribute to a greater decrease in choice determinism than other designs. To examine whether the effects of ASD and DSD on the scale factor vary across designs, we use again the heteroskedastic logit model. Estimation results from these are presented in section 5.5.3. The next section shows the relationship between attribute dispersion and entropy of the choice tasks across the three EDs examined here.

Table 5.1: Average standard deviation (ASD) of attribute levels across alternatives in a choice task

	Orthogonal (ORD)				Bayesian D-efficiency (BDD)				Optimal Orthogonal (OOD)				Pooled
	Block1	Block2	Block3	All	Block1	Block2	Block3	All	Block1	Block2	Block3	All	
Mean	1.07	1.12	1.11	1.10	1.19	1.13	1.17	1.17	1.19	1.17	1.17	1.18	1.15
Standard deviation	0.28	0.11	0.15	0.19	0.15	0.11	0.11	0.13	0.05	0.10	0.16	0.11	0.15
Minimum	0.48	0.92	0.81	0.48	0.98	1.01	1.01	0.98	1.05	1.05	0.73	0.73	0.48
Maximum	1.41	1.29	1.32	1.41	1.46	1.32	1.35	1.46	1.23	1.32	1.29	1.32	1.46
Range (Max less Min)	0.93	0.37	0.51	0.93	0.48	0.31	0.34	0.48	0.18	0.27	0.56	0.59	0.98
Number of choice tasks	27	27	27	81	27	27	27	81	27	27	27	81	243

Table 5.2: Dispersion of standard deviation (DSD) of attribute levels across alternatives in a choice task

	Orthogonal (ORD)				Bayesian D-efficiency (BDD)				Optimal Orthogonal (OOD)				Pooled
	Block1	Block2	Block3	All	Block1	Block2	Block3	All	Block1	Block2	Block3	All	
Mean	0.25	0.12	0.19	0.19	0.15	0.16	0.16	0.16	0.17	0.17	0.21	0.19	0.18
Standard deviation	0.12	0.10	0.11	0.12	0.08	0.06	0.09	0.08	0.06	0.14	0.15	0.12	0.11
Minimum	0.05	0.02	0.07	0.02	0.02	0.06	0.03	0.02	0.10	0.07	0.02	0.02	0.02
Maximum	0.39	0.30	0.43	0.43	0.31	0.23	0.36	0.36	0.29	0.53	0.53	0.53	0.53
Range (Max less Min)	0.34	0.28	0.36	0.41	0.29	0.17	0.33	0.34	0.19	0.46	0.51	0.51	0.51
Number of choice tasks	27	27	27	81	27	27	27	81	27	27	27	81	243

Figure 5.1a: Histogram and kernel density of ASD for ORD

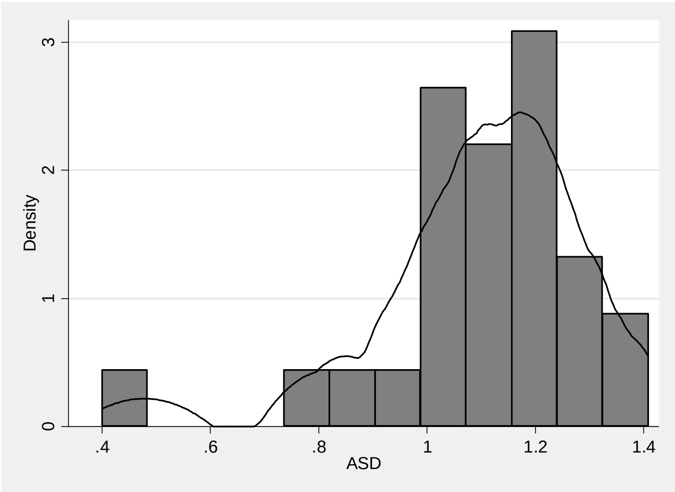


Figure 5.1b: Histogram and kernel density of ASD for BDD

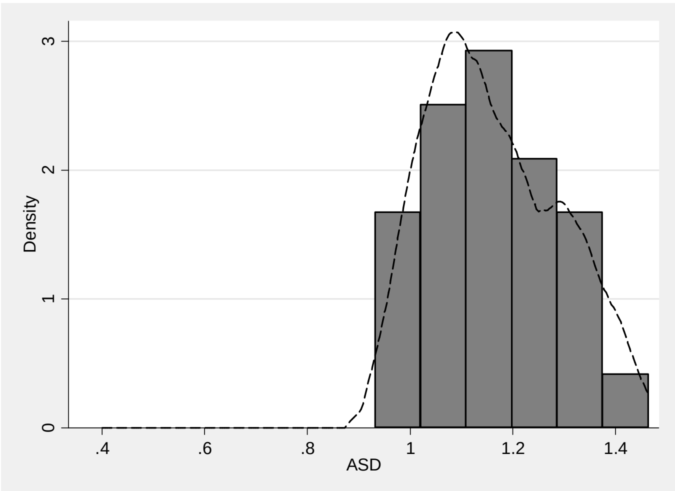


Figure 5.1c: Histogram and kernel density of ASD for OOD

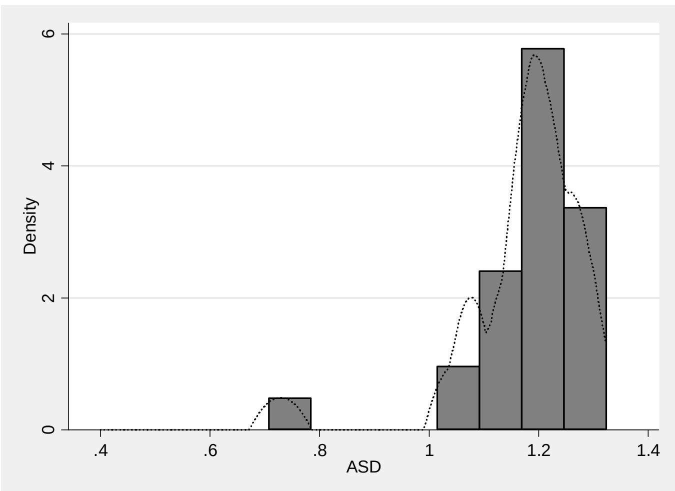


Figure 5.2: Kernel density of ASD by design

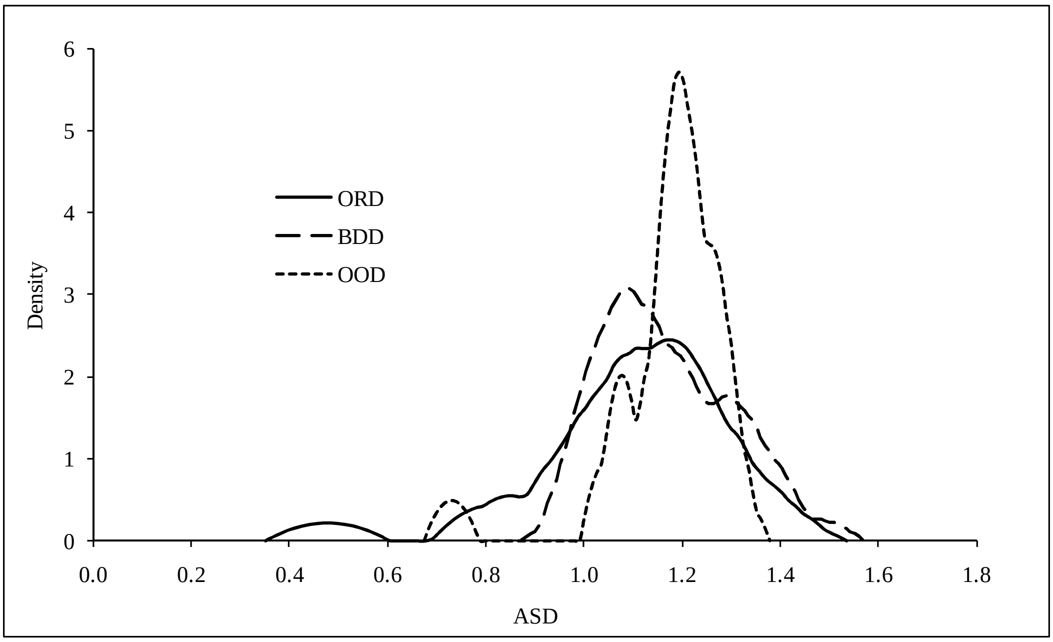


Figure 5.3a: Histogram and kernel density of DSD for ORD

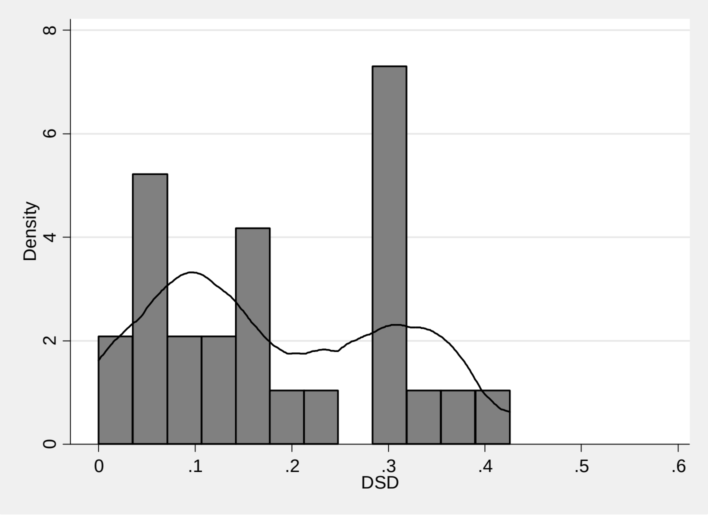


Figure 5.3b: Histogram and kernel density of DSD for BDD

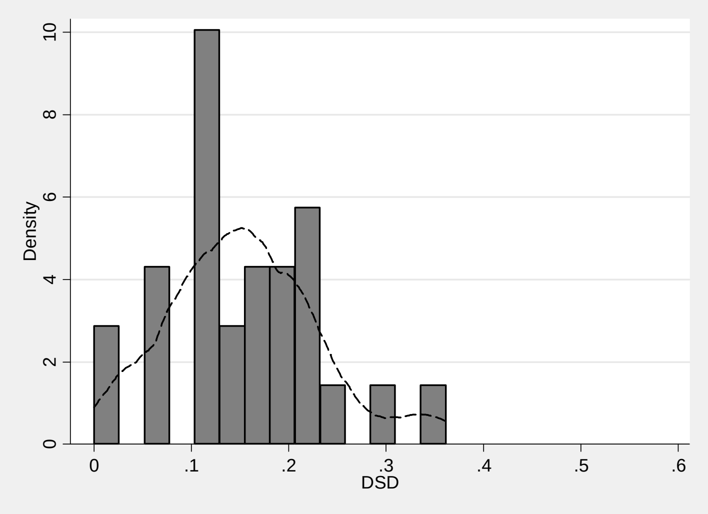


Figure 5.3c: Histogram and kernel density of DSD for OOD

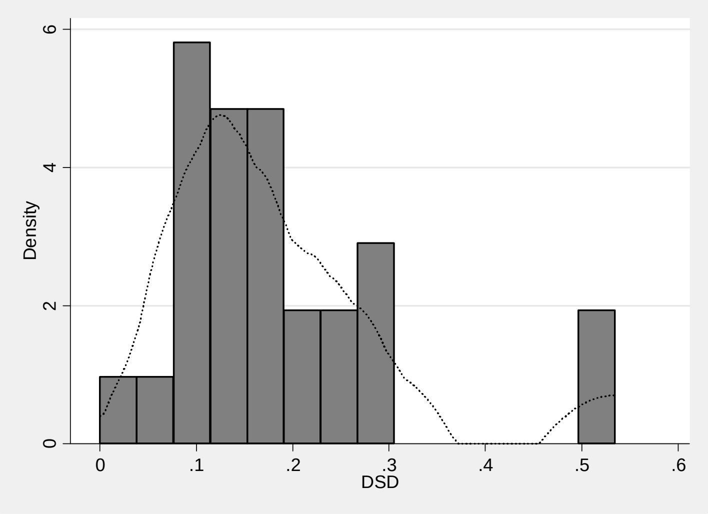
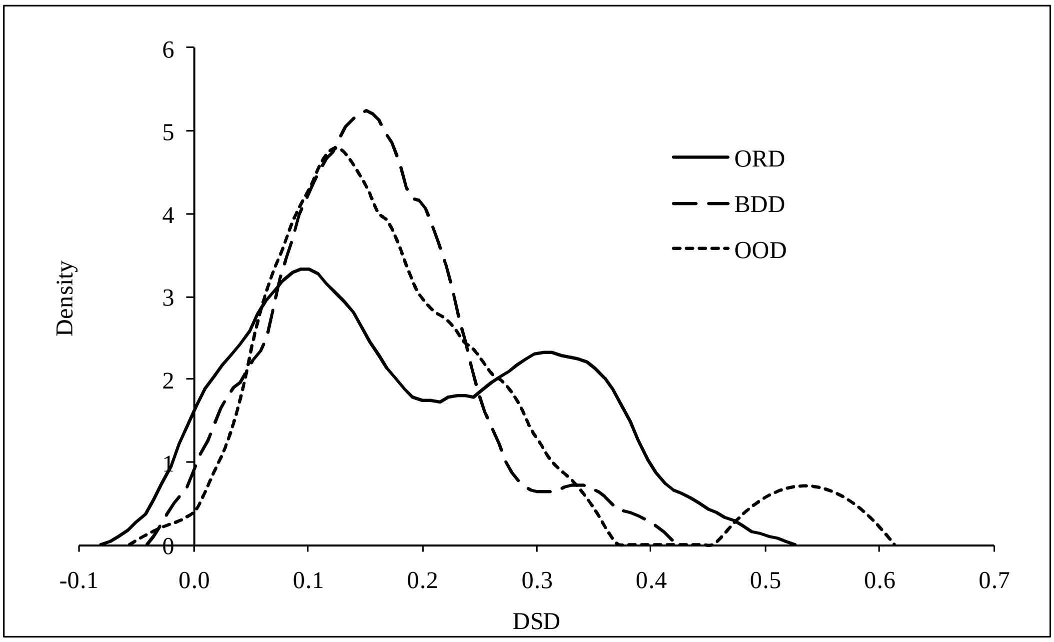


Figure 5.4: Kernel density of DSD by design



5.5.2 Relationship of ASD and DSD with entropy by design

We mentioned earlier that although entropy and attribute dispersion are two different components of choice complexity, the two are associated because both affect choice probabilities (Swait and Adamowicz, 2001b). As entropy and attribute dispersion are both choice task specific measures of complexity, we can examine the relationship between them for each ED. For each choice task we calculated the entropy value that we used to undertake our analysis on entropy in Chapter 4 of this thesis. For the attribute dispersion measures, we calculated the ASD and DSD values of each choice task which are summarised in Tables 5.1 and 5.2. To explore whether correlation applies to our design data sets we calculated pairwise correlation coefficients between entropy and attribute dispersion using those choice task specific values. Column 5 of Table 5.3 shows that the pooled set of designs indicate that entropy values and attribute dispersion values are somewhat positively correlated in a significant way. ASD and entropy are significantly negatively correlated because the increase in the dispersion of attribute levels of alternatives across a choice task increases the difference between alternatives in terms of contribution to utility. This in turn increases the difference in choice probabilities of alternatives in a choice task. For example, a choice task with three alternatives sq, a1, a2 with a low ASD value (e.g., 0.4) would have choice probabilities of sq=0.32, a1=0.33, a2=0.35 while one with high ASD value (e.g., 1.5) would have sq=0.05, a1=0.65, a2=0.35.[25] As mentioned in Chapter 4, a choice task with alternatives having very similar contributions to utility would have a very high entropy value and vice versa. Thus, the above supports the assertion that ASD is negatively correlated with entropy. For the pooled set of designs, DSD and entropy are significantly positively correlated. This is because as DSD increases, the contributions to utility of alternatives in a choice task become similar leading to an increase in the entropy level of a choice task.

Columns 2 and 4 of Table 5.3 show that the correlation between attribute dispersion and entropy for ORD and OOD are consistent with the pooled design set. However, for BDD, entropy and DSD have weaker positive correlation and lack statistical significance. This suggests that the increase in DSD in the Bayesian D-efficient design contributes less to increasing entropy compared to the two other EDs. A possible reason for this is that, as Table 5.2 shows, BDD choice tasks have relatively lower mean DSD (and narrower range of DSD) compared to the other EDs.

We have also calculated the correlation coefficients for ASD and DSD for each design sets and pooled design set. Table 5.3 shows a relatively strong negative correlation in BDD while the two other EDs have weak positive correlations. This

---

[25] Please note that those values of choice probabilities were arbitrarily selected to distinguish a choice task with high ASD from one with low ASD.

implies that in choice tasks generated using the BDD criterion, one can expect that as ASD increases DSD decreases while the opposite applies for the two other EDs. This demonstrates that BDD has a different attribute level dispersion property compared with ORD and OOD.

Table 5.3: Summary of correlation coefficients showing the association between attribute level dispersion and entropy of groups of choice tasks

	ORD	BDD	OOD	Pooled
ASD and Entropy	-0.327	-0.270	-0.239	-0.227
	(0.003)	(0.015)	(0.032)	(0.000)
DSD and Entropy	0.438	0.139	0.387	0.327
	(0.000)	(0.215)	(0.000)	(0.000)
ASD and DSD	0.049	-0.203	0.037	-0.021
	(0.666)	(0.069)	(0.742)	(0.746)
Number of choice tasks observations	27	27	27	81

Note: Figures in parentheses represent *p*-values.

5.5.3 Effects of ASD and DSD on the scale factor by design

To examine the effects of ASD and DSD on the scale factor, we analyse the balanced data set described in detail in Chapter 2. Columns 3 and 4 of Table 5.4 show the log-likelihood values from conditional (or homoskedastic) logit and heteroskedastic logit models of the three design treatment and the pooled sample. The homoskedastic logit approach assumes that scale is not a function of components of complexity (via the ASD and DSD proxies) while the heteroskedastic logit approach allows the scale to vary based on ASD and DSD values of each choice task. Column 5 presents the chi-square statistics for the hypothesis that both ASD and DSD terms in the scale function are zero, suggesting that preferences are not a function of complexity. Since the critical value is 5.991 at the 95% confidence level, we reject the null hypothesis that the scale factor is not a function of complexity in the BDD treatment. Since the analysis of ORD and OOD data indicates that we lack sufficient evidence to reject homoskedasticity, this suggests that the complexity levels in ORD and OOD choice tasks do not influence choice determinism. We also fail to reject homoskedasticity in the pooled sample. However, it is important to note that, the heteroskedastic logit models used in Table 5.4 suffer from a problem of having non-significant estimates of utility coefficients. This could have implications for the robustness of the conclusions drawn from the effects of designs on choice behavior. Estimated utility coefficients for the joint effects of attribute dispersion for the split design and pooled samples are reported in Appendix Table 5.

Column 6 of Table 5.4 shows significantly positive scale coefficient estimates of ASD in the BDD sample suggesting that higher ASD leads to increase in scale. This indicates that higher ASD is actually advantageous as it contributes to a decrease in the cognitive cost of respondents in evaluating choice tasks. As Table 5.3 shows somewhat a negative correlation (-0.203) between ASD and DSD for the BDD choice tasks, we tested what would happen to the *p*-values of scale coefficients if we estimated ASD and DSD separately. The 10^{th} row of Table 5.4 shows a lower *p*-value and a higher magnitude of the ASD scale coefficient for BDD when estimated separately with DSD. This suggests that ASD in the BDD treatment, is the only measure of dispersion that positively influences scale. This is corroborated by the result reported in the third to the last row where the DSD coefficient remains not significant in terms of effect on scale.

Column 7 row 4 of Table 5.4 shows a negative scale coefficient for DSD (significant at the 88% confidence level) in the ORD sample indicating that higher DSD contributes to an increase in cognitive cost. This is consistent with the findings in DeShazo and Fermo (2002), which also generated choice tasks using the ORD criterion. Although we know from Table 5.3 that ASD and DSD are very weakly associated (or not associated) in the ORD treatment, we still separately estimated the two scale coefficients as shown in rows 9 and 14 in Table 5.4. Separate estimation shows an increase in most *p*-values of the ASD and DSD scale coefficients indicating that the joint effect of the two on scale is stronger than their individual or separate effects. However, the BDD sample is an exception where the 10th row shows that the *p*-value for ASD scale coefficient is lower and the magnitude of the coefficient is also higher indicating a stronger effect with separate scale estimation.

Although Table 5.4 shows that only the BDD sample has a statistically significant scale coefficient estimate for ASD (as indicated by a *p*-value of 0.06), three other scale coefficient estimates are nearly significant. We plotted scale coefficient estimates for ASD and DSD in Figures 5.5 and 5.6 to illustrate their association to the scale factor for the three designs. Figure 5.5 shows the variation of the effect of ASD on the scale factor for BDD, OOD and Pooled samples. The graph for ORD was not included because ASD in that sample did not lead to any statistically significant decrease (increase) in scale (variance) as shown in Table 5.4. Figure 5.5 shows that the higher ASD in BDD and OOD samples result in an increase in the scale factor which implies that higher ASD, especially in BDD, contributes to increasing the consistency of choices made by respondents. In terms of the impact of DSD on the scale factor, Figure 5.6 shows that higher DSD in the ORD sample would likely contribute to greater choice complexity (or lower choice determinism) as indicated by a decreasing trend in the scale factor as DSD increases.

Table 5.4: Conditional and heteroskedastic logit model estimates

(1)	(2)	(3)	(4)	(5)	(6)	(7)
Experi-mental Design (ED)	No. of observed choice sets	Log-likelihood Conditio-nal Logit	Log-likelihood Heteroskedastic Logit	Likelihood Ratio Test (-2*(LL_0-LL_1))	Estimated Scale Coefficient of ASD (Robust *p*-value)	Estimated Scale Coefficient of DSD (Robust *p*-value)
ASD and DSD						
ORD	503	-459.28	-457.51	3.54	-0.55 (0.33)	-1.49 (0.11)
BDD	503	-497.66	-493.85	7.62*	2.59 (0.06)	-1.72 (0.41)
OOD	503	-469.62	-468.77	1.70	1.42 (0.20)	-0.76 (0.56)
Pooled	1509	-1460.32	-1458.18	3.02	-0.65 (0.14)	-1.15 (0.13)
ASD Only						
ORD	503	-459.28	-458.98	0.60	-0.43 (0.45)	
BDD	503	-497.66	-494.28	6.76*	2.79 (0.05)	
OOD	503	-469.62	-468.98	1.28	1.27 (0.20)	
Pooled	1509	-1460.32	-1459.57	1.49	-0.55 (0.21)	
DSD Only						
ORD	503	-459.28	-458.01	2.54		-1.40 (0.13)
BDD	503	-497.66	-496.80	1.72		-2.25 (0.23)
OOD	503	-469.62	-469.47	0.30		-0.68 (0.62)
Pooled	1509	-1460.32	-1459.24	2.16		-1.40 (0.13)

* Significant at the 95% confidence level as it exceeds critical Chi-square value with 2 degrees of freedom $\chi_2^{0.95} = 5.99$ (One degree of freedom = 3.084)

Figure 5.5: Effect of ASD on the scale factor by design

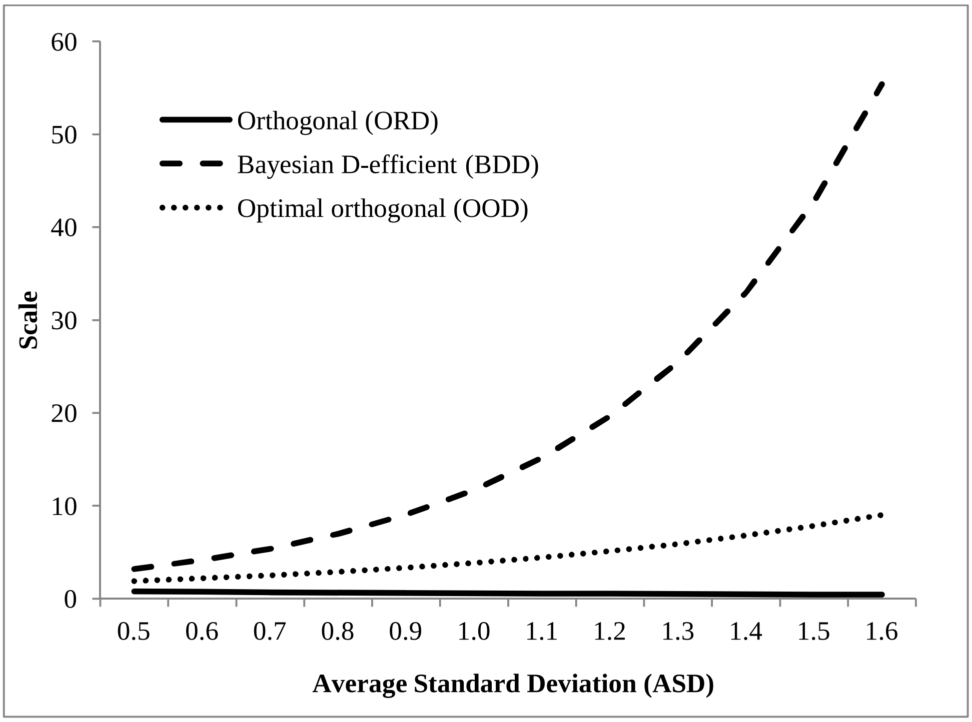
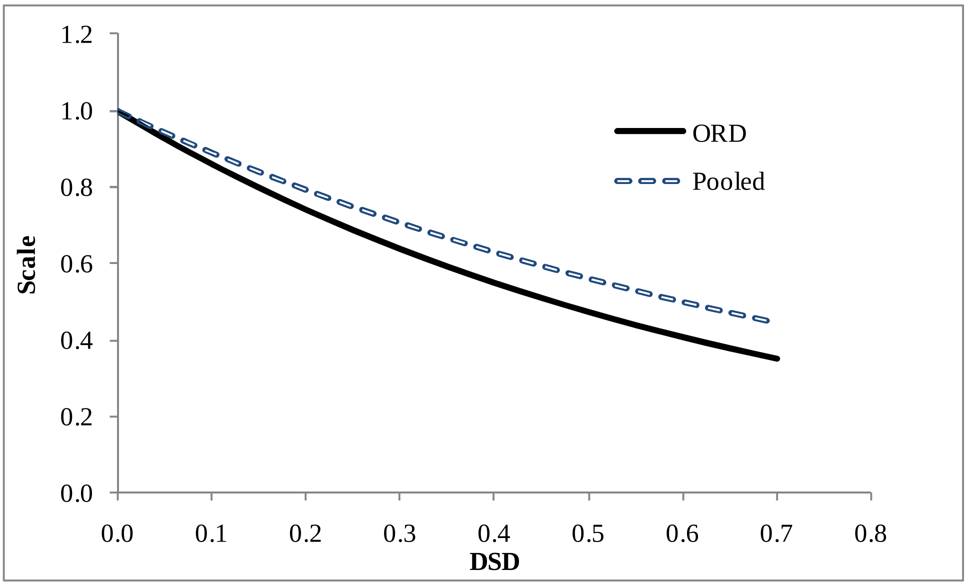


Figure 5.6: Effect of DSD on the scale factor by design



5.5.4 Choice task order effects by design

We present here the estimation results on the effects of choice task order on the scale factor. Estimates of scale coefficients for task order effects are presented in Table 5.5. All three heteroskedastic logit regression models converged with ORD taking only 30 iterations to converge while BDD took 298 iterations with both starting the search from pre-estimated initial values. The highly non-linear-in-parameters specification might have contributed to making some of the utility coefficients to become non-significant compared to the estimates of the homoskedastic conditional logit model. Another possible reason is that higher statistical efficiency of BDD has contributed to a lower variation in both DSD and ASD and this might have affected the model estimation process.

Table 5.5 shows that the scale coefficients for choice task order in the ORD sample are all non-significant, suggesting that this ED neither exhibits learning nor fatigue effects. The set of scale coefficients for BDD, instead, exhibits learning effects as shown by a steadily increasing coefficient values from 2^{nd}&3^{rd} choice task to the 8^{th} choice task. There seems to be strong evidence of learning especially in the 8^{th} choice task where the scale coefficient increases from 1.13 in the 7^{th} to 1.84 in the 8^{th} choice task. Although the coefficient for the 9^{th} choice task is slightly lower than the 8^{th}, the former remains higher than the coefficient for the 7^{th}. This shows a pattern of continuous and sustained learning in the case of the BDD which may indicate that we could have increased the number of choice task in our survey from nine to maybe sixteen choice tasks to build upon the favourable learning effect brought about by using BDD. In terms of the OOD sample, the set of scale coefficients show a sign of significant learning effect in the 6th choice task. However, there is neither a clear sign of a continuous learning nor a clear indication of fatigue effects.

Figure 5.7 illustrates the effects of task order on the scale factor by design. BDD exhibits the best pattern of learning beginning from the 4th choice task to the 8th.[26] This pattern of continuous learning might indicate that respondents tended to become increasingly interested in evaluating the sequence of choice tasks as they progress through the first eight choice tasks. It might also indicate a pattern that respondents exerted more effort to understand the succeeding choice tasks as demonstrated by a trend of increasing scale coefficients (increasing choice determinism) up to the 8th task order. The graph for OOD exhibited a distinctive learning effect on the 6th choice task while ORD does not show any learning effect. All EDs do not exhibit any significant fatigue effects which are consistent with findings in Ohler, et al., (2000); Savage and Waldman (2008); and Brazell and Louviere (1997).

---

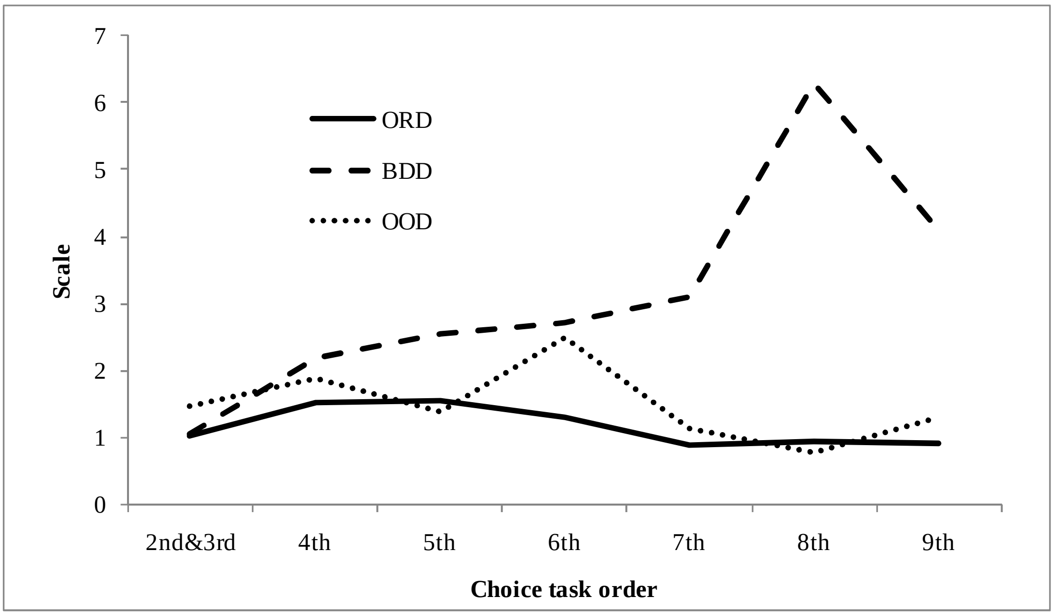
[26] We used the first choice task as reference since it is likely to be the most difficult task to evaluate because respondents would typically exert the greatest effort to learn how to properly choose the preferred alternative.

Table 5.5: Heteroskedastic logit model estimates of choice task order effects on the scale factor

	ORD			BDD			OOD		
	Coef	Robust Std Err	Robust *p*-value	Coef	Robust Std Err	Robust *p*-value	Coef	Robust Std Err	Robust *p*-value
Utility coefficients									
Brown kiwi 1	*0.447*	*0.275*	*0.10*	0.164	0.156	0.29	*0.486*	*0.292*	*0.10*
Brown kiwi 2	**0.618**	**0.353**	**0.08**	*0.225*	*0.141*	*0.11*	**0.469**	**0.230**	**0.04**
Native fish 1	**0.372**	**0.214**	**0.08**	0.246	0.273	0.37	0.137	0.162	0.40
Native fish 2	0.225	0.232	0.33	-0.004	0.101	0.97	0.186	0.150	0.21
Native plant 1	0.259	0.283	0.36	-0.145	0.255	0.57	0.092	0.161	0.57
Native plant 2	-0.184	0.298	0.54	0.033	0.115	0.77	0.068	0.162	0.68
Green gecko 1	0.095	0.185	0.61	0.066	0.118	0.57	-0.116	0.152	0.44
Green gecko 2	0.379	0.248	0.13	-0.007	0.068	0.92	0.030	0.155	0.85
Bush falcon 1	0.469	0.341	0.17	0.261	0.339	0.44	0.148	0.236	0.53
Bush falcon 2	0.787	0.508	0.12	0.374	0.326	0.25	**0.320**	**0.169**	**0.06**
Cost to resp	**-0.025**	**0.015**	**0.08**	-0.008	0.008	0.26	**-0.023**	**0.009**	**0.02**
Indicator for non-SQ	-0.509	0.414	0.22	0.003	0.162	0.98	0.296	0.321	0.36
Scale coefficients									
2nd & 3rd task order	0.015	0.558	0.98	0.045	2.190	0.98	0.384	0.459	0.40
4th	0.416	0.692	0.55	0.777	1.090	0.48	0.628	0.571	0.27
5th	0.439	0.766	0.57	0.937	1.090	0.39	0.314	0.495	0.53
6th	0.254	0.680	0.71	1.000	1.400	0.47	**0.916**	**0.490**	**0.06**
7th	-0.136	0.609	0.82	*1.130*	*0.704*	*0.11*	0.134	0.697	0.85
8th	-0.058	0.773	0.94	**1.840**	**1.050**	**0.08**	-0.251	0.765	0.74
9th	-0.091	0.648	0.89	1.410	1.020	0.17	0.257	0.552	0.64
Log-likelihood value	-456.46			-490.59			-465.36		
Number of observations	503			503			503		
Number of iterations	30			298			37		
Converged?	Yes			Yes			Yes		

Note: Figures in boldface font are significant at the 90% confidence level; those in italics are nearly significant.

Figure 5.7: Choice task order effects on scale by experimental design



It is important to note that, in the survey questionnaire, we provided as much information as possible to the respondents about the five environmental attributes to illustrate a clear picture of the environmental good that they valued. We have also provided them with some coaching on how to properly evaluate each choice task. On page 6 of the survey questionnaire, we provided respondents with a description of each attribute and the coaching was represented by a demonstration of how a person might think her way through to evaluate each choice task. Given the amount of details that we provided them, it seems that, based on the results of choice task order effects, the sequence of choice tasks generated from the BDD criterion has sustained respondents' enthusiasm to answer the sequence of choice tasks.

5.5.5 Heteroskedastic logit regressions on pooled sample

To jointly test for the effects of attribute dispersion and task order on the scale factor, we analyse the pooled balanced data set using the heteroskedastic logit model described in Equation 5.4. The pooled data set has 1509 choice observations as each of the three design treatments has 503 observations. Estimates for the three heteroskedastic logit models are presented in Table 5.6. The columns under Model 1 are estimates for the conditional logit model, where the estimated utility coefficients are consistent with economic theory as sign that environmental improvements are positive, while the marginal utility of income exhibits a negative sign. The utility coefficients for additional native birds (i.e., kiwi and falcon) and a Level 1 increase in the number of native fish are significantly positive indicating that those improvements are valued by respondents and they would be willing to pay to support the increase in population of those threatened species. However, the coefficients for the increase in the number of geckos and the native plant kakabeak are not significant. As mentioned earlier, one might believe these attributes were irrelevant to the population and that they should not have been included in the choice task.

Model 2 is a heteroskedastic logit model where we investigate the effects of task order, experimental designs BDD and OOD, and attribute dispersion (ASD and DSD) on the scale factor.[27] In a side regression, where we evaluated the task order effects on the scale factor, we found that orders 4 to 6 and orders 7 to 9 contributed similar magnitudes of positive increases in the scale factor. From this result, we elected to use an indicator variable for two groups of task orders in Model 2. The indicator variable for *Task orders 4 to 6* is significantly positive at the 99.9% confidence level. This indicates the presence of learning effect that increases (decreases) scale (error variance) thereby increasing choice determinism. The scale

---

[27] Model 2 has a higher log-likelihood value (-1446) compared to Model 1 (-1460). Likelihood ratio test shows a Chi-square value of 28.40 that exceeds the critical $\chi_5^{99.9} = 20.52$ .Thus, the null hypothesis that scale is not a function of choice inconsistency is strongly rejected.

coefficients for BDD and OOD are both negative with BDD having a *p*-value less than 0.01. This suggests that, with ORD as the reference design, BDD contributes significantly to reducing choice determinism while OOD does not. This result contradicts the findings and discussion in sections 5.5.3 and 5.5.4 suggesting that the BDD criterion produces choice tasks that contribute to increasing the scale factor. Section 5.5.4 suggests that BDD exhibits the most favourable learning effect as indicated by the steady increase in the scale factor from task orders 3 to 8 while the two other EDs do not show any pattern of continuous learning. There could possibly be some joint effects between EDs, attribute correlations and task orders. To account for joint effects, we include interaction variables in the set of scale coefficients in Model 3.

Accounting for joint effects, Model 3 estimates suggest that BDD does not have any net effect on scale, but OOD has a highly significant negative net effect on the scale factor of -3.2 (with ORD as the reference design). Although the interaction between OOD and ASD (or the joint effect of OOD and ASD) has a significantly positive scale coefficient of 2.0, the magnitude of this still does not sufficiently compensate for the greater negative magnitude of the net effect of OOD on scale. By controlling for the joint effect of ED and attribute dispersion, we find the result that supports the findings in section 5.5.3 where, relative to ORD, BDD does not contribute to decreasing choice determinism while OOD significantly contributes to decreasing choice determinism. In terms of joint effects of ED and task order, the interaction between *BDD* and *Task Orders 7 to 9* is positive and significant at the 98% confidence level. This supports the finding in 5.5.4 that the greater learning effect in BDD, relative to the other EDs, contributes to increasing choice determinism.

In terms of the net effect of attribute dispersion, Model 3 shows that higher levels of DSD significantly reduce choice determinism. These results are consistent with the findings in DeShazo and Fermo (2002, 2004) that suggest that greater attribute dispersion is positively associated with an increase in choice inconsistency. With regard to the net effect of *Task_orders_4_to_6*, the magnitude of its effect has diminished from Model 2 to Model 3 as it was interacted with the design variables. However, it remains positive and significant at the 92% confidence level.

With regard to the joint effect of EDs and attribute dispersion, the scale coefficient for the interaction of BDD and DSD is significantly positive suggesting that higher DSD in choice tasks generated using the BDD criterion would likely increase the scale factor relative to choice tasks from ORD criterion with high DSD. OOD $\times$ ASD is also positively significant indicating that higher ASD in choice task generated from OOD criterion would lead to higher scale factor relative to ORD choice task with higher ASD. Given that BDD$\times$DSD provided the highest magnitude of increase in the scale factor, a choice analyst would be better off employing the BDD design criterion because it contributes the most to increasing choice determinism relative to the two other design criteria.

Table 5.6: Heteroskedastic logit model estimates for the pooled sample

	Model 1 Conditional Logit Model			**Model 2 HMNL no interactions Order, Designs, ASD and DSD**			**Model 3 HMNL with interactions Order, Designs, ASD and DSD**		
	Coef	Robust Std Err	Robust P-value	Coef	Robust Std Err	Robust P-value	Coef	Robust Std Err	Robust P-value
Brown kiwi 1	**0.495**	**0.111**	**<0.01**	**0.990**	**0.554**	**0.07**	*1.270*	*0.907*	*0.16*
Brown kiwi 2	**0.654**	**0.105**	**<0.01**	**1.230**	**0.679**	**0.07**	*1.620*	*1.240*	*0.19*
Native fish 1	**0.318**	**0.101**	**<0.01**	**0.650**	**0.383**	**0.09**	0.903	0.700	0.20
Native fish 2	0.134	0.104	0.20	0.351	0.268	0.19	0.530	0.441	0.23
Native plant 1	**0.179**	**0.102**	**0.08**	0.302	0.248	0.22	0.324	0.347	0.35
Native plant 2	**0.228**	**0.105**	**0.03**	0.322	0.257	0.21	0.219	0.363	0.55
Green gecko 1	0.020	0.103	0.85	-0.022	0.196	0.91	0.064	0.285	0.82
Green gecko 2	0.098	0.101	0.33	0.198	0.208	0.34	0.279	0.301	0.35
Bush falcon 1	**0.481**	**0.107**	**<0.01**	**0.930**	**0.538**	**0.08**	*1.260*	*0.945*	*0.18*
Bush falcon 2	**0.720**	**0.101**	**<0.01**	**1.380**	**0.758**	**0.07**	*1.780*	*1.350*	*0.19*
Cost to respondent	**-0.026**	**0.002**	**<0.01**	**-0.050**	**0.028**	**0.08**	*-0.070*	*0.053*	*0.19*
Indicator for SQ	-0.159	0.176	0.37	-0.336	0.379	0.38	-0.198	0.492	0.69
Task orders 4 to 6				**0.467**	**0.156**	**<0.01**	**0.452**	**0.252**	**0.07**
Task orders 7 to 9				*0.231*	*0.170*	*0.17*	0.036	0.245	0.88
Indicator for BDD				**-0.483**	**0.164**	**<0.01**	1.230	2.000	0.54
Indicator for OOD				-0.117	0.128	0.36	**-2.660**	**1.290**	**0.04**
ASD				-0.492	0.442	0.27	-0.571	0.622	0.36
DSD				-0.830	0.673	0.22	**-1.630**	**0.881**	**0.06**
BDD * ASD							*-2.440*	*1.720*	*0.16*
BDD * DSD							**3.650**	**1.850**	**0.05**
OOD * ASD							**2.040**	**1.100**	**0.06**
OOD * DSD							1.200	1.780	0.50
BDD * Ord 4-6							0.464	0.445	0.30
BDD * Ord 7-9							**1.110**	**0.453**	**0.01**
OOD * Ord 4-6							-0.153	0.391	0.70
OOD * Ord 7-9							-0.102	0.419	0.81
Log-likelihood	-1460.32			-1446.12			-1436.93		
Adj rho-square	0.112			0.117			0.118		
Observations	1509			1509			1509		
No. of iterations	7			25			31		

Note: Figures in boldface font are statistically significant at the 90% confidence level.

5.6 Conclusions

Based on the data studied here, overall, our results indicate that BDD is the superior design criterion, as it generates choice tasks that would likely reduce the cognitive burden to respondents while also providing respondents with a better learning experience in evaluating of a sequence of choice tasks. These features thereby contribute to the enhancement of behavioural efficiency of a CE respondent who is required to evaluate relatively complex sets of alternatives. Since a major issue in CE is the complexity of choice tasks, choosing an ED that could contribute to increasing choice determinism and enhancing the learning effects would be very helpful in improving the quality of choice data collected from expensive CE surveys. This research has addressed the problem faced by many choice analysts as to what particular ED would be most appropriate for their CE exercise, given the several ED criteria to choose from. However, our results might be specific to the data studied here. Our sample size is also small making it more difficult to be generalised. In future studies, this problem might be addressed by using a larger sample as well as using an experimental design algorithm that allows the increase or decrease of the range of attribute dispersion.

We consider this study as being one of the few works to empirically investigate how different experimental designs affect choice behaviour in general and choice determinism in particular. We hope this study opens the door for more studies that would provide useful suggestions for choice analysts to further improve the method of collecting choice data. Whilst many would consider that using an ED with statistically higher efficiency (i.e., minimal D-error) is important, identifying an ED that could help improve behavioural efficiency of respondents is equally important. In this study, we found that choice tasks generated from an ED criterion with the lowest D-error contributed the most to enhancing respondents' behavioural efficiency relative to the two other EDs with higher D-errors. We therefore conclude that the Bayesian D-efficient criterion not only contribute to the provision of choice tasks with higher statistical efficiency, but also increase behavioural efficiency relative to the other designs studied here. However, there are still other areas that should be examined further. We suggest that future studies should examine different EDs based on strengths, weaknesses and potential of respondents to attend to complex choice tasks (relative to contingent valuation scenarios). It would be interesting to see how the Bayesian D-efficient Design criterion (under the conditional logit model assumption) compares in behavioural efficiency achieved in other types of EDs which are not examined here (e.g., Bayesian S-efficient design under model averaging approach (Scarpa and Rose, 2008); heterogenous design (Sandor and Wedel, 2005)). We also support the suggestion of Louviere et al. (2011) to undertake a concerted effort to transparently examine different EDs for the benefit of everybody in this field (e.g., applied researchers, respondents, academicians, policy analysts).

Chapter 6: Conclusions and future research

6.1 Thesis summary and conclusions

Chapter 1 provides an overview and main research questions of the thesis. Chapter 2 is a generic chapter that describes most of the models that have been estimated. It also describes some of measured of design efficiency. Furthermore, it also provides a description of how the data was collected and shows a summary of the data sets that were constructed and analysed using various logit models. Chapter 3 has shown using choice experiments that a typical New Zealander would be willing to financially support biodiversity enhancement in the country's 1.8 million hectares of planted forests. Accounting for hypothetical and aggregation biases, New Zealand taxpayers would be willing to pay an aggregated national value of approximately NZ$26.5 million per year for five years to support a national biodiversity enhancement initiative coordinated by the Department of Conservation (DOC) with forest companies, environmental NGOs and community groups. Using Random Effects Panel Regression Analysis, the factors identified to positively influence WTP include being a volunteer to conservation organisations such as DOC and Forest and Bird, being a female, having higher education, having appreciation of native birds and residing in a place with large planted forests within the 10-kilometre radius. We find the spatial factor to be very useful for planning the country's afforestation programme where native biodiversity in planted forests are valued by people.

Chapter 4 provides an overview of attribute non-attendance (ANA) and choice task complexity (via the entropy proxies) in CE. We tested the hypothesis that the selection of ED criterion does not influence ANA and choice variability. The analysis has examined three ED criteria which are Orthogonal Design (ORD), Optimal

Orthogonal Design (OOD) and Bayesian D-efficient Design (BDD). Estimated Latent Class Logit models with panel specification accounting for ANA indicate that the selection of ED criterion matters. We find that higher complexity levels in choice tasks derived from the BDD criterion do not increase choice variability unlike ORD and OOD. This could explain why BDD choice tasks are more attended than choice tasks derived from the two other designs. However, these results might be specific to the data studied here.

Chapter 5 provides an overview of another component of choice task complexity called *attribute dispersion*. DeShazo and Fermo (2002) have shown evidence that higher ASD leads to a decrease in choice determinism based on their choice data collected using ORD. In contrast, this present study shows the opposite where estimation results from our choice data collected using BDD choice tasks indicate that higher ASD leads to increasing choice determinism. This shows an important implication of selecting a design criterion in the study of task complexity. This is because as higher ASD may have a negative impact on choice determinism in one design criterion, this may have a totally different impact on choice determinism in another design. Overall our empirical results in Chapter 5 indicate that in choosing a design criterion, an analyst would be better to select the BDD criterion, as it generates choice tasks that would likely reduce the cognitive burden of respondents while providing them with a better learning experience in evaluating a sequence of choice tasks. This research has therefore addressed the problem faced by many choice analysts as to what particular ED would be most appropriate for their CE exercise given the several ED criteria to choose from.

6.2 Implications and future directions

On the policy side, we have shown evidence that a proposed biodiversity enhancement programme in planted forests is valued by New Zealand taxpayers with an aggregate amount of NZ$26.5 million per year for five years. This shows that enhancing the provision of an environmental service in planted forests would benefit society and the taxpayers would be willing to financially support such initiative. This corroborates the report of the New Zealand Department of Conservation which mentions that New Zealanders place a high value on indigenous species, as they form a basis of national identity (DOC, 2000). The estimated national value of biodiversity enhancement also sheds light on the true value of planted forests. This is because, at present, the value of planted forests is regarded mainly in terms of forest products such as timber, pulp and paper, and, to a certain extent, carbon sequestration service. But in fact, in addition to providing habitat for threatened native species, they also provide other ecosystem services such as erosion control, flood mitigation, water quality improvement and recreation. The present situation shows that planted forests are highly under valued in terms of their contribution to species conservation and habitat creation. One major reason is that the economic value of other ecosystem services they provide is poorly understood and not accounted for in policy decision making. Estimating the value of habitat provision as we have done here is an initial step towards defining the true value of planted forests. We therefore suggest that future studies should estimate other ecosystem services provided by planted forests in addition to developing market mechanisms to sustain and further enhance these services. Furthermore, markets for biodiversity services are now being established by groups of large multinationals in coordination with universities and government institutions (Corporate Ecosystem Valuation) (WBCSD, 2011).

On the research dimension, we contributed by extending previous studies that provide empirical evidence that the Bayesian D-efficient design (BDD) criterion produces significantly improved results, in a statistical sense of relative efficiency (e.g., Rose et al. 2008). This study provides evidence that in addition to improved statistical efficiency, as indicated by lower Bayesian D-error, the BDD criterion might also contribute to improving behavioural efficiency of respondents thereby contributing to higher quality data collected from a choice survey. The sequence of choice tasks derived from the BDD criterion has been found to be more attended to and demonstrated a pattern of continuous learning. Data collected has lower choice variability that could somehow indicate that respondents have found the choice tasks more coherent compared to OOD and ORD. This study therefore provides some evidence as to what ED criterion an analyst should choose given three different ED criteria. It would also be interesting to examine the preferences of analysts in choosing a particular design given that each of them is caught in different situations in terms of budget, time, software and number of available respondents. It is therefore suggested that choice analysts, especially those who have been involved in several choice experiments exercises, should be interviewed given that they may have different preferences in choosing an experimental design based on their situation. Analysts might choose to trade off between statistical efficiency, behavioural efficiency, or maybe avoid constructing choice tasks with dominant alternatives.

It is important to note that the sample size of the three design data sets that we analysed in this study is relatively small. This limited us to the use of the basic heteroskedastic logit model and not the heteroskedastic mixed logit model that could account for individual heterogeneity. We suggest that future studies aiming to compare the behavioural efficiency of different designs should have a relatively large sample size and use more advance heteroskedastic logit models (e.g., heteroskedastic panel mixed logit model with error components, and generalised mixed logit models).

It is worth mentioning that the study of ANA and entropy in this exercise would have benefited from knowing the amount of time it took a respondent to complete each choice task. The amount of time spent in evaluating each choice task would likely provide an indication whether a respondent had either thoroughly processed the information in a choice task or made random choices. We suggest that future studies on choice task complexity and/or ANA should account for the time spent responding to each choice task. Several online survey packages (e.g., Qualtrics[28]) allow the recording of the number of seconds and/or minutes it took a respondent to browse through certain pages of the online questionnaire. We believe that incorporating task response time as described in Rose and Black (2006) would cast additional light on this area of research.

---

[28] Accessed on 10 May 2011 at http://www.qualtrics.com/

References

Arentze, T., Borgers, A., Timmermans, H. and Del Mistro, R. 2003. Transport stated choice responses: effects of task complexity, presentation format and literacy. Transportation Research 39E, 229–244.

Auckland Council. 2011. Rodney Local Board Auckland (spatial) plan. Accessed on 2 May 2012 at http://www.aucklandcouncil.govt.nz/SiteCollectionDocuments/aboutcouncil/localboards/rodneylocalboard/rodneylocalboardaucklandplanpresentation.pdf

Baltagi, B. 2008. Econometric Analysis of Panel Data, 4th Edition. Wiley.

Baillie, B. 2009. Personal communication via face to face meeting. Rotorua, New Zealand.

Bateman, I.J., Carson, R.T., Day, B., Dupont, D., Louviere, J.J., Morimoto, S., Scarpa, R. and Wang, P. 2008. Choice Set Awareness and Ordering Effects in Discrete Choice Experiments. Paper presented at the 2008 EAERE conference. Gothenburg, Sweden.

Bateman, I.J., Day, B.H., Georgiou, S. and Lake, I. 2006. The aggregation of environmental benefit values: welfare measures, distance decay and total WTP. Ecological Economics 60(2): 450–460.

Bateman, I.J., Langford, I.H., Nishikawa, N. and Lake, I. 2000. The Axford debate revisited: A case study illustrating different approaches to the aggregation of benefits data. Journal of Environmental Planning and Management 43(2): 291–302.

Ben-Akiva, M. and Lerman, S.R. 1985. Discrete Choice Analysis. Cambridge: MIT Press.

Ben-Akiva, M. and Morikawa, T. 1990. Estimation of travel demand models from multiple data sources. In: Transportation and Traffic Theory (M. Koshi, ed), Elsevier, New York, pp.461–476.

Berndt, L.A., Brokerhoff, E.G. and Jactel, H. 2008.Relevance of exotic pine plantations as a surrogate habitat for ground beetles (Carabidae) where native forest is rare. Biodiversity and Conservation 17, 1171–1185.

Bienabe, E. and Hearne, R.R. 2006. Public preferences for biodiversity conservation and scenic beauty within a framework of environmental services payments. Forest Policy and Economics 9(4): 335–348.

Bioweb. 2009. Bioweb Herpetofauna. Available online at http://dataversity.org.nz/guide/systems/bh/.

Bierlaire, M. 2009. Estimation of discrete choice models with BIOGEME 1.8. Accessed on 6 April 2011 at http://transp-or2.epfl.ch/biogeme/doc/tutorial.pdf

Bierlaire, M. 2008. A lecture note in summer school on discrete choice model held on 7 to 18 July 2008 in Bologna, Italy.

Bishop, R. and Heberlein, T. 1979. Measuring values of extramarket goods: Are indirect measures biased?. American Journal of Agricultural Economics 61, 926–930.

Bliemer, M.C.J., Rose, J.M., 2005. Efficiency and Sample Size Requirements for Stated Choice Studies, Working Paper: ITLS-WP-05-08.

Bliemer, M.C.J. and Rose, J.M. 2006. Designing Stated Choice Experiments: State-of-the-Art. Paper presented to the 11th International Conference on Travel Behaviour Research, Kyoto, 16–20 August 2006.

Bliemer, M.C.J. and Rose, J.M. 2009a. Designing stated choice experiments: state-of-the-art. In: Kitamura, R., Yoshii, T., Yamamoto, T. (Eds.), The Expanding Sphere of Travel Behaviour Research. Emerald, UK, pp. 499–537.

Bliemer, M.C.J. and Rose, J.M. 2009b. Efficiency and sample size requirements for stated choice experiments. Paper presented to the Transportation Research Board 88th Annual Meeting, Washington D.C.

Bliemer, M.C.J. and Rose, J. M. 2011. Experimental design influences on stated choice outputs: An empirical study in air travel choice. Transportation Research Part A: Policy and Practice, 45(1), 63–79.

Borkin, K. M. and Parsons, S. 2009. Long-tailed bats' use of a *Pinus radiata* stand in Kinleith Forest: implications for monitoring. New Zealand Journal of Forestry 53 (4): 38–43

Boxall, P., Adamowicz, W. 2002. Understanding heterogeneous preferences in Random Utility Models: a latent class approach. Environmental and Resource Economics 23, 421–446.

Bradley, M. and Daly, A.J. 1994. Use of the logit scaling approach to test rank-order and fatigue effects in stated preference data. Transportation 21, 167–184.

Brazell, J. and Louviere, J. 1997. Respondent's help, learning and fatigue. Paper presented at the 1997 INFORMS Marketing Science Conference, University of California, Berkeley.

Brockerhoff, E.G., Ecroyd, C.E., Leckie, A.C., Kimberley, M.O. 2003. Diversity and succession of adventive and indigenous vascular understorey plants in *Pinus radiata* plantation forests in New Zealand. Forest Ecology and Management 185: 307–326.

Brockerhoff, E.G., Shaw, W.B., Hock, B., Kimberley, M., Paul, T., Quinn, J., Pawson, S. 2008. Re-examination of recent loss of indigenous cover in New Zealand and the relative contribution of different land uses. New Zealand Journal of Ecology 32, 115–126.

Bulman, L. 2009. Personal communication via face to face meeting. Rotorua.

Burgess, L. and Street, D. 2003. Optimal designs for 2k choice experiments. Communications in Statistics: Theory and Methods, 32, 2185–206.

Burgess, L. and Street, D. J. 2005. Optimal designs for choice experiments with asymmetric attributes. Journal of Statistical Planning and Inference, 134, 288– 301.

Cameron, T.A. and DeShazo, J.R. 2010. Differential attention to attributes in utility-theoretic choice models. Journal of Choice Modelling 3(3): 73–115.

Campbell, D. 2007. Willingness to pay for rural landscape improvements: Combining mixed logit and random-effects models. Journal of Agricultural Economics 58 (3): 467–483.

Campbell, D., Hutchinson, W. G. and Scarpa, R. 2008a. Incorporating discontinuous preferences into the analysis of discrete choice experiments. Environmental and Resource Economics 41, 401–417.

Campbell, D., Scarpa, R. and Hutchinson, W.G. 2008b. Assessing the Spatial Dependence of Welfare Estimates Obtained from Discrete Choice Experiments. Letters in Spatial and Resource Sciences 1: 117–126.

Campbell, D., Hutchinson, W.G. and Scarpa, R. 2009. Using choice experiments to explore the spatial distribution of willingness to pay for rural landscape improvements. Environment and Planning A 41: 97–111.

Carlsson, F. and Martinsson,P. 2003. Design techniques for stated preference methods in health economics. Health Economics 12(4): 281–294.

Carlsson, F., Kataria, M. and Lampi, E. 2010. Dealing with ignored attributes in choice experiments on valuation of Sweden’s environmental quality. Environmental and Resource Economics 47: 65–89.

Carnus, J.M., Parrotta, J., Brockerhoff, E., Arbez, M., Jactel, H., Kremer, A., Lamb, D., Ó Hara, K., Walters, B. 2006. Planted forests and biodiversity. Journal of Forestry 104: 65–77.

Caussade, S., de Dios Ortuzar, J., Rizzi, L.I., and Hensher, D. 2005. Assessing the influence of design dimensions on stated choice experiment estimates. Transportation Research Part B 39(7): 621–640.

Chaloner, K. and Verdinelli, I. 1995. Bayesian experimental design: A review. Statistical Science 10(3): 273–304.

ChoiceMetrics. 2011. Ngene 1.1 User Manual and Reference Guide. Accessed on 10 July 2011 at http://www.choice-metrics.com/documentation.html

Christie, M. 2007. An examination of the disparity between hypothetical and actual willingness to pay for Red Kite conservation using the contingent valuation method. Canadian Journal of Agricultural Economics 55, 159–169.

Christie, M., Hanley, N., Warren, J., Murphy, K., Wright, R. and Hyde, T. 2006. Valuing the diversity of biodiversity. Ecological Economics 58(2): 304–317.

Colbourne, R., Bassett, S., Billing, T., McCormick, H., McLennan, J., Nelson, A. and Robertson, H. 2005. The development of Operation Nest Egg as a tool in the conservation management of kiwi. Science and Technical Publishing, Department of Conservation, Wellington , p. 24.

Cummings, R.G. and Taylor, L.O. 1999. Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. American Economic Review 89, 649–665.

Czajkowski, M., Buszko-Briggs, M. and Hanley, N. (2008) Valuing Changes in Forest Biodiversity. Stirling Economics Discussion Paper. Department of Economics, University of Stirling

Daly, A., Hess, S. and Train, K. 2011. Assuring finite moments for willingness to pay in random coefficient models. Transportation. Accessed on 14 July 2011 at http://elsa.berkeley.edu/~train/DHT_WTP.pdf

Day, B. and Pinto-Prades, J.L. 2010. Ordering anomalies in choice experiments. Journal of Environmental Economics and Management 59, 271–285.

Day, B., Bateman, I., Carson, R., Scarpa, R., Louviere, J., Wang, P. and Dupont, D. 2010. Task Independence in Stated Preference Studies: A Test of Order Effect Explanations. The 4th World Congress of Environmental and Resource Economics, 29th June-2nd July 2010

DeSarbo, W. S., Duncan F. H., Fong, J.L. and Coupland, J.C. 2005. Evolutionary preference/utility functions: A dynamic perspective. Psychometrika 70(1): 179–202.

DeSarbo, W.S., Lehman, D.R. and Hollman, F.G. 2004. Modeling dynamic effects in repeated-measures experiments involving preference/choice: An illustration involving stated preference analysis. Applied Psychological Measurement 28(3): 186–209.

DeShazo, J.R. and Fermo G. 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.

DeShazo, J. R., and Fermo, G. 2004. Implications of Rationally-Adaptive Pre-choice Behavior for the Design and Estimation of Choice Models. Working paper, School of Public Policy and Social Research, University of California at Los Angeles.

Dhakal, B., Yao, R., Turner, J. and Barnard, T. 2012. Recreational users' willingness to pay and preferences for changes in planted forest features. Forest Policy and Economics 17, 34-44.

DOC (Department of Conservation). 2000. The New Zealand Biodiversity Strategy. Department of Conservation, Wellington.

DOC. 2010. Kiwi Recovery Group Annual Report 2008–2009. Kiwi Recovery Group. Internal report, Department of Conservation, Wellington. Accessed on 4 April 2011 at http://www.doc.govt.nz/upload/documents/conservation/native-animals/birds/kiwi-recovery-group-annual-report-2008-09.pdf

Dyck, B. 2003. Benefits of planted forests: social, ecological and economic. Presented at the United Nations Forum on Forests (UNFF) Intersessional Experts Meeting on the Role of Planted Forests in Sustainable Forest Management, 24–30 March 2003, Wellington, New Zealand.

FAO. 2010. Planted forests in sustainable forest management – A statement of principles. Rome. Accessed on 17 June 2011 at www.fao.org/forestry/plantedforests

Ferrini, S. and Scarpa, R. 2007. Designs with *a priori* information for nonmarket valuation with choice experiments: A Monte Carlo study. Journal of Environmental Economics and Management 53, 342–363.

Fisher, B., Turner, R.K., Zylstra, M., Brouwer, R., de Groot, R., Farber, S., Ferraro, P., Green, R., Hadley, D., Harlow, J., Jefferiss, P., Kirkby, C., Morling, P., Mowatt, S., Naidoo, R., Paavola, J., Strassburg, B., Yu, D., and Balmford, A., (2008) Ecosystem Services and Economic Theory: Integration For Policy-Relevant Research. Ecological Applications 18(8): 2050–2067.

Folke, C., Holling, C.S., Perrings, C. 1996. Biological diversity, ecosystems and the human scale. Ecological Applications 6:1018–24.

Foster, V., Bateman, I. and Harley, D. 1997. Real and hypothetical willingness to pay for environmental preservation: a non-experimental comparison. Journal of Agricultural Economics 48(2): 123–138.

Foster, V. and Mourato, S. 2003. Elicitation format and sensitivity to scope: Do contingent valuation and choice experiments give the same results? Environmental and Resource Economics 24: 141-160.

Frykblom, P. 1997. Hypothetical question modes and real willingness to pay. Journal of Environmental Economics and Management 34, 275–287.

Getis, A. and K. Ord. 1992. The analysis of spatial association by use of distance statistics. Geographical Analysis 24: 189–206.

Goldberg, I., Roosen, J., 2007. Scope insensitivity in health risk reduction studies: A comparison of choice experiments and the contingent valuation method for valuing safer food. Journal of Risk and Uncertainty 34, 123-144.

Greene, W.H. 2008. Econometric analysis. 6th ed., Upper Saddle River, N.J. Prentice Hall.

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

Greene, W.H., Hensher, D.A. and Rose, J.M. 2005. Using classical simulation-based estimators to estimate individual WTP values. Chapter 2, pages 17-34, In Scarpa R. and Alberini, A. (eds.). Applications of simulation methods in environmental and resource economics. Springer Publisher, Dordrecht, The Netherlands.

Grossmann, H., Holling, H., and Schwabe, R. 2002. Advances in optimum experimental design for conjoint analysis and discrete choice models. In: Advances in Econometrics, Econometric Models in Marketing, Vol. 16, Philip H. Franses and Alan L. Montgomery, eds. Amsterdam: JAI Press, 93–117.

Hahn, G. J., and Shapiro, S. S. 1966. A catalog and computer program for the design and analysis of orthogonal symmetric and asymmetric fractional factorial experiments. General Electric Research and Development Center Technical Report No. 66-C-165, Schenectady. N.Y.: Research and Development Center.

Haaijer, M. E. 1999. Modeling conjoint choice experiments with the Probit model. PhD thesis, University of Groningen, Labyrint Publications, The Netherlands.

Haaijer, M. E., Kamakura, W.A., and Wedel, M. 2001. The no-choice alternative in conjoint choice. International Journal of Market Research 43(1): 93–106.

Hanchet, S.M. 1990. Effect of land use on the distribution and abundance of native fish in tributaries of the Waikato River in the Hakarimata Range, North Island, New Zealand. New Zealand Journal of Marine and Freshwater Science 24, 159–171.

Hanley, N., Wright, R.E. and Koop, G. 2002. Modelling recreation demand using choice experiments: rock climbing in Scotland. Environmental and Resource Economics 22: 449–466.

Hanley, N., Schläpfer, F., and Spurgeon, J. 2003. Aggregating the benefits of environmental improvements: Distance-decay functions for use and non-use values. Journal of Environmental Management 68, 297–304.

Hartman, R.S., Doane, M.J. and Woo, C.K. 1990. Status quo bias in the measurement of value of service. Resources and Energy 12, 197–214.

Heckman, J. and Singer, B. 1984. Econometric duration analysis. Journal of Econometrics 24, 63–132.

Hensher, D. A. 2006. How do respondents process stated choice experiments? Attribute consideration under varying information load. Journal of Applied Econometrics 21, 861–878.

Hensher, D. A. 2010. Hypothetical bias, choice experiments and willingness to pay, Transportation Research B 44, 735–752.

Hensher, D.A. and Greene, W.H. 2003. The mixed logit model: the state of practice. Transportation 30, 133–176.

Hensher, D.A. and Rose, J.M. 2008. Simplifying choice through attribute preservation or non-attendance: Implications for willingness to pay. Transportation Research Part E: Logistics and Transportation Review 45, 583-590.

Hensher, D.A., Louviere, J.J. and Swait, J. 1999. Combining sources of preference data. Journal of Econometrics 89, 197–221.

Hensher, D.A., Rose, J.M. and Greene, W.H. 2005. Applied Choice Analysis: A Primer. Cambridge University Press.

Hess, S. and Rose, J.M. 2009. Allowing for intra-respondent variations in coefficients estimated on stated preference data. Transportation Research B 43(6): 708–719.

Hess, S., Bolduc, D. and Polak, J. 2010. Random covariance heterogeneity in discrete choice models. Transportation 37(3): 391–411.

Hofler, R. and List, J.A. 2000. Valuation on the frontier: calibrating actual and hypothetical statements. University of Arizona, Working Paper.

Holmes, T. and Adamowicz, W. L. 2003. Attribute based methods. In: A Primer on Non-Market Valuation. ed. K.J. Boyle and P.A. Champ. Boston: Kluwer Academic Publishers.

Holmes, T. and Boyle, K. J. 2005. Learning and context-dependence in sequential, attribute-based, stated-preference valuation questions. Land Economics 81(1): 114–126.

Holzapfel, S., Robertson, H.A., McLennan, J.A., Sporle, W., Hackwell, K. and Impey, M. 2008. Kiwi (Apteryx spp.) Recovery Plan 2008–2018. Threatened species recovery plan 60. Department of Conservation. 71 p. Wellington. Accessed on 4 April 2011 at http://www.doc.govt.nz/upload/documents/science-and-technical/tsrp60entire.pdf

Huber, J. and Zwerina, K. 1996. The importance of utility balance in efficient choice design. Journal of Marketing Research 33, 307–17.

Humprey, J., Ferris, R. and Quine, C.P. eds. 2003. Biodiversity in Britain's Planted Forests Results from the Forestry Commission's Biodiversity Assessment Project. Forestry Commission, Edinburgh. 118 pp.

Islam, T., Louviere, J. J. and Burke, P. F. 2007. Modeling the effects of including/excluding attributes in choice experiments on systematic and random components. International Journal of Market Research 24: 289–300.

Johnson, F.R., Kanninen, B., Bingham, M., and Ozdemir, S. 2007. Experimental design for stated choice studies. In: Kanninen, B. (ed) Valuing Environmental Amenities Using Stated Choice Studies. Springer, Dordrecht.

Johnston, R.J., Ramachandran, M., Schultz, E.T., Segerson, K. And Basedin, E.Y. 2011. Characterizing spatial pattern in ecosystem service values when distance decay doesn't apply: Choice experiments and local indicators of spatial association. Selected Paper prepared for presentation at the Agricultural & Applied Economics. Association's 2011 AAEA & NAREA Joint Annual Meeting, Pittsburgh, Pennsylvania, July 24-26, 2011.

Jukes, M. and Peace, A. 2003. Invertebrate communities in plantation forests. In: Biodiversity in Britain's Planted Forests: Results from the Forestry Commission's Biodiversity Assessment Project (Eds. J. Humphrey, R. Ferris and C. Quine) pp. 75–91.

Kanninen, B. 2002. Optimal design for multinomial choice experiments. Journal of Marketing Research 39, 214–227.

Keller, K.L. and Staelin, R. 1987. Effects of quality and quantity of information on decision effectiveness. Journal of Consumer Research 14, 200–213.

Kennedy, P. 2008. A Guide to Econometrics 6E. Blackwell Publishing.

Kessels, R., Goos, P., and Vandebroek, M. 2006. A comparison of criteria to design efficient choice experiments. Journal of Marketing Research 43, 409–419.

Kerr, G. N. and Sharp, B. M. H. 2010. Choice experiments adaptive design benefits: a case study. Australian Journal of Agricultural and Resource Economics 54(4): 407–420.

Kjær, T., Bech, M., Gyrd-Hansen, D., and Hart-Hansen, K. 2006. Ordering effect and price sensitivity in discrete choice experiments: Need we worry? Health Economics 15: 1217–1228.

Kuhfeld, W. F. 2004, Marketing research methods. In: SAS Experimental Design, Choice, Conjoint, and Graphical Techniques. SAS Institute, Cary, NC, USA.

Kuhfeld, W. F., Tobias, R. D. and Garratt, M. 1994. Efficient experimental design with marketing research applications. Journal of Marketing Research 31, 545–557.

List, J.A. 2001. Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sports cards. American Economic Review 91, 1498–1507.

List, J.A. and Gallet, C.A. 2001. What experimental protocol influence disparities between actual and hypothetical stated values? Environmental and Resource Economics 20, 241-254.

List, J. A. and Shogren, J. 1998. Calibration of the difference between actual and hypothetical valuations in a field experiment. Journal of Economic Behavior and Organization 37(2): 193–205.

Louviere, J.J. 2001. What if consumer experiments impact variances as well as means? Response variability as a behavioural phenomenon. Journal of Consumer Research 28, 506–511.

Louviere, J., Street, D., Carson, R., Ainslie, A., DeShazo, J.R., Cameron, T., Hensher, D., Kohn, R. and Marley, T., 2002. Dissecting the random component of utility. Marketing Letters 13, 177–193.

Louviere, J.J. and Hensher, D.A. 1983. Using discrete choice models with experimental design data to forecast consumer demand for a unique cultural event. Journal of Consumer Research 10, 348–361.

Louviere, J.J. and Woodworth, G. 1983. Design and analysis of simulated consumer choice or allocation experiments: An approach based on aggregate data. Journal of Marketing Research 20, 350–367.

Louviere, J.J., Hensher, D.A. and Swait, J.D. 2000. Stated Choice Methods—Analysis and Application. Cambridge University Press.

Louviere, J.J., Islam, T., Wasi, N., Street, D. and Burgess, L. 2008. Designing discrete choice experiments: Do optimal designs come at a price? Journal of Consumer Research, 35(2): 360–375.

Louviere, J.J., Pihlens, D., and Carson, R. 2011. Design of discrete choice experiments: A discussion of issues that matter in future applied research. Journal of Choice Modelling, 4(1): 1–8.

Lusk, J. and Norwood, F. B. 2005. Effect of experimental design on choice-based conjoint valuation estimates. American Journal of Agricultural Economics 87(3): 771–785.

MAF (Ministry of Agriculture and Forestry). 2010. National Exotic Forest Description as at 1 April 2009 (26th Edition). Ministry of Agriculture and Forestry, Wellington.

Maunder, C., Shaw, W., Pierce, R. 2005. Indigenous biodiversity and land-use – what do exotic plantation forests contribute? New Zealand Journal of Forestry 49, 20–26.

Maunder, C. 2008. Personal communication via face to face meeting. Rotorua.

Mazzotta, M. J., and Opaluch, J. J. 1995. Decision making when choices are complex: a test of Heiner's hypothesis. Land Economics, 71(4): 500–515.

McFadden, D. 1974. Conditional Logit Analysis of Qualitative Choice Behaviour.” In P. Zarembka (ed) Frontiers in Econometrics, pp. 105–42. New York: Academic Press.

McFadden, D. L. and Train, K. E. 2000. Mixed MNL models for discrete response. Journal of Applied Econometrics 15, 447–470.

Meyerhoff, J. and Liebe, U. 2009. Status quo effect in choice experiments: Empirical evidence on attitudes and choice task complexity. Land Economics 85(3): 515–528.

MfE. 2011. The New Zealand Land Cover Database. Accessed on 31 May 2011 at http://www.mfe.govt.nz/issues/land/land-cover-dbase/index.html

Norton, D.A. 1998. Indigenous biodiversity conservat ion and plantation forestry: options for the future. N.Z. Forestry 43 (2) : 34–39.

Ohler, T., Le, A., Louviere, J.J., Swait, J.D. 2000. Attribute range effects in binary response tasks. Marketing Letters 11, 249–260.

Ortúzar, J. de D., Roncagliolo, D.A. and Velarde, U.C., 2000. Interactions and independence in stated preference modelling. In: Ortúzar, J.de D.(Ed.), Stated Preference Modelling Techniques Perspectives 4, PTRC, London.

Özdemir, S. Johnson, F.R. and Hauberb, A.B. 2009. Hypothetical bias, cheap talk, and stated willingness to pay for health care. Journal of Health Economics 28, 894–901.

Pate, J., Loomis, J.B., 1997. The effect of distance on willingness to pay values: a case study of wetlands and salmon in California. Ecological Economics 20, 199–207.

Pawson, S.M., Ecroyd. C.E., Seaton, R., Shaw, W.B., Brockerhoff, E.G. 2010. New Zealand’s exotic plantation forests as habitats for threatened indigenous species. New Zealand Journal of Ecology 34(3): 342–355.

Pawson, S., Brockerhoff, E.G., Norton, D., Didham, R., 2006. Clear-fell harvest impacts on biodiversity: past research and the search for harvest size thresholds. Canadian Journal of Forest Research 36: 1035–1046.

Pérez, P.E., Martínez, F.J., Ortúzar, J., de D, 2003. Microeconomic formulation and estimation of a residential location choice model: implications for the value of time. Journal of Regional Science 42, 771–789.

Pierce, R.J., Shaw, W.B., Bycroft, C., Kimberly, M. 2002. An assessment of understorey and avifauna biodiversity in old growth and harvest age radiata pine and Douglas fir plantations in the central North Island. Wildland Consultants Contract Report No. 470. Prepared for Fletcher Challenge Forests.

Pimentel, D., Stachow, U., Takacs, D.A., Brubaker, H.W., Dumas, A.R., Meaney, J.J., O'Neil, J., Onsi, D.E., Corzilius, D.B. 1992. Conserving biological diversity in agricultural/forestry systems. BioScience 42: 354–362.

Rolfe, J. and Windle, J. 2010. Testing for geographic scope and scale effects with choice modelling: Application to the Great Barrier Reef. Environmental Economics Research Hub Research Reports 1069, Environmental Economics Research Hub, Crawford School of Public Policy, The Australian National University.

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

Rose, J. M. and Bliemer, M. C. J. 2004, The design of stated choice experiments: The state of practice and future challenges. Working Paper ITS-WP-04-09, Institute of Transport Studies, University of Sydney and Monash University.

Rose, J.M. and Bliemer, M.C.J. 2007. Stated preference experimental design strategies. In: D.A. Hensher and K. Button, Editors, Transport Modelling (second ed.), Handbooks in Transport vol. 1, Elsevier Science, Oxford (Chapter 8).

Rose, J.M. and Bliemer, M.C.J. 2008. Stated Preference Experimental Design Strategies, in Hensher, D.A. and Button, K.J. (eds) Handbook of Transport Modelling, Elsevier, Oxford, Ch 8, 151–180.

Rose, J.M., Bain, S., and Bliemer, M.C.J.2011. Experimental Design strategies for Stated Preference Studies Dealing with Non Market Goods. In: The International Handbook on Non-Market Environmental Valuation. edited by J. Bennett. Edward Elgar Publishing, Cheltenham, Ch 14, 273–299.

Rose, J. M., Bliemer, M. C. J., Hensher, D. A. and Collins, A. T. 2008. Designing efficient stated choice experiments in the presence of reference alternatives. Transportation Research Part B: Methodological 42(4): 395–406.

Rose, J.M., Scarpa, R. and Bliemer, M.C.J. 2009. Incorporating model uncertainty into the generation of efficient stated choice experiments: A model averaging approach. ITLS Working Paper 09-08. Institute of Transport and Logistics Studies. The University of Sydney.

Ross, A. 2011. Face to face communication during tea time (or coffee break) of the Biodiversity Policy Futures workshop in Wellington in March 2011.

Rowlands, R.P.V. 1989. New Zealand Geckos: A Guide to Captive Maintenance and Breeding. (2nd Edition) New Zealand Herpetological Society.

Ryan, M. and Wordsworth, S. 2000. Sensitivity of willingness to pay estimates to the level of attributes in discrete choice experiments. Scottish Journal of Political Economy 47 (5): 504–524.

Samuelson,W. and Zeckhauser, R. 1988. Status-quo bias in decision-making. Journal of Risk and Uncertainty 24(1): 7–59.

Sandor, Z. and Wedel, M. 2001. Designing conjoint choice experiments using managers' prior beliefs. Journal of Marketing Research 38, 430–444.

Sandor, Z. and Wedel, M. 2002. Profile construction in experimental choice designs for mixed logit models. Marketing Science 21, 455–475.

Sandor, Z. and Wedel, M. 2005. Heterogeneous conjoint choice designs. Journal of Marketing Research 42, 210–218.

Scarpa, R. 2003. The Recreational Value of Woodland in Great Britain. Report to the Forestry Commission, Edinburgh, by Centre for Research in Environmental Appraisal and Management, University of New Castle.

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

Scarpa, R., Campbell, D. Hutchinson, W. G. 2007. Benefit estimates for landscape improvements: sequential Bayesian design and respondents' rationality in a choice experiment study. Land Economics 83(4): 617–634.

Scarpa, R., Ferrini, S. and Willis, K. 2005. Performance of error component models for status-quo effects in choice experiments. In: Applications of Simulation Methods in Environmental and Resource Economics, ed. Riccardo Scarpa and Anna Alberini. Dordrecht: Springer.

Scarpa, R., Gilbride, T., Campbell, D., Hensher, D.A. 2009. Modelling attribute non-attendance in choice experiments for rural landscape valuation. European Review of Agricultural Economics 36(2): 151–174.

Scarpa, R., Raffaelli, R., Notaro, S., and Louviere, J. 2011a. Modelling the effects of stated attribute non-attendance on its inference: An application to visitors benefits from the alpine grazing commons. Paper Selected for presentation at the 2nd International Conference of Choice Modelling, Leeds, Oulton Hall, 4-6 July 2011.

Scarpa, R., Notaro, S., Louviere, J. and Raffaelli, R. 2011b. 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): 813–828;

Scarpa, R., Thiene, M., and Marangon, F. 2007. The value of collective reputation for environmentally friendly production methods: the case of Val di Gresta carrots. Journal of Agricultural and Food Industrial Organization, 5(1): article7.

Scarpa, R., Thiene, M., and Train, K. 2008. Utility in willingness to pay space: a tool to address confounding random scale effects in destination choice to the Alps. American Journal of Agricultural Economics 90(4): 994–1010.

Scarpa, R., Willis, K., Acutt, M. 2007. Valuing externalities from water supply: status-quo, choice complexity and individual random effects in panel kernel logit analysis of choice experiments. Journal of Environmental Planning and Management, 50(4):449–466.

Schroth, G. and da Mota, M.S.S. 2004. The role of agroforestry in biodiversity conservation in the tropics: a synthesis of evidence. Paper presented at the International Ecoagriculture Conference and Practitioners' Fair held in Nairobi, Kenya. The World Agroforestry Centre (ICRAF), Nairobi, September–October 2004

Seaton, R. 2006. The ecological requirements of the New Zealand falcon (Falco novaseelandiae) in plantation forestry. Unpublished PhD thesis. Massey University, Palmerston North, New Zealand.

Seaton, R., Minot, E.O., and Holland, J.D. 2010. Variation in bird species abundance in a commercial pine plantation in New Zealand. NZ Journal of Forestry 54 (4): 3–11.

Shaw, W.B. 1993. Kowhai ngungutuka (Clianthus puniceus) recovery plan. Threatened Species Recovery Plan Series No. 8. Threatened Species Unit, Department of Conservation, Wellington.

Shaw, W.B., Burns. B.R. 1997. The ecology and conservation of the endangered endemic shrub, kōwhai ngutukākā Clianthus puniceus in New Zealand. Biological Conservation 81: 233–245.

Sinden, J. A. 1988. Empirical tests of hypothetical biases in consumers' surplus surveys. Australian Journal of Agricultural Economics 32(2&3): 98–112.

Slui, B. 2011. Kakabeak: Maungataniwha's Hidden Species. A brief report available online at http://www.matarikiforests.co.nz/news/15/56/Kakabeak-Maungataniwha-s-Hidden-Species/

Spellerberg, I.F. and Sawyer, J.W.D. 1995. Multiple-use, biological diversity and standards. New Zealand Forestry 39(4): 21–25.

Sporle, W. and Bliss, T. 2008. Forestry management guidelines for North Island Brown Kiwi in exotic plantation forests. A guideline published by the BNZ Save the Kiwi Trust and Bay of Plenty Regional Council.

Stewart, D. and Hyde, N. 2004. New Zealand falcons (Falco novaeseelandiae) nesting in exotic plantations. Notornis 51, 119–121.

Stewart, D. 2012. Personal communication (face to face) in April 2012 at Wingspan Birds of Prey Centre in Rotorua.

Stopher, P.R., Hensher, D.A., 2000. Are more profiles better than fewer. Searching for parsimony and relevance in stated choice experiments. Transportation Research Record 1719, 165–174.

Street, D.J. and Burgess, L. 2004. Optimal and near-optimal pairs for the estimation of effects in 2-level choice experiments. Journal of Statistical Planning and Inference 118, 185–199.

Street, D.J., Burgess, L., and Louviere, J.J. 2005. Quick and easy choice sets: Constructing optimal and nearly optimal stated choice experiments. International Journal of Research in Marketing 22, 459–470.

Street, D.J. and Burgess, L. 2007. The Construction of Optimal Stated Choice Experiments: Theory and Methods, Hoboken, NJ: Wiley.

Swait, J. 2001. A non-compensatory choice model incorporating attribute cutoffs. Transportation Research Part B 35(10): 903–928.

Swait, J. 1994. A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data. Journal of Retailing and Consumer Services 1, 77-89.

Swait, J. and Louviere, J. 1993. The role of the scale parameter in the estimation and use of multinomial logit models. Journal of Marketing Research 30, 305–314.

Swait, J. and Adamowicz, W. 2001a. Choice environment, market complexity, and consumer behavior: A theoretical and empirical approach for incorporating decision complexity into models of consumer choice. Organizational Behavior and Human Decision Processes, 86(2): 141–167.

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

TEEB. 2010. The Economics of Ecosystems and Biodiversity: Mainstreaming the Economics of Nature: A synthesis of the approach, conclusions and recommendations of TEEB. Accessed on 12 April 2011 at http://www.teebweb.org/TEEBSynthesisReport/tabid/29410/Default.aspx.

Thiene, M. and Scarpa, R. 2009. Deriving and testing efficient estimates of WTP distributions in destination choice models. Environmental and Resource Economics 44(3): 379-395.

Tilahun, N.Y., Levinson, D.M. and Krizek, K.J. 2007. Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. Transportation Research Part A 41 (4): 287–301.

Toubia, O., Hauser, J.R. and Garcia, R. 2007. Probabilistic polyhedral methods for adaptive choice-based conjoint analysis: theory and application. Marketing Science 26, 596–610.

Toner, J.P., Clark, S.D., Grant-Muller, S.M. and Fowkes, A.S. 1999. Anything you can do, we can do better: a provocative introduction to a new approach to stated preference design, WCTR Proceedings, 3, Antwerp, 107–120.

Train, K. 2003. Discrete choice methods with simulation. 1st ed. Cambridge University Press.

Train, K. 2009. Discrete choice methods with simulation. 2nd ed. Cambridge University Press.

Train, K. and Weeks, M. 2005. Discrete choice models in preference space and willingness-topay space. In: Applications of Simulation Methods in Environmental and Resource Economics, ed. Riccardo Scarpa and Anna Alberini. Dordrecht: Springer.

Train, K. and Wilson, W. 2008. Estimation on stated-preference experiments constructed from revealed-preference choices. Transportation Research Part B 40: 191–203.

Tversky, A. and Shafir, E. 1992. Choice under conflict: the dynamics of deferred decision, Psychological Science 3, 358–361.

UNCED. 1992. The Forest Principles. Agenda 21, Clause 6d, Chapter 11. Rio de Janeiro, Brazil, United Nations Conference on Environment and Development. World Business Council for Sustainable Development (WBCSD). Accessed on 12 April 2011 at http://www.wbcsd.org/templates/TemplateWBCSD5/layout.asp?type=p&MenuId=MTc3Ng&doOpen=1&ClickMenu=LeftMenu

Van der Waerden, P., Borgers, A., Timmermans, H.J.P. and Berenos, M. 2006. Order effects in stated-choice experiments: Study of transport mode choice decisions. Transportation Research Board 1985, 12–18.

Vermuelen, B., Goos, P., Scarpa, R., and Vandebroek, M. 2011. Bayesian conjoint choice designs for measuring willingness to pay. Environmental and Resource Economics 48(1): 129–149.

Villas-Boas, M. and Winer, R. 1999. Endogeneity in brand choice models. Management Science 45, 1324–1338.

Viney, R., Savage, E. and Louviere, J. 2005. Empirical investigation of experimental design properties of discrete choice experiments in health care. Health Economics 14, 349–362.

Wang, D. and Li, J. 2002. Handling large numbers of attributes and/or levels in conjoint experiments. Geographical Analysis 34, 350–362.

Watt, M.S., Palmer, D.J. and Höck, B.K. 2011. Spatial description of potential areas suitable for afforestation within New Zealand and quantification of their productivity under Pinus radiata. New Zealand Journal of Forestry Science 41, 115–129.

WBCSD. 2011. Guide to Corporate Ecosystem Valuation (CEV). World Business Council for Sustainable Development.

Weir, P. 2010. Personal communication on 26 October 2010 during the EIANZ conference in Wellington, New Zealand.

Widlert, S. 1998. Stated preference studies: the design affects the results. In: Ortuzar, J. de D., Hensher, D.A., Jara-Díaz, S.R. (Eds.), Travel Behaviour Research: Updating the State of Play. Pergamon, Oxford.

Yao, R. and Kaval, P. 2010. Valuing biodiversity enhancement in New Zealand. International Journal of Ecological Economics and Statistics 16(10): 26–42.

Appendix A: Example of a survey instrument used in the study

Threatened Native Animals and Plants in New Zealand's Planted Forests: What Do You Think?

Adult female New Zealand falcon. D. Stewart 2003.

scion
Next generation biomaterials

THE UNIVERSITY OF WAIKATO
Te Whare Wānanga o Waikato

The University of Sydney

Threatened Native Animals and Plants in New Zealand's Planted Forests: What Do You Think?

Scion (NZ Forest Research Institute), in collaboration with the University of Waikato and the University of Sydney, is conducting a study on the management of **exotic planted forests** for the conservation of threatened native animals and plants (e.g., kiwi, kakabeak). We would like to know your views on the role of exotic forests for NZ native plants and animals. While about 90% of these forests consist of one type of **exotic (foreign) tree,** which is **Radiata pine *(Pinus radiata*),** the remaining 10% include other foreign trees, such as **Douglas-fir (*Pseudotsuga menziesii*)** and **Gum tree (*Eucalyptus nitens/ fastigata*)**. Your views on planted forests are important to us. If you report your honest views to us they will help guide future decision making.

<u>The Survey</u>
There are no right or wrong answers to this survey. We are only interested in your honest views. All your answers will be kept confidential in compliance with the Privacy Act of 1993.

<u>Where you live</u> (If you own or manage more than one property, please answer these questions in relation to the property you live at for most of the year)

1. Is your home located in (please tick one):
   ☐ Whangarei ☐ Auckland ☐ Hamilton ☐ Tauranga
   ☐ New Plymouth ☐ Gisborne ☐ Napier-Hastings ☐ Rotorua
   ☐ Wanganui ☐ Palmerston North ☐ Kapiti Coast ☐ Wellington
   ☐ Nelson ☐ Christchurch ☐ Dunedin ☐ Invercargill
   ☐ Other: specify city, town or nearby town ____________________________________

2. How many years have you lived at this property? ___________

3. Approximately how large is your property? _____ hectares or _____ acres or ______ sq metres

4. Do you own or rent this property? (Please tick one)
   ☐ Own ☐ Rent ☐ Other: please specify ___________________________

5a. Is your home located close (less than 10 km) to an exotic planted forest? ☐ Yes ☐ No
   5b. If Yes, can you please specify the name of this planted forest? ___________________________
   5c. If Yes, approximately how far is your home from this planted forest? ____________ km

6. Have you previously lived in a property close to a planted forest? ☐ Yes ☐ No

7. Are you aware that New Zealand's planted forests could provide habitat for rare native plants and animals even though the trees are non-native in New Zealand? ☐ Yes ☐ No

8. Since this survey is about threatened plants and animals that can be found in New Zealand's exotic planted forests, can you please indicate your level of familiarity with the species in the table below by ticking the box "☐":

	Never heard of	Heard of	Read about	Seen in zoo/ garden/aquarium	Seen in the bush	Sought in the bush
NZ Bush Falcon (bird)	☐	☐	☐	☐	☐	☐
Giant Kokopu (fish)	☐	☐	☐	☐	☐	☐
Long-tailed Bat	☐	☐	☐	☐	☐	☐
Brown Kiwi (bird)	☐	☐	☐	☐	☐	☐
Auckland Green Gecko	☐	☐	☐	☐	☐	☐
Kakabeak (plant)	☐	☐	☐	☐	☐	☐

Your Views About Planted Forest

9. This next set of questions indicates some things that an exotic planted forest in New Zealand can help provide. Please indicate the extent to which you agree or disagree by circling the appropriate number. If you are unsure, please circle "U".

Do you agree that NZ planted forests provide…	Strongly disagree	Slightly disagree	Neutral	Slightly agree	Strongly agree	Unsure
9a. habitat for <u>threatened native</u> plants (e.g., kakabeak, orchids)	1	2	3	4	5	U
9b. habitat for <u>threatened native</u> fish (e.g., banded kokopu, giant kokopu, inanga)	1	2	3	4	5	U
9c. habitat for <u>threatened native</u> mammals (e.g., long-tailed bat, short-tailed bat)	1	2	3	4	5	U
9d. habitat for <u>threatened native</u> birds (e.g., kiwi, bush falcon)	1	2	3	4	5	U
9e. habitat for <u>non-threatened native</u> birds (e.g., tui, bellbirds, whitehead, tomtit)	1	2	3	4	5	U
9f. habitat for <u>threatened native</u> reptiles (e.g., frogs, skinks, geckos)	1	2	3	4	5	U
9g. habitat for <u>non-threatened native</u> insects (e.g., tree weta, huhu beetles)	1	2	3	4	5	U
9h. connectivity between <u>native</u> forest patches (e.g., movement of native species and shelter)	1	2	3	4	5	U
9i. maintenance of existing <u>native</u> bush (e.g., rimu, kauri, kahikatea)	1	2	3	4	5	U
9j. a rich understorey of <u>native</u> plants (e.g., ponga, kanono)	1	2	3	4	5	U
9k. The maintenance of water quality (e.g., clean streams)	1	2	3	4	5	U
9l. Recreation (e.g., walking, fishing, biking, horse riding, camping, hunting)	1	2	3	4	5	U
9m. Storage of carbon in forests to mitigate climate change	1	2	3	4	5	U

10. The NZ Bush Falcon in Kaingaroa Forest in the Central North Island

We would now like to provide you with some background information about the NZ bush falcon in the Kaingaroa Forest. Please read the following information before answering the survey questions.

The NZ bush falcon, known by Maori as "Karearea", is unique to NZ. NZ bush falcons are the country's fastest bird. They achieve speeds of up to 97 km/hour and their eyesight is eight times more powerful than our own.

Despite these extraordinary abilities, the NZ bush falcon is a "threatened" native bird. The word "threatened" means that their population is very low, with a risk of becoming extinct, particularly when no conservation effort is undertaken. Recent estimates revealed that there are only 3,000 left, and they are rarer than some species of kiwi. The reasons for their demise include habitat loss and introduced predators. The fact that bush falcons nest on the ground, coupled with their inability to see well at night, makes falcon eggs and chicks vulnerable to attack by predators (ferret, stoat and weasel) introduced to New Zealand.

Planted forests offer a good habitat for bush falcons. 'Cutover' areas remaining after harvesting provide hunting grounds and suitable nesting sites. But, bush falcons are still at risk from forestry operations. Forest managers can help protect these birds by controlling predators and reducing the impact of harvesting and planting operations in known nesting areas.

Such initiatives have helped to conserve the bush falcon in Kaingaroa Forest in the Central North Island. This large forest currently has the highest bush falcon concentration in the country and successful control of predators has enabled the local population of NZ falcon to slightly increase. NZ bush falcons may be observed frequently in different sections of the forest. Between 2005 and 2006, 36 bush falcon nests were found in the forest.

Adult female New Zealand falcon. D. Stewart 2003.

Project

A government-coordinated conservation programme will be undertaken over the next five years to increase and sustain the bush falcon population in Kaingaroa Forest. This conservation programme needs public support.

Project Aim

This study aims to measure how much members of the public would value the conservation of the NZ bush falcon in Kaingaroa Forest.

Please select one answer

☐ I have read fully the description above

☐ I have partly read or skimmed through the description above

THE IMPORTANT QUESTION

We would like to know if you would contribute some money to support a 5-year conservation programme to increase and sustain the bush falcon population in Kaingaroa Forest. The money would be paid in the form of an additional amount in your annual income tax for 5 years only. The collected money would be given to the Department of Conservation, which—in collaboration with the forest corporations—would use it to undertake the programme. (If you do not pay any income tax now, please answer the question as if you did pay income tax). Please, note that all funding will go directly towards this programme and none will be used for administrative fees. In answering the following questions please consider that you could use this money for alternative uses. For example, to pay for other needs in your household and for other activities you enjoy.

Sometimes when people are asked these types of questions they do not pay sufficient attention to the dollar amount as they think that we are dealing with an imaginary situation. However, it is very important to obtain your honest response to these questions. It is perfectly fine if you would not be willing to pay any amount to support the conservation of the NZ bush falcon in Kaingaroa Forest.

10a. We would now like to know if you prefer to pay an additional amount in your income tax to support the programme to increase and sustain the bush falcon population in Kaingaroa Forest. Please respond just *exactly* as you would if you were really going to commit an additional amount in your income tax over the next five years. Now, would you be willing to pay **$10 per year for five years**?

☐ **Yes** ☐ **No**

10b. If you ticked "Yes" above, would you pay **$30**? *or* If you ticked "No", would you pay **$5**?

☐ **Yes** ☐ **No** ☐ **Yes** ☐ **No**

10c. If you are not prepared to pay any amount, please explain why. (Tick one only)

☐ I did not want to place a dollar value
☐ I object to the way the question is presented
☐ I am opposed to a further increase in income tax
☐ Forest companies should pay
☐ The government should pay
☐ Not enough information provided
☐ Other: specify ____________________
____

10d. Please rate your <u>understanding</u> of the <u>background information</u> on the NZ bush falcon (circle a number below):

Did not Understand at all (1) — *Moderately Understood* (5) — *Completely Understood* (10)

1 2 3 4 5 6 7 8 9 10

Valuing Threatened Animals and Plants in New Zealand's Planted Forests

When undertaken, conservation programmes in planted forests of NZ can also benefit other threatened species apart from the NZ bush falcon. In 2008, exotic planted forests represented 22% of New Zealand's total forest area. These forests provide habitat to **over 100 *threatened* native animals** and **plants (including the NZ bush falcon)**. The word "threatened" means that their population is very low, with a risk of becoming extinct, particularly when no conservation effort is undertaken. Some of these threatened animals and plants in planted forests include:

	***Brown Kiwi*** Throughout New Zealand, the **brown kiwi** population has been declining at a rate of 5% per year, which implies their population halves every decade. Conservation initiatives have started to ensure that the brown kiwi continues to live in a few exotic forests. They can be found in planted forests in **Northland**, **Coromandel**, **Central North Island**, **Bay of Plenty** and **Hawke's Bay** that also contain remnants of native trees, stream edges with trees, clearfell and stands of various ages. The brown kiwi is nocturnal and can be heard calling after dark.
	***Native Fish*** The **giant kokopu** is a rare native fish whose populations are gradually declining throughout New Zealand. They can be found in suitable waterways in planted forests in **Bay of Plenty**, **East Coast**, **Waikato**, **southern North Island, West Coast** and **Southland**. They can be seen at night in gently flowing streams with overhanging native vegetation.
	***Native Shrub*** The **kakabeak** is a widely cultivated shrub, however, natural populations are extremely rare in the wild. Kakabeak has been found in planted forests on the **East Coast** and **Hawke's Bay,** where they are found in stream edges with trees and in steep gullies.
	***Native Lizard*** Populations of the **Auckland green gecko** are in gradual decline. Populations have been found in planted forests in **Northland**, **Waikato** and **Bay of Plenty** regions. They have well developed vocal cords and can bark or chirp by clicking their tongues against the roof of the mouth. They can be seen in tree branches, foliage and open ground. Although they hunt by night for insects, they also like to sunbathe.
	***NZ Bush Falcon*** The **NZ bush falcon** is classified as vulnerable to extinction. Very few bush falcons can be sighted on native bush but many can be found in large planted forests in North Island which include **Kaingaroa Forest** in the **Central North Island** and in **South Island** planted forests including the **Golden Downs in Nelson**. They can be sighted in forest stand edges between clearfell and mature stands.

We are now going to present you with a number of choice situations. These describe the outcomes of conservation policies that could be undertaken by the Department of Conservation in partnership with concerned organisations (e.g., forest corporations). Ecologists suggest that over the next five years, planted forests could be managed to provide better habitat for threatened species. These species include the above four threatened animals and one plant species. For each choice situation we present you, we will ask you to select the alternative with the conservation outcomes you prefer. Some outcomes will require a contribution to the Department of Conservation through an additional amount in your annual income tax for five years. In each choice situation, there is also the possibility of taking no conservation action ("Current Condition") and paying no money.

11. An example of answering the choice situation on the next few pages:

Below is an example of a choice situation that provides you with three options. The column heading "current condition" represents an alternative with no change. In this case, there is no enhancement of habitat for threatened species in planted forests. There is no increase in the occurrence of threatened species in planted forests, and this has no cost to you. If you chose **Option 1**, the associated additional annual cost in your income tax is **$50 for five years only**. Choosing **Option 1** would guarantee an increase in hearing **kiwi calls** from the current condition of **1 out of 200 planted forests** to a changed condition of **20 out of 200 planted forests**, a 10% increase. Forest ecologists suggested that this 10% increase is feasible. Choosing **Option 1** also corresponds to an increase in **Bush falcon sightings** when driving through plantations from **1 out of 8** to **3 out of 8 occasions**. Alternatively, if you chose **Option 2,** you would pay less than Option 1, only **$25**. However, this option would have less increase in the number of **kiwi calls** heard (**10 out of 200**), but more **bush falcons** would be seen when driving through the forests (**5 out of 8** occasions). **Option 2** also provides a greater increase in **Kakabeak** in 20% of the planted forest areas on the East Coast and Hawkes Bay. In this case, and after considering the change in the condition of the **Auckland green gecko** and **Giant kokopu**, if you prefer to hear more **Kiwi**, you can tick (✓) the box under **Option A** and this also indicates that you would be willing to pay $50 in extra income tax for five years.

Threatened Animal/Plant	***Current Condition***	***Option 1***	***Option 2***
<u>Brown Kiwi</u> (Frequency of hearing calls in planted forests in North Island) 	Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 20 out of 200** planted forests	Kiwi calls heard **in 10 out of 200** planted forests
<u>Giant Kokopu</u> (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand) 	Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 5 out of 10** suitable streams	Kokopu seen **in 3 out of 10** suitable streams
<u>Kakabeak</u> (Occurrence in 20% of the planted forests on the East Coast and Hawke's Bay) 	At least **3 naturally occurring** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs	At least **20 actively managed** Kakabeak shrubs
<u>Auckland Green Gecko</u> (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions) 	Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 5 out of 50** walks	Gecko sighted **in 3 out of 50** walks
<u>NZ Bush Falcon</u> (Bush falcon sightings while driving through pine forests in Central North Island and Nelson) 	Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 3 out of 8** drives	Bush falcon sighted **in 5 out of 8** drives
Additional amount to be paid yearly in your income tax for five years	**$0**	**$50**	**$25**
I would choose (please tick)	☐	☑	☐

Now that we drove you through an example on the previous page, we would like you to make the next choices on your own. Please remember to consider the payment as if it was real and give honest answers so as to inform conservation policy.

11a. Which of the three options below would you prefer most? Read the description and tick the box "□" that corresponds to your most preferred option.

Threatened Animal/Plant		*Current Condition*	*Option A*	*Option B*
Brown Kiwi (Frequency of hearing calls in planted forests in North Island)		Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 20 out of 200** planted forests	Kiwi calls heard **in 1 out of 200** planted forests
Giant Kokopu (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand)		Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 3 out of 10** suitable streams	Kokopu seen **in 1 out of 10** suitable streams
Kakabeak (Occurrence in 20% of the planted forests on the East Coast and Hawke's Bay)		At least **3 naturally occurring** Kakabeak shrubs	At least **3 naturally occurring** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs
Auckland Green Gecko (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions)		Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 5 out of 50** walks	Gecko sighted **in 1 out of 50** walks
NZ Bush Falcon (Bush falcon sightings while driving through pine forests in Central North Island and Nelson)		Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 3 out of 8** drives	Bush falcon sighted **in 1 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only		**$0**	**$30**	**$60**
I would choose (please tick)		□	□	□

11b. Below is a choice situation different to that of 11a, taking into account your income constraints and household needs, etc., which of the three options below would you prefer most? Tick the box “□” at the bottom that corresponds to your most preferred option.

Threatened Animal/Plant		*Current Condition*	*Option C*	*Option D*
Brown Kiwi (Frequency of hearing calls in planted forests in North Island)		Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 10 out of 200** planted forests	Kiwi calls heard **in 1 out of 200** planted forests
Giant Kokopu (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand)		Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 3 out of 10** suitable streams	Kokopu seen **in 1 out of 10** suitable streams
Kakabeak (Occurrence in 20% of the planted forests on the East Coast and Hawke’s Bay)		At least **3 naturally occurring** Kakabeak shrubs	At least **3 naturally occurring** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs
Auckland Green Gecko (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions)		Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 3 out of 50** walks
NZ Bush Falcon (Bush falcon sightings while driving through pine forests in Central North Island and Nelson)		Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 3 out of 8** drives	Bush falcon sighted **in 5 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only		**$0**	**$90**	**$30**
I would choose (please tick)		□	□	□

11c. Looking at a different choice situation below, taking into account your income constraints and household needs, etc., which of the three options below would you prefer most? Tick the box "□" that corresponds to your most preferred option.

Threatened Animal/Plant	***Current Condition***	***Option E***	***Option F***
Brown Kiwi (Frequency of hearing calls in planted forests in North Island) 	Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 10 out of 200** planted forests	Kiwi calls heard **in 1 out of 200** planted forests
Giant Kokopu (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand) 	Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 5 out of 10** suitable streams
Kakabeak (Occurrence in 20% of the planted forests on the East Coast and Hawke's Bay) 	At least **3 naturally occurring** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs	At least **3 naturally occurring** Kakabeak shrubs
Auckland Green Gecko (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions) 	Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 5 out of 50** walks
NZ Bush Falcon (Bush falcon sightings while driving through pine forests in Central North Island and Nelson) 	Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 5 out of 8** drives	Bush falcon sighted **in 1 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only	**$0**	**$60**	**$60**
I would choose (please tick)	□	□	□

11d. Looking at another choice situation below, taking into account your income constraints and household needs, etc., which of the three options below would you prefer most? Tick the box “□” that corresponds to your most preferred option.

Threatened Animal/Plant	*Current Condition*	*Option G*	*Option H*
Brown Kiwi (Frequency of hearing calls in planted forests in North Island) 	Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 20 out of 200** planted forests	Kiwi calls heard **in 10 out of 200** planted forests
Giant Kokopu (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand) 	Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 3 out of 10** suitable streams	Kokopu seen **in 5 out of 10** suitable streams
Kakabeak (Occurrence in 20% of the planted forests on the East Coast and Hawke’s Bay) 	At least **3 naturally occurring** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs	At least **3 naturally occurring** Kakabeak shrubs
Auckland Green Gecko (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions) 	Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 5 out of 50** walks	Gecko sighted **in 1 out of 50** walks
NZ Bush Falcon (Bush falcon sightings while driving through pine forests in Central North Island and Nelson) 	Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 3 out of 8** drives	Bush falcon sighted **in 5 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only	**$0**	**$30**	**$60**
I would choose (please tick)	□	□	□

11e. Looking at the fifth choice situation below, taking into account your income constraints and household needs, etc., which of the three options below would you prefer most? Tick the box "□" that corresponds to your most preferred option.

Threatened Animal/Plant	*Current Condition*	*Option I*	*Option J*
<u>Brown Kiwi</u> (Frequency of hearing calls in planted forests in North Island) 	Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 10 out of 200** planted forests	Kiwi calls heard **in 10 out of 200** planted forests
<u>Giant Kokopu</u> (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand) 	Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 5 out of 10** suitable streams	Kokopu seen **in 1 out of 10** suitable streams
<u>Kakabeak</u> (Occurrence in 20% of the planted forests on the East Coast and Hawke's Bay) 	At least **3 naturally occurring** Kakabeak shrubs	At least **3 naturally occurring** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs
<u>Auckland Green Gecko</u> (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions) 	Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 3 out of 50** walks	Gecko sighted **in 5 out of 50** walks
<u>NZ Bush Falcon</u> (Bush falcon sightings while driving through pine forests in Central North Island and Nelson) 	Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 3 out of 8** drives	Bush falcon sighted **in 1 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only	**$0**	**$60**	**$90**
I would choose (please tick)	□	□	□

11f. Looking at the sixth choice situation below, taking into account your income constraints and household needs, etc., which of the three options below would you prefer most? Tick the box "□" that corresponds to your most preferred option.

Threatened Animal/Plant		***Current Condition***	***Option K***	***Option L***
Brown Kiwi (Frequency of hearing calls in planted forests in North Island)		Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 10 out of 200** planted forests	Kiwi calls heard **in 20 out of 200** planted forests
Giant Kokopu (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand)		Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 3 out of 10** suitable streams	Kokopu seen **in 1 out of 10** suitable streams
Kakabeak (Occurrence in 20% of the planted forests on the East Coast and Hawke's Bay)		At least **3 naturally occurring** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs
Auckland Green Gecko (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions)		Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 3 out of 50** walks	Gecko sighted **in 5 out of 50** walks
NZ Bush Falcon (Bush falcon sightings while driving through pine forests in Central North Island and Nelson)		Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 5 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only		**$0**	**$90**	**$90**
I would choose (please tick)		☐	☐	☐

11g. Looking at the seventh choice situation below, taking into account your income constraints and household needs, etc., which of the three options below would you prefer most? Tick the box "□" that corresponds to your most preferred option.

Threatened Animal/Plant		***Current Condition***	***Option M***	***Option N***
Brown Kiwi (Frequency of hearing calls in planted forests in North Island)		Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 20 out of 200** planted forests	Kiwi calls heard **in 1 out of 200** planted forests
Giant Kokopu (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand)		Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 5 out of 10** suitable streams	Kokopu seen **in 1 out of 10** suitable streams
Kakabeak (Occurrence in 20% of the planted forests on the East Coast and Hawke's Bay)		At least **3 naturally occurring** Kakabeak shrubs	At least **20 actively managed** Kakabeak shrubs	At least **20 actively managed** Kakabeak shrubs
Auckland Green Gecko (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions)		Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 5 out of 50** walks	Gecko sighted **in 1 out of 50** walks
NZ Bush Falcon (Bush falcon sightings while driving through pine forests in Central North Island and Nelson)		Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 5 out of 8** drives	Bush falcon sighted **in 1 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only		**$0**	**$60**	**$30**
I would choose (please tick)		□	□	□

11h. Looking at the eight and last choice situation below, taking into account your income constraints and household needs, etc., which of the three options below would you prefer most? Tick the box "□" that corresponds to your most preferred option.

Threatened Animal/Plant		*Current Condition*	*Option O*	*Option P*
Brown Kiwi (Frequency of hearing calls in planted forests in North Island)		Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 20 out of 200** planted forests
Giant Kokopu (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand)		Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 5 out of 10** suitable streams	Kokopu seen **in 3 out of 10** suitable streams
Kakabeak (Occurrence in 20% of the planted forests on the East Coast and Hawke's Bay)		At least **3 naturally occurring** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs	At least **3 naturally occurring** Kakabeak shrubs
Auckland Green Gecko (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions)		Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 3 out of 50** walks	Gecko sighted **in 5 out of 50** walks
NZ Bush Falcon (Bush falcon sightings while driving through pine forests in Central North Island and Nelson)		Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 3 out of 8** drives	Bush falcon sighted **in 1 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only		**$0**	**$90**	**$60**
I would choose (please tick)		□	□	□

11i. Looking at the eight and last choice situation below, taking into account your income constraints and household needs, etc., which of the three options below would you prefer most? Tick the box "□" that corresponds to your most preferred option.

Threatened Animal/Plant	***Current Condition***	***Option Q***	***Option R***
Brown Kiwi (Frequency of hearing calls in planted forests in North Island) 	Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 20 out of 200** planted forests
Giant Kokopu (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand) 	Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 5 out of 10** suitable streams	Kokopu seen **in 3 out of 10** suitable streams
Kakabeak (Occurrence in 20% of the planted forests on the East Coast and Hawke's Bay) 	At least **3 naturally occurring** Kakabeak shrubs	At least **20 actively managed** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs
Auckland Green Gecko (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions) 	Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 5 out of 50** walks	Gecko sighted **in 1 out of 50** walks
NZ Bush Falcon (Bush falcon sightings while driving through pine forests in Central North Island and Nelson) 	Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 5 out of 8** drives	Bush falcon sighted **in 3 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only	**$0**	**$60**	**$90**
I would choose (please tick)	□	□	□

11j. In answering questions 11a to 11h, what feature/s attracted you most when you made your selection? (tick all that apply)

- ☐ Would like to hear more kiwi calls
- ☐ Would like to have more sightings of bush falcons
- ☐ Would like to see a kakabeak in the wild
- ☐ Would like to see and/ or hear more green geckos
- ☐ More giant kokopu in the wild
- ☐ Location sites of the programme
- ☐ Other: specify ___________________ _________________________________

11k. Assuming that you have considered the proposed cost, are there any species that you ignored when making your selection? If so, which? (tick all that apply)

- ☐ NZ bush falcon
- ☐ Giant kokopu
- ☐ Auckland green gecko
- ☐ Kakabeak
- ☐ Brown kiwi

11l. Are there any other animals or plants that we did not include in the above options that you would prefer to pay for?

- ☐ Long-tailed bat
- ☐ Hochstetter's Frog
- ☐ Blue duck
- ☐ Ponga
- ☐ Tui
- ☐ Other: specify __________________ _______________________________

11m. Please rate your <u>understanding</u> of the choice Questions in 11a to 11h (circle a number below):

Did not Understand at all				*Moderately Understood*					*Completely Understood*
1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

11n. Please rate <u>how easy</u> it was to choose your most preferred option in Questions 11a to 11h (circle a number below):

Very Difficult				*Neither Easy nor Difficult*					*Very Easy*
1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*

11o. If you did not choose any changed option in Questions 11a to 11h (e.g., more kiwi heard, increase in the number of falcon sightings), please explain why? (tick all that apply)

- ☐ I didn't want to place a dollar value
- ☐ I object to the way the question is presented
- ☐ I am opposed to further increases in income tax
- ☐ I am unsure at the moment, need to ask someone first
- ☐ The government should pay
- ☐ Not enough information provided
- ☐ Other: specify ____________________ _________________________________

<u>About You</u>

ALL YOUR ANSWERS IN THIS QUESTIONNAIRE ARE STRICTLY CONFIDENTIAL AND WILL ONLY BE USED FOR THE ANALYSIS OF THIS STUDY. NAMES AND ADDRESSES WILL NOT BE DISCLOSED.

12. Are you (Please tick one box): ☐ Male ☐ Female

13. Were you born in New Zealand? ☐ Yes ☐ No

14. Which age group do you belong to?
- ☐ Under 25 years old
- ☐ 25 to 34
- ☐ 35 to 44
- ☐ 45 to 54
- ☐ 55 to 64
- ☐ 65 to 74
- ☐ 75 and above

15. How many people live in your home (including yourself)?
Adults (18 and above) _____ Children (under 18) ______ Elderly dependent(s) ______

16. What is your highest level of formal schooling? (Tick one box)
- ☐ Primary
- ☐ Secondary/High School/College
- ☐ Trades certificate/Post-school diploma
- ☐ Tertiary/Undergraduate/University
- ☐ Post Graduate/Masters/PhD
- ☐ Other: specify ____________________________

17. Which one of the following best describes your current employment status?
- ☐ Employed full time
- ☐ Employed part time
- ☐ Self employed
- ☐ Not employed, but seeking work
- ☐ Not in the labour force (retired, student, etc)
- ☐ On ACC or sickness benefit
- ☐ Other: specify ____________________________
____

18. What is your current main occupation? ____

19. Type of employer (Agriculture, Healthcare, Sales, etc) _________________________________

20a. Have you ever driven through an exotic planted forests? ☐ Yes ☐ No

20b. Have you ever visited an exotic planted forests? ☐Yes ☐No

20c. If Yes, please specify the name/s of the exotic planted forests you have visited?
____

20d. If you have visited at least one planted forest, what activities did you participate in? (Tick all that apply)
- ☐ Bird watching
- ☐ Walking
- ☐ Camping
- ☐ Photography
- ☐ Horse riding
- ☐ Jogging/Running
- ☐ Picnicking
- ☐ Hunting
- ☐ Fishing
- ☐ Nature observation
- ☐ Cycling
- ☐ Other/s: specify
____

21a. Would you be willing to volunteer (e.g. bird counting, habitat restoration, giving talks) to help in the conservation of threatened animals and plants in planted forests? ☐ Yes ☐ No

21b. If Yes, how many days in the next 365 days would you be willing to volunteer? ______ days

21c. If No, please explain why.
- ☐ Prefer to volunteer in native forests
- ☐ Lack of time to volunteer
- ☐ Involved in other voluntary endeavour
- ☐ Other: specify __________________
____

22. Do you or does any of your household members own a share of a planted forest or stock in a forestry company in New Zealand?

☐ Yes ☐ No

23. What ethnic group do you belong to:

☐ New Zealand Māori	☐ Pacific Islands origin	☐ Asian
☐ New Zealand born European	☐ African	☐ Latin American
☐ European immigrant	☐ Other: specify ______________________________	

24. Your last year's personal income before taxes (tick one)

☐ $10,000 and below	☐ $50,001 to $60,000	☐ $100,001 to $110,000
☐ $10,001 to $20,000	☐ $60,001 to $70,000	☐ $110,001 to $120,000
☐ $20,001 to $30,000	☐ $70,001 to $80,000	☐ $120,001 and above
☐ $30,001 to $40,000	☐ $80,001 to $90,000	☐ Not Applicable
☐ $40,001 to $50,000	☐ $90,001 to $100,000	

25. Your spouse's personal income before taxes last year (tick one)

☐ $10,000 and below	☐ $50,001 to $60,000	☐ $100,001 to $110,000
☐ $10,001 to $20,000	☐ $60,001 to $70,000	☐ $110,001 to $120,000
☐ $20,001 to $30,000	☐ $70,001 to $80,000	☐ $120,001 and above
☐ $30,001 to $40,000	☐ $80,001 to $90,000	☐ Not Applicable
☐ $40,001 to $50,000	☐ $90,001 to $100,000	

26. Do you participate in any conservation organisations? (Tick all that apply).

☐ Bird conservation member (e.g., Wingspan)	☐ Green Peace member
☐ Forest and Bird member	☐ Department of Conservation volunteer
☐ Care group (e.g., Kokako Trust)	☐ Other/s: specify ____________________
☐ None	____________________________________

27. If you participate in one or more community organisations, please tick all that apply.

☐ Federated Farmers or Young Farmers	☐ Rural Women NZ
☐ Service Club (e.g., Lions, Rotary)	☐ Church group
☐ Playgroup, Kindergarten or Kohanga Reo	☐ Family recreation group (e.g., cards)
☐ Sports, Hunting or Fishing Club	☐ Professional organisation
☐ None	☐ Other/s: specify ____________________
	____________________________________

Thank you very much for completing the survey. We greatly appreciate your input in helping us study the importance of threatened species in New Zealand's exotic planted forests.

<u>Although not required: Feel free to use this space for notes, comments, etc.</u>

Appendix Table 1

Estimates of normalised AICs of 20 latent class logit model specifications (full sample)

LC Model Number	Latent classes (LC#s) – Attributes ignored	Normalised AIC (AIC/N)	
		Cross section specification	Panel specification
1	LC1 - Ignored SQ, LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Ignored all attributes	1.932	1.256
2	LC1 - Ignored SQ, LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Ignored Cost	1.940	1.272
3	LC1 - Ignored SQ, LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Full attendance	Did not converge	1.266
4	LC1 - Ignored SQ LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Full attendance LC4 – Ignored all attributes	1.941	1.256
5	LC1 – Ignored cost LC2 - Ignored SQ LC3 – Ignored Gecko, Kakabeak and Kokopu LC4 – Ignored all attributes	1.941	1.154
6	LC1 – Ignored cost LC2 - Ignored SQ LC3 – Ignored Gecko, Kakabeak and Kokopu LC4 – Ignored Falcon	1.941	1.273
7	LC1 – Ignored cost LC2 - Ignored SQ LC3 – Ignored Gecko, Kakabeak and Kokopu LC4 – Ignored Kiwi	1.935	1.273
8	LC1 - Ignored SQ LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Full attendance LC4 – Ignored Kiwi	1.935	1.267
9	LC1 - Ignored SQ LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Full attendance LC4 – Ignored Falcon	1.936	Did not converge
10	LC1 - Ignored SQ LC2 – Ignored Gecko, Kakabeak and Kokopu LC3 – Ignored Kiwi LC4 – Ignored Falcon	1.936	1.268

Appendix Table 2

Estimates of logit models using piecewise linear coded attributes

Item	**Model 1** Conditional Logit			**Model 2** Latent Class Logit Panel			**Model 3** Random Parameters Logit (RPL)			**Model 4** RPL with Error Components		
Attributes and SQ	Coef	Std Err	P-value	Coef	Std Err	P-value	Coef	Std Err	P-value	Coef	Std Err	P-value
Brown kiwi 1a	**0.504**	**0.098**	**<0.01**	**0.669**	**0.121**	**<0.01**	**0.894**	**0.139**	**<0.01**	**0.914**	**0.136**	**<0.01**
Brown kiwi 2	0.118	0.090	0.19	0.150	0.121	0.22	0.115	0.133	0.38	0.150	0.142	0.29
Native fish 1a	**0.287**	**0.093**	**<0.01**	0.163	0.131	0.21	**0.331**	**0.133**	**0.01**	0.299	0.156	0.05
Native fish 2	-0.144	0.091	0.11	-0.139	0.130	0.28	-0.212	0.134	0.11	-0.179	0.179	0.32
Native plant 1a	**0.287**	**0.093**	**<0.01**	0.181	0.136	0.18	-0.025	0.136	0.85	-0.057	0.135	0.68
Native plant 2	*0.065*	*0.090*	*0.47*	*-0.053*	*0.127*	*0.68*	*-0.020*	*0.135*	*0.88*	*-0.013*	*0.162*	*0.94*
Green gecko 1a	0.017	0.093	0.86	-0.115	0.135	0.40	-0.025	0.136	0.85	-0.057	0.135	0.68
Green gecko 2	*0.076*	*0.093*	*0.41*	*0.053*	*0.128*	*0.68*	*0.181*	*0.152*	*0.23*	*0.178*	*0.172*	*0.30*
Bush falcon 1a	**0.453**	**0.098**	**<0.01**	**0.476**	**0.120**	**<0.01**	**0.964**	**0.144**	**<0.01**	**0.949**	**0.153**	**<0.01**
Bush falcon 2	***0.248***	***0.089***	***0.01***	***0.438***	***0.120***	***<0.01***	*0.285*	*0.148*	*0.05*	*0.285*	*0.181*	*0.11*
Status Quo Indicator	0.177	0.158	0.26	-5.864	0.504	<0.01	**-3.591**	**0.284**	**<0.01**	**-1.473**	**0.578**	**0.01**
Cost	***-0.025***	***0.002***	***<0.01***	***-0.123***	***0.011***	***<0.01***	***-0.183***	***0.010***	***<0.01***	***-0.064***	***0.004***	***<0.01***
Attribute non-attendance												
Ignoring cost				**0.347**	**0.038**	**<0.01**						
Ignoring status quo				**0.369**	**0.043**	**<0.01**						
Ignoring non-iconics				**0.227**	**0.038**	**<0.01**						
Ignoring all attributes				**0.057**	**0.020**	**<0.01**						
Random parameters												
Bush falcon 2							**1.905**	**0.482**	**<0.01**	**1.770**	**0.639**	**0.01**
Native plant 2							1.057	0.597	0.08	**1.490**	**0.508**	**<0.01**
Cost							**0.183**	**0.010**	**<0.01**	**0.064**	**0.004**	**<0.01**
Green gecko 2							**2.040**	**0.547**	**<0.01**	**1.888**	**0.532**	**<0.01**
Error components										**7.793**	**1.017**	**<0.01**
Log-likelihood	-1785.14			-1052.57			-1139.32			-990.50		
Normalised AIC	1.943			1.154			1.248			1.088		
McFadden Pseudo R^2	0.122			0.482			0.439			0.513		
No. of observations	1850			1850			1850			1850		

Appendix Table 3

Model 4 estimates of RPL-EC models using full and split samples (with dummy coded attribute levels)

Item	Model 4 Full Sample			Model 4L Large planted forest sample			Model 4S Small planted forest sample		
Attributes and SQ	Coef	Std err	*p*-value	Coef	Std err	*p*-value	Coef	Std err	*p*-value
Brown kiwi 1	**0.898**	**0.137**	**<0.01**	**1.063**	**0.198**	**<0.01**	**0.715**	**0.291**	**0.01**
Brown kiwi 2	***1.048***	***0.128***	***<0.01***	***1.365***	***0.186***	***<0.01***	***0.669***	***0.326***	***0.04***
Native fish 1	**0.307**	**0.153**	**0.04**	0.243	0.210	0.25	0.399	0.290	0.17
Native fish 2	*0.138*	*0.145*	*0.34*	*0.017*	*0.200*	*0.93*	*0.324*	*0.324*	*0.32*
Native plant 1	**0.343**	**0.163**	**0.04**	0.326	0.200	0.10	0.240	0.426	0.57
Native plant 2	***0.329***	***0.161***	***0.04***	*0.280*	*0.210*	*0.18*	*0.297*	*0.348*	*0.39*
Green gecko 1	-0.053	0.135	0.70	-0.060	0.173	0.73	-0.027	0.306	0.93
Green gecko 2	*0.124*	*0.159*	*0.43*	0.011	0.212	0.96	0.388	0.370	0.29
Bush falcon 1	**0.909**	**0.147**	**<0.01**	**0.824**	**0.202**	**<0.01**	**1.046**	**0.335**	**<0.01**
Bush falcon 2	***1.188***	***0.147***	***<0.01***	***1.334***	***0.200***	***<0.01***	***1.029***	***0.302***	***<0.01***
Status Quo Indicator	**-1.594**	**0.637**	**0.01**	**-2.024**	**0.806**	**0.01**	-0.814	1.453	0.58
Cost	***-0.063***	***0.004***	***<0.01***	***-0.077***	***0.007***	***<0.01***	***-0.041***	***0.006***	***<0.01***
Random Parameters									
Bush falcon 2	**1.606**	**0.658**	**0.01**	1.179	0.903	0.19	1.652	1.228	0.18
Native plant 2	**1.446**	**0.557**	**0.01**	1.291	0.866	0.14	**1.752**	**0.855**	**0.04**
Cost	**1.369**	**0.520**	**0.01**	**0.077**	**0.007**	**<0.01**	**0.041**	**0.006**	**<0.01**
Green gecko 2	**0.063**	**0.004**	**<0.01**	**1.944**	**0.768**	**0.01**	0.480	2.165	0.82
Error Component	**7.652**	**1.005**	**<0.01**	**8.050**	**1.207**	**<0.01**	**7.652**	**2.272**	**<0.01**
Log-likelihood	-990.68			-664.84			-313.69		
Normalised AIC	1.088			1.056			1.177		
McFadden Pseudo R^2	0.512			0.531			0.490		
Number of observations	1850			1290			560		
No. of groups	209			145			64		

Note1: Values in *italics* represent estimates for random parameters
Note2: Values in **boldface font** represent estimates statistically significant at 5% level.

Appendix Table 4

Heteroskedastic logit model (scale as a function of entropy) estimates for split and the pooled samples

	ORD			BDD			OOD			Pooled		
	Coeff	**Std Err**	***p*-value**	**Coeff**	**Std Err**	***p*-value**	**Coeff**	**Std Err**	***p*-value**	**Coeff**	**Std Err**	***p*-value**
Utility coefficient												
Brown kiwi 1	9.540	10.400	0.37	0.082	1.290	0.96	17.400	73.500	0.36	4.120	6.080	0.50
Brown kiwi 2	14.700	15.000	0.35	0.121	1.870	0.96	20.000	83.700	0.33	5.730	8.730	0.51
Native fish 1	5.120	6.040	0.42	0.106	1.650	0.96	3.470	15.100	0.54	2.620	3.920	0.50
Native fish 2	3.230	4.810	0.52	-0.040	0.610	0.96	6.140	26.400	0.45	0.864	1.440	0.55
Native plant 1	2.310	3.560	0.51	-0.042	0.653	0.96	5.370	23.000	0.52	1.490	2.450	0.54
Native plant 2	-3.070	4.860	0.53	0.155	2.380	0.96	2.150	10.400	0.72	1.700	2.500	0.50
Green gecko 1	2.780	4.440	0.53	0.003	0.084	0.98	-2.010	10.600	0.75	0.296	0.994	0.77
Green gecko 2	5.490	5.600	0.30	-0.039	0.609	0.96	3.640	15.600	0.53	0.545	1.070	0.61
Bush falcon 1	9.640	10.100	0.35	0.171	2.640	0.96	7.030	30.200	0.45	4.130	6.260	0.51
Bush falcon 2	16.300	16.300	0.35	0.231	3.570	0.96	13.400	56.500	0.37	5.870	8.720	0.50
Cost to respondent	-0.497	0.486	0.33	-0.006	0.097	0.96	-0.927	3.890	0.35	-0.215	0.320	0.50
Indicator for SQ	-13.600	14.600	0.38	0.060	0.910	0.96	12.800	54.800	0.47	-1.300	2.400	0.59
Scale coefficient												
Entropy	-11.100	3.660	<0.01	5.670	33.900	0.90	-5.860	9.150	0.03	-4.820	3.860	0.21
Entropy squared	9.460	3.190	<0.01	-4.750	18.500	0.84	2.410	5.060	0.21	2.670	2.450	0.28
Model statistics												
Adjusted Rho-square	0.151			0.102			0.126			0.120		
Log-likelihood value	-455.09			-496.14			-468.92			-1458.86		
Number of choice observations	503			503			503			1509		

Appendix Table 5

Heteroskedastic logit model (scale as a function of attribute dispersion) estimates for split and the pooled samples

	ORD			BDD			OOD			Pooled		
	Coeff	**Std Err**	***p*-value**	**Coeff**	**Std Err**	***p*-value**	**Coeff**	**Std Err**	***p*-value**	**Coeff**	**Std Err**	***p*-value**
Utility coefficient												
Brown kiwi 1	0.971	0.654	0.14	0.012	0.023	0.59	0.126	0.196	0.52	1.260	0.689	0.07
Brown kiwi 2	1.600	1.060	0.13	0.020	0.030	0.51	0.163	0.240	0.50	1.640	0.865	0.06
Native fish 1	0.741	0.647	0.25	0.021	0.031	0.49	0.045	0.075	0.55	0.849	0.499	0.09
Native fish 2	0.625	0.584	0.28	-0.015	0.020	0.46	0.045	0.075	0.55	0.362	0.316	0.25
Native plant 1	0.355	0.472	0.45	-0.016	0.021	0.43	0.051	0.084	0.54	0.447	0.345	0.19
Native plant 2	-0.492	0.660	0.46	0.025	0.037	0.50	0.029	0.054	0.59	0.558	0.369	0.13
Green gecko 1	0.029	0.506	0.95	-0.008	0.014	0.55	-0.005	0.042	0.91	0.016	0.260	0.95
Green gecko 2	0.622	0.502	0.21	-0.014	0.020	0.49	0.031	0.057	0.58	0.117	0.264	0.66
Bush falcon 1	1.460	1.070	0.17	0.042	0.054	0.43	0.051	0.092	0.58	1.270	0.715	0.08
Bush falcon 2	2.070	1.390	0.13	0.054	0.071	0.45	0.113	0.171	0.51	1.810	0.956	0.06
Cost to respondent	-0.065	0.047	0.17	-0.001	0.002	0.44	-0.007	0.010	0.51	-0.067	0.037	0.07
Indicator for SQ	-1.160	0.767	0.13	0.023	0.033	0.48	0.055	0.114	0.63	-0.311	0.449	0.49
Scale coefficient												
ASD	-0.554	0.555	0.32	2.590	1.040	0.01	1.420	1.260	0.26	-0.653	0.438	0.14
DSD	-1.490	0.903	0.10	-1.720	1.880	0.36	-0.763	1.240	0.54	-1.150	0.722	0.11
Model statistics												
Adjusted Rho-square	0.147			0.081			0.126			0.120		
Log-likelihood value	-457.51			-493.85			-468.77			1458.18		
Number of choice observations	503			503			503			1509		

Appendix Figure 1

A sample choice task derived from an orthogonal design with overlaps in three attributes. Many ORD choice tasks have at least one overlapping attribute levels.

Threatened Animal/Plant	***Current Condition***	***Option I***	***Option J***
Brown Kiwi (Frequency of hearing calls in planted forests in North Island) 	Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 10 out of 200** planted forests
Giant Kokopu (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand) 	Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 1 out of 10** suitable streams
Kakabeak (Occurrence in 20% of the planted forests on the East Coast and Hawke's Bay) 	At least **3 naturally occurring** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs
Auckland Green Gecko (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions) 	Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 5 out of 50** walks	Gecko sighted **in 3 out of 50** walks
NZ Bush Falcon (Bush falcon sightings while driving through pine forests in Central North Island and Nelson) 	Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 5 out of 8** drives	Bush falcon sighted **in 5 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only	**$0**	**$90**	**$60**
I would choose (please tick)	☐	☐	☐

Appendix Figure 2

A choice task derived Bayesian D-efficient design with an overlap in cost attribute. One would occasionally find overlaps in BDD choice tasks.

Threatened Animal/Plant		*Current Condition*	*Option I*	*Option J*
Brown Kiwi (Frequency of hearing calls in planted forests in North Island)		Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 20 out of 200** planted forests
Giant Kokopu (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand)		Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 3 out of 10** suitable streams	Kokopu seen **in 1 out of 10** suitable streams
Kakabeak (Occurrence in 20% of the planted forests on the East Coast and Hawke's Bay)		At least **3 naturally occurring** Kakabeak shrubs	At least **20 actively managed** Kakabeak shrubs	At least **3 actively managed** Kakabeak shrubs
Auckland Green Gecko (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions)		Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 3 out of 50** walks	Gecko sighted **in 1 out of 50** walks
NZ Bush Falcon (Bush falcon sightings while driving through pine forests in Central North Island and Nelson)		Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 5 out of 8** drives	Bush falcon sighted **in 1 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only		**$0**	**$30**	**$30**
I would choose (please tick)		☐	☐	☐

Appendix Figure 3

A choice task derived from optimal orthogonal design with no attribute overlap. No overlapping attribute levels were found in all pairs of changed alternatives in OOD choice tasks.

Threatened Animal/Plant		***Current Condition***	***Option I***	***Option J***
Brown Kiwi (Frequency of hearing calls in planted forests in North Island)		Kiwi calls heard **in 1 out of 200** planted forests	Kiwi calls heard **in 20 out of 200** planted forests	Kiwi calls heard **in 1 out of 200** planted forests
Giant Kokopu (Occurrence in slow moving streams with overhanging native vegetation in planted forests throughout New Zealand)		Kokopu seen **in 1 out of 10** suitable streams	Kokopu seen **in 5 out of 10** suitable streams	Kokopu seen **in 1 out of 10** suitable streams
Kakabeak (Occurrence in 20% of the planted forests on the East Coast and Hawke's Bay)		At least **3 naturally occurring** Kakabeak shrubs	At least **10 actively managed** Kakabeak shrubs	At least **20 actively managed** Kakabeak shrubs
Auckland Green Gecko (Gecko sightings in open grounds in planted forests in Northland, Waikato and Bay of Plenty regions)		Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 1 out of 50** walks	Gecko sighted **in 3 out of 50** walks
NZ Bush Falcon (Bush falcon sightings while driving through pine forests in Central North Island and Nelson)		Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 1 out of 8** drives	Bush falcon sighted **in 3 out of 8** drives
Additional amount to be paid yearly in your income tax for five years only		**$0**	**$30**	**$60**
I would choose (please tick)		☐	☐	☐