Valuing Benefits of Finnish Forest Biodiversity Conservation: Fixed and Random Parameter Logit Models for Pooled Contingent Valuation and Contingent Rating/Ranking Survey Data

Juha Siikamäki
Ph.D. Candidate
Department of Environmental Science and Policy
University of California, Davis
Email: juha@primal.ucdavis.edu

Presented at the Annual Meeting of Western Regional Project W-133
Miami, Florida
February 26-28, 2001

Acknowledgements: Doug Larson and Dave Layton are acknowledged for their comments and suggestions.

Abstract: This paper examines the use of contingent valuation and contingent rating/ranking valuation methods (CV and CR methods) in measuring willingness-to-pay (WTP) for nonmarket goods. Random parameter models are applied to CV and CR data and their performance is evaluated in comparison to conventionally used fixed parameter models. A framework for using data pooling techniques to test for invariance between separate sources of data is presented and applied to combined CV and CR data. The empirical application deals with measuring the WTP for conserving biodiversity hotspots in Finnish non-industrial private forests. Results suggest that the random coefficient models perform statistically well in comparison to the fixed parameter models that sometimes violate the assumptions of the conditional logit model. Based on the pooled models on combined data, parameter invariance between the CV and CR data cannot be uniformly accepted or rejected. Rejecting pooling of the data becomes more likely as more detailed response models are applied.

Introduction

This paper examines the use contingent valuation and contingent rating/ranking (CV and CR) methods in measuring willingness-to-pay (WTP) for nonmarket goods. Recent developments in discrete choice econometrics using random parameter models are applied to CV and CR data, and their performance is evaluated in comparison to conventionally used fixed parameter econometric models. Further, invariance between the CV and CR data is examined by data pooling techniques.

Stated preference methods (SP methods) are widely used in measuring economic values related to the environment. Standard SP applications include conducting surveys, in which respondents are described hypothetical alternatives, usually policy options. Each policy option results a certain supply of nonmarket good, such as environmental quality, for certain costs to respondents. The respondents are asked to evaluate the alternatives and state their preferences regarding them. The CV is based on asking for acceptance/refusal of hypothetical payment for implementing a policy alternative; the CR relies on asking respondents to rate or rank the available alternatives, at the simplest by choosing a preferred alternative. Obtaining responses for a variety of cost-environmental quality combinations, data with implicit information on individual tradeoffs between money and environmental quality are collected. The tradeoffs can be quantified by using discrete choice econometric models, that explain the observed choices by attributes of policy alternatives and respondents. In essence, the econometric models are used to measure an individual level exchange rate between a nonmarket good and money. Willingness to pay (WTP) for changes in the environmental quality can then be calculated using the estimation results.

Although several stated preference methods are currently in use, their performance and consistency has not been exhaustively studied. Examples of studies on differences between SP methods include Desvouges and Smith 1983, Magat et al. 1988, Boxall et al. 1996, and Stevens et al. 2000. They all suggest substantial differences between the various SP methods. However, the objective of all the methods is to measure essentially the same tradeoffs between money and changes in the environmental quality, and their results should therefore be very similar.

Previous studies on differences across the SP methods are typically based on fixed parameter discrete choice models, usually logit models. The assumptions and properties of fixed logit models are restrictive, but more flexible models with random parameters have been practically unavailable due to limitations in computing power and simulation based econometric techniques. Both constraints have recently been greatly relaxed and random parameter models are now possible to be employed in modeling the discrete choice SP data. Both Train's (1998) analysis of recreational fishing site choice and Layton's (2000) work on rankings data demonstrate that random parameter formulation can substantially improve the statistical performance of the econometric models typically used for valuing natural resources.

As mentioned, the differences between the SP methods have been previously studied by relatively restrictive models. Hence, it is justified to re-examine the differences of the SP methods by using less restrictive random parameter models. More flexible models let us evaluate if the previous conclusions have resulted from actual inequalities between different SP data sources, or perhaps from using overly restrictive econometric models.

Adamoviz et al. (1994) tested for differences between the observed and stated choices by estimating models for combined data on observed and stated choices. Recently, Hensher et al. (1999) provide a general framework for applying data pooling techniques to test for the invariance between separate sources of data. The data pooling approach is adopted here and used in testing for the equality of the CV and CR data. The approach can be easily extended into different settings.

The data pooling approach enables comparing different data sources already in estimation. Several benefits follow in comparison to the traditional approach of estimating separate models for different data sets and comparing their results afterwards: First, likelihood ratio based tests for data source invariance become available. Second, if data from different SP sources can in fact be considered equal, the pooled econometric models provide practical means to utilize all the information in data collected. This in turn can result in more reliable model estimates than the unpooled models.

The empirical application deals with measuring WTP for conserving especially valuable habitats (biodiversity hotspots) in Finnish non-industrial private forests. According to ecologists, protection of the biodiversity hotspots is particularly important for biodiversity conservation in Finland. The hotspots cover a total of 1.1 million hectares, that is some 6 % of the Finnish forests. Current regulations protect some 110,000 hotspot hectares and extending their protection is currently debated. This study evaluates the potential conservation policy alternatives by examining public preferences for them.

Forest conservation in Finland is an inexhaustible source of public debates and policy conflicts. Clearly, management and harvesting of forests are the primary reasons for species extinction. Rather intensive forest management practices over a long period of time have provided country with more timber resources than ever in the known past. At the same time, substantial losses of old forests and other important habitats for many currently threatened species have resulted. On the other hand, a big share of the country's exports consist of forest products such as paper- and sawmill products. Economic interests related to forests are therefore evident. Noting further that forests consist mostly (65-75 %) of small holdings (avg. size 100 acres), owned by private households, and that almost 10 % of the Finnish population owns some areas forests, it is clear that forest conservation policies are both of considerable public and regulatory interest.

The specific objectives of this paper are to

1) Review and to discuss current logit models for the CV and CR data.
2) Examine the random parameter modeling approach in comparison to fixed parameter models.
3) Test for differences between SP methods by using data pooling methods.
4) Analyze WTP estimates for both fixed and random parameter, and unpooled and pooled models for CV and CR data.

The rest of the paper is organized as follows: The first section represents the econometric models for CV and CR data, including fixed and random coefficient models. The next section explains how data pooling techniques can be used to test for data invariance between different stated preference data sources. The empirical section starts with a description of the public survey for preferences for biodiversity conservation in Finland. Results start with fixed logit models and continue with results for random parameter models. After separate estimation of the CV and CR data, the two data sets are pooled and invariance between the CV and CR data tested. The results section is concluded with the WTP estimates for different models. Last, the results of the study are discussed and concluded.

Econometric Models for Contingent Valuation and Contingent Rating/Ranking Survey Responses

Econometric models for stated preference surveys are typically based on McFadden's (1974) random utility model (RUM). The following section uses the RUM as a point of departure for explaining various econometric models for CV and CR survey responses. The CV section draws from works by Hanemann (1984), Hanemann et al. (1991), and Hanemann and Kanninen (1996); the CR section relies on McFadden (1974), Beggs et al. (1981), Chapman and Staelin (1982), Hausman and Ruud (1987), and on recent works by Train (e.g. 1998), Train and McFadden (2000) and Layton (2000a).

Random Utility Theoretic Framework for Modeling Individual Choices

Typical stated preference surveys measure individual tradeoffs between changes in environmental quality q and costs A of implementing them. This is accomplished by asking respondents to state their choices between the status quo with zero cost and one or more hypothetical policy alternatives with altered environmental quality and its costs. Consider an individual i choosing a preferred alternative from a set of m alternatives, each alternative j providing utility U_{ij}, that can be additively separated into an unobserved stochastic component ε_{ij} and a deterministic component $V_{ij}(q_j,y\text{-}A_j)$ i.e the restricted indirect utility function that depends only on individual's income y and environmental quality q. The utility of alternative j can then be represented as

$$U_{ij}=V_{ij}(q_j,y-A_j)+\varepsilon_{ij} \tag{2.1}$$

The stochastic ε_{ij} represents the unobserved factors affecting the observed choices. They can be related to individual tastes, choice task complicity, or any other factors with significant influence on choices. They are taken into consideration by individual j choosing between the alternatives, but to an outside observer, ε_{ij} remains unobserved and stochastic in the econometric modeling. From the viewpoint of individual making a choice, utility has no stochastic nature.

Choices are based on utility comparisons between the available alternatives, and the alternative providing the highest utility becomes the preferred choice. The probability of person i choosing alternative j among all the m the alternatives therefore equals the probability that the alternative j provides person i with greater utility U_{ij} than any other available alternative with U_{ik}. It is determined as

$$P_{ij} = P\,(U_{ij} > U_{ik},\ k = 1, \ldots, m, \quad k \quad j), \tag{2.2}$$

Denoting the difference of random components between alternatives j and k as $\varepsilon_{ijk} = \varepsilon_{ij} - \varepsilon_{ik}$, and the difference between the deterministic components as $\Delta V_{ijk}\,(.)= V_{ik}\,(q,y\text{-}A_k)- V_{ij}(q,y\text{-}A_j)$, the probability P_{ij} can be presented as probability

$$P_{ij} = P\,(\varepsilon_{ijk} > \Delta V_{ijk}(.),\ k = 1, .., m, \quad k \quad j) \tag{2.3}$$

Estimating parametric choice models requires specification of both the distribution of ε_i and the functional form of V_{ij}. Specification of ε_{ijk} determines the probability formulas for the observed responses; the functional form of V_{ij} is employed in estimating the unknown parameters of interest. Denoting all the exogenous variables of alternative j for the ith person as a vector X_{ij}, and the unknown parameters as a vector β, V_{ij} is typically specified as linear in parameters $V_{ij}=X_{ij}\beta$.

The following sections describe response probability formulas for different contingent valuation (CV) and contingent ranking (CR) models. Response probability formulas can be thought of as a likelihood function for the *i*th person. Since observations are independent, the likelihood function for the total sample is simply a sum of individual likelihood functions. The total maximum likelihood function can then be employed in estimating the V_{ij}.

Logit models for CR data

Assume in the following that random terms ε_j and ε_k are independently and identically distributed, type I generalized extreme value (GEV) random variables. It follows that their difference ε_{ijk} is logistically distributed. Under these assumptions, McFadden (1974) showed that choice probability P_{ij} in (2.5) is determined as a conditional logit model

$$P_{ij} = \frac{e^{\mu X_{ij}\beta}}{\sum_{k=1}^{k=m} e^{\mu X_{ik}\beta}} \tag{2.4}$$

The log-likelihood function for conditional logit model is

$$\ln L = \sum_{i}^{N} \ln P_{ij} \tag{2.5}$$

The parameter μ is a scale factor that appears in all the choice models based on RUM. It links the structure of random terms and the parameter estimates of $V_{ij}=X_{ij}\beta$. With data from a single source, the scale factor is typically set equal to one, left out, and parameter vector β estimated given the restricted scale factor. This is necessary for identification; without the imposed restriction on μ, neither μ nor β could be identified. However, in combining data from different sources, the scale factor plays an essential role. Since pooling of CV and CR data plays an important role in the analysis, scale parameters are included in all the following models. The role of the scale factor in pooling different sources of data will be discussed more in section 2.5.

Beggs et al (1981) and Chapman and Staelin (1982) extended the conditional logit model to modeling ranking of alternatives. A rank-ordered logit model treats ranking as *m-1* consecutive conditional choice problems. It assumes that ranking results from *m-1* utility comparisons, where the highest ranking is given to the best alternative (the preferred choice from the available alternatives), the second highest ranking to the best alternative from the remaining *m-1* alternatives, third from the remaining *m-2* alternatives, and so on. The probability of the observed ranking *r* for the person *i* is given by

$$P_{ir} = \prod_{j=1}^{m-1} \frac{e^{\mu X_{ij}\beta}}{\sum_{k=j}^{k=m} e^{\mu X_{ik}\beta}} \tag{2.6}$$

Hausman and Ruud (1987) developed a rank-ordered heteroscedastic logit model that is flexible enough to take into account possible increases (or decreases) in variance of the random term in the RUM as the ranking task continues. It is based on formulation with a rank-specific scale parameter that accounts systematic changes in the variance of the random term. By its structure, a rank-ordered heteroscedastic logit model can identify *m-2* scale parameters.

$$P_{ir} = \prod_{j=1}^{m-1} \frac{e^{\mu_j X_{ij} \beta}}{\sum_{k=j}^{k=m} e^{\mu_j X_{ik} \beta}} \tag{2.7}$$

As with the conditional logit model, the log-likelihood function for rank-ordered logit models (2.6) and (2.7) is the sum over individual probabilities over the whole sample.

Logit models for CV data

A single bounded discrete choice CV method is based on asking respondents if they would or would not be willing to pay certain reference amount *Bid* of money for altering the environmental quality q. Data consist of binary responses that result from yes/no answers to CV questions, asking for refusal/acceptance of paying an amount *Bid* for some policy alternative. In essence, the CV-method asks respondents to choose between status quo with utility $U_{i0}(q_0) = V_{i0}(q_0) + e_{i0}$ and an alternative providing utility $U_{i1}(q_1) = V_{i1}(q_1, y\text{-}Bid) + e_{i1}$. Given a logistically distributed stochastic term in the RUM, the probability of individual i choosing the alternative with costs *Bid* and environmental quality q_1 is the probability of obtaining a *Yes*-answer from person i. Expressing the observed parts of utilities as $V_{i0} = X_{i0}$ and $V_{i1} = X_{i1}$, the probability of a Yes-answer is given by the conditional logit model with two alternatives.

In double bounded CV, respondents are asked a follow-up question based on the first response. The objective is to gather more information on WTP than is possible by asking just a single question. Respondents who answered *Yes* to the first question (*FirstBid)* are asked a similar second question, this time with *HighBid* > *FirstBid.* Respondents who answered *No* get a second question with *LowBid* < *FirstBid.* Second responses provide more detailed data on individual preferences between the two alternatives and the choice probabilities can now be determined based on responses to two separate questions. Four possible response sequences can be observed: *Yes-Yes*, *Yes-No*, *No-Yes* and *No-No*. Using the conditional logit model, and denoting the exogenous variables for questions with *FirstBid*, *HighBid* and *LowBid* by X_{iFB}, X_{iHB} and X_{iLB}, the probabilities of the different responses are given by:

$$\begin{aligned}
\text{P(Yes-Yes)} &= P_i\,(YY) = \frac{e^{\mu X_{iHB}\beta}}{e^{\mu X_{iHB}\beta}+1} \\
\text{P(Yes-No)} &= P_i\,(NY) = \frac{1}{1+e^{\mu X_{iHB}\beta}} - \frac{1}{1+e^{\mu X_{iFB}\beta}} \\
\text{P(No-Yes)} &= P_i\,(NY) = \frac{1}{1+e^{\mu X_{iFB}\beta}} - \frac{1}{1+e^{\mu X_{iLB}\beta}} \\
\text{P(No-No)} &= P_i\,(NN) = \frac{1}{1+e^{\mu X_{iLB}\beta}}
\end{aligned} \tag{2.9}$$

Using dummy variables I_{yy}, I_{yn}, I_{ny}, I_{nn} to indicate *Yes-Yes*, *Yes-No*, *No-Yes* and *No-No* responses, the log-likelihood function for double-bounded CV is

$$L = \sum_{i=1}^{n} \ln[I_{yy}P_i(YY) + I_{yn}P_i(YN) + I_{ny}P_i(NY) + I_{nn}P_i(NN)] \tag{2.10}$$

Random parameter logit models

Although typically applied to SP data, some undesirable properties and assumptions are embodied in the fixed parameter logit models. First, they overestimate the joint probability of choosing close substitutes. This is known as the Independence of Irrelevant Alternatives (IIA) property (McFadden 1974). Second, they are based on the assumption that the random terms ε_{ij} are independently and identically distributed, although in practice it is likely that individual specific factors influence evaluation of all the available alternatives and make random terms correlated instead of independent. Third, assuming homogeneous preferences alone is restrictive. Any substantial variation in individual tastes conflicts with this assumption, possibly resulting in violations in many applications.

Random parameter logit (RPL) models have been proposed to overcome possible problems of the fixed parameter choice models (e.g. Revelt and Train 1998, Train 1998, Layton 2000). The RPL is specified similarly as the fixed parameter models, except that the parameters β now vary in the population rather than stay the same for everybody. Utility is expressed as a sum of population mean b, individual deviation   , which accounts for differences in individual taste from the population mean, and an unobserved i.i.d. random term ε. Total utility for person i from choosing the alternative j is determined as

$$U_{ij} = X_{ij}b + X_{ij}\quad_{i} + \varepsilon_{ij} \tag{2.11}$$

where $X_{ij}b$ and $X_{ij}\quad_{i} + \varepsilon_{ij}$ are the observed and unobserved parts of utility. Utility can also be expressed in form $X_{ij}(b + \quad_{i}) + \varepsilon_{ij}$, which is easily comparable to fixed parameter models. The only difference is that previously fixed β now varies across people as $\beta_i = b + \quad_{i}$.

Although the RPL models account for heterogeneous preferences via parameter $\quad_{i}$, individual tastes deviations $\quad_{i}$ are neither observed nor estimated. The RPL models aim at finding the different moments, for instance the mean and the deviation, of the distribution of β, from which each β_i is drawn. Parameters β vary in population with density $f(\beta|\quad)$, with   denoting the parameters of density. Since actual tastes are not observed, the probability of observing a certain choice is determined as an integral of the appropriate probability formula over all the possible values of β weighted by its density. Probability for choosing alternative j out of m alternatives can now be written as

$$P_{ij} = \int \left[\frac{e^{\mu X_{ij}\beta_i}}{\sum_{k=j}^{k=m} e^{\mu X_{ik}\beta_i}} \right] f(\beta \mid \Omega) d\beta \tag{2.12}$$

Equation (2.12) is the random parameter extension of the conditional logit model (2.4). Random parameter models for the rank-ordered logit models and the double bounded CV are defined similarly. Extension is straightforward and not replicated here. It suffices to note that they are formulated by replacing the bracketed part of (2.12) by the appropriate probability formula.

Integral (2.12) cannot be analytically calculated and must be simulated for estimation purposes. Therefore, exact maximum likelihood estimation is not available and simulated maximum likelihood is to be used instead. Train has developed a method that is suitable for simulating (2.12) and its many extensions needed in this study. His simulator is smooth, strictly positive and unbiased (Brownstone and Train 1999), and can be easily modified to allow for non-negative/positive random parameters. That is particularly practical in CV and CR studies in which theoretical considerations often suggest restrictions for parameter values. Simulating (2.12) is carried out simply by drawing a random β_i, calculating the bracketed part of the equation, and repeating the procedure over and over again. Although Train's simulator is unbiased for just one draw of β_i, its accuracy is increased with the number of draws. Using R draws of β_i from $f(\beta|\ )$, the simulated probability of (2.12) is

$$SP_{ij} = \frac{1}{R}\sum_{r=1}^{r=R}\frac{e^{\mu X_{ij}\beta_{ir}}}{\sum_{k=j}^{k=m} e^{\mu X_{ik}\beta_{ir}}} \tag{2.13}$$

Simulator (2.13) can be extended to rank-ordered logit model and to logit models for single and double bounded CV. The only required change is replacing the R times summed portion of (2.13) with the rank-ordered or double bounded CV probability formulas, as expressed by (2.6), and (2.9). In estimating mean and variance for the distribution of β, (2.17) can be employed by defining $X_{ij}\beta_{ir}=X_{ij}(b+\ \ e_{ir})$, where b and are estimated mean and deviation parameters and e_{ir} a standard normal deviate for rth replication for individual i. Estimation of parameters is carried out by maximizing the simulated likelihood function, determined by the appropriate simulated response probability formula, in much the same way as for fixed logit models. The simulated log-likelihood function for the random parameter conditional logit model (2.13) is

$$SL = \sum_{i}^{N} \ln SP_{ij} \tag{2.18}$$

It is worth noting that RPL models are also flexible in approximating the response probabilities generated by other than Type 1 GEV distributed random terms in the RUM, such as normally distributed random terms. McFadden and Train (2000) show that any discrete choice model derived from random utility maximization can be approximated arbitrarily close by random parameter multinomial logit model.

Pooling Data

The scale of estimated parameters in all the choice models based on the RUM is related to the magnitude of the random component in the RU model. The scale factor μ relates the estimates with the random component, being inversely related to the variance of the random component in the RUM. Using a single source of data, μ is typically set equal to one since it cannot be identified. The estimated vector of coefficients β is therefore confounded with constant μ. This in turn makes absolute values of the parameter estimates incomparable between different data sets; only the ratios of coefficients are comparable across different sources of data (Swait and Louviere 1993).

Consider n separate sources of stated preference data, such as survey data using CV and CR. Normalizing scale factors equal to one in estimation of separate data sources, each data $q=1,...,n$ provides us with parameter estimates β_q. Denoting the scale parameters of different data sources with μ_q, n vectors $\mu_q\beta_q$ of parameter estimates results. Pooling n sources of data, it is possible to identify $n-1$ scale parameters for different data sources. Fixing one scale factor, say $\mu_1=1$, the rest $n-1$ estimated scale parameters are inverse variance ratios relative to the reference data source (Hensher et al. 1999).

Denote the vector of CV and CR estimates by $\mu_{CV}\beta_{CV}$ and $\mu_{CR}\beta_{CR}$. Pooling the CV and CR models, fixing $\mu_{CV}=1$, and estimating μ_{CR}, then accounts for possible differences in the variance of random terms between the CV and CR data. To test for the parameter invariance between the CV and CR data, models with and without restriction $\beta_{CV}=\beta_{CR}$ need to be estimated. Likelihood ratio tests can then be applied to accept/reject the imposed parameter restriction. If the null hypothesis cannot be rejected, the data generation processes can be considered generated by the same taste parameters but still have variance differences. Restricting both $\beta_{CV}=\beta_{CR}$ and $\mu_{CR}=1$ provides an even stricter test of data invariance, testing for both parameter and random component invariance. If not rejected, the two data sets can be considered similar and absolute parameter estimates comparable across the source of data.

Data

Data were collected using a mail survey, sent out in spring 1999 to a sample of 1740 Finns between 18-75 years of age. The sample was randomly drawn from the official census register, and divided into two random sub-samples of 840 and 900 respondents. The first sub-sample received a double bounded CV questionnaire and the second sub-sample a CR questionnaire.

Questionnaires started with questions about respondents attitudes on how important the different aspects of forests, such as their economic importance and different uses (timber production, recreation, nature conservation etc) should be in formulating forest policy. Next, respondents were asked to state how important issues such as public healthcare, education, employment, economic growth, nature conservation and equal income distribution should be in formulating public policies in general. Thereafter, respondents were asked still a number of attitude questions about forest conservation, landowners' and public's responsibilities in conservation, and the acceptability of forcing landowners to protect forests by regulatory approaches. The next section of the questionnaire included the valuation questions, described in more detail later. The questionnaire concluded with questions on the respondent's socioeconomic background.

While designing the survey, questionnaire versions went through several rounds of modifications and reviews by accustomed SP-practitioners, as well as other economists, foresters and ecologists with expertise in survey methods and/or biodiversity conservation. After hearing their comments, questionnaires were tested by personal interviews and a pilot survey (n=100), and modified based on the results. The final survey was mailed out in May 1999. A week after the first mailing, everyone in the sample was sent a reminder card. Two more rounds of reminders with a complete questionnaire were sent to non-respondents in June-July. The CV and CR surveys resulted in 48.9 % and 50 % response rates, respectively. After censoring for all the missing answers to valuation questions, 376 CV and 391 CR responses were available for further examination.

WTP is measured for three hypothetical conservation programs: Increasing conservation from the current 120,000 hectares to (1) 275,000 hectares, (2) 550,000 hectares and (3) 825,000 hectares. The new alternatives correspond to protection of 25 %, 50 % and 75 % of all the available biodiversity hotspots. In designing the survey, special attention was paid to formulating conservation policy scenarios so that they were policy relevant, credible and easy to understand. A one page easy to read section in the questionnaire explained different conservation programs and their details.

The CR survey described to respondents the status quo and all three hypothetical programs of setting aside additional 155,000, 430,000 and 705,000 hectares of hotspots for 30-years[1]. Table 1 describes how the conservation programs were summarized in the CR questionnaire. Using a scale from 0 to 10, each respondent was asked to rate the four programs. Note here that the respondents are not asked to hypothetically buy any forest areas; they are simply asked to express their preferences regarding different conservation programs that would each result in different conservation levels and costs to their households. The three hypothetical programs were assigned costs using the same variation across the respondents and the conservation programs as in the CV survey, described in more detail later.

Table 1. Four possible conservation projects presented for the CR respondents

Conservation Project	Total area under conservation	Proportion of conserved of all the Finnish forests
1. Current regulation, no new conservation	120,000 hectares	0.6 percent
2. Increasing conservation to cover one fourth (25%) of the biodiversity hotspots	275,000 hectares	1.5 percent
3. Increasing conservation to cover half (50%) of the biodiversity hotspots	550,000 hectares	3 percent
4. Increasing conservation to cover three fourths (75%) of the biodiversity hotspots	875,000 hectares	4.5 percent

The respondents of the CV questionnaires were divided into two groups. The first group was asked to state their WTP for the first two policy alternatives, i.e. 275,000 and 550,000 hectares as described in Table 1, and the second group for the 550,000 and 825,000 hectare alternatives. Each respondent was asked two separate CV questions, and responses for 50 % conservation were therefore collected by both the first and the second WTP questions, depending on the respondent's sub-sample. The CV method was applied using a double bounded format. The bid vector in the CV survey consisted of first bids between US$ 4-500, and the follow-up bids between US$ 2-800, with seven different starting bids. The same bid amounts appeared a first and second bids for different respondents, and the bids for different levels of conservation were randomly chosen from across the full vector of bids.

The final survey consisted of 29 different questionnaire versions; 14 were CV surveys and 15 CR surveys. In both types of surveys, WTP was measured as an increase in the annual tax

[1] The length of protection is determined as 30 years because the current policy programs for voluntary conservation are based on 30 year protection.

burden of the household. Except for the valuation question, the CV and CR questionnaires were similar. The set up for the data collection is such that only the choice task in the valuation question vary between the CV and CR respondents.

Results

The next sections report and discuss the results of fixed and random parameter logit models separately for the CV and CR data. After separate estimation of the CV and CR models, the two data sets are combined and invariance between the CV and CR data tested. All results are based on maximum likelihood estimation of the models described earlier in this paper. They were programmed and estimated in GAUSS.

Estimated models use a dummy specification for the conservation programs. In other words, conservation was modeled as three different conservation programs, not as a continuous variable of conserved hectares under each policy alternative. This results in several benefits: First, the specification is very flexible and does not restrict the value function for conservation to follow any certain functional form, letting it take practically any form instead. Second, the WTP for different conservation programs can now be calculated simply as ratios of the estimated parameters. Third, separate dummies can be used in measuring the variances of taste parameters for different extents of conservation. The observed part of RUM is estimated as

$$V_{ij} = \beta_{BID} BID_{ij} + \beta_{D25} D25_{ij} + \beta_{D50} D50_{ij} + \beta_{D75} D75_{ij} \tag{2.19}$$

where BID is the annual cost to the respondent's household from implementing policy alternative j, and D25, D50 and D75 dummy variables that indicate the extent (25 %, 50 % and 75 %, respectively) of conservation in policy alternative j. The specification of the V_{ij} stays the same throughout the reported models.

Fixed Parameter Logit Models

Contingent Valuation

The CV data with 376 observations was first censored for missing responses. The remaining 306 observations with complete double bounded responses to the both WTP questions were employed in the estimation.

As mentioned, the CV sample was divided into two groups, with 50 % conservation program in either the first or the second WTP question. Differences between the responses from these two groups were studied by first estimating separate parameters for 50 % conservation for the two groups. The parameters were then restricted equal, and a constrained model with a single parameter for the 50 % alternative was estimated. Based on the unconstrained and constrained model results, a likelihood ratio test was formulated to test for the similarity of responses to 50 % conservation program between the two groups.

Table 2. Model estimates for contingent valuation data

Model:	Fixed logit CV	
Estimate (|*t*-statistic|)	Unconstrained	Constrained (*D50_1*=*D50_2*)
BID	-0.3527	-0.3529
	(13.055)	(13.065)
D25	1.5996	1.6002
	(6.509)	(6.511)
D50_1	1.3864	
	(5.287)	
D50_2	1.3066	
	(5.225)	
D50		1.3445
		(7.048)
D75	0.9640	0.9645
	(3.610)	(3.612)
Mean LL	-380.771	-380.799
LL at 0 [a]	-446.775	-446.775
Pseudo R^2	0.148	0.148

[a] Pseudo R^2 is calculated as 1-LLU/LLR, where LLU and LLR are the log-likelihood values for the estimated model and model with only a constant.

Note: Number of observations 306. Table cells for structurally non-identified parameters are shaded in this and the following tables.

Table 2 reports the results of the fixed logit models for CV data. Dependent variable in the models is the probability of Yes-answer to the dichotomous choice WTP question. The estimated parameters are defined as follows: *BID*[2] is the household's annual cost from implementation of the suggested conservation program; *D25* and *D75* are dummies that indicate the 25 % and 75 % levels of conservation in the WTP question; *D50_1* stands for the 50 % conservation program as the first WTP question; *D50_2* stands for the 50 % conservation program as the second WTP question; *D50* is a dummy that pools 50 % conservation programs by restricting *D50_1* = *D50_2*.

Both double bounded models result in highly significance parameter estimates. Estimates for the *D25* are significantly greater than zero, suggesting that the 25 % conservation program is preferred to status quo. The estimates for *D50* and *D75* are positive and greater than zero. Therefore, they are also preferred to status quo. However, the estimate for *D50* is systematically lower than the estimate for *D25*, and the estimate of *D75* in turn lower than the one for *D50*. This suggests a conservation policy preference order (25 % > 50 % > 75 % > status quo), thereby an increasing WTP from status quo to 25 % conservation, an possibly a negative marginal WTP for the higher levels of conservation.

No statistically significant differences between the responses to the 50 % conservation program are found based on the values of the maximized log-likelihood functions. The likelihood

[2] Note that in all the results reported in this chapter, variable BID is divided by 100 to facilitate estimation.

ratio testing[3] for the constraint *D50_1=D50_2* results in LR test static value 0.05. The test therefore rejects the null hypothesis of different parameter estimates for the *D50_1* and *D50_2*, and suggest accepting the constrained model. Based on strong rejection of the null hypotheses, the constrained models are used in the further analysis. In practical terms, this means pooling the responses to the 50 % conservation program and estimating *D50* as a single parameter.

Contingent Rating/Ranking

All the models for the CR data are based on rankings that were obtained by transforming respondents' ratings for policy alternatives into a preference ordering, assuming that preferred alternatives were rated higher than the less preferred ones. Rankings utilize only ordered information on preferences. Respondents with ratings sequences (3,2,1,0) and (10,9,3,1) are therefore considered similar responses with the same preference ordering A>B>C>D. In building the ranking data, observations with ties or missing ratings were censored, leaving a total of 270 observation left for the estimation. The results are therefore based on data with full and unique rankings of all four policy alternatives.

The specification of the CR models is the same as for the CV data. The following models were estimated: (1) a conditional logit model for the highest ranked alternative out of all the alternatives i.e. conditional logit model for preferred choice, (2) rank-ordered logit models for 2 and 3 ranks, both as rank-homoscedastic (ROL) and rank-ordered heteroscedastic models (ROHL). The models for 2 ranks explain the first two preferred alternatives; the model for 3 ranks a full ranking of the four alternatives.

Several rank-ordered logit models were estimated in order to examine the consistency of rankings. Information on more than only the preferred alternative is valuable, but beneficial only if the rankings are consistent and generated by the same parameters (e.g. Layton 2000). It is known that the variance of stochastic term in RU model tends to change (typically increase) as the ranking continues. This has been suggested to result from the respondents ranking the preferred alternatives with more care than the less preferred alternatives, causing data on 2nd ranks to be more noisy than data on 1st rank, data on 3rd rank to be more noisy than data on 2nd rank, and so forth.

The changing variance of random term between the ranks violates the i.i.d. assumption of rank-ordered logit model. Inconsistent rankings reveal violations of the assumption. Models with violations should be rejected and models for fewer ranks used instead. If variance of the random term changes sufficiently systematically and similarly over the rankings, the problems caused by the inconsistency of rankings could be solved by employing a Hausman–Ruud rank-ordered heteroscedastic logit model. Testing for the consistency of for instance the first and the second ranks can be carried out by estimating separate models for the first and the second ranks. These models result in two maximized log-likelihood values, denoted by LL_1 and LL_2. Constrained model is then estimated as a rank-ordered logit model for two ranks, resulting in a maximized log-likelihood value LL_R. A LR-test statistic with degrees of freedom equal to the number of constrained parameters is calculated as $-2*(LL_R\text{-}LL_1+LL_2)$. If the test statistic is insignificant, the ranks can be pooled and a rank-ordered logit model for two ranks used. If the LR-test fails to accept the pooling of the ranks, a Hausman-Ruud style rank-heteroscedastic logit model can be estimated and similar LR-test procedure carried out using it; this time with one less degrees of freedom because an additional parameter is estimated.

[3] LR-test statistic is calculated as –2(LLR-LLU), where LLR and LLU are the values of maximized log- likelihood function for constrained and unconstrained models, respectively (e.g. Amemiya 1983).

Table 3. Fixed logit models for contingent rating/ranking data

Model:	1 Rank	2 Ranks		3 Ranks	
Estimate (|t-statistic|)		ROL	ROHL	ROL	ROHL[a]
BID	-0.0888	-0.0779	-0.0484	-0.0592	-0.00099
	(-3.153)	(-3.904)	(-2.275)	(-3.635)	(-1.340)
D25	0.3537	0.7561	0.6217	1.1588	0.4519
	(1.328)	(3.783)	(3.321)	(6.282)	(2.083)
D50	0.2760	0.5559	0.4358	0.9959	0.3445
	(0.818)	(2.242)	(2.226)	(4.522)	(2.000)
D75	0.2055	0.1155	0.0672	0.1473	0.0480
	(0.507)	(0.368)	(0.297)	(0.531)	(0.578)
$\mu_{Rank2,3}$			1.8146		4.6160
			(2.765)		(1.980)
LL	-150.709	-261.938	-260.515	-319.496	-307.919
LL at zero	-162.556	-291.379	-291.379	-372.657	-372.657
Pseudo R^2	0.073	0.101	0.106	0.143	0.174

Note: Number of observations 270.

[a] Hausman-Ruud rank-heteroskedastic model for three ranks is estimated with common scale factor for the second and third rank. Estimating separate scale factors for second and third ranks is structurally possible but they could not be identified.

Table 3 reports the results of fixed parameter CR models. The signs, relative magnitudes and statistical significance of parameter estimates are rather well in line with the estimates for the CV models. The parameter estimates for *D25* are always greater than the estimates for *D50* and *D75*. The CR models also suggest uniformly that in the average, the 25 % conservation policy is preferred over the other policy alternatives, including status quo. Moreover, the relation between the *D50* and *D75* is similar as in the CV results, with the *D75* estimates having the smallest absolute but still positive estimates.

The insignificance of *D75* estimates in all the CR models is distinctive in comparison with the CV results. This could be related to the questionnaire design; the 75 % conservation alternative was always presented as the last policy alternative, possibly resulting in less careful rating than for the first, second and third policy alternatives. Another possibility is that preferences regarding the 75 % conservation policy are simply so heterogeneous that the identification of parameters is troublesome. Conditional logit model for the first rank results in statistically insignificant estimates for all the parameters except for the *BID*. In addition to the noise in the data, this can be related to relatively small number of observations.

Examining next the homoscedastic rank ordered logit models for 2 and 3 ranks, it is noted that all the statistically significant parameter estimates are greater in absolute magnitude than their counterparts in the first rank model. This is logical; by utilizing more information on the individual preferences, relative magnitude of the stochastic term in the RUM is decreased and substituted by higher parameter estimates and therefore higher proportion of observed variation. Further, all the parameter estimates except *D75* are now statistically significant both in the 2 and 3 rank models. The pseudo R^2 measures are still relatively low, although higher than for the first rank model.

Although exploiting information on more than only the first rank first seems to provide improvements compared to first rank model, the consistency of rankings is necessary to be examined before accepting the rank ordered models. The LR-test results for the consistency of rankings are reported in Table 4. The tests suggest that both the homoscedastic logit models (ROL) for 2 ranks can be accepted and the first two ranks pooled. Evidence is not particularly strong but the LR-test statistics are insignificant at the 1 % level. Consistency of three ranks is rejected with strong statistical evidence.

Table 4. Hypothesis tests on pooling different ranks

	2 Ranks		3 Ranks	
	ROL	ROHL	ROL	ROHL
LL sum of separate ranks	-255.655	-255.655	-275.441	-275.441
LL with pooled ranks	-261.938	-260.516	-319.496	-307.919
^{2} (nonpooled vs. pooled ranks)[a]	12.57	9.72	88.11	64.96
Pooling of ranks	Accepted	Accepted	Rejected	Rejected

[a] At 1 % significance level, critical values for ^{2} test statistic for 3 and 4 degrees of freedom are 11.34 and 13.28, respectively

Table 3 also reports the results of Hausman-Ruud style rank-ordered heteroscedastic logit models (ROHL). They are obtained by fixing the scale factor for the first rank and estimating relative scale factors for the second and third ranks. The estimates of scale factors are significantly greater than one in the both ROHL models. Being inversely related to the magnitude of random term in the RU model, the magnitude of random term seems to decrease as ranks are added. The consistency of three ranks is strongly rejected with ROHL model for 3 ranks, suggesting that the heterogeneity of responses is not sufficiently systematically related to the ranks for the model to be consistent. Based on these findings, the fixed parameter logit model for three ranks is rejected.

Random Parameter Logit Models

At least two important aspects of modeling strategy must be considered carefully before estimating random parameter models. First, parameters with and without heterogeneity must be selected, preferably by using some prior information. Second, distributions for random coefficients must be specified, typically based on theoretical considerations.

In choosing the heterogeneous and homogeneous coefficients, it is of course possible to allow all the parameters to vary in the population. This strategy relies on the results of the flexibly specified model itself to suggest which parameters are heterogeneous and which not. Following this approach should be supported by prior expectations about parameter heterogeneity. Not only the methodological but also time considerations suggest that. Even with the recent improvements in the computing power, estimating random parameter models can be very time consuming[4]. For instance, the speed of the Train's simulator mainly depends on the number of estimated heterogeneous parameters, and including irrelevant heterogeneous parameters should be avoided. Further, identification is always an issue in estimating random parameters, especially with non-negative/positive coefficients. As it only gets harder with increasing number of random parameters, careful selection of the heterogeneous parameters is recommended.

Random parameters are typically estimated as normally distributed parameters. The normally distributed parameters β_n can get both negative and positive values. They are estimated as $\beta_n=(b_n+\ \ _n e)$, where b and   are the estimated mean and deviation parameters of the β_n, and e a standard normal deviate (Train 1998).

Both the theory and common sense often suggest that some random coefficients are non-negatively/positively distributed. In this case, the *BID* coefficient is assumed to be non-positively distributed. For non-positive values of the *BID*, increasing the costs of a policy alternative always decreases its probability to become chosen.

Train (1998) suggests that the non-positive/negative random parameters can be estimated as log-normally distributed, and provides a method for incorporating them into his simulator. Each log-normal β_k can be estimated by expressing them as $\beta_k=exp(b_k+\ \ _k e)$, where b and   are estimated mean and deviation parameters of $\ln(\beta_k)$, and e an independent standard normal deviate. Log-normal non-positive parameters are estimated with entering the appropriate exogenous variables as their negative. For the disadvantage of the log-normally distributed random parameters, they are often very hard to estimate and identify (e.g. McFadden and Train 2000).

Alternatively, Layton (2001) proposes employing distributions determined by a single parameter in estimating the non-negative/positive random parameters. While the RP models typically estimate the mean and variance of the RP distribution, the one-parameter distributions (such as Rayleigh-distribution) allow finding all the moments of the RP distribution by estimating just a single parameter. A non-negative parameter *BID* with Rayleigh distribution has a cumulative density function $F(BID)=1-exp[-BID^2/(2b^2)]$ and a probability density function $f(x)=(BID/b^2)exp[-BID^2/(2b^2)]$, where b is the scale parameter fully determining the shape of the distribution. Using the inverse transformation method, the Rayleigh distributed *BID* can be obtained as $BID = (-2b^2 ln(1-u))^{1/2}$, where u is a random uniform deviate and b the estimated parameter. The mean, variance, median and mode of the Rayleigh-distributed *BID* are $b(\Pi/2)^{1/2}$, $(2-\Pi/2)b^2$, $b(log4)$, and b, respectively (Layton 2001).

In the case at hand, both the *BID* and policy alternative dummies were modeled as random parameters. The previous RP applications have typically modeled either the BID or alternative specific dummies as random parameters, not both. With these data, the heterogeneity of preferences for policy alternatives with extensive conservation levels was possible to appear, and random parameter formulation of policy alternative dummies is therefore of specific interest. On the other hand, previous studies suggest that the heterogeneity of preferences is often related to the *BID* coefficient. It was therefore also estimated as a random parameter. Since it essentially represents the negative of the marginal utility of income, it was estimated as a non-positively distributed parameter. Despite continuous and substantial efforts, all the necessary models for

---

[4] For instance, many of the models reported in this chapter took more than half a day to converge with an up-to-date processor. In addition, several runs with different starting values are often needed to find the global maximum, since the log-likelihood functions of the RP models, unlike their fixed parameter counterparts, are not necessarily globally convex and can therefore have multiple local maximum (McFadden & Train 2000).

this study were impossible to be estimated with the log-normal BID[5]. The BID was therefore expressed as a Rayleigh-distributed random parameter BID – RAYLEIGH[6]. With this specification, convergence was reached much easier and estimation considerably faster.

Table 5 reports the random parameter model results for both the CV and CR data. The pseudo-R^2 of the CR models for 2 and 3 ranks is increased from 0.106 and 0.143 of the fixed logit models to 0.231 and 0.296 of the random parameter models. The explanatory power of the models therefore more than doubled as a result of incorporating unobserved preference heterogeneity.

The CR model for the first rank does not converge with random dummies and its results cannot be reported; a model with fixed dummies and random bid parameter is reported instead. The CR model for first rank performs poorly also in terms of explanatory power. In the CV model, policy alternative dummies are not significant. Explanatory power of the random parameter CV models is substantially higher than for the fixed models; pseudo-R^2 is increased from 0.148 to 0.267.

Table 2.4.4. Random coefficient logit models for CV and CR data

Model: Estimate (*t*-statistic)	CV[a]	1 Rank[b]	2 Ranks	3 Ranks
Bid-Rayleigh	1.1309 (3.014)	0.0760 (2.606)	0.1257 (2.396)	0.1578 (3.116)
D25-mean	3.6657 (2.905)	0.3879 (1.393)	1.1133 (3.506)	1.7359 (5.742)
D50-mean	3.5006 (2.882)	0.2965 (0.828)	0.8660 (1.422)	1.4423 (2.313
D75-mean	3.1339 (2.779)	0.1961 (0.458)	-1.8904 (1.332)	-0.2733 (0.281)
D25-dev	-1.6448 (1.304)		0.0122 (0.022)	0.0349 (0.043)
D50-dev	1.5736 (1.290)		3.3713 (4.650)	3.6318 (5.251)
D75-dev	1.7742 (1.417)		6.2289 (4.130)	5.7319 (5.754)
LL	-333.18	-151.11	-224.04	-262.19
LL at zero	-446.78	-162.56	-291.38	-372.66
Pseudo R^2	0.254	0.0704	0.231	0.296

Note: Number of CV and CR observations is 306, and 270, respectively. Simulator with 200 draws were used in estimating the models.

[5] The estimation of CV models with log-normally distributed *BID* was generally more successful, although excessively time consuming. The estimation of CR models did not succeed. Especially with the models for 2 and 3 ranks, iteration first proceeded seemingly fine for some 80 iterations but then failed to reach the convergence. This could be due to multiple local maxima of the log-likelihood function. However, even the specifications with the *BID* as a sole random parameter did not lead to the convergence.

[6] Dave Layton is acknowledged for suggesting this.

[a] Model results for the CV with fixed and random normal dummies are similar and not statistically significantly different. Specification with random normal dummies is chosen for CV data to make results directly comparable to CR, necessary for the data pooling purposes, reported in the next section of the paper.
b Deviations cannot be identified in the model for the first rank, and the results for it are therefore based on model that restricts them as zero.

The estimates of the deviations for the *D50* and *D75* (*D50–dev*, *D75-dev*) are strikingly large and significant, suggesting that the preference heterogeneity for policy alternatives is considerable and should be taken into account while modeling these data. Parameter heterogeneity is also one possible explanation for insignificance of the mean estimates of the *D75*. With highly variable preferences for 75 % policy alternative, estimation cannot provide a significant estimate of the location parameter for distribution of *D75*. Note that the same phenomena was observed in fixed parameter logit models, where significance of *D75* estimate was lower than in the RP models.

The random parameter formulation provided estimation of both CV and CR data with significant improvements. They are next applied together with fixed parameter logit models to combined CV and CR data that estimate pooled models for the CV and CR data.

Pooled models for CV and CR data

Pooled models were estimated using a combined CV and CR data. The estimation can be implemented in several ways; the main concern is to make sure that appropriate likelihood functions are applied to each of the respondents. An indicator variable for the CR data can facilitate estimation. Defining I_{iCR} with a value 1 for CR respondents and a value 0 for CV respondents, the pooled log-likelihood function for individual i is determined as $LL_i = I_{iCR} * P_{iCR} + (1 - I_{iCR}) * Pi_{iCV}$, where P_{iCR} and P_{iCV} are the appropriate CV and CR response probabilities of the model. Similarly as with the unpooled models, the pooled total log-likelihood function is a sum of the individual likelihoods over the whole sample[7].

Table 6 reports the models results for the combined CV and CR data. A variety of pooled models were estimated to examine the effects of modeling choices on accepting/rejecting pooling the data. The same specification as in unpooled models was applied for pooled models. The unpooled counterparts of all the pooled models can be found from the previous sections of this paper. LR-tests are used for accepting/rejecting the pooling hypothesis; the respective LR-test statistics are reported in the second last row of the Table 6. The test statistics follow ^{2} distribution with degrees of freedom equal to difference in number of estimated parameters between pooled and unpooled models. Estimating a scale parameter in pooled model versions, degrees of freedom for fixed and random parameter logit models with all parameters random equal to 3 and 6, respectively. The respective critical values are 11.34 and 18.48. The LR-test statistic for random parameter models with random BID and fixed policy alternative dummies also has 3 degrees of freedom. If the LR-test statistic is smaller than the critical value, the pooling of data cannot be rejected.

All the results include an estimate of parameter μ_{CR}. It is a scale factor for the CR data, accounting for possible differences in the variance of random term of the RUM between the CV and CR data. As noted before, only the parameter relations are comparable between the different sources of data. Estimating a scale factor allows for direct comparisons of the estimates.

[7] Considerable time savings, especially in estimating random parameter models, can be obtained by structuring the program so that unnecessary calculations are avoided in calculating the log-likelihood function. Calculations of CR response probabilities are uncessary for CV respondents, vice versa.

Logically, if no differences in the random term variance exist between the CV and CR data, the estimate of μ_{CR} is not statistically different from one. Since the scale factor is inversely related to the variance of the random component of the RU-model, an estimate $\mu_{CR} < 1$ suggests that the CR data is noisier than the CV data, and $\mu_{CR} > 1$ the opposite.

Models "CV & 1 Rank" pool the CV model with a CR model for first rank, using both the fixed (FL) and the random coefficient (RCL) formulation. The estimates of the μ_{CR} are statistically significant and smaller than one in both models, suggesting that CR data is noisier. The random parameter model provides a significantly higher explanatory power than the fixed parameter counterpart. Both LR tests statistics for pooling hypothesis are insignificant. Therefore, both fixed and random parameter models provide support for accepting pooling of the CV and CR data.

Table 6. Logit Models for Pooled CV and CR Data (n=576)

Model:	CV & 1 Rank		CV & 2 Rank		CV & 3 Rank	
Estimate (*t*-statistic)	FL	RCL[a]	FL	RCL	FL	RCL
Bid – Fixed	-0.3532 (13.077)		-0.3484 (12.963)		-0.3402 (12.677)	
Bid-Rayleigh		0.8516 (6.351)		1.1286 (2.994)		0.8380 (3.123)
D25 – *mean*	1.5928 (6.669)	2.7152 (7.684)	1.7335 (7.436)	3.7174 (2.876)	1.8513 (7.961)	3.1829 (3.094)
D50 – *mean*	1.3410 (7.120)	2.5373 (8.015)	1.3727 (7.398)	3.5140 (2.853)	1.4679 (7.985)	3.7130 (3.409)
D75 – *mean*	0.9672 (3.697)	2.2661 (5.545)	0.8606 (3.373)	3.1244 (2.757)	0.7131 (2.848)	3.3504 (3.128)
D25 – *dev*				-1.6906 (1.326)		-0.8670 (0.738)
D50 – *dev*				1.6001 (1.297)		2.6625 (2.564)
D75 – *dev*				1.7806 (1.414)		4.2451 (2.925)
μ_{CR}	0.2549 (3.781)	0.0905 (2.902)	0.2617 (5.395)	0.0688 (2.218)	0.2490 (5.996)	0.9103 (3.747)
LL pooled	531.54	485.43	-645.89	613.33	-719.34	-633.91
LL unpooled	531.50	484.29	-642.74	557.22	-700.30	-595.37
LL at zero	609.33	609.33	738.15	738.15	819.432	819.432
LR-test of pooling	0.063	2.28	6.3	112.22	38.08	77.08
Pseudo R^2	0.128	0.203	0.125	0.169	0.122	0.226

Note: Number of CV and CR observations is 306, and 270, respectively. Simulator with 200 draws were used in estimating the models.

[a] Since the CR model for 1 rank with random dummies could not be estimated, a pooled model for CV and 1 Rank CR is also based on model with fixed dummies and Rayleigh bid

"CV & 2 Rank" models pool the fixed and random parameter models for the CV model and the CR model for 2 ranks. Similarly as the previous pooled models, these models result in highly significant estimates with expected signs. Comparing the unpooled and unpooled fixed parameter models results in LR-test statistic 6.3 and accepting the pooling hypothesis. However, the pooled random parameter model strongly rejects pooling of the CV and CR data. Despite rejection of pooling, the random parameter model results in substantially higher pseudo-R^2 than the fixed parameter model.

Models "CV & 3 Rank" pool fixed and random parameter model for 3 ranks and the CV data. Results from these models are similar with the pooled models with CR model for two ranks. However, pooling of CV and CR data is now strongly rejected for both the fixed and the random parameter models. Rejecting the pooling with the random parameter models, together with the results of pooled models with CR model for 2 ranks, suggest that differences in parameter heterogeneity between the two data are a possible source of their inequality.

The overall similarity of the parameter estimates across all the models is distinctive for the pooled model results. This is likely to have resulted from more precise estimates of the CV models data "dominating" the identification of estimates for the pooled models. With lower variability of the random term of RUM, as suggested uniformly smaller than one estimates of the μ_{CR}, the CV data plays a relatively more important role in identifying the estimates, even with almost equal number of CV and CR observations.

Although not reported in Table 6, the pooling hypothesis was further tested with the restriction $\mu_{CR}=1$ that imposes equal variances of the RUM random terms for the CV and CV models. Using the completely pooled model, the pooling of the CV and CR data is rejected using all the models. The pooled fixed model for CV and first rank CR data with a CR scale factor provides the strongest support for accepting pooling, and is therefore the likeliest candidate to provide support for the complete invariance hypothesis. Estimating the pooled model for it results in a value 45.5 for the LR test statistic, strongly rejecting the complete pooling of the CV and CR data. Other models are less likely to provide support for complete pooling hypothesis, and complete invariance of CV and CR data is therefore uniformly rejected.

Willingness to Pay Estimates

Including policy implementation costs and policy specific dummies in estimated models allows capturing WTP for different policy scenarios indirectly from the results. The mean WTP for policy alternative x_j is calculated as (e.g. Goett et al. 2000)

$$\frac{\partial U / \partial x_j}{\partial U / \partial y} \tag{2.21}$$

The $\partial U / \partial x_j$ is measured by the alternative specific dummies *D25*, *D50* and *D75*, and the $\partial U / \partial y$ by the *BID* estimate. The mean WTP estimates for the fixed logit estimates are calculated as *D25/BID*, *D50/BID* and *D75/BID*. The means of normally distributed random parameters equal their estimates and calculation of WTP is similar as with the fixed parameter models. The means for Rayleigh distributed *BID* must be calculated as described earlier in this paper.

Table 7 reports the mean WTP estimates for the estimated models. The results are divided into fixed and random parameter models for unpooled and pooled data. The estimates for the CR 1 rank model were not statistically significant, and the WTP estimates are therefore not presented for them. In the previous analyses, the following models were clearly rejected because of inconsistency of rankings or failure to accept data pooling hypothesis: (a) Fixed logit model "CR-3 rank data", (b) Pooled fixed logit model "CV & CR – 3 rank", (c) Pooled random parameter model "CV & CR-2 rank" and (d) Pooled random parameter model "CV & CR-3 rank". Their results are expressed in *(italics)*.

Table 2.4.6. Willingness to pay estimates (US$) [a, b,c]

Policy alternative	Fixed Parameter Logit			Random Parameter Logit		
Model type	25 %	50%	75 %	25 %	50 %	75 %
CV	**73**	**61**	**44**	**42**	**40**	**36**
CR – 2 Ranks	**156**	**115**	*insign.*	**95**	**92**	**53**
CR – 3 Ranks	*(315)*	*(271)*	*insign.*	**142**	**118**	*insign.*
Pooled models						
CV + CR – 1 Rank	**73**	**61**	**44**	**41**	**38**	**34**
CV + CR – 2 Rank	**80**	**64**	**40**	*(42)*	*(40)*	*(34)*
CV + CR – 3 Rank	*(88)*	*(70)*	*(34)*	*(49)*	*(57)*	*(51)*

[a] Insignificant estimates are expressed with *insig*.
[b] US$=FIM 6.2.
[c] WTP estimates for the rejected models are reported in *italics* in brackets.

Results for the 75 % conservation alternative are very similar across all the models providing it with a significant estimate. The WTP for it varies between US$ 36 and 53. Estimates for the mean WTP for 50 % alternative cover a considerably wider range from US$ 38 and 118, similarly as the mean WTP for 25 % conservation alternative with values between US$ 41 and 156 .

All the significant and accepted WTP models result consistently higher WTP for the 25 % alternative than for the 50 % alternative, and in turn for higher WTP for the 50 % alternative than for the 75 % alternative.

Figure 1 graphs the WTP estimates. Each set of bars graphs estimates for one set of models. The first set "CV" stands for all the CV models; its first three bars in the left are the fixed model results for 25%, 50%, and 75 % programs. Three bars on the right are the estimates based on the random parameter models. The "CR1" has no bars since none of the estimated WTP figures were statistically significant. The "CR2" and "CR3" graph the WTP estimates for the CR models for 2 and 3 ranks. Note that the fixed 3 rank model for the CR data was rejected in testing for the consistency of rankings. "CV+CR1", "CV+CR2", and "CV+CR3" graph the WTP estimates of the pooled models. Note also that the random parameter model for the pooled CV and CR data on 2 ranks was rejected. Further, both the fixed and random parameter models for pooled CV and 3-rank CR data were rejected.

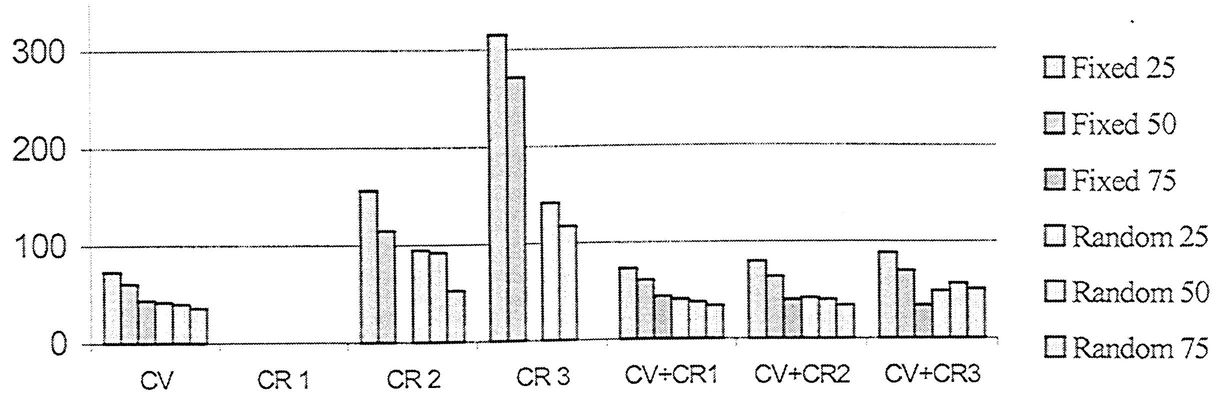

Figure 1. Estimates of the WTP for fixed and random parameter models for the CV and CR data

The results consists of variety of WTP estimates based on different models. The question becomes which results are preferred and chosen for further purposes, such as policy evaluation. Clearly, all the rejected models can be screened out first. The rest of the models can be evaluated by using success in pooling the CV and CR data as criteria. The models that successfully pool the CV and CR data are not fully contingent on a single survey method, and can therefore considered most general.

Using the pooling success criteria leaves us with a choice from the following three models: (1) Pooled fixed parameter model for the CV data and the CR data on 1 rank, (2) Pooled random parameter model for the CV data and the CR data on 1 rank, and (3) Pooled fixed parameter model for the CV data and the CR data on 2 ranks. The fixed model with 1 rank CR data can be screened out as more restrictive than the random parameter model with 1 rank, and as less detailed than the fixed parameter model with 2 ranks. Both remaining models have certain advantages. The random parameter model with 1 rank is less restrictive than the fixed model with 2 ranks, but the fixed parameter model with 2 ranks utilizes the data in more detail than the model with 1 rank. However, the random parameter models statistically outperformed the fixed models throughout the analysis, and the pooled random parameter model for the CV data and the CR data on 1 rank is therefore chosen as the preferred approach for modeling these data.

Discussion

This study examined different econometric modeling strategies for CV and CR survey data. Both conventional fixed logit models and recently developed random parameter logit models were reviewed and applied to the data at hand. The results provided another confirmation that considerable care must be practiced in applying fixed parameter logit models. Especially the fixed parameter models for the CR data on full rankings of four alternatives violated assumptions of the conditional logit model.

Applying data pooling techniques in testing for equality between CV and CR was another objective of this study. Successful pooling of the CV and CR data required estimation of scale factors for separate data sources, which then made the parameter estimates comparable between separate sources of data. Without the scale parameter, pooling was uniformly rejected. With scale factors in the estimated model, pooling of the CV and CR data could not be uniformly rejected or accepted. The more detailed models for the CV and CR responses are likely to reject the pooling hypothesis. Less detailed models such as the CV and conditional choice logit models generally did not provide sufficient statistical evidence for rejecting the pooling hypothesis.

The random parameter models do not seem to provide a miracle in terms of solving differences between the CV and CR data, although there is some evidence that fixed logit models could exaggerate the differences between the WTP results for the CV and CR data. Especially the results of the rejected fixed logit models for the three ranks CR data suggest that. This could be related to the model specification, since the WTP results are typically sensitive to the chosen specification. Examining different specifications is one of the objectives for future research.

Data pooling techniques provide a powerful approach to tests for invariance between the different sources of data. The analysis of pooled data that has been presented in this paper highlights only some of the possible uses of data pooling methods in examining SP survey methods. For instance, sources of differences between different SP sources can be further examined using the same framework.

Issue of negative marginal WTP for conservation after reaching a certain conservation levels, is of course very interesting and worth some further investigation. Given the opposition of some of the public for increasing the forest conservation in Finland, the result is not necessarily a surprising finding. Further, a major conservation program called Natura 2000 was under preparation around the time of the survey. The Natura 2000 is part of the European network of conservation areas, which in turn is the core policy for nature conservation in the EU. Preparation of the Natura 2000 was carried out without public hearings until the first official proposal was released. The program covers all the existing conservation areas, but also introduces a large number of new areas with various restrictions in their use. Many of the new areas are on private lands, and releasing the first proposal for the Natura 2000 resulted a public outcry, including literally thousands of appeals. The public opinion also criticized the way in which the program was prepared. Therefore, it is a logical finding that most Finns support moderate increases in conservation, but are reluctant to show support to extensive conservation programs, that they may view a regulatory “overshooting”.

References

Adamoviz, W., Louviere, J. & Williams, M. 1994. Combining revealed and stated preference methods for valuing environmental amenities. *Journal of Environmental Economics and Management* 26:271-292.

Amemiya, T. 1985. *Advanced Econometrics*. Harvard University Press. 521 p.

Beggs, S., Cardell, S. & Hausman, J. 1981. Assessing the potential demand for electric cars. *Journal of Econometrics* 16:1-19.

Boxall, P., Adamoviz, W., Swait, J., Williams, M. & Louviere, J.1996. A comparison of stated preference methods for environmental valuation. *Ecological Economics* 18:243-253.

Brownstone, D. & Train, K. 1999. Forecasting New Product Penetration with Flexible Substitution Patterns. *Journal of Econometrics*, Vol. 89(1):109-129.

Chapman, R.G. & Staelin, R. 1982. Exploiting rank ordered choice set data within the stochastic utility model. *Journal of Marketing Research* 19: 288-301.

Desvouges, W.H. & Smith, V.K. 1983. Option price estimates for water quality improvements: a contingent valuation study for Monogahela river. *Journal of Environmental Economics and Management* 14:248-267.

Goett, A., Hudson, K. & Train, K. 2000. Customers' Choice Among Retail Energy Suppliers: The Willigness-to-Pay for Service Attributes. Department of Economics, University of California, Berkeley.

Hanemann, M. W. 1984. Welfare Evaluations in Contingent Valuation Experiments with Discrete Responses. *American Journal of Agricultural Economics* 66:332-341.

Hanemann, M. W., Loomis, J. & Kanninen, B. 1991. Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation. *American Journal of Agricultural Economics* 1255:1263.

Hanemann, M.W. & Kanninen, B. The Statistical Analysis of Discrete-Response CV Data. In Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation (ed. Bateman, I & Willis, K.).

Hausman, J.A. & Ruud, P.A. 1987. Specifying and testing econometric models for rank-ordered data. *Journal of Econometrics* 34:83-104.

Hensher, D, Louviere, J.J. & Swait, J. 1999. Combining sources of preference data. *Journal of Econometrics* 89:197-221.

Layton, D.F. 2000. Random Coefficient Models for Stated Preference Surveys. *Journal of Environmental Economics and Management* 40:21-36.

Layton, D.F. 2001. Notes on One Parameter Family of Distributions for Random Coefficient Models. Department of Environmental Policy and Science. University of California Davis.

Louviere, J.J. & 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.

Magat, W.A., Viscusi, W.K., Huber, J. 1988. Paired comparison and contingent valuation approaches to morbidity risk valuation. *Journal of Environmental Economics and Management* 15:395-411.

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

McFadden, D. & Train, K. 2000. Mixed MNL Models for Discrete Response. Forthcoming in *Journal of Applied Econometrics*.

Revelt, D. & Train, K. 1998. Mixed Logit with Repeated Choice: Households' Choices of Appliance Efficiency Level. *The Review of Economics and Statistics*. Vol. 80: 647-657.

Roe, B., Boyle, K.,J. & Teisl, M.,F. Using Conjoint Analysis to Derive Estimates of Compensating Variation. *Journal of Environmental Economics and Management* 31:145-159.

Stevens, T.H., Belkner, R., Dennis, D., Kittredge, D. & Willis, C. 2000. Comparison of contingent valuation and conjoint analysis in ecosystem management. *Ecological Economics* 32: 63-74.

Swait, J. & Louviere, J.J. 1993. The role of scale parameter in the estimation and use of generalized extreme utility models. *Journal of Marketing Research* 10:23-45.

Train, K. 1998. Recreation Demand Models with Taste Differences Over People. *Land Economics* 74(2):230-39.