Environmental and Resource Economics **14:** 95–117, 1999.

The Validity of Environmental Benefits Transfer: Further Empirical Testing

ROY BROUWER[1] and FRANK A. SPANINKS[2]
[1]*Centre for Social and Economic Research on the Global Environment (CSERGE), University of East Anglia, Norwich and University College London, UK;* [2]*Institute for Environmental Studies (IVM), Vrije Universiteit, Amsterdam, The Netherlands (email: r.brouwer@uea.ac.uk.)*

Accepted 30 June 1998

Abstract. This paper provides further empirical evidence of the validity of environmental benefits transfer based on CV studies by expanding the analysis to include control factors which have not been accounted for in previous studies. These factors refer to differences in respondent attitudes. Traditional population characteristics were taken into account, but these variables do not explain why respondents from the same socio-economic group may still hold different beliefs, norms or values and hence have different attitudes and consequently state different WTP amounts. The test results are mixed. The function transfer approach is valid in one case, but is rejected in the 3 other cases investigated in this paper. We provide further evidence that in the case of statistically valid benefits transfer, the function approach results in a more robust benefits transfer than the unit value approach. We also show that the equality of coefficient estimates is a necessary, but insufficient condition for valid benefit function transfer and discuss the implications for previous and future validity testing.

Key words: benefits transfer, contingent valuation, environmental valuation, validity testing

JEL classification: C12, C42, D64, H41, Q21

1. Introduction

Environmental benefits transfer, i.e. transposing monetary environmental values estimated at one site (study site) to another (policy site), has been discussed and applied extensively. A recent application is the transfer exercise by Costanza et al. (1997) in which monetary point estimates from environmental valuation studies carried out at different sites in the Developed and Developing World using different valuation techniques are extrapolated to the world's ecosystem services and natural capital.

However, very little published evidence exists about the validity of transferring point estimates of the monetary value of environmental benefits or complete benefit functions from one site to another (see Table I for a literature overview). Moreover, no single study, based on travel cost (TC) or contingent valuation (CV) data, has unequivocally confirmed the validity of benefits transfer, not even under seemingly ideal circumstances for valid benefits transfer (e.g. Bergland et al. 1995). But,

Table I. Previous Studies Testing the Validity of Benefits Transfer

Study	Valuation technique	Environmental good	Transfer samples	Amount of control in terms of explanatory variables included	Transfer error (%)[a]
Loomis (1992)	zonal TC model	sport fishing	10	– travel distance – fish harvest – fishing quality at substitutes	5–40 5–15
Parsons and Kealy (1994)	random utility TC model	water quality improvements	2	– travel costs (including time) – area size – depth of lake – area accessibility – water quality – main recreational use – household income (dummy)	4–34 1–75
Loomis et al. (1995)	zonal TC model	reservoir based recreation	3 districts 10, 8, 8 sites	– travel costs – area size – availability of substitutes – population size – median age	— 1–475
Bergland et al. (1995)	iterative bidding CV	water quality improvements	2	– bid levels – main recreational use – education (2 dummies) – age (dummy) – user (dummy)	25–45 18–41
O'Doherty (1996)	open-ended CV	green space	2	– green space under development – place of residence	unknown

Table I. Continued

Study	Valuation technique	Environmental good	Transfer samples	Amount of control in terms of explanatory variables included	Transfer error (%)
				– income group – awareness of substitutes – main recreational use	
Downing and Ozuna (1996)	dichotomous choice CV	saltwater fishing	8	– bid level – time period (dummy)	1–34[c] —
Kirchhoff et al. (1997)[b]	payment card CV	white water rafting	2 pairs of sites	– perceived flow level[1)] – expenses/person[1)] – use intensity (dummy)[1+2)] – main reason for visit[1+2)] – place of residence[1+2)] – household income[1+2)] – sex[1)] – age[2)] – education[2)]	24–56 6–228

[a] Minimum and maximum absolute transfer errors found in the studies. The first range refers to the difference between observed average unit values at the study and policy site, and the second range to the difference between the observed average value at the policy site and predicted average value at the policy site based on the benefit function transferred from the study site.
[b] The superscripts [1)] and [2)] refer to the 2 different bid functions that were tested.
[c] Own calculations. In the case of dichotomous choice CV, no observed average sample mean exists. The observed average unit value has to be calculated from the bid function. In the case of Bergland et al., the average ‘observed’ unit values at the study and policy sites were calculated based on the complete bid functions, including other explanatory variables besides the bid levels. This means that the calculated ‘observed’ average Unit values control for significantly influencing factors at ’the specific sites and hence do not correspond with the usual average unit value. In the case of Downing and Ozuna, the range refers to min-max errors within the same year; over years, the maximum error increases to 55%.

as more information about factors influencing environmental valuation outcomes becomes available, for example from the increase in meta-analytical research in the field of environmental valuation (see Brouwer et al. 1997 for an overview), transfers across populations and sites could become more reliable, using secondary information only or supplementing secondary information with primary data.

Even though the transfer of benefit functions has been claimed to be more robust than the transfer of point estimates in terms of transfer errors (e.g. Loomis 1992; Kirchhoff et al. 1997), as Table I shows, the errors in transferring monetary value estimates for seemingly similar environmental goods over sites can be as large as 56% in the case of average unit value transfer and 475% in the case of benefit function transfer. In the former case, the ‘*observed*’ average unit values at two sites are compared with each other. In the latter case, the ‘*observed*’ average unit value at the policy site is compared with the ‘*calculated*’ average unit value at the policy site based on the benefit function transferred from the study site to the policy site and the average values of the explanatory variables included in the transferred benefit function at the policy site.[1]

Downing and Ozuna (1996) and Kirchhoff et al. (1997) found that even if benefit functions are the same at the study and policy site, a prerequisite for valid benefit function transfer, the ‘*observed*’ and ‘*calculated*’ unit values may not. In both studies this was explained by the non-linear models used, which may have introduced possible ‘*asymmetries leading to the divergence between statistically similar benefit functions and their respective benefit estimates*’ (Downing and Ozuna, 1996).[2]

In this paper, we will provide further empirical evidence of the validity of environmental benefits transfer based on CV studies and show that the equality of benefit functions is a necessary, but insufficient condition for valid benefit function transfer. We will furthermore expand the analysis to include explanatory factors which have not been accounted for in previous studies. These factors refer to possible differences in respondent attitudes and behaviour and tell us a little bit more about their social-psychological and perhaps even cultural background. Traditional population characteristics accounted for in previous studies such as a respondent’s (household) income, education level, age or sex do not explain why respondents from the same socio-economic group may hold different beliefs, norms or values and hence have different attitudes and consequently state different willingness to pay (WTP) amounts. CV measures behavioural intentions, not behaviour itself, and an individual’s intention to behave in a certain way depends, among other things, upon his or her attitude towards this behaviour (Fishbein and Ajzen, 1975).

Hence, besides traditional socio-economic variables such as respondent household income, household structure, education level and sex, we will also include variables such as respondents’ expressed feelings and actual behaviour towards paying for environmental protection in general, their familiarity with the areas involved and their perception of the importance of carrying through the proposed environmental protection or enhancement schemes in these areas. In this way, our main aim is to further assess the viability of benefit function transfer based on CV data.

The two studies included in the analysis presented in this paper were carried out independently of each other, but comply with the five selection criteria outlined by Boyle and Bergstrom (1992) and Desvousges et al. (1992) in the special benefits transfer issue of the Water Resources Research journal for selecting among studies for reliable benefits transfer. These criteria focus on the environmental goods involved, the sites in which the goods are found, their beneficiaries and study quality.[3]

The studies in this paper provide adequate data collected on the basis of surveys reflecting the CV state-of-the-art, meeting many of the guidelines set out by the influential NOAA Panel (Arrow et al. 1993). Both studies collected information on a number of important socio-economic and demographic characteristics of randomly selected Dutch citizens. Most favourable for valid benefits transfer is the fact that both studies concentrated on the same type of environmental goods in similar areas, i.e. the amenities found on agricultural peat meadow land in the Netherlands. Moreover, both studies used the same reference and target levels in the valuation scenarios for the relevant attributes of the amenities involved and constructed a similar CV market, which is rarely the case in practical benefits transfer exercises based on CV. Hence, the amount of control necessary for valid benefits transfer is expected to be found primarily in population sample characteristics, not in differences in the environmental goods and benefits involved, site characteristics or research designs.

In the remainder of this paper, we will first outline the statistical procedures for testing the validity of benefits transfer (Section 2). This will be followed by a description and discussion of the two studies used in the analysis (Section 3). Section 4 presents the results and Section 5 concludes.

2. Validity Tests

The validity of environmental benefits transfer across populations from one site to another can be tested in a number of ways. The transfer of benefit functions is preferred to the transfer of average unit values, because effectively more information can be transferred (Pearce et al. 1994). Bergland et al. (1995) recommend that both valuation functions and point estimates should be transferable and advocate the testing of 4 hypotheses which are summarised in Table II.

First, the equality of average WTP amounts ($\overline{\text{WTP}}$) at the policy site (the site to which the benefits have to be transferred) and study site (the site from which they have to be transferred) can be tested with the help of parametric or non-parametric tests. Which test is most appropriate depends upon the (assumed) underlying distribution of the WTP amounts. For example, the *t*-test assumes that the WTP amounts are drawn from a normal distribution, whereas the Mann-Whitney test,

Table II. Validity Tests for Benefits Transfer

Null hypothesis	Test
1a. $\overline{\text{WTP}}_p = \overline{\text{WTP}}_s$	*t*-test/Mann-Whitney test
1b. Distribution $\text{WTP}_{p,i}$ = Distribution $\text{WTP}_{s,i}$	Kolmogorov-Smirnov test
2a. $b_p = \hat{b}_s$	Lagrange Multiplier/Wald test
2b. $\hat{\sigma}_p^2 = \hat{\sigma}_s^2$	Chow/Likelihood Ratio test
3a. $\hat{b} = \hat{b}_p = \hat{b}_s$	Lagrange Multiplier/Wald test
3b. $\hat{\sigma}^2 = \hat{\sigma}_p^2 = \hat{\sigma}_s^2$	Chow/Likelihood Ratio test
4a. $\overline{\text{WTP}}'_p = f(\hat{b}_s, X_p) = \overline{\text{WTP}}_p$	*t*-test/Mann-Whitney test
4aa. Distribution $\text{WTP}'_{p,i}$ = Distribution $\text{WTP}_{p,i}$	Kolmogorov-Smirnov test
4b. $\overline{\text{WTP}}'_s = f(\hat{b}_p, X_s) = \overline{\text{WTP}}_s$	*t*-test/Mann-Whitney test
4bb. Distribution $\text{WTP}'_{s,i}$ = Distribution $\text{WTP}_{s,i}$	Kolmogorov-Smirnov test

Clarification:
$\overline{\text{WTP}}$ – Average WTP at policy (p) or study (s) site.
WTP_i – Observed WTP amounts.
$\text{WTP}'_{p,i}$ – WTP amounts calculated at the policy site by multiplying the vector of coefficient estimates ($\hat{b}_s$) at the study site by the matrix of explanatory variable values (X_p) observcd at the policy site.
$\hat{\sigma}^2$ – Variance explained.

being a non-parametric test, requires no specific assumption regarding the distribution of the WTP amounts.[4] Equality of the distribution of WTP amounts at the two sites is an even more rigorous null hypothesis than the equality of average WTP amounts proposed by Bergland et al. (1995). This will be tested with the help of the Kolmogorov-Smirnov test statistic.

The second and third hypotheses refer to the equality of benefit functions and are tested by comparing the estimated parameters of the benefit function at the policy and study site. For valid benefit function transfer, we are looking for equality of coefficient estimates. If the coefficients of explanatory variables are not the same at different sites, their impact will be different at these sites and hence the model estimated at one site can not be used to predict WTP at another site.

In the second hypothesis, the coefficient estimates at the study and policy site are compared directly. The third hypothesis compares the estimated models at the study and policy site with their pooled model and tests whether the estimated benefit functions originate from a common underlying function. As also noted by Bergland et al. (1995), the third hypothesis is a weaker form for testing the equality of benefit functions than the second. The former does not guarantee that the estimated coefficients at the study and policy site are the same, only that they coincide with the estimates of their pooled model.

Statistical tests used in previous studies to test the second hypothesis are the Chow test (Loomis 1992; Loomis et al. 1995; O'Doherty 1996), the Likelihood Ratio (LR) test (Parsons and Kealy 1994, Kirchhoff et al. 1997) and the Lagrange

Multiplier (LM) test (Bergland et al. 1995), also known as the Score test. The third hypothesis has been tested with the help of the LR test (Bergland et al. 1995; Loomis et al. 1995).

However, the LR and Chow test are tests for a model's global goodness of fit and hence test the equality of explained variance of two or more models.[5] Moreover, the Chow test tests the equality of the sum of squared residuals from two samples compared to their pooled model, which means that it actually tests the weaker hypothesis 3. Equality of explained variance does not necessarily imply equality of coefficient estimates. It only tells us that the explanatory variables included in the model explain the same amount of variability in stated WTP amounts, not which variables contribute to what extent. At two or more sites, different factors may explain different proportions of the variability and, merely by chance, add up to similar explanatory power. On the other hand, if the factors influencing WTP are the same and so is the size of their impact, it is very likely that the explained variance is also the same. A possible reason why this would not be the case could be the presence of outliers affecting the model's goodness of fit. Moreover, in case we extrapolate a benefit function estimated at one site to another site, the variability explained at the former site may differ from the variability explained at the latter with the help of the extrapolated coefficient estimates because of the presence of site specific explanatory variables which are not included in the benefits transfer model.

In this paper, we will rely upon the Wald test, which tests directly for the equality of coefficient estimates and is hence considered an appropriate test statistic for the second and third hypothesis. The test statistic of this test assumes a Chi-squared distribution of the error terms. For a more detailed discussion of the Wald test or any of the other tests, see for example Greene (1990).

Finally, based on the estimated benefit functions, average WTP at the policy site can be calculated using the estimated parameters of the benefit function at the study site and the actual values of the explanatory variables at the policy site. In the fourth hypothesis, this '*calculated*' average WTP at the policy site (or distribution of calculated WTP amounts) is compared with the '*observed*' average WTP at the policy site (or distribution of observed WTP amounts), using the tests mentioned before.

3. The Studies

Contingent valuation was used to estimate the benefits of agricultural wildlife management on Dutch peat meadow land. This technique is believed to be capable of taking into account, in principle, the whole bundle of varying attributes in a spatial area and to measure their use and non-use values. Especially the latter are expected to make up an important part of the economic value of the amenities found on peat meadow land.

Questionnaires complying with Dillman's (1978) '*total design method*' for mail surveys were sent to randomly selected households.[6] Since management agreements in peat meadow areas usually concentrate on the protection of meadow birds and ditch-side vegetation, these elements received most attention in the questionnaires. Except for some minor differences in wording, both studies used the same valuation scenarios.

3.1. THE VALUATION SCENARIOS

The scenarios in the questionnaires describe two future states that could emerge in the agricultural areas involved over the next 10 years, depending on whether or not a number of wildlife protection measures are taken by farmers. The first scenario describes what happens if the current situation continues to exist (the reference situation). Rare meadow bird species will disappear from the area and populations of common species will decline. The flowery ditch-side vegetation will be replaced by common grass. The second scenario reflects an improvement in the situation of meadow birds and plants (the target situation). Rare meadow bird species will still be found in the agricultural areas and the populations of common species will increase. A variety of flowery vegetation will be found in ditch-sides.

Respondents were asked for their WTP to reach the situation as described in the second scenario instead of the first '*business as usual*' scenario. The payment vehicle employed was an annual donation to a private fund. Respondents were guaranteed that the financial resources from this fund would be used exclusively to compensate farmers for their higher costs when implementing the measures needed to reach the situation as described in the second scenario. The scenarios and the WTP questions are available from the authors.

3.2. THE FIRST STUDY

The first study (Spaninks 1993) concentrated on a small peat meadow area (about 500 ha) just to the south of the city of Sneek, located in the province Friesland in the north of the Netherlands. Approximately 40 percent of all Dutch peat meadow land is found in this province (Terwan 1988). The study's main objective was to test for possible range bias, resulting from the use of different payment cards, and temporal embedding.

Three versions of the questionnaire were used. In the first version the highest amount on the payment card was 200 Dutch guilders (D*f*) and respondents receiving this version were asked how much they were willing to contribute to the fund every month. In the second version the highest amount on the card was D*f* 100. The third version again used a payment card with a highest amount of D*f* 200, but respondents were now asked for their annual WTP. All payment cards included an open-ended possibility for those respondents who were willing to pay more than the highest amount shown on the cards.

In May 1993, a total of 1,000 questionnaires were sent to randomly selected households in the city of Sneek; 500 households received the first version of the questionnaire, 250 the second and 250 the third version. Another 250 copies of the first version were sent to randomly selected households in the city of Leeuwarden, the province's capital located approximately 15 kilometres to the north-east of Sneek. This last group was used to test possible distance-decay effects on WTP. Before coming to the WTP question for wildlife preservation on peat meadow land, all versions of the questionnaire first asked respondents for their WTP to solve all environmental problems.

3.3. THE SECOND STUDY

The second study (Brouwer 1995) was carried out in the Alblasserwaard (15,660 ha), situated in the province South Holland in the west of the Netherlands. South Holland is the most densely populated province in the Netherlands. Together with three other adjacent peat meadow areas (Lopikerwaard, Krimpenerwaard and Vijfherenlanden), the Alblasserwaard constitutes the so-called '*Green Heart*' of Holland where 21 percent of Dutch peat meadow land is found (Terwan 1988). The Alblasserwaard is the largest of these areas. The study's main aim was to produce a conservative estimate of the economic value of the amenities found on Dutch peat meadow land for inclusion in cost-benefit analysis.

In May 1994, 2,500 identical questionnaires were sent to randomly selected households in South Holland province. In order to test possible distance-decay effects, another 600 questionnaires were sent to the two most remotely located provinces Friesland (300) and Limburg (300). As mentioned, a substantial amount of Dutch peat meadow land is found in Friesland. Limburg has a distinct landscape and flora and fauna in which no peat meadow land is found at all. It was hypothesised that WTP would be highest in South Holland and lowest in Friesland given the abundant availability of the amenities in Friesland.

Contrary to the first survey, the second survey used colour photographs besides a verbal description to depict the scenarios for the amenities found on peat meadow land. Furthermore, the questionnaire in the second study contained three embedded WTP questions, narrowing down from (1) environmental problems in general to (2) the disappearance of characteristic rural areas in the Netherlands and (3) the preservation of the amenities found in one such characteristic rural area. In each WTP question, respondents were asked for their annual WTP. The first two WTP amounts were elicited using a payment card and the last WTP amount in an open-ended format.

4. Results

4.1. METHODOLOGICAL RESULTS AND BENEFIT TRANSFER SAMPLES

This section briefly reports some individual study results. First, the original response in both studies was about 30 percent. This is reasonable compared with the response rates found in the few other CV studies carried out in the Netherlands (e.g. Hoevenagel 1994). However, the usable response rate was considerably lower (17 percent), the main reason for this being the large number of protest bidders found in both studies: 36 and 32 percent of the original response in studies 1 and 2 respectively. Protest bidders in the two studies mentioned similar reasons why they refused to pay. Most of them considered the preservation of wildlife on peat meadow land a national instead of a local or regional issue, or considered it a government task and hence preferred national taxation instead of a private fund as the most appropriate way of paying.

Secondly, for the purpose of valid benefits transfer, the testing of distance-decay effects in CV is considered equally as important as the identification of the environmental benefits involved or accounting for the characteristics of the beneficiaries. Testing for distance-decay effects in CV helps us to define the relevant population that benefits from the environmental goods or services involved, as in the zonal TC method, and should, in principle, precede the characterisation of the population involved. Information about the geographical demarcation of the relevant population size and consequently the availability and quality of substitutes as one moves further away from the study site is expected to result in more robust benefits transfer modelling.

In the first study, no significant distance-decay effects were found. Significant differences in annual WTP were found between the first and third version of the questionnaire, but not between the first and second version. Hence, range bias could not be detected, only a temporal embedding effect (for more details see Spaninks and Hoevenagel 1995).

In the second study, significant differences were found in stated WTP amounts between the three provinces. As expected, average WTP was highest in South Holland and lowest in Friesland. No significant differences were detected in stated WTP amounts from South Holland respondents living in the research area and at different distances from the area (for more details see Brouwer, 1995).

Given these methodological results, we used three different samples to test the validity of benefits transfer: the usable response (i.e. without protest bidders and non-response) from (1) the first three subsamples in the first study ($n = 220$), (2) the subsample of directly stated annual payments in the first study ($n = 56$) and (3) the South Holland sample ($n = 455$) in the second study. The South Holland sample from the second study was used in order to avoid distance-decay effects. Methodologically speaking, the first study's subsample of directly stated annual payments is the most appropriate sample to use for comparison with the annual payments from the second study. However, the small number of observations and consequently the

Table III. Summary Statistics

	Study 1		Study 2
	Sample 1A Based on monthly and annual payments	*Sample 1B* Based on annual payments	*Sample 2* Based on annual payments
Untruncated Mean[a]	96.6	54.5	74.2
Standard Error	8.9	11.3	5.8
Truncated Mean[b]	78.6 (7)	46.4 (1)	65.3 (5)
Median	60.0	25.0	40.0
Minimum	0	0	0
Maximum	900	500	1000
Skewness[c]	2.8	3.3	4.2
n	220	56	455

[a]In 1995 D*f*.
[b]Between brackets the number of truncated observations.
[c]A positive skewness indicates that the distribution is skewed to the higher values. Zero skewness corresponds to a symmetric distribution such as the normal distribution.

problem of drawing statistically meaningful conclusions led us to use the complete data set from the first study too and to control for temporal embedding effects by the inclusion of a dummy variable to distinguish between respondents who directly stated annual amounts and respondents who stated monthly amounts and whose annual WTP was calculated from these monthly payments.

4.2. TEST RESULTS

4.2.1. *Equality of Unit Values*

Both studies revealed a large number of legitimate zero bidders: 23.6 and 20.6 percent in studies 1 and 2 respectively. Legitimate zero bidders are defined here as those respondents who display a negative attitude towards the preservation of the amenities found on peat meadow land or consider their income (temporary) as inadequate. They could be identified by simply asking respondents in the questionnaire why they refuse to pay.

Although the range of stated WTP amounts is approximately the same in samples 1A and 2, average WTP is D*f* 22 higher in the first study than in the second (Table III). Truncating both data sets, i.e. omitting the observed top WTP values (WTP $\geq$ D*f* 600/year), the difference between mean WTP is considerably smaller. Average WTP is lowest in sample 1B, i.e. the annual stated WTP amounts in the first study. The same applies for the median values: median WTP is highest in sample 1A and lowest in sample 1B. In all three samples the distribution of WTP amounts is skewed to the right.

Table IV. Hypothesis 1: Test Results for Unit Value Transfer under Uncontrolled Conditions

Null hypothesis	K-S test statistic[a]	2-tailed $p \leq$	M-W test statistic[b]	2-tailed $p \leq$
H_0: Distribution $WTP_{1A,i}$ = Normal	3.88	0.001		
H_0: Distribution $WTP_{1B,i}$ = Normal	1.95	0.001		
H_0: Distribution $WTP_{2,i}$ = Normal	5.89	0.001		
H_0: $\overline{WTP}_{1A} = \overline{WTP}_2$			–2.57	0.010
H_0: $\overline{WTP}_{1B} = \overline{WTP}_2$			–1.53	0.125
H_0: Distribution $WTP_{1A,i}$ = Distribution $WTP_{2,i}$	2.89	0.001		
H_0: Distribution $WTP_{1B,i}$ = Distribution $WTP_{2,i}$	1.06	0.210		

[a]Kolmogorov-Smirnov test statistic.
[b]Mann-Whitney test statistic.

The null hypothesis that the stated WTP amounts are normal distributed is rejected by the outcome of the Kolmogorov-Smirnov test statistic (Table IV). Consequently, we used the non-parametric Mann-Whitney test to test the equality of means in the samples. The null hypothesis of equal means is rejected at the 1 percent significance level for samples 1A and 2, but not for samples 1B and 2.[7] The more rigorous null hypothesis of equal distributions is rejected by the Kolmogorov-Smirnov test statistic for samples 1A and 2, but can also not be rejected for samples 1B and 2.

4.2.2. *Regression Results*

For WTP data sets with large numbers of zero bids as in this paper, the Tobit model is a theoretically correct model to explain the variance in stated WTP amounts (Halstead et al. 1991). The Tobit model is a regression model for censored distributions, i.e. for distributions where there are no observations beyond a certain point (in this case below zero). If there is a large portion of observations at this censoring point, ordinary least squares (OLS) estimation techniques may result in biased estimates. The Tobit model assumes that any observation for which the dependent variable is zero or negative is an observation for which the dependent variable is not observed. The model can be written as follows:

$$\mathrm{WTP} = \alpha + \beta X + \varepsilon \quad \text{if WTP} > 0 \tag{1}$$

$$\mathrm{WTP} = 0 \quad \text{otherwise} \tag{2}$$

where X denotes the matrix of explanatory variables and the error term ε is assumed to be normal distributed with zero mean and variance σ^2. The estimates for the regression coefficients α and β are obtained through maximum likelihood (ML) techniques (Maddala 1983; Cramer 1986). These estimates cannot be compared straightforwardly with the OLS estimates (Halstead et al. 1991). The Tobit regression results for each data set are presented in Table V.

Table V. Tobit Regression Results

	Sample 1A		Sample 1B		Sample 2	
	coefficient estimate	standard error	coefficient estimate	standard error	coefficient estimate	standard error
Constant	–2.7477***	0.4578	–2.7244***	0.7934	–1.4906***	0.3260
Sex	0.0580	0.1358	0.3569*	0.2594	–0.0356	0.0868
Education	–0.0427	0.1374	0.1010	0.2476	0.0623	0.0874
# members $<$ 5 yrs	0.0339	0.1134	–0.1197	0.2097	–0.1341**	0.0697
# members 6–18 yrs	0.0808	0.0712	0.0939	0.1181	0.0167	0.0511
# members 19–65 yrs	–0.0007	0.1055	–0.2070*	0.1501	–0.0684	0.0604
# members $>$ 65 yrs	0.3647**	0.1859	–0.0953	0.2805	–0.3597***	0.0968
Household income	0.0001***	0.0000	0.0001*	0.0000	0.0001***	0.0000
Specific attitude	0.5379***	0.0798	0.6050***	0.1483	0.3656***	0.0621
General attitude	0.2321***	0.0733	0.4051***	0.1371	0.1433***	0.0603
Knowledge	0.0397	0.0547	–0.2255**	0.1092	0.0872***	0.0368
Membership	0.1656	0.1320	0.3071*	0.2322	0.1535**	0.0844
Annual/Monthly	0.3359***	0.1400	—	—	—	—
σ	0.8427***	0.0503	0.6913***	0.0796	0.7901***	0.0323
Log Likelihood	–241.930		–53.532		–480.046	
χ^2 (d.f.)	154.7 (12)		44.0 (11)		269.8 (11)	
n	204		54		414	
Positive *n*	159		42		334	

$^{*}p < 0.1$
$^{**}p < 0.05$
$^{***}p < 0.01$
Note: the number of observations in each sample does not correspond to the ones presented in Table III, because of missing values which are automatically deleted by the statistical package we used for the Tobit regression (Time Series Processor (TSP) 4.3).
Clarification:

Variable	Definition	Characteristic
Sex	respondent's sex	dummy (1 = man)
Education	respondent's highest completed education level	dummy (1 = higher education level)
# members $<$ 5 yrs	number of household members aged less than 5	ratio scale (0–5)
# members 6–18 yrs	number of household members aged between 6 and 18	ratio scale (0–4)
# members 19–65 yrs	number of household members aged between 19 and 65	ratio scale (0–4)
# members $>$ 65 yrs	number of household members aged over 65	ratio scale (0–2)
Household income	monthly net household income	midpoint estimates of 8 income groups
Specific attitude	respondent's attitude to flowery ditch-sides and meadow birds	Likert scale (1–5)
General attitude	respondent's attitude to paying for public environmental goods	Likert scale (1–5)
Knowledge	respondent's knowledge of the area involved	Likert scale (1–5)
Membership	member of nature conservation organisation or not	dummy (1 = yes)
Annual/Monthly	annual or monthly payments	dummy (1 = monthly)

Given the skewed distribution of WTP amounts (Table III) and the normality assumption underlying the Tobit regression, the dependent variable has been transformed into common (base 10) logarithmic form. In order not to lose any zero bidders in the analysis, the untransformed dependent variable was raised by one first.

The three models are statistically significant. The outcomes of the likelihood ratio test (χ^2 in Table V) convincingly reject the null hypothesis of zero slopes for all explanatory variables. As expected, standard socio-economic variables such as a respondent's sex or household income play a significant role in explaining differences in stated WTP amounts.

The estimated income coefficient is significant and the same in every sample. A respondent's education level does not play a significant role in the three samples. A respondent's sex is only significant in sample 1B.[8] In the first study (sample 1A and 1B), men appear to be willing to pay more for the conservation of the amenities found on peat meadow land than women, while the reverse is the case in the second study. Respondent's age was not included in the analysis, because of its high correlation (> 0.5) with the variables representing the age structure of the respondent's household. In the second study, the number of household members aged less than 5 years and over 65 has a significant negative effect on WTP. The number of household members aged between 19 and 65 is significant and negative in sample 1B and in sample 1A the number of household members aged over 65. Hence, the more members falling in these age categories, the less a respondent is willing to contribute.

As expected, also a respondent's attitude and actual behaviour towards paying for nature conservation significantly help to explain differences in WTP in both studies. The more positive someone's attitude to paying for public environmental goods in general and the more positive someone ranked his or her attitude to the preservation of flowery ditch-sides and meadow birds in agricultural areas, the higher his or her WTP. Furthermore, respondents who are already paying for nature conservation through membership of a nature conservation organisation are willing to contribute to this specific nature conservation scheme significantly more than non-members in sample 1B and study 2.

Also a respondent's knowledge of the areas involved plays a significant role. In the second study, the coefficient estimate has the expected sign: the more familiar someone is with the area, the more he or she is willing to pay. In sample 1B a respondent's knowledge of the area has an unexpected significant negative effect on stated WTP. Although not statistically significant, a respondent's knowledge is also positively related to WTP in sample 1A. Finally, as mentioned in section 4.1., in sample 1A annual WTP depends significantly on whether the respondent was asked for monthly or annual payments.

Table VI. Hypothesis 2 and 3: Test Results for Benefit Function Transfer

Null hypothesis	Wald test[a]	LR test[b]	$\chi^2_{critical}$[c]	'*Observed*' SSR[d]	'*Calculated*' SSR[e]
2a. H^0: $b_{1A} = \hat{b}_2$	28.23		21.03 (12)	121.3	158.4
2b. H^0: $b_2 = \hat{b}_{1A}$	34.09		21.03 (12)	70.1	268.7
2c. H^0: $b_{1B} = \hat{b}_2$	16.59		21.03 (12)	121.3	30.3
2d. H^0: $b_2 = \hat{b}_{1B}$	158.58		21.03 (12)	16.3	348.6
3a. H^0: $\hat{b}_{1A} = \hat{b}_{1A+2}$	5.87	22.36 (13)			
3b. H^0: $\hat{b}_{1B} = \hat{b}_{1B+2}$	13.00	21.03 (12)			
3c. H^0: $\hat{b}_2 = \hat{b}_{2+1A}$	3.48	21.03 (12)			
3d. H^0: $\hat{b}_2 = \hat{b}_{2+1B}$	1.69	21.03 (12)			
3e. H^0: $\hat{\sigma}^2 = \hat{\sigma}^2_{1A} = \hat{\sigma}^2_2$	9.84	21.03 (12)			
3f. H^0: $\hat{\sigma}^2 = \hat{\sigma}^2_{1B} = \hat{\sigma}^2_2$	15.72	21.03 (12)			

[a]Wald test statistic = $(\hat{b}_p - \hat{b}_s)' \sum_p^{-1} (\hat{b}_p - \hat{b}_s)$ where $\sum_p^{-1}$ is the variance-covariance matrix estimated at the policy site.
[b]LR test statistic = $-2(LL_{pooled} - (LL_p + LL_s))$.
[c]Critical value at 0.05 significance level (degrees of freedom between brackets).
[d]Sum of Squared Residuals of the original estimated benefit function.
[e]Sum of Squared Residuals of the transferred benefit function.

4.2.3. *Equality of Benefit Functions*

The outcome of the second hypothesis testing tells us whether or not there is any statistical legitimacy in transferring the estimated benefit functions from one study to the other. The test results are presented in Table VI. The null hypotheses 2a to 2d have to be read as follows: the coefficients on the left-hand side of the equation represent the estimated coefficients at the policy site and the coefficients on the right-hand side of the equation the estimated coefficients at the study site. Thus, the benefit function is transferred from the sample on the right-hand side of the equation to the sample on the left hand-side.

The second hypothesis of equal regression coefficients is rejected in 3 of the 4 cases. Only the transfer from the second study to sample 1B is justified on the basis of the outcome of the Wald test statistic, which is lower than the critical value at the 0.05 significance level. However, conform results found by Kirchhoff et al. (1997), transferring the benefit function from sample 1B to study 2 results in a rejection of the null hypothesis.

The third hypothesis of equal coefficient estimates in the individual and pooled sample (3a to 3d) can not be rejected. This confirms that equality of regression coefficients of the individual models and their pooled model does not guarantee equality of the regression coefficients of the individual models.

In order to see what valid benefit function transfer means in terms of goodness of fit, we also compared the sum of squared residuals of the transferred function to the original estimated function. As said, if the null hypothesis of equal coefficient estimates can not be rejected, it will be likely that also the goodness of fit of the functions at two sites will be the same. However, as Table VI shows, the goodness of fit differs considerably in all transfer cases, including the statistically valid transfer from the second study to sample 1B. In this latter case, the transfer results in a considerable reduction of the sum of squared residuals. Using the F-test, the reduction can furthermore be shown to be statistically significant (F = 165.9 with 12 and 663 degrees of freedom). The null hypothesis of equal goodness of fit is rejected at the 0.01 significance level, providing statistical evidence that the equality of coefficient estimates does not necessarily imply equality of explained variance.

Finally, to complete the validity testing procedure outlined in section 2, the test results for the hypothesis of equal goodness of fit of the individual and pooled samples are also presented in Table VI. The null hypothesis 3e and 3f can not be rejected by means of the LR test.[9]

4.2.4. *Errors when Transferring 'Observed' and 'Calculated' Unit Values from One Site to Another*

Table VII presents the errors when transferring unit values from one Site to another under uncontrolled ('*observed*' unit values) and controlled ('*calculated*' unit values) conditions. The statistical significance of the errors under uncontrolled conditions was reported in Table IV in section 4.2.1. The hypothesis of equal means can not be rejected for samples 1B and 2. As Table VII shows, using the untruncated mean WTP from these samples would result in an error of +36% when transferring from sample 2 to sample 1B and –27% when transferring from sample 1B to sample 2.

The Tobit coefficient estimates in Table V can be used to calculate WTP under controlled conditions. However, before these '*calculated*' WTP amounts can be compared with the '*observed*' ones, they have to be adjusted for the fact that (1) the regression was performed on a censored data set and included positive bids only and (2) a log transformation was carried out. The adjusted value of Log (WTP + 1) at the policy site which includes both zero and non-zero bidders is (Halstead et al. 1991):

$$\text{Calculated Log (WTP} + 1) = \mathrm{X}'\hat{\mathrm{b}} * F(\mathrm{z}) + \sigma * f(\mathrm{z}) \qquad (3)$$

where X is the matrix of explanatory variable values at the policy site, $\hat{\mathrm{b}}$ the vector of estimated Tobit coefficients at the study site, z the standardised normal variable, F(z) the value of the cumulative normal distribution at z, σ the standard error of the Tobit regression at the study site and f(z) the value of the standard normal distribution at z.

Table VII. Transfer Errors

	Sample 1A	Sample 1B	Sample 2	
			Based on Sample 1A	Based on Sample 1B
Observed Values[a]				
Mean WTP	96.6	54.5	74.2	
Median	60.0	25.0	40.0	
Calculated Values[a]				
Mean WTP	—[c]	42.3	44.6[d]	53.4
Median	—	16.4	16.3	23.5
Transfer Error under uncontrolled conditions				
Mean WTP	—[b]	36.1%	30.2%	–26.6%
Median	—	60.0%	50.0%	–37.5%
Transfer Error under controlled conditions				
Mean WTP	—	–22.4%	–39.9%	–28.0%
Median	—	–34.4%	–59.2%	–41.2%

[a] In 1995 D*f*.
[b] Omitted because of statistical evidence of incomparability.
[c] Mean and median WTP can not be calculated in this case, because the benefit function at study site 2 does not control for the difference between annual and monthly based payments (only annual payments were asked).
[d] Mean and median WTP have been calculated in this case by substituting the value zero for the dummy variable distinguishing between monthly and annual based payments.

Subsequently, median WTP has to be adjusted for the log transformation:

$$\text{Median WTP} = 10^{(x'\hat{b})} - 1 \quad (4)$$

while mean WTP also has to be adjusted for the estimated standard error of the regression:

$$\text{Mean WTP} = 10^{(x'\hat{b})} * 10^{(\sigma/2)} - 1 \quad (5)$$

The median and mean are obtained by substituting the relevant sample mean values for the explanatory variables in equations 4 and 5 (see Table VII).

Table VII shows that the function approach to benefits transfer results in a smaller error in absolute terms than the unit value approach when transferring the benefit function from the second study to sample 1B, which was shown to be statistically legitimate in Table VI. Only in this case the transfer of unit values

Table VIII. Hypothesis 4: Test Restilts for Unit Value Transfer under Controlled Conditions

Null hypothesis	K-S test statistic[a]	2-tailed $p \leq$	M-W test statistic[b]	2-tailed $p \leq$
H_0: Distribution $WTP'_{1A,i}$ = Normal	5.08	0.001		
H_0: Distribution $WTP'_{1B,i}$ = Normal	2.74	0.001		
H_0: Distribution $WTP'_{2A,i}$ = Normal	7.34	0.001		
H_0: Distribution $WTP'_{2B,i}$ = Normal	7.85	0.001		
H_0: $\overline{WTP}'_{1A} = \overline{WTP}_{1A}$			–6.82	0.001
H_0: $\overline{WTP}'_{1B} = \overline{WTP}_{1B}$			–2.24	0.025
H_0: $\overline{WTP}'_{2A} = \overline{WTP}_{2}$			–6.67	0.001
H_0: $\overline{WTP}'_{2B} = \overline{WTP}_{2}$			–5.31	0.001
H_0: Distribution $WTP'_{1A,i}$ = Distribution $WTP_{1A,i}$	4.93	0.001		
H_0: Distribution $WTP'_{1B,i}$ = Distribution $WTP_{1B,i}$	1.91	0.001		
H_0: Distribution $WTP'_{2A,i}$ = Distribution $WTP_{2,i}$	5.47	0.001		
H_0: Distribution $WTP'_{2B,i}$ = Distribution $WTP_{2,i}$	5.25	0.001		

[a]Kolmogorov-Smirnov.
[b]Mann-Whitney.
Clarification:
$WTP'_{2A,i}$: Calculated WTP for study 2 based on the coefficient estimates from sample 1A.
$WTP'_{2B,i}$: Calculated WTP for study 2 based on the coefficient estimates from sample 1B.

under controlled conditions leads to a more robust result than transferring unit values under uncontrolled conditions. '*Calculated*' mean WTP is 22 percent lower than '*observed*' mean WTP. In the other two transfer cases, the benefit function approach results in (slightly) higher errors. In case the benefit function estimated at the first study site is transferred to the second study site, average WTP is 28 or 40 percent lower than real WTP. The transfer of median values produces considerably larger errors than the transfer of mean values under uncontrolled and controlled conditions.

Finally, the equality of '*observed*' and '*calculated*' means and hence the statistical significance of the errors of benefit function transfer is tested by means of the Mann-Whitney test (Table VIII). The computed test statistics reject the null hypothesis of equal means (and distributions) in all transfer test cases, including the valid benefit function transfer from study 2 to sample 1B. These test results provide another indication of the goodness of fit of the estimated benefit function at one site to the observations at another site, which we presented more formally in Table VI.

5. Discussion and Conclusions

In this case study, we followed the comprehensive testing procedure for valid benefits transfer recommended by Bergland et al. (1995). This procedure provides the technical background against which criteria for valid benefits transfer can be checked for their significance. We extended the analyses presented in previous studies by including control factors not accounted for before, but which are expected to play a significant role in explaining differences in stated preferences across individuals, groups and populations. The studies included in the analysis focused on the same environmental goods in similar areas using the same valuation scenario and CV market construct. Hence, the amount of control needed for valid benefit transfer across sites and populations was expected to be found primarily in population characteristics.

The transfer of benefit functions is generally considered the most appropriate procedure to environmental benefits transfer, because it enables someone to control for site and environmental good characteristics, population characteristics or procedural research differences related to specific ‘market properties’ such as how environmental values are or would be elicited (e.g. through income taxation or private funds, monthly or annual payments). In order for benefit function transfer to be a valid monetary valuation procedure, a necessary condition is that factors significantly influencing stated (or revealed) preferences at different sites should be accounted for in the benefit function and the size of their impact should coincide. If their impact differs across sites, the model estimated at one site can not be used to predict WTP at another site.

Since accounting for socio-economic group or class solely does not fully explain why respondents hold different beliefs, norms, values and attitudes to environmental problems in general, the specific environmental problem at hand or paying for public environmental services, we also included attitude variables in the analysis. These attitudes are believed to provide an important key to the question why respondents from similar socio-economic backgrounds still come up with different WTP amounts. Especially possible non-use values, if present, are expected to complicate the transfer of environmental benefits based on expressed preference techniques, because of the need to carefully analyse the motivations underlying stated preferences.

However, if accounting for attitude variables provides a valid basis for benefits transfer, one could argue that this is bad news for the practical viability of CV-based transfers since it suggests the need for data collection for such variables at the policy site. In that case, instead of relying upon previous and perhaps outdated CV study results, one may just as well carry out an original CV study at the policy site.

As expected, respondent attitude to the preservation of the specific amenities involved and his or her willingness to pay for public environmental goods in general play a significant role in the explanation of differences in stated WTP amounts. Nonetheless, accounting for these variables proved insufficient to provide a valid transfer of benefit functions in 3 of the 4 benefit transfer cases investigated in this paper. Moreover, in the case where function transfer was successful, we found that the transferred model’s goodness of fit was significantly lower at the study site than at the policy site, suggesting that we may have overlooked one or more study site or population specific explanatory factors.

Accounting for the difference in procedural research design in the two studies in terms of temporal embedding, the statistically valid transfer of unit values resulted in a minimum absolute error of 27 percent and a maximum of 36 percent. The statistically valid benefit function transfer led to an absolute error of 22 percent. Thus, in this study we provide further empirical evidence that in the case of statistically valid benefits transfer, the function approach results in a slightly more robust estimation than the unit value approach. In those cases where the hypothesis of valid benefit function transfer was rejected, the function transfer results in a higher percentage error in absolute terms than the unit value transfer for both mean and median values.

We also showed that the equality of coefficient estimates does not necessarily imply equality of the variance explained at the study and policy site. Even though the null hypothesis of equal coefficient estimates can not be rejected, this does not mean that the variables included explain enough of the variance in stated WTP amounts at different sites for the unit values under controlled circumstances to be the same. Hence, the equality of coefficient estimates is a necessary, but not sufficient condition for valid benefit function transfer.

This has some important consequences for the validity tests carried out in previous studies. The Chow and Likelihood Ratio test, used in 5 of the 6 previous studies which tested the equality of coefficient estimates, are tests for a model's global goodness of fit. These tests only tell us that the explanatory variables included in the model explain the same amount of the variability in stated WTP amounts, but not which variables and to which extent. Moreover, the Chow test tests the equality of regression coefficients of the individual models and their pooled models, which does not guarantee, as we also saw, equality of the regression coefficients of the individual models.

The Wald test directly tests the equality of coefficient estimates and is hence considered an appropriate test statistic for the hypothesis of equal coefficient estimates. But testing this hypothesis only is not sufficient to guarantee valid benefit function transfer. A necessary next step is to test the goodness of fit of the transfer model, which can be done with the help of the Chow or the LR test, depending on the underlying assumptions regarding the distribution of '*observed*' and '*calculated*' WTP amounts.

Finally, the results in this paper have to be considered with the necessary care. First of all, only two case specific studies were included in the analysis. Using multiple samples will provide a more robust validity testing procedure. Secondly, the sites in the two studies differ considerably in size, a fact we were unable to control for. Thirdly, the number of observations in one of the samples which provided for a valid benefit function transfer was relatively low (n = 56).

Acknowledgements

We gratefully acknowledge the support provided by the Department of Agricultural Economics and Policy of the Wageningen Agricultural University in the Netherlands for carrying out the second study. We furthermore wish to thank the two anonymous referees, and Richard Dubourg, Ian Bateman and Ian Langford for their helpful comments on an earlier draft. The usual disclaimer applies.

Notes

1. '*Calculated*' average WTP at the policy site equals $\hat{b}_s\bar{X}_p$, where $\hat{b}_s$ is the vector of coefficient estimates from the study site (s) and $\bar{X}_p$ the vector of average values of the explanatory variables at the policy site (p).
2. Downing and Ozuna (1996) did not actually test the validity of benefit function transfer statistically in terms of equality of coefficient estimates. They investigated whether or not the dummy variable for the different time periods in which the CV survey was carried out is significant in the estimated benefit functions. If it was in both functions, they concluded that the functions are not transferable. If it was not in either of the two functions, they concluded that the functions are transferable. If the dummy was significant in one of the two functions, the transfer was considered questionable.
3. First of all, studies considered for inclusion must be based on adequate data, sound economic methods and correct empirical techniques. Moreover, studies should contain regression results which describe WTP as a function of socio-economic characteristics. Secondly, the sites must have similar populations. Thirdly, the environmental good and its different provision levels at the different sites should be similar. Fourthly, the sites in which the goods are found should be similar as well. Finally, the constructed markets, including the distribution of property rights, have to be the same at each site.
4. The Mann-Whitney or Wilcoxon rank sum test is a non-parametric test for differences in the location of two distributions, e.g. the mean or median (Diamond et al. 1993). If the estimates of the means of two distributions are different and the MW test rejects the null hypothesis of no differences in the location of these two distributions, we interpret this as evidence for a statistically significant difference in the means of the two distributions.
5. The test statistics of these tests assume a similar a-symptotic distribution of the error terms. The Chow test is a special form of the F-test, whereas the LR test statistic is Chi-square distributed.
6. Dillman (1978) provides guidelines for increasing the response rate of mail surveys. The method prescribes, for example, the use of a booklet, to illustrate the front and not to have questions on the backside of the questionnaire.
7. Amongst other things as a result of the higher monthly based payments stated in sample 1A.
8. Respondent education level was also included as a categorical (ordinal) variable in the analysis, but also this variable did not have a significant impact on stated WTP. No correlation (> 0.25) was detected between education (as a categorical or dummy variable) and household income level.
9. The LR test is easy to compute with the help of the log likelihood values presented in Table V and the log likelihood function of the pooled model (–726.89 for sample 1A and 2 and –541.44 for sample 1B and 2). In order to keep the number of tables to a minimum, the Tobit regression results of the pooled samples are not presented in this paper, but are available from the authors.

References

Arrow, K., R. Solow, P. R. Portney, E. E. Leamer, R. Radner and H. Schuman (1993), *Report of the NOAA Panel on Contingent Valuation*. Federal Register, January 15, Vol. 58 no. 10: 4601–4644.

Bergland, O., K. Magnussen and S. Navrud (1995), *Benefit Transfer: Testing for Accuracy and Reliability*. Discussion Paper, #D-03/1995, Department of Economics and Social Sciences, Agricultural University of Norway.

Boyle, K. J. and J. C. Bergstrom (1992), 'Benefit Transfer Studies: Myths, Pragmatism, and Idealism', *Water Resources Research* **28**(3), 675–683.

Brouwer, R. (1995), *The Measurement of the Non-Marketable Benefits of Agricultural Wildlife Management: The Case of Dutch Peat Meadow Land*. Wageningen Economic Paper, 1995-1, Wageningen Agricultural University.

Brouwer, R., I. H. Langford, I. J.. Bateman, T. C. Crowards and R. K. Turner (1997), *A Meta-Analysis of Wetland Contingent Valuation Studies*. GEC Working Paper 97-20, CSERGE, University of East Anglia and University College London.

Costanza, R., R. d'Arge. R. de Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Nacem, R. V. O'Neill, J. Paruelo, R. G. Raskin, P. Sutton, and M. van den Belt (1997), 'The Value of the World's Ecosystem Services and Natural Capital', *Nature* **387**, 253–260.

Cramer, J. S. (1986), *Econometric Applications of Maximum Likelihood Methods*. Cambridge: Cambridge University Press.

Desvousges, W. H., M. C. Naughton and G. R. Parsons (1992), 'Benefit Transfer: Conceptual Problems in Estimating Water Quality Benefits Using Existing Studies', *Water Resources Research* **28**(3), 675–683.

Diamond, P. A., J. A. Hausman, G. Leonard and M. A. Denning (1993), 'Does Contingent Valuation Measure Preferences? Experimental Evidence', in J. A. Hausman, ed., *Contingent Valuation: A Critical Assessment*. Contributions to Economic Analysis 220. Amsterdam: North Holland.

DilIman, D. A. (1978), *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley and Sons.

Downing, M. and T. Ozuna Jr (1996), 'Testing the Reliability of the Benefit Function Transfer Approach', *Journal of Environmental Economics and Management* **30**, 316–322.

Fishbein, M. and I. Ajzen (1975), *Belief, Attitude, Intention and Behaviour: An Introduction to Theory and Research.* Reading, Massachusetts: Addison-Wesley Publishing Company.

Greene, W. H. (1990), *Econometric Analysis*. New York: Macmillan Publishing Company.

Halstead, J. M., B. E. Lindsay and C. M. Brown (1991), 'Use of the Tobit Model in Contingent Valuation: Experimental Evidence from the Pemigewaset Wilderness Area', *Journal of Environmental Management* **33**, 79–89.

Hoevenagel, R. (1994), *The Contingent Valuation Method: Scope and Validity*. PhD- thesis, Vrije Universiteit Amsterdam.

Kirchhoff, S., B. G. Colby and J. T. LaFrance (1997), 'Evaluating the Performance of Benefit Transfer: An Empirical Inquiry', *Journal of Environmental Economics and Management* **33**, 75–93.

Loomis, J. B. (1992), 'The Evolution of a More Rigorous Approach to Benefit Transfer: Benefit Function Transfer', *Water Resources Research* **28**(3), 701–705.

Loomis, J. B., B. Roach, F. Ward, and R. Ready (1995), Testing Transferability of Recreation Demand Models Across Regions: A Study of Corps of Engineer Reservoirs', *Water Resources Research* **31**(3), 721–730.

Maddala, G. S. (1983), *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.

O'Doherty, R. K. (1996), *Contingent Valuation and Benefit Transfer: An Assessment in a Local Planning Context*. Working Papers in Economics, No. 19, Faculty of Economics and Social Science, University of the West of England, Bristol.

Parsons, G. R. and M. J. Kealy (1994), 'Benefits Transfer in a Random Utility Model of Recreation', *Water Resources Research* **30**(8), 2477–2484.

Pearce, D. W., D. Whittington and S. Georgiou (1994), *Project and Policy Appraisal: Integrating Economics and Environment*. Paris: OECD.

Spaninks, F. A. (1993), *Een Schatting van de Sociale Baten van Beheersovereenkomsten met behulp van de Contingent Valuation Methode*. MSc-thesis, Department of Agricultural Economics and Policy, Wageningen Agricultural University.

Spaninks, F. A. and R. Hoevenagel (1995), *Temporal Embedding in Contingent Valuation*. Paper presented at the 6th annual meeting of the European Association of Environmental and Resource Economists (EAERE), Umeå, Sweden, June 17–20.

Terwan, P. (1988), *Landbouw en Natuur in Veenweidegebieden, Perspectieven voor Verweving*. Centrum voor Landbouw en Milieu en Landelijk Overleg van Boerenwerkgroepen in Relatienotagebieden. Utrecht: Drukkerij Elinkwijk B.V.