Do Hypothetical and Actual
Willingness to Pay Differ in Choice Experiments?
- Application to the Valuation of the Environment

Fredrik Carlsson[A]
Peter Martinsson[B]

Abstract

In this paper we test the validity of choice experiments with donations for environmental projects. In particular, we test whether or not willingness to pay for projects differs between a hypothetical and an actual choice experiment. Our results do not indicate that choice experiments suffer from overstatement in hypothetical willingness to pay; and this contrasts the results found in external tests of validity in Contingent Valuation. In addition, internal tests of validity indicate transitive and stable preferences in both experiments.

Key words: *Choice experiments, External validity, Internal validity, Hypothetical and actual WTP, Public goods.*

JEL classification: H41, Q20.

A) Department of Economics
Göteborg University
Box 640
SE-40530 Göteborg
Sweden

Tel: +46 31 7734174
Fax: +46 31 7731043
e-mail: Fredrik.Carlsson@economics.gu.se

B) Department of Economics
Lund University
Box 7082
SE-220 07 Lund
Sweden

Tel +46 46 2220000

e-mail: Peter.Martinsson@nek.lu.se

1. Introduction[1]

The contingent valuation method (CVM) is the most frequently used method to elicit individuals' willingness to pay (WTP) for environmental goods.[2] There is currently a debate however about whether or not CVM can measure individuals' maximum WTP accurately as the method has failed in many validity tests. More recently, choice experiments have been used as an alternative to CVM in order to elicit individuals' WTP. In brief, a choice experiment asks individuals to choose one alternative from a choice set, where each alternative is described by a number of attributes; several choice sets are presented to each individual. Hanley *et al*. (1998) argue that choice experiments have several advantages over CVM such as easiness to estimate values of single attributes, avoidance of yea-saying, and built-in tests of sensitivity to scope. In this paper we test the validity of choice experiments.

Stated preference methods[3] can be validated in two ways: (1) internal tests and (2) external tests. In internal tests of validity the standard assumptions on individuals' preferences such as transitivity, stability of preferences and monotonicity (usually referred to as test of sensitivity to scope in the CVM literature) are tested. External tests focus on whether or not hypothetical WTP differs from actual WTP. The most preferable and reliable test is, of course, an external test of validity. A reason for the debate over CVM is that some tests have indicated poor performance in internal and especially external tests of validity. The main critique of CVM, based on tests of internal validity, has concentrated on its inability to report different WTP for different sizes of an offered programme (insensitivity to scope)[4], and that WTP for a particular effect depends on whether it is valued on its own or as a part of a more inclusive package (embedding effect) (see *e.g.* Diamond and Hausman 1994).[5]

---

[2] See for example Carson *et al.* (1995) for a comprehensive list of CVM papers until 1995.
[3] We define for instance CVM, choice experiments, contingent ranking, and conjoint analysis as stated preference methods.
[4] Carson (1997) re-examines some previous studies which have indicated insensitivity to scope, and he argues that the results were due to poor research and/or poorly performed surveys.
[5] For a general discussion on potential biases see for example Mitchell and Carson (1989).

Marketable private goods have been used frequently to test for external validity in CVM and the results have indicated that individuals overstate their WTP in a hypothetical setting.[6] However, as argued in Carson *et al*. (1996), there is a fundamental difference between marketable private and public/quasi-public goods concerning the incentives for truthful preference revelation. When a good is provided privately, the incentives for truthful preference revelation differ from the incentives associated with public goods in both an actual and a hypothetical context. For example, in a hypothetical setting, individuals may state what they believe is the market price or the fair price. Hence, private market goods should not be used in order to validate CVM. Some previous studies on donations to environmental projects using CVM have also indicated overstatement of hypothetical WTP (*e.g.* Brown *et al*. 1996, Navrud 1992, and Seip and Strand 1992). Another way to test external validity is to compare the results of CVM studies with revealed preference studies. Carson *et al*. (1996) perform a meta-analysis, including 83 studies, allowing 616 comparisons and they find that CVM estimates are slightly lower than their revealed preferences counterparts, with a mean ratio between CVM and revealed preferences of 0.89.

In this paper we examine the appropriateness of choice experiments as a method to elicit individuals' WTP for an environmental good by performing tests of internal and external validity. Choice experiments are now being used more frequently in valuation of environmental goods and so far the results are promising (Adamowicz *et al*. 1994, 1995, Johnson *et al*. 1997, and Layton and Brown 1998). Previous tests of validity of choice experiments have focused on internal validity using private goods, for example tests of consistency and rationality of responses for different response formats and designs (*e.g.* Ben-Akiva *et al*. 1992, and Leigh *et al*. 1984).

---

[6] Example of previous studies are hunting permits (Bishop and Heberlein 1979, 1990), irradiated meat (Fox *et al*. 1995), paintings (Neill *et al*. 1994, Loomis *et al*. 1996), books (Frykblom 1997), chocolates (Johannesson *et al*. 1998), and electric juice makers and chocolates (Cummings *et al*. 1995). Note, the degree of overstatement of hypothetical WTP varies substantially between the studies.

In transport economics, validity tests are either comparative studies with both hypothetical choice/ranking data and revealed preference data (*e.g.* Benjamin and Sen 1983, Louviere *et al*. 1981), or comparisons of predicted market shares from hypothetical choice/ranking studies with observed market shares (*e.g.* Bradley and Gunn 1990, Wardman 1988). To our best knowledge, there have not been any previous tests of external validity of choice experiments using environmental goods.

The objective of this paper is to test the feasibility of applying choice experiments to donations for environmental projects. We concentrate on two issues: (i) if there is a difference with respect to the preferences between the hypothetical and the actual experiment (external validity), (ii) if preferences are stable and transitive in both the hypothetical and the actual choice experiment (internal validity).

The paper is organised in the following way. In Section 2, we describe the theoretical and the empirical model, and the design of the experiment. In Section 3 we present and discuss the results. Finally, in Section 4 we make recommendations for future research.

2. Method

Our theoretical model is based on the premise that an individual faces a choice set with I alternatives. Let us assume that the utility function for a given alternative i consists of a systematic part, V_i, and a stochastic part, ε_i. If the choice probability of alternative i is equal to or greater than the utilities for all other alternatives in the choice set we have:

$$P\{i \mid I\} = P\{V_i + \varepsilon_i \geq V_j + \varepsilon_j; \forall j \in I\} = P\{\varepsilon_i - \varepsilon_j \geq V_j - V_i; \forall j \in I\}. \tag{1}$$

Let us assume that the error terms are independently and identically distributed type I extreme value with scale parameter μ. Given an additive utility function and the distributional assumptions on the error terms, the probability of choosing alternative i becomes:

$$P\{i \mid I\} = \frac{\exp(\mu V_i)}{\sum_{j \in I} \exp(\mu V_j)}. \tag{2}$$

In order to test for external validity, both a hypothetical and an actual choice experiment are conducted. In each experiment, the respondents make repeated pair-wise choices. Each alternative in the choice set is characterised by three attributes; the amount of money the respondent received, donation paid to an environmental project, and the type of environmental project. The advantage of using environmental projects is that they are public goods, or quasi-public goods, at least as long as they are not linked to membership in an environmental organisation. In the experiment it was therefore stressed that donations would be paid anonymously. We used donations to three different environmental projects: (1) The Rainforest, (2) The Mediterranean and (3) The Baltic Sea. All projects are currently managed by the World Wildlife Foundation (WWF), and we described each project and their main purpose to the respondents both orally and in the questionnaire (see Appendix 1 for a description of the projects).

In the experiments, each of the three attributes has three possible levels, and a full factorial design has 27 combinations of attribute levels. Given this candidate set, we wish to design experiments where each respondent makes 14 pair-wise choices. There are several methods available for designing choice experiments (*e.g.* Anderson and Wiley 1992, Krieger and Green 1991, and Louviere and Woodworth 1983). However, these methods rely on the theory of efficient design for linear models. Discrete choice-models are non-linear in the parameters. Consequently, if we wish to adopt any of the methods above, we have to assume that an efficient design for a linear model is an efficient design for a discrete choice-model. Huber and Zwerina (1996) identify four principles for efficient choice designs: (i) orthogonality, (ii) level balance, (iii) minimal overlap, and (iv) utility balance. However, as Zwerina *et al*. (1996) note, for most cases it is impossible to create choice designs that satisfy all of these principles. Instead they develop a computerised search approach aiming at minimising the D-error.[7] This design-approach for non-linear models also has the advantage of allowing incorporation of anticipated parameter vectors. We apply the search algorithm presented in Zwerina *et al*. (1996) for the design of our experiments. In the final designs, based on the results of

---

[7] Where D-error = $\left| \Sigma \right|^{1/K}$, Σ is the covariance matrix of the maximum likelihood estimator, and K is the number of parameters.

a pilot study, we assume a marginal rate of substitution of five between money the respondent received and donation paid. Further, we assume no interaction effects between the variables and thus the designs are only based on main effects. There is a risk of compliance bias in this type of experiment. That is, the respondents may wish to show that they are consistent by reproducing their answers from the hypothetical experiment in the actual experiment, even if these are not based on their true preferences. In order to reduce the risk of such behaviour, we use another but equally efficient design in the actual experiment. Thus, we create two different designs, one for the hypothetical and one for the actual experiment, where both designs have the same D-efficiency (see Appendix 2). Four of the situations are the same in the two designs. Furthermore, the responses were made anonymously; thus there were no possibility for the respondents to reveal to us that they were consistent.

In addition to the 14 sets created by the optimal design technique we add two choice sets in each experiment since we want to test for transitivity. We test for transitivity by choice sets 6, 15 and 16 in the hypothetical experiment and choice sets 13, 15 and 16 in the actual experiment (see Appendix 2).[8] There is also a question as to whether or not preferences are stable throughout the experiment. The responses may depend on the order of the questions, or the preferences may change during the experiment. In order to test for this, we divide the 16 choice sets in two sequences {A,B}, with eight choice sets in each sequence. Half of the respondents received the sets in the order {A,B}, and half received the sets in the order {B,A}. This split allows us to test for stability of the preferences. We use only the answers in sequence B, choice sets 7-14 in Appendix 2, to test for stability because the choice sets created for the test of transitivity (sets 15 and 16) are allocated in sequence A.

---

[8] Note that the order of the sets presented in the Appendix is not necessarily the order of the sets presented in the questionnaires to respondents since we have changed the order across questionnaires in order to reduce order effects and to allow for test of stability.

The respondents were instructed to make 16 hypothetical pair-wise choices (see Appendix 1 for wordings of the questionnaire). An example of a choice situation is presented in Table 1 below.

Table 1. *Description of a choice situation.*

Your Choice		
	Alternative 1	**Alternative 2**
Money	35 kr[9]	50 kr
Donation	175 kr	150 kr
Environmental Project	Rainforest	Baltic Sea

After they had completed the first part of the experiment (the hypothetical part), we asked the respondents to make 16 pair-wise choices of the same type as before (the actual part). The difference in this part of the experiment was that one of the choices they had made would be drawn randomly as the actual choice set. The respondents' choice in this randomly drawn choice set resulted in a donation made anonymously by us into the postal account of the chosen environmental project and the money income chosen was then paid to the respondents in cash. It should be noted that the respondents were not aware of the actual experiment until they had finished the hypothetical part.

3. Data and Results

We recruited respondents for the experiment among undergraduate students in economics at Göteborg University and Karlstad University, Sweden. The experiment was announced at a lecture immediately prior to the lecture during which the experiment took place. It was clearly stated that participation was voluntary. In total, all 35 students approached participated, thus the data analyses are based on 490 observations. We conducted the choice experiment in several steps and all respondents answered and returned the questionnaires at each step before proceeding to the next step; (1) the respondents answered a questionnaire about socio-economic factors, attitude and knowledge questions about the environment, (2) hypothetical experiment and (3) actual experiment.

---

[9] 1 US dollar corresponds to 7.90 Swedish kronor using the exchange rate of May 1998.

In the experiment, two of the respondents did not trade between the variables donation and money, *i.e.* they had lexicographic preferences at least in the intervals presented in the experiments. One respondent opted for the alternative with highest income in both experiments, whilst another opted for the alternative with the highest donation in both experiments. It should be noted that neither of these two respondents changed their behaviour between experiments. Explanations for lexicographic preferences in choice experiments may be that it is a simple way of solving this type of exercise, an indication that the levels are not sufficiently differentiated to ensure them to trade or that their preferences are genuinely lexicographic. However, the latter seems unlikely to be true for all intervals.

3.1 Test of External Validity

In the econometric analyses of the data we use a conditional logit model, with choice-specific constants, given an indirect utility of alternative *i* of the following form:

$$V_i = \alpha_i + \beta_1 Money + \beta_2 Donation + \varepsilon_i , \tag{3}$$

where the choice specific constant for The Baltic Sea is set to zero (reference case). Using Limdep 7.0 we estimate equation (3) for the hypothetical experiment, the actual experiment and both experiments together (pooled). In the estimations we exclude the two choice sets we created in order to test for transitivity. A heteroskedastic extreme value logit model (Bhat 1995) was also estimated, but the conditional logit model could not be rejected in any of the estimated models at the 5% significance level. Table 2 provides the estimated results for all three models.[10]

---

[10] We estimated several models with socio-economic variables included, but these variables were not significant. This might be explained by the homogeneity of the sample.

Table 2. *Estimated utility functions for hypothetical and actual experiment, and pooled data.*

Variable	Hypothetical		Actual		Pooled	
	Coefficient (s.e.)	P-value	Coefficient (s.e.)	P-value	Coefficient (s.e.)	P-value
Money	0.033 (0.010)	0.000	0.025 (0.009)	0.006	0.028 (0.007)	0.000
Donation	0.021 (0.003)	0.000	0.014 (0.003)	0.000	0.017 (0.002)	0.000
Mediterranean	-0.885 (0.148)	0.000	-0.923 (0.149)	0.000	-0.891 (0.103)	0.000
Rainforest	-0.088 (0.145)	0.558	0.074 (0.141)	0.599	0.027 (0.10)	0.786
Log-likelihood	-282.840		-293.164		-579.316	
No. of obs.	490		490		980	
Marginal WTP-donation	0.629 (0.112)		0.559 (0.121)		0.602 (0.085)	

The choice specific constant for the project The Mediterranean is negative and significant in both experiments indicating that in both experiments this alternative was less preferred than The Baltic Sea project. The choice specific constant for the project The Rainforest has different signs in the two experiments, but it is insignificant in both cases. We perform three tests of external validity. First we test the hypothesis if the coefficients of the hypothetical and the actual experiment can be restricted to be equal, and in a likelihood ratio test, we cannot reject this hypothesis (p-value = 0.157).[11] Second, we test whether or not the marginal WTP differs between the hypothetical and actual experiment. We calculate the marginal WTP for donations as the ratio between the coefficient for donation and the coefficient for money. We estimate the marginal WTP and its standard errors by a jackknife procedure (see *e.g.* Shao and Tu 1995), where we treat each respondent as one observation. We cannot reject the hypothesis that marginal WTP is equal between the experiments using a t-test where the experiments are assumed to be independent (p-value = 0.403).

---

[11] Likelihood ratio: $\lambda = -2[\ln L_{Pooled} - (\ln L_{Hypo} + \ln L_{Actual})] = 6.624 \sim \chi_4^2$.

Third, a perhaps weaker of test of external validity is to compare the choices in the four choice situations which were exactly the same in both the hypothetical and the actual experiment. In total there are 140 comparable choice situations. In Table 3 we present the inconsistent responses.

Table 3. *Inconsistent answer between hypothetical and actual experiment.*

No. of responses changed	Changed to more money in actual	Changed to higher donation level in actual	Changed to both higher income and donation
1	7 respondents	5 respondents	
2	1 respondent	1 respondent	2 respondents

Eighteen of the respondents answered in exactly the same way in both experiments, while 16 of them changed in at least one choice set. Four of the respondents had inconsistent choices on two choice sets, while twelve respondents had inconsistent answers in only one choice set. There is no clear pattern to the inconsistency as some respondents went for higher income while others went for higher donation in the actual experiment. In total, there are 20 inconsistent choices. In a one-tailed binomial test we reject the hypothesis of random consistency, i.e. we reject the hypothesis that the proportion of inconsistent responses is 0.5 (p-value = 0.00).[12] In a one-tailed binomial test however we can also reject the hypothesis that the proportion of inconsistent responses is zero (p-value = 0.00). Thus, the experiments indicate a potential problem with inconsistent responses.[13]

3.2 Test of Internal Validity

In the experiments we built in two tests of internal validity: (1) transitivity and (2) stability. In the test of transitivity, only one respondent failed in both experiments and one respondent failed in the actual experiment. Thus, there does not seem to be any large problem of intransitivity in the experiments.

As described in Section 2 we test for stability of preferences by estimating the utility functions on the choices in subsequence {B}.

---

[12] See *e.g.* Siegel and Castellan (1988).
[13] Note that the significance levels in the binomial tests are sensitive in the case of few observations.

The test for stability is performed by comparing the choices when the subsequence {B} was given as the first eight sets and as the last eight sets. We test if the preferences from the first and the last eight sets are significantly different from the pooled set. The results of these estimations are presented in Table 4.

Table 4. *Test for stability of preferences in hypothetical and actual experiment based on subsequence {B}.*

	Hypothetical			Actual		
	First 8	Last 8	Pooled	First 8	Last 8	Pooled
Variable	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)	Coefficient (p-value)
Money	0.007 (0.698)	0.067 (0.000)	0.038 (0.002)	0.019 (0.320)	0.017 (0.336)	0.018 (0.170)
Donation	0.017 (0.002)	0.029 (0.000)	0.023 (0.000)	0.014 (0.010)	0.012 (0.023)	0.013 (0.001)
Mediterranean	-0.898 (0.000)	-0.909 (0.000)	-0.863 (0.000)	-0.727 (0.021)	-1.290 (0.000)	-1.037 (0.000)
Rainforest	-0.2376979 (0.416)	-0.069 (0.800)	-0.144 (0.465)	0.094 (0.725)	-0.082 (0.732)	-0.005 (0.977)
Log-likelihood	-81.189	-87.397	-173.015	-81.750	-98.410	-181.460
No. of obs.	245	245	490	245	245	490

We cannot in a likelihood-ratio test reject the hypothesis of stable preferences in the experiments (in the hypothetical experiment p-value = 0.065 and in the actual p-value = 0.627).[14] These results indicate that there may be a problem of instability in the hypothetical experiment. An explanation for this may be that respondents are inexperienced in answering this type of excercise. For example Sælensminde (1998) finds that inconsistent answers in hypothetical choice experiments usually occur in the first choice sets. Another explanation may be that preferences are not stable in a hypotethical experiment.

---

[14] Likelihood-ratio hypothetical: $\lambda = -2[\ln L_{Pooled} - (\ln L_{First8} + \ln L_{Last8})] = 8.858 \sim \chi^2_4$.

Likelihood-ratio actual: $\lambda = -2[\ln L_{Pooled} - (\ln L_{First8} + \ln L_{Last8})] = 2.6 \sim \chi^2_4$.

4. Discussion

In this paper we test for both external and internal validity of a choice experiment with donations to environmental projects. The result indicates that we cannot reject the hypothesis that the preferences are the same in the two experiments. Further, the results indicate an insignificant discrepancy between marginal WTP for donations in the hypothetical and the actual experiments. Thus, our tests of external validity indicate that choice experiments do not seem to suffer from problems with overstatement in a hypothetical context. Further, internal tests indicate stability and transitivity of preferences. However, one problem in our experiments was that a number of responses were inconsistent over the choice experiments.

Our results suggest that individuals need some "warm-up" choices since there is some instability of preferences in the first 8 choice sets in the hypothetical experiment. There is of course a risk of compliance bias, in the sense that the respondents wish to be consistent throughout the experiments. In order to minimise the risk of simply reproducing the same choices, most of the choice sets differed between the hypothetical and the actual experiment. One problem with conducting experiments is to define a suitable public good. Using donations may result in a problem with "warm glow" (Andreoni 1990) or "purchase of moral satisfaction" (Kahnemann and Knetsch 1992). This can be a source for bias if the degree of "warm glow" differs between the hypothetical and the actual context. The degree of "warm glow" may be higher in the hypothetical context since it is costless in that situation, but on the other hand, it can be higher in the actual context, since donation is actually paid in that situation. Note that in our experiment donations were paid anonymously.

Overall, our results indicate that choice experiments seem to be a promising method for eliciting preferences for public goods, especially as there was no indication of overstatement in the hypothetical experiment. This contrasts the empirical results on external validity of CVM, where in general overstatement of WTP has been found. There are other strengths with a choice experiment such as reduction of the risk of yea/nay-saying and a built-in test of sensitivity to scope (Hanley *et al*. 1998). However, further research on choice experiments is needed especially with emphasis on tests of external validity.

References

Adamowizc, W., P. Boxall and J. Louviere (1995), 'Stated preference approaches for measuring passive use values: Choice experiments versus contingent valuation', Staff Paper 95-03, Department of Rural Economy, University of Alberta.

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

Anderson, D. and J. Wiley (1992), 'Efficient choice set designs for estimating cross-effects models', *Marketing Letters* **3**, 357-370.

Andreoni, J. (1990), 'Impure altruism and donations to public goods: A theory of warm glow giving', *Economic Journal* **100**, 464-77.

Ben-Akiva, M., T. Morikawa and F. Shiroishi (1992), 'Analysis of the reliability of preference ranking data', *Journal of Business Research* **24**, 149-164.

Benjamin, J. and L. Sen (1982), 'Comparison of the predictive ability of four multiattribute approaches to attitudinal measurement', *Transportation Research Record* **890**, 1-6.

Bishop, R. and T. Heberlein (1979), 'Measuring values of extra-market goods: Are indirect measures of value biased?', *American Journal of Agricultural Economics* **61**, 926-930.

Bishop, R. and T. Heberlein (1990), 'The contingent valuation method', in Johnson, R. and G. Johnson (eds.), *Economic Valuation of Natural Resources: Issues, Theory and Applications*, Boulder: Westview Press.

Bhat, C. (1995), 'A heteroskedastic extreme-value model of travel mode choice', *Transportation Research* **29**, 471-483.

Bradley, M. and H. Gunn (1990), 'Stated preference analysis of values of travel time in the Netherlands', *Transportation Research Record* **1285**, 78-88.

Brown, T., P. Champ, R. Bishop and D. McCollum (1996), 'Which response format reveals the truth about donations to a public good?', *Land Economics* **72**, 152-166.

Carson, R. (1997), 'Contingent valuation surveys and tests of insensitivity to scope', in Kopp, R., W. Pommerhene and N. Schwartz (eds.), *Determining the Value for Non-Marketed Goods: Economic, Psychological and Policy Relevant Aspects of*

Contingent Valuation Methods, Boston: Kluwer Academic Press.

Carson, R., N. Flores, K. Martin, and J. Wright (1996), 'Contingent valuation and revealed preference methodologies Comparing the estimates for quasi-public goods' *Land Economics* **72**, 80-99.

Carson, R., J. Wright, N. Carson, A. Alberini, and N. Flores (1995), *A Bibliography of Contingent Valuation Studies and Papers*, La Jolla: NRDA Inc.

Cummings, R., G. Harrison and E. Rutstrom (1995), 'Home-grown values and hypothetical surveys: Is the dichotomous choice approach incentive compatible?', *American Economic Review* **85**, 260-266.

Diamond, P. and J. Hausman (1994), 'Contingent valuation: Is some number better than no number?', *Journal of Economic Perspectives* **8**, 45-64.

Fox, J., J. Shogren, D. Hayes and J. Kliebenstein (1995), 'CVM-X: Calibrating contingent values with experimental auction markets', Working paper, Department of Economics and Finance, University of Wyoming.

Frykblom, P. (1997), 'Hypothetical question modes and real willingness to pay', *Journal of Environmental Economics and Management* **34**, 275-287.

Hanley, N., R. Wright, and V. Adamowicz (1998), 'Using choice experiments to value the environment: Design issues, current experience and future prospects', *Environmental and Resource Economics* **3-4**, 413-428.

Huber, J. and K. Zwerina (1996), 'The importance of utility balance in efficient choice designs', *Journal of Marketing Research* **33**, 307-317.

Johannesson, M., B. Liljas and P.-O. Johansson (1998), 'An experimental comparison of dichotomous choice contingent valuation questions and real purchase decisions', *Applied Economics* **30**, 643-647.

Johnson, F. M. Ruby and W. Desvousges (1997), 'Valuing stated preferences for health benefits of improved air quality: Results of a pilot study', Working Paper No. T-9702, Triangle Economic Research.

Kahneman, D. and J.L. Knetsch (1992), 'Valuing public goods: The purchase of moral satisfaction', *Journal of Environmental Economics and Management* **22**, 57-70.

Krieger, A. and P. Green (1996), 'Designing Pareto optimal stimuli for multiattribute choice experiments', *Marketing Letters* **2**, 337-348.

Leigh, T., D. MacKay and J. Summers (1984), 'Reliability and validity of conjoint analysis and self-explicated weights: A comparison', *Journal of Marketing Research* **21**, 456-462.

Layton, D.F. and G. Brown (1998), 'Heterogenous preferences regarding global climate change', Working paper, Department of Economics, University of Washington.

Loomis, J., T. Brown, B. Lucero and G. Peterson (1996), 'Improving validity experiments of contingent valuation methods: results of efforts to reduce the disparity of hypothetical and actual willingness to pay', *Land Economics* **72**, 450-461.

Louviere, J., D. Henley, G. Woodworth, R. Meyer, I. Levin, J. Stoner, D. Curry and D. Anderson (1981), 'Laboratory simulation versus revealed preference methods for estimating travel demand models', *Transportation Research Record* **794**, 42-51.

Louviere, J. and G. Woodworth (1983), 'Design and analysis of simulated consumer choice of allocation experiments: A method based on aggregate data', *Journal of Marketing Research* **20**, 350-367.

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

Navrud, S. (1992), 'Willingness to pay for preservation of species: An experiment with actual payments, in S. Navrud (ed.), *Pricing the European Environment*, Oslo: Scandinavian University Press.

Neill, H., R. Cummings, P. Ganderton, G. Harrison and T. McGuckin (1994), 'Hypothetical surveys and real economic commitments', *Land Economics* **70**, 145-154.

Sælensminde, K. (1998), 'The impact of choice inconsistencies on the valuation of travel time in stated choice studies', Paper presented at World Congress of Environmental and Resource Economics in Venice, June 25-27 1998.

Seip, K. and J. Strand (1992), 'Willingness to pay for environmental goods in Norway: A contingent valuation study with real payments', *Environmental and Resource Economics* **2**, 91-106.

Siegel, S. and N. Jr. Castellan (1988), *Nonparametric Statistics for the Behavioural Sciences*, New-York: McGraw-Hill.

Shao, J. and D. Tu (1995), *The Jackknife and Bootstrap*, Berlin: Springer-Verlag.

Wardman, M. (1988), 'A comparison of revealed and stated preference models of travel behaviour', *Journal of Transport Economics and Policy* **22**, 71-91.

Zwerina, K., J. Huber and W. Kuhfeld (1996), 'A general method for constructing efficient choice designs', Working paper, Fuqua School of Business, Duke Univ.

Appendix 1. Questionnaire

<u>Hypothetical experiment</u>

In this part of the experiment we are interested in people's preferences for different types of environmental projects, and donations to these. We will present sixteen different choice sets. For each choice set we want you to tick which of the two alternatives you would prefer. Each alternative has three attributes:

Your money: The amount of money which you could receive, and in this case the money would be freely available.

Donation: The amount of money which would be paid anonymously into a postal giro account for a certain environmental project.

Environmental project: There are three possible environmental projects. All projects are run by World Wildlife Foundation:

> **Rainforest:** Preservation of tropical rainforest. The fund supports, among other things, projects for protection of land areas, education of staff in the national parks, and development of environmentally friendly forestry methods.

> **Mediterranean:** Preservation of the Mediterranean area. The fund supports projects with focus on environmental education, sea, coast, forest and endangered species. Among other things, projects for the sea turtle and new environmental protection areas are supported.

> **Baltic Sea:** Environmental work in the Baltic Sea. The fund supports, among other things, projects for new environmental protection areas and works for reduction of emissions of nitrogen oxides and phosphorus, and waste oil in the area.

In this part you will answer questions similar to those in the previous section. We will present sixteen different choice sets. Afterwards one of the choice sets will be drawn randomly as a so called actual choice set. It means that you will receive in cash the amount of money you have chosen, and the donation you have chosen will be paid into the postal giro account of the environmental project.

For each choice set we want you to tick which of the two alternatives you would prefer. Each alternative has the same attributes as before, that is:

Your money: The amount of money which you could receive, and in this case the money would be freely available.

Donation: The amount of money which would be paid anonymously into a postal giro account for a certain environmental project.

Environmental project: There are three possible environmental projects. All projects are run by World Wildlife Foundation:

> **Rainforest:** Preservation of tropical rainforest. The fund supports, among other things, projects for protection of land areas, education of staff in the national parks, and development of environmentally friendly forestry methods.

> **Mediterranean:** Preservation of the Mediterranean area. The fund supports projects with focus on environmental education, sea, coast, forest and endangered species. Among other things, projects for the sea turtle and new environmental protection areas are supported.

> **Baltic Sea:** Environmental work in the Baltic Sea. The fund supports, among other things, projects for new environmental protection areas and works for reduction of emissions of nitrogen oxides and phosphorus, and waste oil in the area.

Appendix 2. Designs

	Hypothetical			Actual		
Set	**Money**	**Donation**	**Project**	**Money**	**Donation**	**Project**
1	65	150	Baltic Sea	50	200	Mediterranean
1	50	200	Rainforest	65	100	Rainforest
2	35	150	Mediterranean	65	150	Mediterranean
2	65	100	Rainforest	35	200	Rainforest
3	50	100	Baltic Sea	35	200	Mediterranean
3	35	200	Mediterranean	50	150	Baltic Sea
4	65	100	Mediterranean	35	200	Baltic Sea
4	50	150	Rainforest	65	150	Rainforest
5	50	200	Baltic Sea	50	100	Mediterranean
5	65	150	Rainforest	35	150	Baltic Sea
6	35	200	Baltic Sea	65	100	Mediterranean
6	50	150	Rainforest	50	200	Rainforest
7	65	100	Rainforest	65	100	Rainforest
7	50	150	Mediterranean	35	150	Mediterranean
8	65	150	Mediterranean	65	150	Rainforest
8	35	200	Rainforest	50	200	Baltic Sea
9	65	150	Baltic Sea	50	200	Rainforest
9	35	200	Mediterranean	65	150	Baltic Sea
10	50	200	Baltic Sea	35	200	Rainforest
10	65	100	Mediterranean	50	100	Mediterranean
11	50	100	Baltic Sea	35	150	Rainforest
11	35	150	Rainforest	65	100	Baltic Sea
12	65	100	Baltic Sea	50	150	Mediterranean
12	50	200	Mediterranean	65	100	Baltic Sea
13	35	200	Baltic Sea	65	100	Mediterranean
13	50	100	Mediterranean	50	150	Baltic Sea
14	35	150	Baltic Sea	35	200	Mediterranean
14	65	100	Mediterranean	50	100	Rainforest
15	35	200	Baltic Sea	35	200	Mediterranean
15	65	100	Mediterranean	65	100	Rainforest
16	50	150	Rainforest	50	150	Baltic Sea
16	65	100	Mediterranean	65	100	Rainforest
D-efficiency			4.23331			4.23331