WATER RESOURCES RESEARCH, VOL. 41, W03017, doi:10.1029/2004WR003466, 2005

# Temporal stability and transferability of models of willingness to pay for flood control and wetland conservation

Roy Brouwer
Institute for Environmental Studies (IVM), Vrije Universiteit, Amsterdam, Netherlands

Ian J. Bateman[1]
Centre for Social and Economic Research on the Global Environment (CSERGE), School of Environmental Sciences, University of East Anglia, Norwich, UK

Received 3 July 2004; revised 6 December 2004; accepted 20 December 2004; published 18 March 2005.

[1] This study investigates the temporal stability and transferability of dichotomous choice willingness to pay responses and their determinants from two large-scale contingent valuation surveys in the area of flood control and wetland conservation. The study considers a time period between surveys which is more than double that considered in previous test-retest analyses. Whereas such previous studies have reported stable values over relatively short time periods, the present study finds a statistically significant decrease in real willingness to pay over this more extended time period. Analyses of model transfer between the two survey periods indicate that models derived solely from economic-theoretic determinants pass transferability tests. However, expanding these models to include more ad hoc, transitory factors yields nontransferable models. This provides a guide for future analyses.

**Citation:** Brouwer, R., and I. J. Bateman (2005), Temporal stability and transferability of models of willingness to pay for flood control and wetland conservation, *Water Resour. Res.*, *41*, W03017, doi:10.1029/2004WR003466.

## 1. Temporal Stability of Contingent Valuation Results

[2] This paper addresses the reliability of contingent valuation (CV) estimates of willingness to pay (WTP) for nonmarket goods through time. Although the National Oceanic and Atmospheric Administration (NOAA) Panel on Contingent Valuation [*Arrow et al.*, 1993] raises some concern about the temporal stability of CV estimates, to date test-retest studies have only considered relatively short periods, ranging from 2 weeks [*Kealy et al.*, 1988, 1990] to 2 years [*Carson et al.*, 1997]. These have supported the replicability of findings and stability of values across such modest periods. An overview of studies investigating the reliability of CV estimates is given by *McConnell et al.* [1998].

[3] The present paper examines the temporal stability of incentive compatible dichotomous choice (DC) CV models across a 5-year period, i.e., a period more than twice as long as the longest considered previously. The issue of temporal stability over extended periods is one of more than academic interest. Benefit-cost analyses (BCA) frequently employ values estimated some considerable time prior to those analyses. Temporal stability is therefore implicitly assumed rather than explicitly tested. Yet there is no reason to suppose that values for nonmarket goods should remain constant over extended periods.

[1]Also at Department of Agricultural and Resource Economics, University of Western Australia, Perth, Western Australia, Australia.


0043-1397/05/2004WR003466$09.00

[4] This study addresses the issue of temporal stability through the application of two matching surveys, concerning the same case study area (the Norfolk Broads, in the UK), focusing on the same environmental good and valuation scenarios (flood protection and conservation of freshwater wetland habitat and associated recreational amenities), using the same payment vehicle (coercive taxation), the same sampling frame (random in-person interviews) applied to the same sample population (visitors to the area), but sampling at different points in time, namely, in the summers of 1991 and 1996. The study's main objective is to test the transferability of resultant models of WTP and the stability of their determinants across this more extended time period.

## 2. Survey Data

[5] The Norfolk Broads are a large freshwater wetland area located in East Anglia, UK. The area consists of a system of shallow lakes, marshes, and fens, linked by low-lying rivers. The site is of national and international wildlife importance, being a designated Environmentally Sensitive Area (ESA) and containing 24 Sites of Special Scientific Interest (SSSI), including two sites notified under the international Ramsar convention. The area is also a major focus for recreation, attracting more than 1 million visitors a year, of which 200,000 spend their holidays on boats hired for a week or longer [*Broads Authority*, 1997].

[6] The character of the low-lying landscape of the Broads depends upon 210 km of reinforced river embankments for protection from saline tidal waters. However, at the time of the surveys these flood defenses were increasingly at risk from failure, because of their age, erosion from

**Table 1.** Summary Statistics of Respondent Characteristics

| Respondent Characteristic | 1991 | 1996 |
|---|---|---|
| Average age | 47 | 48 |
| Average real household income level (£1991 prices) | 22,300 | 18,725 |
| Average household size (number of persons) | 3 | 3 |
| Average number of children | 1 | 1 |
| Percentage living or working in the area | 16 | 18 |
| Percentage on holiday | 65 | 65 |
| Average (one way) distance traveled from home, km | 219 | 226 |
| Percentage visiting for the first time | 25 | 18 |
| Average number of previous visits | 11 | 10 |
| Percentage staying on a boat | 23 | 19 |
| Percentage member of environmental group | 45 | 35 |

boat wash, and the sinking of the surrounding marshlands. Thus the standard of flood protection afforded by these man-made defences was decreasing over time. If flood defences were breached, the ensuing saline inundation would fundamentally and enduringly alter the nature of the area both in terms of its habitat capabilities and in respect of the recreational opportunities currently afforded.

[7] In 1991 the National Rivers Authority (NRA), later named the Environment Agency (EA), initiated a wide-ranging Flood Alleviation Study to develop a cost-effective strategy to alleviate flooding in the Norfolk Broads for the next 50 years [*Bateman et al.*, 1995]. The flood alleviation study consisted of five main components: hydraulic modeling; engineering; benefit-cost assessment; environmental assessment; and public consultations. The item of most relevance here is the benefit-cost assessment, which compared benefits of undertaking a scheme to provide a particular standard of flood protection with the corresponding costs. Although market benefits of flood alleviation have been considered in terms of agriculture, industry, living conditions, and infrastructure [*Turner and Brooke*, 1988], the value of the nonmarket benefits from the area was uncertain.

[8] As part of the benefit-cost assessment, a large CV study was mandated in 1991 [*Bateman et al.*, 1994, 1995] and a follow-up was carried out in 1996 [*Powe*, 1999; *Powe and Bateman*, 2003, 2004], in order to assess user valuations of conserving the area in its current state. The studies aimed, among other things, to provide a valid and reliable monetary estimate of the current recreational and amenity benefits enjoyed by visitors to the Broads. Findings were used to inform a BCA of various flood defense options [*Brouwer et al.*, 2001]. The benefit-cost ratio found ranged between 0.98 and 1.94 [*National Rivers Authority*, 1993]. The results, including the findings from the 1991 CV study, were submitted to the relevant Ministry of Agriculture, Food and Fisheries as part of an application of central government funding support for the proposed flood alleviation strategy. Following lengthy consideration of this application, the EA received conditional approval for a program for bank strengthening and erosion protection in 1997 [*Environment Agency*, 1997]. The actual scheme was taken forward in 2000 on the basis of a long-term private-public partnership scheme between the EA and relevant government support ministries and a private engineering firm consortium.

[9] The 1996 questionnaire was based upon the one used in 1991. The textual information, valuation scenario (WTP for a flood prevention scheme), elicitation technique (a DC referendum), and payment vehicle (taxation) were identical. The questionnaires were designed in line with the guidelines set out by the influential NOAA Blue-Ribbon Panel on CV [*Arrow et al.*, 1993]. These design guidelines were implicated through an extensive process of focus group sessions and pilot surveys described elsewhere [*Bateman et al.*, 1995; *Brouwer et al.*, 1999]. Both the 1991 and 1996 studies consisted of a number of separate and common design treatments (including tests of differing elicitation techniques and part/whole analyses). However, this paper restricts consideration to that treatment which was common across the two studies. This consisted of a large DC referendum sample concerning WTP for the flood prevention scheme outlined above.

[10] After thorough pretesting (The 1991 questionnaire was pretested in 433 face-to-face on-site interviews, and the 1996 questionnaire was tested in three focus group sessions and 100 face-to-face on site interviews.), the questionnaires were applied through on-site interviews with visitors at representative sites around the Norfolk Broads based on a random "next to pass" basis. The response rate in the 1991 survey was 78% and in the 1996 survey 68% (see *Bateman et al.* [1995] and *Powe* [1999] for more details). Nonresponse (including possible protest responses) to the valuation questions was consistently low across the two surveys, being less than 5 percent of the total number of respondents. Excluding these, a total of 1747 responses were obtained from the common treatment element of the 1991 survey and 1108 responses were obtained from the 1996 survey.

[11] A summary of the respondent characteristics in the 1991 and 1996 survey is presented in Table 1. Respondents in the two surveys have a fairly similar demographic and socioeconomic background. Compared with 1991, a significantly larger number of randomly selected visitors in the 1996 sample fall in lower income categories. Furthermore, in the 1996 sample more people had visited the Broads before, and fewer people are a member of an environmental group or organization.

## 3. Statistical Testing Procedures

[12] Temporal reliability of DC CV models is tested by examining the statistical equality of unadjusted average WTP values (hypothesis 1) and the DC WTP functions (hypothesis 2). Comprehensive statistical testing procedures were originally proposed by *Bergland et al.* [1995]. Testing the first hypothesis is relatively straightforward [e.g., *Greene*, 1993]. The second hypothesis can be broken down into two different subhypotheses: equality of coefficient estimates (hypothesis 2a) and equality of variances (hypothesis 2b). It has been shown that equality of coefficients (hypothesis 2a) is a necessary (if not, on its own, sufficient) condition for function transfer [*Brouwer and Spaninks*, 1999]. The statistical significance of individual explanatory variables and the overall goodness of fit of estimated models (hypothesis 2b) may also be important, as these reflect the stability of the predictive power of the models over time.

[13] The LR and Chow tests are most widely applied in studies testing the second hypothesis [e.g., *Loomis et al.*,

**Table 2.** Mean Real WTP Values From the 1991 and 1996 Surveys (Pounds per Annum in 1991 Prices) Obtained From the Parametric Logistic Model and (Lower Bound) Nonparametric Turnbull Model

| | Parametric Linear-Logistic | | Nonparametric Turnbull | |
|---|---|---|---|---|
| | 1991 | 1996 | 1991 | 1996 |
| Mean WTP, £ | 248.1 | 215.8 | 54.2 | 37.8 |
| Standard error | 23.3 | 29.3 | 2.9 | 2.4 |
| 95% CI {1996–1991} | . . . | {−34.3; −30.3} | . . . | {−16.6; −16.2} |
| Minimum-maximum values | $-\infty$ to $+\infty$ | $-\infty$ to $+\infty$ | 0–200 | 0–200 |
| *N* | 1747 | 1108 | 1747 | 1108 |

1995; *Kirchhoff et al.*, 1997; *McConnell et al.*, 1998]. In this case study, the LR test is used to assess the equality of slopes and variances for the 1991 and 1996 models and a model based upon the pooled data sets. However, this test examines the equality of the parameter estimates and their underlying population parameters. Testing the equality of coefficient estimates may not be guaranteed when pooling data sets. Showing that two samples originate from the same underlying distribution does not guarantee equality of either their coefficient estimates or their explained variance. All it shows is that the estimated models based on the individual samples equal the pooled model, not that the individual models themselves are the same.

[14] Therefore also the Wald test is used in this case study in order to be able to compare the estimated coefficients of the 1991 and 1996 model directly. The use of the Wald test has been proposed as a more appropriate test for benefit function transfer than the tests mentioned above [*Brouwer and Spaninks*, 1999].

## 4. Analytical Methods for WTP Estimation and Testing Model Transfer

[15] Mean and median WTP measures for DC responses are inferred from the underlying statistical distribution of the probability that respondents say yes or no to the different bid amounts used [*Hanemann and Kanninen*, 1999]. Different mean WTP values can be calculated depending on the statistical specification of the valuation function and the truncation strategies used. The calculation procedures for mean and median values based on binary WTP response data were first outlined by *Hanemann* [1984]. Various statistical models exist to estimate underlying distributions. These include parametric approaches such as logit or probit models, and distribution-free or nonparametric approaches such as the one developed by *Kriström* [1990] or the Turnbull estimator [*Haab and McConnell*, 1997]. These models can produce significantly different results, with the Turnbull estimator yielding a lower bound on WTP. In this study, both parametric (logistic regression) and nonparametric (Turnbull estimation) approaches are used to test the temporal reliability of WTP values.

[16] Turning to consider tests of model transferability, a novel iterative approach is developed in order to see how much control is needed to produce transferable models of WTP. These models are generated by progressively blending theoretically expected determinants of WTP with additional ad hoc variables, which may be more transitory in their effect. This approach involves a gradual expansion in the number of explanatory variables added to a model of WTP. We begin with models solely utilizing variables which economic theory suggests should influence WTP (initially a one-variable model relating WTP to the DC bid amount, then extending this by adding household income). The resulting model is then extended through the progressive addition of various other ad hoc control variables (such as whether the respondent lives locally, their rating of the scenery in the Norfolk Broads, etc.). At each addition of a variable we assess temporal transferability by applying the model to both the 1991 and 1996 data and undertaking various tests as described subsequently. This progressive expansion approach allows us to identify the optimal level of control for transferability. We compare this approach to that obtained by estimating a statistical best fit model for a given data set (as reported in Table 4) and transferring this to the other survey period and vice versa.

[17] For each model we assess transferability both forward in time (from 1991 to 1996) and back (from 1996 to 1991) using the Wald test for coefficient stability as per *Brouwer and Spaninks* [1999]. A further test of the transferability of each specification is obtained by pooling the data and assessing transferability through application of the likelihood ratio (LR) test as per *Downing and Ozuna* [1996] or *Carson et al.* [1997]. For this latter test we pool data from the two surveys and simply include a dummy variable representing the year in which the study was undertaken. If study year has a significant impact on respondent WTP, this implies that the study results are not transferable. The pooled regression results are the same as the outcomes of the LR test.

## 5. Results

### 5.1. Temporal Reliability of the Mean WTP Values

[18] Mean WTP values based on the parametric and nonparametric estimation approaches discussed in section 4 are presented in Table 2. In order to be able to compare the 1991 and 1996 WTP values, the 1996 values are corrected

**Table 3.** Estimated Univariate Linear-Logit Models

| | Probability ($y_i$ = Yes) | |
|---|---|---|
| | 1991 | 1996 |
| Constant | 1.985[a] | 2.158[a] |
| | (0.082) | (0.107) |
| Bid level | −0.008[a] | −0.010[a] |
| | (0.0005) | (0.001) |
| Log likelihood | −778.15 | −915.23 |
| $R^2$ | 0.285 | 0.243 |
| Predictive power | 0.794 | 0.819 |
| *N* | 1747 | 1108 |

[a]Significant at 0.01.

**Table 4.** Best Fit Multivariate Linear-Logit Models for the 1991 and 1996 Surveys

| Explanatory Factors | Value Range[a] | 1991, Probability ($y_i$ = Yes)[b] | Value Range[a] | 1996, Probability ($y_i$ = Yes)[b] |
|---|---|---|---|---|
| Constant | | 0.506 (0.400) | | 0.768[c] (0.407) |
| Bid (the DC bid level presented To respondents) | 1–500 | −0.009[d] (0.0005) | 1–412[e] | −0.008[d] (0.0008) |
| Income (Annual household income, £) | 2500–62500 | 0.249[c] $10^{-4}$(d) (0.564[c] $10^{-5}$) | 2060–51500[e] | 0.193[c] $10^{-4}$(f) (0.833[c] $10^{-5}$) |
| Size (number of persons in the household) | 1–9 | −0.143[f] (0.056) | 1–12 | ... ... |
| Distance (number of kilometers traveled to reach the site) | 0–580 | −0.002[d] (0.0007) | 0–650 | 0.002[c] (0.001) |
| Visits (number of previous visits per annum) | 0–305 | 0.009[f] (0.004) | 0–356 | ... ... |
| Scenic (appreciation of scenery) | 1–4 | 0.513[d] (0.112) | 1–4 | 0.386[d] (0.108) |
| Holidaymaker (respondent was on holiday when interviewed) | 0–1 | ... ... | 0–1 | −0.757[d] (0.269) |
| Log likelihood | | −705.9 | | −426.5 |
| Likelihood ratio test ($\chi^2$) | | 533.3 ($p < 0.01$) | | 145.9 ($p < 0.01$) |
| Pseudo *R*-square, % | | 32.0 | | 15.7 |
| Predictive power, % | | 80.8 | | 81.9 |
| *N*[g] | | 1665 | | 1015 |

[a]Minimum and maximum values.
[b]Standard errors between brackets.
[c]Significant at 0.10.
[d]Significant at 0.01.
[e]Corrected for inflation.
[f]Significant at 0.05.
[g]The number of observations is lower than in Table 3 because of missing values for some of the explanatory factors.

for intervening differences in purchasing power. The 15 DC bid levels used in 1996 (1, 5, 10, 15, 20, 25, 30, 40, 50, 60, 75, 100, 150, 200, 500 pounds) were first converted to the eight bid levels used in 1991 (1, 5, 10, 20, 50, 100, 200, 500 pounds) and subsequently adjusted for differences in purchasing power. (The increase in inflation and purchasing power between 1991 and 1996 is 21 percent (UK National Statistical Office, www.statistics.gov.uk.) The estimated linear-logistic models are presented in Table 3. The standard errors in the Turnbull models are estimated using nonparametric bootstrapping [e.g., *Efron and Tibshirani*, 1993].

[19] The results from the linear-logistic and Turnbull models suggest that visitor valuation of the recreational and amenity benefits provided by the Broads has decreased across the period between the two surveys. In constant prices, mean WTP calculated from the linear-logistic model is 13 percent lower in 1996 than in 1991, and 30 percent in the case of the Turnbull model. The observed difference in income levels between the 1991 and 1996 visitors in Table 1 is one possible explanation for this decrease.

[20] Although the Turnbull model is known to provide a lower bound for mean WTP, the large difference between the Turnbull and linear-logistic model is striking. The parametric estimates are about 5 times higher than the nonparametric estimates. No big differences exist in terms of the accuracy of the estimates. In relative terms the standard errors of the linear-logistic estimates are only slightly higher than the standard errors of the Turnbull estimates.

[21] The differences in mean WTP are statistically significant, as can be seen from the 95 percent confidence interval (CI) constructed around their difference based on

**Table 5.** Transfer Test Results From the DC CV Models[a]

| Transfer | Test | Model Specification: Bid | Bid Income | Bid Income Distance | Bid Income Local | Bid Income Distance Scenic | Bid Income Local Scenic | Best Fit 1991 | Best Fit 1996 |
|---|---|---|---|---|---|---|---|---|---|
| Transfer of the estimated 1991 models to 1996 | Wald | (0.93) | (3.71) | 9.70 | (3.51) | 13.20 | (5.88) | 20.50 | 15.03 |
| | $\chi^2_{critical}$ | (5.99) | (7.81) | 9.45 | (9.49) | 11.07 | (11.07) | 14.07 | 12.59 |
| | LR | (0.58) | (2.19) | (6.19) | (2.07) | (7.97) | (3.23) | (11.49) | (10.40) |
| | $\chi^2_{critical}$ | (5.99) | (7.81) | (9.45) | (9.49) | (11.07) | (11.07) | (14.07) | (12.59) |
| Transfer of the estimated 1996 models to 1991 | Wald | (1.64) | (5.31) | 15.98 | (4.98) | 19.92 | (7.45) | 26.35 | 30.61 |
| | $\chi^2_{critical}$ | (5.99) | (7.81) | 9.45 | (9.49) | 11.07 | (11.07) | 14.07 | 12.59 |
| | LR | (0.58) | (2.19) | (6.19) | (2.07) | (7.97) | (3.23) | (11.49) | (10.40) |
| | $\chi^2_{critical}$ | (5.99) | (7.81) | (9.45) | (9.49) | (11.07) | (11.07) | (14.07) | (12.59) |

[a]Critical values at 5%. Values in parentheses indicate that null hypothesis of model equality cannot be rejected (model is transferable).

the standardized normal variable (*z*). The hypothesized value of zero (no difference) does not fall in the 95 percent CI, and hence the null hypothesis of equal mean WTP values is rejected at the 5 percent level in favor of the alternative hypothesis that the values are not the same. The estimated differences indicate that the real value of the recreational amenities in the Broads have decreased by 3–6 percent per annum over the study period. This significant decrease in real WTP is in contrast to the nonsignificant changes noted over shorter periods and may well be a consequence of the longer interval under consideration in our study.

### 5.2. Temporal Reliability of the Multivariate Contingent Choice Models

[22] Results from our various analyses of model transferability are shown in Table 5. The definition of explanatory variables is given in Table 4.

[23] Considering the results presented in Table 5, we observe that using the LR test, all models appear transferable. However, adopting the Wald test (which we argue is more stringent) yields a more mixed result, but one from which a clear pattern emerges. Focusing upon these latter tests, we observe that both of the models relying solely upon variables suggested by economic theory (models using the bid variable alone or those supplementing this with the household income variable) are transferable. However, when such models are extended through the addition of more ad hoc variables, not derived from theory, then transferability becomes sporadic. Here those models using the binary local variable (identifying those respondents who live near to the study site) do transfer, whereas those substituting in the continuous distance variable (the number of kilometers traveled to reach the site) fail Wald tests of transferability, questioning the usefulness of more sophisticated distance-decay relationships in models of WTP for transfer purposes. Statistical best fit models also fail Wald transferability tests. This reflects the differing determinants, which enter each of these models (as detailed in Table 4). (The variables size and visits are included in the best fit model of the 1991 data and excluded from the corresponding model of the 1996 data. The reverse holds for the holidaymaker variable.)

## 6. Discussion and Conclusions

[24] This study investigates the temporal stability of WTP responses from two large-scale CV surveys. The study differs from previous analyses because of the large time span between the two surveys, being more than twice the length of previously considered test-retest periods. While previous studies considering shorter periods have shown no significant difference in real WTP values, our analysis reveals a significant difference across this longer period. However, tests of model transferability indicate that simple models, based solely upon variables derived from economic theory, are transferable across this period. This suggests that underlying relationships for such key determinants are stable even across this longer period. However, expanding models by including theoretically unanticipated factors brings ad hoc and possibly transitory factors into the models, which consequently prove nontransferable.

[25] Using commonly used testing procedures in the benefits transfer literature, we are able to show that also DC models extended with these ad hoc factors are transferable, even though the residual variance in these statistically best fit models is significantly different in the two survey years. Contrary to previous findings, this seems to suggest that "the unobserved determinants of preference embedded in the stochastic components of utility over time" [*McConnell et al.*, 1998, p. 372] is not stable in this study. The 1996 model explains less of the variability in the dependent variable than the estimated 1991 model. Hence we may have overlooked important determinants of WTP, which have stayed unobserved.

[26] Indeed, the questionnaire used in the 1996 survey included a number of additional questions, which may shed more light upon the changing conditions over time and their effect on stated WTP. For instance, the perceived feasibility of implementing the flood alleviation scheme appeared to be a highly significant determinant of stated WTP in the 1996 survey [*Powe*, 1999], especially among local residents after years of consultation over whether or not to implement the flood alleviation scheme in the region [*Brouwer et al.*, 1999]. However, information about respondent perception of the feasibility of implementing the proposed flood alleviation scheme is not available in the 1991 survey, as this was not considered an issue then.

[27] In conclusion, our study suggests that over extended periods real WTP for public goods, such as the flood protection and wetland conservation scheme considered here, can change by statistically significant amounts. However, our analysis suggests that underlying economic-theoretic determinants of WTP remain stable over such periods. Nevertheless, ad hoc changes in determinants other than those predicted by theory can result in nontransferability of extended (and statistically best fit) models. This suggests that transfer exercises might usefully focus upon models with firm theoretical underpinnings rather than incorporating more transitory factors.

## 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, *Fed. Regist.*, *58*(10), 4601–4614.

Bateman, I. J., K. W. Willis, and G. D. Garrod (1994), Consistency between contingent valuation estimates: A comparison of two studies of UK national parks, *Reg. Stud.*, *28*(5), 457–474.

Bateman, I. J., I. H. Langford, R. K. Turner, K. G. Willis, and G. D. Garrod (1995), Elicitation and truncation effects in contingent valuation studies, *Ecol. Econ.*, *12*(2), 161–179.

Bergland, O., K. Magnussen, and S. Navrud (1995), Benefit transfer: Testing for accuracy and reliability, *Discussion Pap., D-03/1995*, Dep. of Econ. and Social Sci., Norw. Univ. of Life Sci., Ås, Norway.

Broads Authority (1997), Broads Plan 1997: The strategy and management plan for the Norfolk and Suffolk Broads, Norfolk, UK.

Brouwer, R., and F. A. Spaninks (1999), The validity of environmental benefits transfer: Further empirical testing, *Environ. Resour. Econ.*, *14*(1), 95–117.

Brouwer, R., N. Powe, R. K. Turner, I. J. Bateman, and I. H. Langford (1999), Public attitudes to contingent valuation and public consultation, *Environ. Values*, *8*(3), 325–347.

Brouwer, R., et al. (2001), Management of a multi-purpose coastal wetland: The Norfolk and Suffolk Broads, England, in *Economics of Coastal and Water Resources: Valuing Environmental Functions*, edited by R. K. Turner et al., pp. 159–213, Springer, New York.

Carson, R. T., W. M. Hanemann, R. J. Kopp, J. A. Krosnick, R. C. Mitchell, S. Presser, P. A. Ruud, V. K. Smith, M. Conaway, and K. Martin (1997),

Temporal reliability of estimates from contingent valuation, *Land Econ.*, *73*(2), 151–163.
Downing, M., and T. Ozuna (1996), Testing the reliability of the benefit function transfer approach, *J. Environ. Econ. Manage.*, *30*, 316–322.
Efron, B., and R. J. Tibshirani (1993), *An Introduction to the Bootstrap*, CRC Press, Boca Raton, Fla.
Environment Agency (1997), Broadland flood alleviation strategy: Bank strengthening and erosion protection, report, 15 pp., Suffolk, UK.
Greene, W. H. (1993), *Econometric Analysis*, 2nd ed., Macmillan, New York.
Haab, T. C., and K. E. McConnell (1997), Referendum models and negative willingness to pay: Alternative solutions, *J. Environ. Econ. Manage.*, *32*, 251–270.
Hanemann, W. M. (1984), Welfare evaluations in contingent valuation experiments with discrete responses, *Am. J. Agric. Econ.*, *66*, 332–341.
Hanemann, W. M., and B. Kanninen (1999), The statistical analysis of discrete-response CV data, in *Valuing Environmental Preferences: Theory and Practice of the Contingent Valuation Method in the US, EU, and Developoing Countries*, edited by I. J. Bateman and K. G. Willis, pp. 302–441, Oxford Univ. Press, New York.
Kealy, M. J., J. F. Dovidio, and M. L. Rockel (1988), Accuracy in valuation is a matter of degree, *Land Econ.*, *64*, 158–171.
Kealy, M. J., M. Montgomery, and J. F. Dovidio (1990), Reliability and predictive validity of contingent values: Does the nature of the good matter?, *J. Environ. Econ. Manage.*, *19*, 244–263.
Kirchhoff, S., B. G. Colby, and J. T. LaFrance (1997), Evaluating the performance of benefit transfer: An empirical inquiry, *J. Environ. Econ. Manage.*, *33*, 75–93.
Kriström, B. (1990), A non-parametric approach to the estimation of welfare measures in discrete response valuation studies, *Land Econ.*, *66*, 135–139.
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 Resour. Res.*, *31*(3), 721–730.
McConnell, K. E., I. E. Strand, and S. Valdes (1998), Testing temporal reliability and carry-over effect: The role of correlated responses in test-retest reliability studies, *Environ. Resour. Econ.*, *12*, 357–374.
National Rivers Authority (1993), A flood alleviation strategy for Broadland: Final report Annex 3—Cost benefit studies, report, 100 pp., Peterborough, UK.
Powe, N. A. (1999), Economic valuation of wetland recreation and amenity values, Ph.D. thesis, School of Environ. Sci., Univ. of East Anglia, Norwich, UK.
Powe, N. A., and I. J. Bateman (2003), Ordering effects in nested "top-down" and "bottom-up" contingent valuation designs, *Ecol. Econ.*, *45*, 255–270.
Powe, N. A., and I. J. Bateman (2004), Investigating insensitivity to scope: A split-sample test of perceived scheme realism, *Land Econ.*, *80*(2), 258–271.
Turner, R. K., and J. Brooke (1988), Management and valuation of an environmentally sensitive area: Norfolk Broadland, England, case study, *Environ. Values*, *12*(2), 193–207.

I. J. Bateman, Centre for Social and Economic Research on the Global Environment (CSERGE), School of Environmental Sciences, University of East Anglia, Norwich, UK.

R. Brouwer, Institute for Environmental Studies (IVM), Vrije Universiteit, De Boelelaan 1087, 1081 HV Amsterdam, Netherlands. (roy.brouwer@ivm.vu.nl)