Logit Models for Pooled Contingent Valuation and Contingent Rating and Ranking Data: Valuing

Benefits from Forest Biodiversity Conservation

by

Juha Siikamäki[1,2]
Ph.D. Candidate
Department of Environmental Science and Policy
University of California at Davis

and

David F. Layton
Assistant Professor
Department of Environmental Science and Policy
University of California at Davis

May 15th 2001

Abstract

Contingent valuation and contingent rating and ranking methods for measuring willingness-to-pay for non-market goods are compared by using random coefficient models and data pooling methods. Pooled models on CV data and CR data on the preferred choice accept pooling if scale differences between the model estimates of CV and CR methods are allowed for. More detailed response models, such as pooled CV model and rank-ordered models for two or three ranks, reject pooling of the data.

[1] Address correspondence to Juha Siikamäki, Department of Agricultural and Resource Economics; University of California at Davis; Davis, CA 95616. Email: juha@primal.ucdavis.edu.

Introduction

Stated preference methods (SP methods) are widely used in measuring economic values related to the environment. They involve conducting surveys, in which respondents are presented with hypothetical alternatives, usually policy options. Each policy option results in the supply of a certain nonmarket good, such as environmental quality, for certain costs to respondents. Survey respondents are asked to evaluate the alternatives and to state their preferences over them. The contingent valuation (CV) method is based on asking for acceptance or refusal of a policy alternative, which has a specified cost. The contingent rating and ranking (CR) methods ask respondents to rate or rank the suggested alternatives. In its simplest form, this is done by asking them to choose a preferred alternative.

Although CV and CR are the most commonly used SP methods, their performance and consistency has not been exhaustively studied. Comparisons between the CV and CR include Desvouges and Smith 1983, Magat et al. 1988, MacKenzie 1993, Ready et al. 1995, Boxall et al. 1996, Adamowicz et al 1998, Stevens et al. 2000, and Cameron et al. 2001. They all suggest that there are differences in the results of the SP methods. All the methods, however, attempt to measure essentially the same tradeoffs between money and changes in the environmental quality, and their results should be very similar when applied to the same policy problem.

The previous studies that find differences between SP methods have been based on fixed coefficient discrete choice models, typically logit models. The assumptions and properties of fixed logit models are restrictive, but more flexible models with random coefficients have been impractical until recently due to limitations in computing power. This constraint has recently been greatly relaxed by the development of simulation-based econometric techniques, and random coefficient models can now be employed in modeling discrete choice data (e.g. Train's 1998, Layton 2000). The results of random coefficient applications suggest that it would be beneficial to examine the differences between SP methods by using

less restrictive econometric models than has been used in the past. More flexible models help determine whether the past conclusions have resulted from actual discrepancies between different SP methods, or perhaps from using overly restrictive econometric models.

Data pooling methods involve combining separate sources of data, such as CV and CR data, and estimating the econometric models by utilizing the pooled data set. This enables to compare different data sources at the estimation stage, and brings several benefits over estimating separate models for different data 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 the data collected. This in turn can result in more reliable estimates than are obtained with the unpooled models.

Fairly few studies have used the data pooling approach in examining SP methods. Adamowicz et al. (1994) combined survey data on choices in the past and stated preferences over hypothetical future alternatives. Adamowicz et al. (1998) combine choice experiment and contingent valuation data and estimate models on the combined data. Cameron et al. (2001) pool data from one actual and six different hypothetical value elicitation experiments regarding the same good. Hensher et al. (1999) provide a general framework for applying data pooling techniques to test for the invariance between separate sources of data. Their approach is adopted here and used in testing for the consistency of CV and CR data.

The empirical application of this study deals with measuring WTP for conserving habitats that are especially valuable ecologically (i.e. biodiversity hotspots) in non-industrial private forests in Finland. According to ecologists, protection of biodiversity hotspots is particularly important for biodiversity conservation in Finland. The hotspots cover a total of 1.1 million hectares, which is some 6 percent of

Finnish forests. Current regulations protect some 120,000 hectares of hotspots, and their extended protection is currently debated. This study evaluates potential conservation policy alternatives for the future by examining the public's preferences for them.

The rest of the paper is organized as follows. Econometric models for CV and CR data are explained first. Then, using the data pooling approach in testing for invariance between different data sources is described. The empirical section starts with a description of the survey on public preferences for biodiversity conservation in Finland. The results section starts with results for separate CV and CR data, and continues with pooled models that are used in testing for the invariance between the CV and CR data.

Discrete Choice Econometric Models for Stated Preference Survey Responses

Econometric models for stated preference survey responses are typically based on McFadden's (1974) random utility model (RUM). It is used here as a point of departure for explaining various econometric models for CV and CR survey responses[3]. Consider an individual i choosing a preferred alternative from a set of m alternatives providing utility U_{ij}, that can be additively separated into an unobserved stochastic component ε_{ij} and a deterministic component $V_{ij}(z_j, y_i - A_{ij})$, i.e., the indirect utility function that depends only on individual's income y and environmental quality z. Denoting the cost of alternative j to person i with A_{ij}, the utility of alternative j can then be represented as:

$$U_{ij} = V_{ij}(z_j, y_i - A_{ij}) + \varepsilon_{ij} \tag{1}$$

The stochastic term ε_{ij} represents the unobserved factors affecting the choices. They are known, and taken into consideration, by individual j choosing between the alternatives, but are not observed by the analyst.

---

[3] The CV section draws from works by Hanemann (1984), Hanemann et al. (1991), and Hanemann and Kanninen (1999); the CR section relies on McFadden (1974), Beggs et al. (1981), Chapman and Staelin (1982), Hausman and Ruud (1987), recent works by Train (e.g. 1998), McFadden and Train (2000), and Layton (2000).

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 from among all the m alternatives therefore equals the probability that alternative j provides person i with greater utility U_{ij} than any other available alternative U_{ik}, i.e.,:

$$P_{ij} = P(U_{ij} > U_{ik}, k = 1, ..., m, \forall k \neq j) \tag{2}$$

Representing $U_{ij} = V_{ij}(._j) + \varepsilon_{ij}$, then rearranging and 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 $\Delta V_{ijk}(.) = V_{ik}(z_{ik}, y_i - A_{ik}) - V_{ij}(z_{ij}, y_i - A_i)$, the probability P_{ij} can be presented as:

$$P_{ij} = P(\varepsilon_{ijk} > \Delta V_{ijk}(.), k = 1, .., m, \forall k \neq j) \tag{3}$$

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

Fixed Coefficient Logit Models

Contingent Rating or Ranking

Assume in the following that the random terms ε_j and ε_k are independently and identically distributed, type I extreme value (TEV) random variables. It follows that their difference ε_{ijk} is logistically distributed. Under these assumptions, McFadden (1974) showed that choice probability P_{ij} in (3) is determined as a conditional logit model:

$$P_{ij} = \frac{e^{\mu X_{ij}\beta}}{\sum\limits_{k=1}^{k=m} e^{\mu X_{ik}\beta}} \tag{4}$$

The parameter μ is a scale factor that appears in all the choice models based on the RUM, because the parameter vector β in the choice probabilities is identified up to a multiplicative constant. With data from a single source, the scale factor is typically normalized to one, and the normalized parameter vector β is estimated. This is necessary for identification; without the restriction imposed on μ, neither μ nor β could be identified. In combining data from different sources, in general *n-1* scale factors can be estimated from *n* sources of data. Since pooling of CV and CR data, and estimating scale factors for different data sources plays an important role here, scale parameters are included in all the following models.

Beggs et al. (1981) and Chapman and Staelin (1982) extended the conditional logit model to modeling the ranking of alternatives. A rank-ordered logit model treats the ranking as *m-1* consecutive conditional choice problems. In other words, it assumes that the 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, the third ranking to the next best alternative, and so on. The probability of the observed ranking *r* for person *i* is given by:

$$P_{ir} = \prod_{j=1}^{m-1} \frac{e^{\mu X_{ij}\beta}}{\sum\limits_{k=j}^{k=m} e^{\mu X_{ik}\beta}} \tag{5}$$

Hausman and Ruud (1987) developed a rank-ordered heteroscedastic logit model that is flexible enough to take into account possible chances in variance of the random term in the RUM as the ranking task continues. It is based on formulation (5) and modified to include a rank-specific scale parameter that accounts for systematic changes in the variance of the random term. By its structure, a rank-ordered heteroscedastic logit model can identify *m-2* scale parameters.

Maximum likelihood method is typically used in estimating the unknown vector of parameters. The earlier presented response probabilities form the likelihood function for each individual. By independence of observations, the total likelihood is then simply the sum of individual likelihoods over the whole sample. Estimation is usually based on a logarithmic transformation of the likelihood function. For instance, a conditional logit model is estimated by maximizing the total of $ln(P_{ij})$ over the whole sample.

Double Bounded Contingent Valuation Responses

In the double bounded CV, respondents are asked a follow-up question based on the first response. 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*. The second responses provide more detailed data on individual preferences between the two alternatives, and the choice probabilities can now be determined on the basis of responses to two separate questions. The four possible response sequences are: *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:

$$\text{P(Yes-Yes)} = P_i\ (YY) = \frac{e^{\mu X_{iHB}\beta}}{e^{\mu X_{iHB}\beta}+1} \tag{6}$$

$$\text{P(Yes-No)} = P_i\ (YN) = \frac{1}{1+e^{\mu X_{iHB}\beta}} - \frac{1}{1+e^{\mu X_{iFB}\beta}} \tag{7}$$

$$\text{P(No-Yes)} = P_i\ (NY) = \frac{1}{1+e^{\mu X_{iFB}\beta}} - \frac{1}{1+e^{\mu X_{iLB}\beta}} \tag{8}$$

$$\text{P(No-No)} = P_i\ (NN) = \frac{1}{1+e^{\mu X_{iLB}\beta}} \tag{9}$$

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 the double-bounded CV is:

$$L = \sum_{i=1}^{N} \ln[I_{yy}P_i(YY) + I_{yn}P_i(YN) + I_{ny}P_i(NY) + I_{nn}P_i(NN)] \tag{10}$$

Random Coefficient Models

Although they are frequently applied to SP data, the fixed coefficient logit models have some undesirable properties and assumptions. First, they overestimate the joint probability of choosing close substitutes, because of 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. In practice, it is more likely that individual specific factors influence the evaluation of all the available alternatives, and make the random terms correlated instead of independent.

Random coefficient logit (RCL) models have been proposed to overcome possible problems of the fixed coefficient logit models (e.g. Revelt and Train 1998, Train 1998, Layton 2000). The RCL model is specified similarly to the fixed coefficient models, except that the parameters β are assumed to vary in the population rather than be fixed at the same value for each person. Utility is expressed as the sum of population mean b, an individual deviation η, which accounts for differences in individual taste from the population mean, and an unobserved i.i.d. random term ε. The total utility for person i from choosing alternative j is therefore:

$$U_{ij} = X_{ij}b + X_{ij}\eta_i + \varepsilon_{ij} \tag{11}$$

where $X_{ij}b$ and $X_{ij}\eta_i + \varepsilon_{ij}$ are the observed and unobserved parts of utility, respectively. Utility can also be written as $X_{ij}(b + \eta_i) + \varepsilon_{ij}$, which shows how the previously fixed β now varies across people as $\beta_i = b + \eta_i$.

Although the RCL models account for heterogeneous preferences via parameter η_i, individual taste deviations are neither observed nor estimated. The RCL 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

7

vary across the population with density $f(\beta|\Omega)$, with Ω 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. The probability

of choosing alternative j out of m alternatives can now be written as:

$$P_{ij} = \int \left[\frac{e^{\mu X_{ij}\beta_i}}{\sum\limits_{k=1}^{k=m} e^{\mu X_{ik}\beta_i}} \right] f(\beta|\Omega)d\beta \tag{12}$$

Equation (12) is the random coefficient extension of the conditional logit model (4). Random coefficient

models for rank-ordered logit models are defined similarly, as well as the RCL model for double-bounded

CV. The extension is straightforward and not replicated here. It suffices to note that in other models the

bracketed part of (12) is replaced by the appropriate probability formula.

Integral (12) cannot be analytically calculated and must be simulated for estimation purposes. Train has

developed a method that is suitable for simulating (12), and its many extensions needed in this study. His

simulator is smooth, strictly positive and unbiased for just one draw of β_i (Brownstone and Train 1999)

and can be easily modified to allow for strictly negative or positive random coefficients. Simulating (12)

is carried out simply by drawing a random β_i, calculating the bracketed part of the equation, and repeating

the procedure many times. Using R draws of β_i from $f(\beta|\Omega)$, the simulated probability of (12) is:

$$SP_{ij} = \frac{1}{R} \left[\sum\limits_{r=1}^{r=R} \frac{e^{\mu X_{ij}\beta_{ir}}}{\sum\limits_{k=1}^{k=m} e^{\mu X_{ik}\beta_{ir}}} \right] \tag{13}$$

The simulator (13) can be easily extended to the rank-ordered logit models and to logit models for the CV

responses. The only change required is to replace the bracketed term with the rank-ordered or double-

bounded CV probability formulas.

8

Data Pooling Method

The scale factor μ of the response models is 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 problem is that estimated vector of parameters β gets confounded with constant μ. This in turn makes absolute values of the parameter estimates incomparable between different data sets; only the ratios of parameters 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 set $q=1,...,n$ 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 result. 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 of the $n-1$ 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 the random terms between the CV and CR data. To test for the parameter invariance between the CV and CR data, models with and without the restriction $\beta_{CV}=\beta_{CR}$ need to be estimated. Likelihood ratio tests can then be applied to accept or reject the imposed parameter restriction. If the null hypothesis cannot be rejected, the data can be considered generated by the same taste parameters but still have scale 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 data sources.

The Survey

Data were collected using a mail survey, sent out in spring 1999 to 1740 Finns between 18 and 75 years of age. The sample was randomly drawn from the official census register of Finland, and is therefore a representative sample of the Finnish population. The sample was randomly divided into two 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.

WTP was 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 25-, 50- and 75-percent protection of all the biodiversity hotspots, respectively. The current legislation already protects 10 percent of them.

In designing the survey, special attention was paid to making the conservation policy scenarios relevant and credible. An easy to read one page section in the questionnaire explained different conservation programs and their details. Suggested new conservation programs were described as extensions to an already existing policy measure that uses incentive payments to encourage landowners to voluntarily set aside biodiversity hotspots. While designing the survey, questionnaire versions went through several rounds of modifications and reviews by experienced SP practitioners as well as other economists, foresters and ecologists with expertise in survey methods or biodiversity conservation. After the expert comments were incorporated, the questionnaires were tested by personal interviews and a pilot survey (n=100), and modified on the basis of the results.

The questionnaires started with questions about the respondents' attitudes on different aspects of forest and public policies. The next section of the questionnaire described the forest management and current

conservation situation in the country. The valuation questions followed next. The questionnaire concluded with questions on the respondents' socioeconomic background.

The respondents to 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, 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-percent conservation were collected by using both the first and the second WTP question, depending on the respondent's sub-sample. The CV method employed a double-bounded format. The bid vector in the CV survey consisted of first bids between $4 - 500 and the follow-up bids between $2 - 800, with seven different starting bids. The same bid amounts appeared as first and second bids for different respondents and different programs. Description of the programs was exactly the same as in the CR surveys, except that each CV-respondent was offered only the status quo and two new policy alternatives, not three new alternatives as in the CR surveys.

The CR survey described to respondents the status quo and all three hypothetical programs of setting aside 275,000, 550,000 and 825,000 hectares of hotspots. Each respondent was asked to rate the four programs on a scale from 0 to 10. The three hypothetical programs were assigned costs using the same variation across the respondents and the conservation programs as in the CV survey. The costs were assigned so that the higher conservation percentages were always assigned higher bids than the lower conservation percentages.

The final survey consisted of 29 different questionnaire versions, of which 14 used the CV and 15 used the CR. In both types of surveys, WTP was measured as the increase in annual tax burden of the household. Except for the valuation question, the CV and CR questionnaires were exactly the same. Therefore, only the choice task in the valuation question varies between the CV and CR respondents.

11

The survey was mailed out in May 1999. A week after the first mailing, the whole sample was sent a reminder card. Two more rounds of reminders with a complete questionnaire were sent to non-respondents in June and July. The CV and CR surveys resulted in response rates of 48.9 percent and 50 percent, respectively. Different methods therefore resulted in almost the same response rates.

Results

The CV data (376 observations) was first censored for missing responses. The remaining 306 observations with complete responses to the both WTP questions were employed in the estimation. The CR data with ratings of alternatives was first transformed into rankings by assuming that preferred alternatives were rated higher than the less preferred ones. Rankings utilize only the ordinal information on preferences. Respondents with, for instance, ratings sequences (3, 2, 1, 0) and (10, 9, 3, 1) for the four policy alternatives 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 observations for estimation. Hence, the results are based on data with full and unique rankings of all four policy alternatives[4].

Several alternative specifications were estimated[5]. They specified the valuation function for conservation either as a continuous linear, logarithmic or a quadratic function, or by using alternative specific dummies. Non-nested model selection test by Pollak and Wales 1991 was then employed in selecting a preferred specification among them[6]. Based on its statistical performance, a quadratic models was chosen as the preferred specification for these data. The observed part of the RUM is estimated as:

---

[4] Ties were distributed as follows: The ratings for the status quo and the 25-percent alternative were tied 15 times, resulting in 3.9 percent of observations being tied. The status quo and the 50-percent alternative had 12 ties (3.1%); the status quo and the 75-percent alternative 19 ties (4.9%); the 25- and 50-percent alternatives were tied 37 times (9.6%); the 25- and 75-percent alternatives 36 times (9.3%); the 50- and 75-percent alternative 67 times (17.4%). Examining the ties, and modeling them within the rank-ordered logit model, is computationally demanding and an objective for future work.

[5] All the models were programmed and run in GAUSS.

$$V_{ij} = \beta_{BID}BID_{ij} + \beta_{Percent}PERCENT_{ij} + \beta_{SqrPercent}SQRPERCENT^2_{ij} \tag{14}$$

where *BID* is the cost of alternative *j* to person *i*, *PERCENT*[7] is the percentage of the biodiversity

hotpots conserved under alternative *j,* and *SQRPERCENT* is its square.

Random coefficients are typically estimated as normally distributed parameters. A normally distributed

parameter β_n can take both negative and positive values. It is estimated as $\beta_n = (b_n + \eta_n e)$, where *b* and η

are the estimated mean and deviation parameters of the β_n, and *e* a standard normal deviate (Train 1998).

The *BID* parameter is assumed to be strictly negative. For the strictly negative *BID*, increasing the costs of

a policy alternative always decreases its probability to become chosen. Train (1998) suggests that the

strictly positive random coefficients can be estimated as log-normally distributed, and provides a practical

method for incorporating them into his simulator. Each log-normal β_k can be estimated by expressing

them as $\beta_k = exp(b_k + \eta_k e)$, where *b* and η are estimated mean and deviation parameters of $ln(\beta_k)$, and *e* an

independent standard normal deviate. Strictly negative parameters are estimated by entering the variable

values in question as their negative. A disadvantage of the log-normally distributed random coefficients is

that 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

strictly negative or positive random coefficients. While the RCL models typically estimate the mean and

variance of the distribution, one-parameter distributions, such as Rayleigh distribution, identify all the

moments of the distribution by estimating just a single parameter. A strictly positive *BID* with Rayleigh

distribution has a cumulative density function $F(BID)=1-exp[-BID^2/(2b^2)]$, and a probability density

function $f(BID)=(BID/b^2)exp[-BID^2/(2b^2)]$, where $b > 0$ is the scale parameter fully determining the shape

[6] Non-nested models selection results are not reported in detail here; their results are available from the authors.

of the distribution. Using the inverse transformation method, the Rayleigh distributed *BID* can be obtained as $BID=(-2b^2ln(1-u))^{1/2}$, where *u* is a random uniform deviate and *b* the random draw for *BID* parameter. Since Rayleigh distributed *BID* is strictly positive, it is estimated by entering the values of the *BID* variable as their negative. 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 (Johnson et al. 1994, Layton 2001).

Previous applications have typically modeled either the *BID* or other alternative characteristics as random coefficients, not both. With these data, heterogeneity of preferences for conservation level could be observed, and *PERCENT* and *SQRPERCENT* were therefore modeled as normally distributed random coefficients[8]. On the other hand, previous studies suggest that heterogeneity of preferences is often related to the *BID* coefficient. It was therefore also estimated as a random coefficient. Since it essentially represents the negative of the marginal utility of income, it was estimated as a non-positively distributed coefficient. The *BID* was estimated using a Rayleigh-distributed random coefficient $BID_{RAYLEIGH}$.[9]

Table 12 reports both the fixed (FCL) and random coefficient (RCL) model results for the quadratic specification. The FCL models result in statistically significant estimates for all the parameters. Further, the BID parameter has an expected negative sign in all the results. The linear term of the value function, PERCENT, and the quadratic term, SQRPERCENT, together reveal how value of conservation changes as the conservation level is increased. In all the models, PERCENT gets a positive, and SQRPERCENT a negative estimate, suggesting that the value of conservation is first increasing as a function of conservation, and then decreasing as conservation is further increased.

---

[7] The CR alternatives are given values 0, 15, 40, and 65 for a measure of increased conservation under each alternative; the CV data is constructed similarly.
[8] No correlation parameters between random coefficients were estimated to ensure identification of coefficients.
[9] Despite substantial efforts, *BID* was not estimable as a log-normal random coefficient for all the models. Employing a Rayleigh distributed *BID* lead to faster iteration and convergence of all the models.

Table 1. Model results

Model: Estimate ($\lvert t\text{-statistic}\rvert$)	CV		CR - 1 Rank		CR - 2 Ranks		CR - 3 Ranks	
	FCL	RCL	FCL	RCL	FCL	RCL	FCL	RCL
BID	-0.3556 (19.850)		-0.0860 (4.746)		-0.0752 (5.803)		-0.0557 (5.232)	
$BID_{RAYLEIGH}$		1.6353 (8.252)		0.0752 (3.311)		0.1905 (2.901)		0.1096 (2.311)
Means								
$PERCENT$	0.9216 (11.309)	2.8384 (9.060)	0.1831 (1.677)	0.2268 (1.056)	0.4180 (5.290)	2.3525 (5.239)	0.6925 (10.124)	1.7012 (7.458)
$SQRPERCENT$	-0.1189 (8.343)	-0.2926 (7.913)	-0.0257 (1.671)	-0.0350 (0.851)	-0.0667 (5.875)	-0.4737 (4.882)	-0.1075 (11.110)	-0.3217 (9.832)
Deviations								
$PERCENT$		0.9266 (6.539)		-0.0350 (0.135)		2.1215 (5.131)		1.8121 (8.383)
$SQRPERCENT$		0.0114 (0.398)		0.0471 (0.272)		0.1765 (2.632)		0.0079 (0.285)
LL	- 846.85	-693.34	-347.82	-348.73	-608.78	-479.99	-746.98	-545.31
LL_0	-848.41	-848.41	-374.30	-374.30	-670.93	-670.93	-858.08	-858.08
Pseudo-R^2	0.018	0.183	0.071	0.068	0.093	0.285	0.129	0.364

Note: Number of the CV and CR responses are 306, and 270, respectively. 1000 replications were employed in simulating the maximum likelihood function. Bid variable values were divided by 100 to facilitate estimation; the PERCENT was divided by 10, and the SQRPERCENT calculated using the scaled PERCENT.

Although exploiting information on more than the first rank seems to improve model performance, one must check the consistency of rankings before accepting the rank ordered models. The LR-tests in the fashion of Hausman and Ruud (1987) for the consistency of rankings suggested that none of the rank-ordered models can be accepted. Consistency of two and three ranks was not accepted, and hence, the fixed coefficient rank-ordered logit models for two and three ranks were rejected. Rejection of the consistency of rank-heteroscedastic models suggests that the heterogeneity of responses is not sufficiently systematically related to the ranks for the rank-ordered models to be consistent[10].

---

[10] Estimating models separately for second and third rank result in log-likelihood values 251.28 and 45.56, respectively. Tests for consistency of rankings for the rank-ordered logit models for two and three ranks result in LR-test statistics 19.36 and 204.64, respectively, and rejection of both rank-ordered models. Also the Hausman and Ruud rank-heteroscedastic models lead to rejection of consistency of rankings.

The RCL models are statistically superior models for the CV data, and the CR data for two and three ranks. The pseudo-R^2 has increased substantially for all these models. Interestingly, the CR 1 Rank model performs statistically slightly worse than the fixed parameter model[11]. Despite the poor performance of the CR 1 Rank model, the other model results confirm the importance of addressing the parameter heterogeneity issue in modeling these data.

Using the formula for the asymptotic mean of Rayleigh distributed *BID* parameter, the estimated $BID_{RAYLEIGH}$ parameters translate into the mean of -1.60 for *BID* parameter, and values for CR models for one, two and three ranks into values -0.37, -0.34, and -0.30 for means of *BID* parameters, respectively. The RCL models for CV data, and for CR data on two and three ranks, produce a statistically significant estimate of PERCENT variable, suggesting that preferences regarding the conservation level are heterogeneous. The curvature parameter, SQRPERCENT, has a statistically significant deviation estimate only in the CR model for two ranks. The estimate for the mean for the SQRPERCENT parameter is significant for all the RCL models except the CR 1 Rank-model, suggesting that the quadratic nature of the value function was not a product of the fixed parameter formulation, but a phenomena that is inherent in the data.

Models for Pooled Contingent Valuation and Contingent Ranking Data

Pooled models were estimated on 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 response. Formulating a dummy variable for the CR data can facilitate estimation. Defining $I_{iCR} = 1$ for CR respondents, and $I_{iCR} = 0$ for CV respondents, the pooled log-likelihood function for individual i is determined as:

---

[11] This is expected to be a result of the properties of Rayleigh distribution. For instance, it has as long right hand tail,

$$LL_i = I_{iCR}*ln(P_{iCR})+ (1- I_{iCR})*ln(*Pi_{iCV})$$ (14)

where P_{iCR} and P_{iCV} are the appropriate CV and CR response probabilities of the model. Logically, the total log-likelihood function is again a sum of the individual likelihoods over the whole sample[12].

Table 2 reports the model results for the combined CV and CR data. Although the previous analysis has shown that the random coefficient formulation is the preferred modeling approach for these data, a variety of pooled models were estimated to examine the effects of modeling choices on accepting or rejecting pooling the data. LR-tests were employed in accepting or rejecting the pooling hypothesis; the respective LR-test statistics are reported in the second last row of the Table 2. The test statistics is distributed χ^2 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 (d.o.f) for fixed and random parameter logit models with all parameters random equal to 2 and 5, respectively. The critical values with 2 d.o.f at 5 percent and 1 percent significance levels are 5.99 and 9.21; the respective values for 5 d.o.f are 11.07, and 15.09. If the LR-test statistic gets a lower value than the critical value, pooling of data cannot be rejected.

All the models estimate parameter μ_{CR}, which is a scale factor for the CR data, and accounts for differences in the variance of random term of the RUM between the CV and CR data. Since only the parameter ratios are comparable between the different sources of data, estimating a scale factor allows for direct comparisons of the estimates. 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.

---

which, despite the benefits due to parameter heterogeneity, can make a RCL model more constrained than the FCL model.
[12] Considerable time savings, especially in estimating random coefficient models, can be obtained by structuring the data so that unnecessary calculations of the log-likelihood function are avoided. Calculations of CR response probabilities are unnecessary for CV respondents, *vice versa*.

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.

Table 2. Model results for pooled CV and CR data

Model:	CV & 1 Rank		CV & 2 Rank		CV & 3 Rank	
Estimate (*t*-statistic)	FCL	RCL[a]	FCL	RCL	FCL	RCL
BID	-0.3558 (19.863)		-0.3522 (19.719)		-0.3439 (19.348)	
BID$_{Rayleigh}$		1.6086 (8.243)		1.2032 (8.383)		1.1947 (8.594)
Means						
PERCENT	0.9204 (11.504)	2.8398 (9.104)	0.9924 (12.718)	2.9765 (9.690)	1.0877 (14.170)	2.9839 (9.841)
SQRPERCENT	-0.1190 (8.569)	-0.2957 (8.026)	-0.1334 (9.939)	-0.3299 (9.421)	-0.1512 (11.508)	-0.3268 (9.698)
Deviations						
PERCENT		0.9199 (6.529)		1.3855 (8.613)		1.4052 (8.738)
SQRPERCENT		0.0114 (0.407)		0.0301 (0.632)		0.0234 (0.911)
μ_{CR}	0.2569 (5.661)	0.0486 (4.140)	0.2579 (7.862)	0.6797 (5.634)	0.2390 (8.836)	0.7152 (7.084)
LL pooled	-1194.89	1043.09	-1462.19	-1265.32	-1631.77	-1355.77
LL unpooled	-1194.67	1042.16	-1455.63	-1173.33	-1593.83	-1238.65
LL$_0$	-1222.71	-1222.71	-1519.34	-1519.34	-1706.49	-1706.49
LR-test	0.44	1.86	13.12	183.98	75.88	234.24
Pseudo-R^2	0.023	0.147	0.038	0.167	0.044	0.206

Note 1: 1000 replications were employed in simulating maximum likelihood function. BID variable was divided by 100 to facilitate estimation; the PERCENT was divided by 10; the SQRPERCENT was calculated from the scaled PERCENT.
Note 2: The values of log-likelihood functions for unpooled models can be obtained simply as sums of separate log-likelihood functions for the CV and CR data. The LL$_0$ is obtained similarly.

Models "CV & 1 Rank" pool the CV model with a CR model for first rank, using both the FCL and the RCL formulation. The estimate of the μ_{CR} is statistically significant and smaller than one in both models, suggesting that CR data is noisier. The random coefficient model provides a substantially higher

explanatory power than the fixed coefficient counterpart. The value 1.6027 of the $BID_{RAYLEIGH}$ parameter translates into the mean of -1.59 for the BID parameter. Both LR test statistics (0.44 for FCL, and 1.86 for RCL) are smaller than the critical values (5.99, 11.07, respectively) and very insignificant. Therefore, pooling the two data sets, and estimating a pooled model did not result in statistically significant decrease in the value for the maximized likelihood function. Hence, both fixed and random coefficient models provide support for accepting pooling of the CV and CR data.

"CV & 2 Rank" models pool the fixed and random coefficient models for the CV model and the CR model for 2 ranks. Similarly as in the previous pooled models, these models result in highly significant estimates with expected signs. The value of $BID_{RAYLEIGH}$ parameter translates into the mean of -1.37 for the BID parameter. Comparing the unpooled and unpooled fixed coefficient models results in a LR test statistic 13.12, thus rejecting the pooling hypothesis at 1 percent significance level. The random coefficient model results in a much higher LR test statistic, and a strong rejection of pooling the CV and CR data. Despite rejection of pooling, the random coefficient model results in substantially higher pseudo-R^2 than the fixed coefficient model. The scale parameter for the CR data gets an estimate 0.6797 with standard error of 0.1206, making the scale parameter estimate statistically significantly smaller than one. It suggests that variance of the random term is higher in the CR data than in the CV data, and that absolute parameter estimates are not directly comparable between the two data sources.

Models "CV & 3 Rank" pool the fixed and random coefficient models for 3 ranks and the CV data. Results are similar with the pooled models involving the CR model for two ranks. The value of $BID_{RAYLEIGH}$ parameter translates into the mean of -1.37 for the BID parameter. Pooling of CV and CR data is strongly rejected for both the fixed and the random coefficient models. The scale parameter for the CR data gets an estimate of 0.7152, and a standard error of 0.1010. Hence, similarly as with the "CV & 2 Rank" model, the scale parameter is statistically significantly smaller than one.

Although not reported in Table 3, 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 was uniformly and strongly rejected. The fixed model for the pooled CV and first rank CR data using a CR scale factor provides the strongest support for accepting pooling, and is therefore likeliest to provide support for the complete invariance hypothesis. Estimating the pooled model for it results in a log-likelihood value 1247.37, hence, a LR-test statistic 105.4, that suggests strong rejection of completely pooling the CV and CR data. All the other models without an estimated scale parameter reject the pooling hypothesis with stronger statistical evidence than this model. Complete invariance of the CV and CR data is therefore uniformly rejected.

Willingness to Pay for Biodiversity Conservation

The results show a 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. Therefore, the fixed coefficient models for CR data on two and three ranks are rejected first. The rest of the models can be evaluated by using success in pooling the CV and CR data as criterion. The models that successfully pool the CV and CR data are not fully dependent on a single survey method, and can therefore considered most general.

Two models were accepted by the pooling success criterion: (1) pooled fixed coefficient model for the CV and the CR data on one rank, and (2) pooled random coefficient model for the CV data and the CR data on one rank. The FCL model can then be screened out as more restrictive than its RCL counterpart. The random coefficient model for the pooled CV data and the CR data on first choice is therefore chosen as the preferred model for evaluation preferences for conservation of biodiversity of forests.

The WTP for policy alternative x_j is calculated as:

$$\partial V / \partial x_j * (\partial V / \partial y)^{-1} \tag{15}$$

where the $\partial V / \partial x_j$ is the utility change from reaching some conservation level, and $(\partial V / \partial y)^{-1}$ is the inverse of the marginal utility of income, i.e., the inverted *BID* parameter. The estimates for *PERCENT* and *SQRPERCENT* parameters are employed in calculating a point estimate for the utility change from increasing conservation from status quo to a conservation level of interest. This change in utility is then transformed into a money measure by multiplying it by inverted marginal utility of income.

The mean estimate multiplied by the aggregate population measures the aggregate WTP, and the mean WTP is therefore chosen as the measure of the WTP. Calculating the mean WTP estimates from the fixed coefficient model results is straightforward. The fixed coefficient estimates equal their expectation, and with linear-in-income specification, the relation between estimates for the utility changes $\partial V / \partial x_j$ and $\partial V / \partial y$ measure the mean WTP. Calculating the WTP estimates for the random coefficient models requires calculating expectation of the inverted *BID* which is Rayleigh distributed. Note that since the *BID^{-1}* is now a non-linear transformation of the *BID*, its expected value $E(BID^{-1})$ cannot be calculated straightforwardly as an inverse of the expected value of *BID*. The $E(BID^{-1})$ can be obtained as $(\Pi/(2b^2))^{1/2}$, where b is the estimated scale parameter parameter for the Rayleigh distribution (Layton 2001). The WTP is then obtained multiplying a point estimate of the utility change from reaching a certain level of conservation by $E(BID^{-1})$.

The mean estimates of WTP for the preferred pooled models are reported in Table 3, as well as their 90-percent confidence intervals. The confidence intervals are based on the Krinsky and Robb (1986), and calcualte them by using a Monte Carlo simulation, programmed and run in GAUSS. Simulating the WTP results was carried out by using the estimated parameter vector and its variance-covariance matrix to

obtain a vector of parameters, which was then use to calculate the WTP. The procedure was repeated 2000 times, resulting in a simulated distribution for the WTP.

Table 3. Willingness to pay estimates

Conservation	WTP								
	25-percent			50-percent			75-percent		
Model	CI 10[b]	**Mean**	CI 90[b]	CI 10	**Mean**	CI 90	CI 10	**Mean**	CI 90
FCL CV + CR 1	*43*	**48.1**	*54*	*64*	**76.9**	*91*	*10*	**41.3**	*74*
RCL CV + CR 1	*36*	**42.7**	*50*	*65*	**78.9**	*94*	*46*	**71.4**	*97*

[a] All estimates in reported in USD=FIM6.5.
[b] 90 percent confidence intervals (the lower limit CI 10 and the upper limit CI 90) are reported in *italics*. They are reported as integers, rounding the lower limit down, and the upper limit up.

The WTP estimates are quite similar between the FCL and RCL models. The WTP for 25-percent conservation for the FCL model is about $48, whereas the RCL model produces a WTP estimate of about $43. The FCL and RCL estimates of WTP for 50-percent conservation are almost equal (some $77 and $79, respectively). The models behave somewhat differently in estimating the WTP for 75-percent conservation. The FCL model gives an estimate of $41, whereas the RCL models produces a considerably higher estimate, $71. The RCL model therefore predicts a higher WTP for 75-percent conservation than for the 50-percent conservation, and the FCL the other way around. Both models estimate the highest WTP for the 50-percent alternative.

Conclusions

This study examined using stated preference (SP) surveys for measuring public's willingness-to-pay (WTP) for conservation of forest biodiversity in Finland. Both fixed and random coefficient logit (FCL, and RCL, respectively) models were described and applied to the data collected. Comparison of the FCL and RCL models suggested that from purely statistical viewpoint, the RCL models are superior to the FCL models in modeling these data.

Successful pooling of the CV and CR data required estimation of scale factors for separate data sources, which made the parameter estimates comparable between the data sources. With scale factors in the estimated model, pooling of CV and CR data could not be uniformly rejected or accepted. Less detailed models such as CV and conditional choice logit models accepted the pooling hypothesis. The more detailed models such as models for pooled CV data and CR data on two and three ranks rejected the pooling hypothesis. The RCL models rejected the pooling hypothesis more strongly than the FCL models.

Two models satisfied the pooling criterion: (1) pooled fixed coefficient model for the CV and the CR data on one rank, and (2) pooled random coefficient model for the CV data and the CR data on one rank. The FCL model with one rank CR data was then screened out as more restrictive than its random coefficient counterpart. The RCL model for the pooled CV data and the CR data on first choice was therefore chosen as the preferred model for evaluation preferences for conservation of biodiversity of forests. Its WTP estimates for the 25-, 50-, and 75- percent conservation programs were $43, $79, and $71.

The random coefficient models do not seem to provide a miracle in terms of solving differences between CV and CR data. Applying the RCL models was strongly supported by the econometric analysis, but their flexibility did not provide an easy means to combining of the CV and CR data. To the contrary, pooling of the CV and CR data was more easily accepted by the less flexible fixed coefficient models. Differences in the choice task in the CR and CV method, especially in the amount of information they collect on individual preferences, seems to be the primary contributor to the differences between the two methods, not the flexibility of the econometric models used for modeling the stated responses.

The issue of negative marginal WTP for conservation after reaching a certain conservation level is of course worth some further investigation. The results suggest that the marginal utility from conservation is

decreasing after a certain level of conservation. This observation at first may seem counterintuitive and contradictory to economic theory. To understand the results, one must interpret them from the viewpoint of the respondents. To many of them, the choice between the policy alternatives is probably more like a choice between moderate and extreme conservation policies than a purely economic consideration with nicely behaving tradeoffs between money and the conservation benefits. The choice between the different policy alternatives is therefore not limited to just choosing between the money and conservation; other important tradeoffs are involved as well. Many people feel it is not justified, and simply not fair to keep increasing the conservation. Some have personal experience of the past conservation efforts and of the rather aggressive takings approach that was favored by the environmental agencies and lead into conflicts and somewhat strong public opposition to conservation.

The strong opinions about conservation became apparent already when the survey was carried out. Many of the phone calls received criticized increases in the conservation. On the other hand, some respondents were perhaps as strong supporters of the conservation as the opponents were against it. Since the preferences for conservation seem to be rather polarized, modeling the WTP distribution as a two peaked bimodal distribution would possibly be fruitful in the future. Further, polarized preferences are not necessarily unique to the Finnish case of forest conservation, but are likely to appear with many other controversial policies; a very typical circumstance for stated preference applications.

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.

Rejecting the pooling of CV and CR data despite very similar survey setting naturally raises some questions about the ability to measure WTP values at all. Although almost everything to facilitate the comparisons of WTP models was done, from sampling the same population to using nearly identical questions in different treatments, yet it proved not fully possible to reconcile the two sources of data within a single estimation framework. If this ultimately proves fruitless, then it raises questions about the ability to measure WTP, because each method seems as credible as the other, yet they give different results and generate incompatible data.

Failing to pool the data could be a result from different methods imposing a sufficiently different cognitive burden on the respondents so that different responses are elicited. However, misspecification of WTP either in the systematic part or the error structure could have also contributed to inability to pool the data. For instance, inadequately accounting for heterogeneity in the population would be sufficient to cause this. As noted, the WTP could be modeled using, for instance, a bimodal distribution. It could also be that there are multiple sets of preferences at work that do not fall into single normal or Rayleigh distribution. Sorting the data, perhaps along the lines of environmental and other attitudinal responses, into groups that have different parameter distributions, is also an interesting issue for further research.

References

Adamowicz, 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.

Adamowicz, W., Boxall, P., Williams, M., & Louviere, J.1998. Stated preference approaches for measuring passive use values: choice experiments and contingent valuation. *American Journal of Agricultural Economics 80*:64-75.

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

Boxall, P., Adamowicz, 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.

Cameron, T. A., Poe G. L., Ethier, R.G., Schulze, W.D. (2001) "Alternative Nonmarket Value-Elicitation Methods: Are Underlying Preferences the Same?" Forthcoming in *Journal of Environmental Economics and Management*.

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.

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. 1999. The statistical analysis of discrete-response CV data. In: Bateman, I. & Willis, K. (eds) *Valuing environmental preferences: theory and practice of the contingent valuation method in the US, EU, and developing countries*. Oxford University Press.

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.

Johnson, N.L. Kotz, S. & Balakrishnan, N. 1994. *Continuous univariate distributions*. 2nd edition. John Wiley.

Krinsky, I. and Robb, A. 1986. Approximating the statistical properties of elasticities. *Review of Economic and Statistics* 68:715-719.

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.

MacKenzie, J. A 1993. Comparison of Contingent Preference Models. *American Journal of Agricultural Economics* 75:593-603.

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*.

Pollak, R.A. & Wales, T.R. 1991. The likelihood dominance criterion: a new approach to model selection. *Journal of Econometrics* 47:227-242.

Ready, R., Whitehead, J. & Blomquist, G. 1995. Contingent valuation when respondents are ambivalent. *Journal of Environmental Economics and Management* 29:181-197.

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.

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. 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.