Survey

Valuing water level changes in reservoirs using two stated preference approaches: An exploration of validity

Nele Lienhoop *, Till Ansmann

Helmholtz-Centre for Environmental Research (UFZ), Department of Economics, Permoser Straße 15, 04318 Leipzig, Germany

ARTICLE INFO

Article history:
Received 22 October 2010
Received in revised form 16 February 2011
Accepted 19 February 2011
Available online xxxx

Keywords:
Water levels changes
Contingent valuation
Contingent behaviour
Travel cost method
Validity
Count data model

ABSTRACT

The combination of travel cost (TCM) and contingent behaviour (CB) methods is a relatively new research avenue in the recreational valuation community. Contrary to simple TCM applications, TCM–CB facilitates the ex ante valuation of marginal welfare effects resulting from environmental quality or quantity changes, similar to the contingent valuation method (CV). Even though TCM–CB is highly policy relevant, i.e. to inform changes in management regimes at recreational sites, the validity of estimates has hardly received any attention and little is known about the performance of TCM–CB compared to CV. In this paper, TCM–CB and CV are explored with respect to several validity tests in a case study on the recreational effects of water level changes in a reservoir. Overall, the findings reveal that TCM–CB and CV perform equally well in terms of theoretical validity, but that the marginal recreational value varies significantly between the two methods. We also observe that both methods face similar internal difficulties with respect to the stability of values when the order of a set of valuation questions is changed.

© 2011 Elsevier B.V. All rights reserved.

1. Introduction

The recreation economics literature is dominated by the application of the travel cost method (TCM) to assess recreational demand for a given recreational site. Although TCM is suited to infer the recreational benefits that a site provides under current conditions, it is very difficult to value marginal benefits or costs of environmental changes with this method (Gillig et al., 2003). From a policy-making point of view, however, the valuation of marginal effects is highly relevant when changes to environmental quality/quantity or management rules at recreational sites are discussed. In order to address this shortcoming of TCM, considerable research effort has focussed on combining TCM with scenario-based stated preference data. A number of authors have developed econometric approaches to combine TCM and contingent valuation (CV) data (Gonzalez et al., 2008; Huang et al., 1997; Loomis, 1997). However, procedures to establish consistency between the two data types (i.e. imposing identical underlying utility functions) have turned out to be very complex and prone to information loss. Another, very recent, research avenue focuses on the combination between TCM and contingent behaviour (CB) data (e.g. Adamowicz et al., 1994; Azevedo et al., 2003; Grossmann, 2010; Prayaga et al., 2010). In addition to eliciting information about current travel distance and frequency such as in standard TCM, CB asks individuals to state their intended visit frequency if a hypothetical change in environmental quality/quantity at the site under investigation occurred (Chakraborty and Keith, 2000; Eiswerth et al., 2000; Grijalva et al., 2002; Hanley et al., 2003; Whitehead et al., 2000).

At first glance, an advantage of combining TCM with CB data is that CB data seems to be directly compatible with TCM data and thus makes the combination of the two data types relatively straightforward and pragmatic. An important prerequisite, however, is that both welfare estimates generated with TCM and CB have the same underlying utility function (Hensher et al., 1999). A range of papers have developed procedures to ensure compatibility of revealed and stated preference data (Adamowicz et al., 1997; Boxall et al., 2003; Hensher et al., 1999; Huang et al., 1997; Whitehead et al., 2000). A further advantage is the improvement of model efficiency to estimate welfare measures (e.g. Huang et al., 1997; Whitehead et al., 2000). Despite these advantages, there are also potential problems associated with the collection of CB data. Asking respondents to state their expected number of trips under a certain scenario is a cognitively demanding task, and hence may lead to failure to anticipate the ‘true’ number of trips. While other stated preference methods, i.e. CV have been a subject to extensive explorations with respect to the validity of
respondent decisions under hypothetical situations, the validity of CB
data has received very little attention even though the method is
equally controversial due to its hypothetical nature. Exceptions are
Grijalva et al. (2002) and Jeon and Herriges (2010) who compared
stated ‘pre-quality change’ trips with real ‘post-quality change’ trips.
Findings are mixed: the former paper supports and the latter rejects
criterion validity.

In terms of policy-making, a combined TCM–CB approach seems to
be very promising. Bearing in mind that CV is a well-established (but
also controversial) method for the valuation of marginal costs and
benefits, it is worthwhile to compare CV and TCM–CB in terms of their
performance in the valuation of recreational costs or benefits. Against
the background of policy decisions, a pressing issue is the validity of
welfare estimates. So far no effort has been spent on comparing the
validity of TCM–CB estimates with CV, and hence little is known about
the advantages of TCM–CB compared to CV in terms of validity or vice
versa. The research described in this paper is intended to fill this
research gap. The overall aim is to explore two valuation methods –
combined TCM–CB, and CV – under policy relevant scenarios with
respect to the validity of welfare estimates they generate. Specifically,
we address four issues:

1) Assess the validity and robustness of TCM–CB and CV in the form of
comprehensive regression analyses.
2) Test the sensitivity of TCM–CB and CV estimates to changes in size
(small, medium and significant environmental change).
3) Test the sensitivity of the two methods towards order effects.
4) Compare mean estimates and test the convergent validity of TCM–CB
and CV data.

The background to our methodological investigation is a study on
reservoir management for flood control and recreation in the Ore
Mountains, located in South Eastern Germany. In this region, the
traditional role of reservoirs has been the provision of water to water
works and to the mining industry. During the past decades a number
of reservoirs have become popular recreational sites, with bathing,
boating, walking and camping being the main recreational activities.
Thus, recreationalists have become a very important beneficiary of
reservoirs. As a result of a devastating flood in the region in 2002 and
anticipating an increase of flood events due to climate change, the
flood detention function of reservoirs has gained considerable
importance. In order to alleviate future floods, an obvious solution
would entail a decrease in water level to increase the flood detention
volume in some reservoirs (Kundzewicz et al., 2007). Such measures are
expected to aggravate the conflict with recreational uses, as water
level changes in the past have led to conflicts between reservoir
managers and recreational parties that prefer high water levels
(Sieber, 2003). In order to inform decisions regarding flood detention
volumes, this study aims to assess the costs that would accrue to
recreational users as a result of decreasing water levels at the
reservoir Pöhl, a typical reservoir in the Ore Mountains. Three
alternative scenarios of water level decrease are valued to enable
integrative consideration of flood control and recreation in the design
of new reservoir management regimes.

2. Survey Design and Validity Tests

Since recreational users are the main losers of flood control
measures, a survey among visitors was undertaken at the reservoir
Pöhl during the summer of 2008. 591 visitors were approached at
beaches, car parks and walking paths. First, respondents were given
an explanation about the purpose of the survey and asked to answer a
few questions regarding their travel behaviour. Respondents were
then provided with a detailed self-administered questionnaire and a
pre-paid envelope. They had the choice to fill in the questionnaire on
site and return it to one of the five collection boxes, or to take it home
and send it back within one week. The questionnaire included the
main components of TCM, CB and CV methods and was structured in
the following way:

1) Questions on socio-economic characteristics and views towards
recreation and flood control.
2) Questions on current travel and recreation behaviour (TCM
component).
3) Background information about flood control in reservoirs.
4) Description of three hypothetical scenarios of water level decrease
(−1, −3.5 and −5 m).
5) Questions on changes to travel behaviour under the three
hypothetical scenarios (CB component).
6) A hypothetical market to elicit WTP to avoid a decrease in water
levels (CV component).

The main part of the questionnaire was the valuation of three
scenarios of water level decrease in terms of hypothetical visit rate
and willingness to pay (WTP). In order to enable respondents to make
judgements regarding the change in water levels, each scenario was
described with pictures showing changes to the shoreline and in
terms of implications on recreational activities. TCM and CB data
were obtained by asking the respondents about their current number
of trips per year and travel distance to the reservoir (TCM) and their
expected number of trips per year under each scenario (CB). In order
to elicit WTP in the CV exercise, a hypothetical market was
constructed. It assumed that alternative flood control measures (e.g.
larger dykes and polders) would be more expensive than flood control
in reservoirs, but would mean that a high water level in the reservoir
could be maintained. Payments to support these alternative flood
control measures and thus maintain the current water level would be
collected in the form of a user fee to be paid at beaches, camp sites, car
parks and boat hires. The user fee was framed as a daily household
payment. WTP was elicited for each water level scenario using a
payment card with fourteen bids ranging from zero to 15 Euros. The bid
range was determined on the basis of a pre-test. Given that respondents
were presented with all scenarios prior to WTP elicitation on a mutually
exclusive list, we felt that the payment card was more suited than the
generally favoured dichotomous choice format, as it enables the
respondents to accurately distinguish between their preferences for
each of the three scenarios. One of the main criticisms of payment cards
is the potential for incentive incompatibility. While we cannot fully
exclude the potential of incentive incompatible behaviour, we have
some evidence based on respondents’ explanations regarding their WTP
decision that bids are well founded (i.e. decisions were based on budget
constraints and perceived importance of water level reductions).
In order to ensure comparability of CV with CB, the hypothetical market
focussed merely on recreational use values. In order to test for order
effects, the sequence in which the three scenarios were presented was
reversed in a split sample: half of the sample was asked to value the
three scenarios in an ascending order (−1, −3.5, and −5 m) and the
other half was offered the scenarios in a descending order, (−5, −3.5,
and −1 m). The respondents were presented with all three scenarios
before they were asked to value them.

In order to explore the two valuation methods (TCM–CB and CV)
in terms of the validity of welfare estimates, several tests of validity
were employed.

- Theoretical validity tests to what degree a measure (visit rate and
WTP) can be explained by factors that are theoretically expected to
explain the measure (Mitchell and Carson, 1989). We tested the
theoretical validity of TCM–CB and CV in two ways:
  1) By testing the functional relationship between individual visit
rate/WTP and independent variables (household income, recre-
national activities, and further socio-demographic characteristics)
that are theoretically expected to explain these measures using
regression analyses.
2) By comparing mean ‘visit rate’ and WTP for different conditions (here water level changes) for which theory suggests different values. In this way, the methods’ sensibility to size under policy relevant scenarios is tested. Economic theory predicts that welfare losses (reduced visit rate/higher WTP to avoid the decrease) should be larger for significant reductions in water levels than for smaller reductions in water level. This is known as size effect. The hypothesis to be tested is that changes in reservoir water levels cause changes in welfare estimates and that this is the case in both CB and CV:

\[ H_0: \text{Costs}_{\text{Sc-1m}} \leq \text{Costs}_{\text{Sc-3.5m}} \leq \text{Costs}_{\text{Sc-5m}} \]

versus

\[ H_1: \text{Costs}_{\text{Sc-1m}} \approx \text{Costs}_{\text{Sc-3.5m}} \approx \text{Costs}_{\text{Sc-5m}} \]

where the alternative hypothesis is the equivalence of all cost measures (insensitivity to changes in size). Failure to reject the null hypothesis provides evidence of theoretical validity.

- **Order effects:** According to economic theory the value for an environmental good is independent on the serial position of the good in a sequence of other environmental goods, because preferences are assumed to be well-defined, stable, consistent and context-independent. Further, order effects should not occur, if goods appear on a mutually exclusive list (Clark and Fiesen, 2008) and when respondents are presented with all levels of an environmental change prior to the elicitation of WTP (Bateman and Langford, 1996). Based on these findings and economic theory, we would expect that the order in which water levels are presented should not influence the value for each scenario of water level decrease. It is therefore hypothesised that welfare estimates for the three scenarios of water level decrease are unaffected by the order in which they are presented:

\[ H_0: \text{Costs}_{\text{Sc-1m}} \ldots \text{Costs}_{\text{Sc-5m}} = \text{Costs}_{\text{Sc-5m}} \ldots \text{Costs}_{\text{Sc-1m}} \]

versus

\[ H_1: \text{Costs}_{\text{Sc-1m}} \ldots \text{Costs}_{\text{Sc-5m}} \neq \text{Costs}_{\text{Sc-5m}} \ldots \text{Costs}_{\text{Sc-1m}} \]

where the alternative hypothesis is the divergence in ‘visit rate’/WTP for a certain scenario depending on the order in which it is presented. Rejection of the null hypothesis for ‘visit rate’ or WTP would indicate the presence of an order effect in the respective method.

- **Convergent validity** assumes that two measures of the same construct should give equal results (Loomis and Gonzalez-Caban, 1997). Convergent validity is frequently used to check the consistency of estimates generated with TCM and CV (e.g. Herath and Kennedy, 2004). If correlation exists, the validity of each method can be confirmed. However, a correspondence between the methods does not prove the accuracy of estimates, but rather contributes to the credibility of both methods (Mitchell and Carson, 1989). We assume that the marginal costs of decreasing water levels are equal between CB and CV.

\[ H_0: \text{Costs}_{\text{CB}} = \text{Costs}_{\text{CV}} \]

versus

\[ H_1: \text{Costs}_{\text{CB}} \neq \text{Costs}_{\text{CV}} \]

where alternative hypothesis is the divergence of cost measures. Failure to reject the null hypothesis confirms convergent validity.
under each of the three scenarios of water level change (CB) showed that there are only marginal differences between the utility functions underlying the different sets of visit rate data. We thus conclude that TCM and CB data are comparable and can be combined.

Further exploration of the data shows that a decrease in reservoir water levels reduces the recreational value perceived by visitors. This is reflected in a decreasing number of trips and increasing WTP with decreasing water levels. Fig. 2 shows the standardised size effect of water level change under the three scenarios of water level decrease. Sampled visitors would on average take one, five and eight fewer trips under the three scenarios, respectively, whereas mean WTP to avoid the scenarios amounts to €0.45, 0.72 and 1.03 per visit day. According to a Mann–Whitney test the difference in trips is not significant between the status quo and scenario 1 (−1 m) (p=0.786 level; z=0.271); however the number of trips is significantly reduced between scenarios 1 and 2 (−1 m versus −3.5 m) (p=0.0186; z=2.354) and scenarios 2 and 3 (−3.5 m versus −5 m) (p=0.0154; z=2.424). This corresponds with the WTP data, where mean WTP increases significantly between scenarios 1 and 2 (p=0.0002, z=−3.772) and scenarios 2 and 3 (p=0.0074, z=−2.405). Additionally, Fig. 2 shows how the standardised effects of a decrease in water levels differs according to two subsamples that were asked to value the three scenarios either in an ascending (Subsample Sc−1−3.5−5 m upward) or in a descending order (Subsample Sc−5−3.5−1 m).

Table 1
Dependent and explanatory variables used in the regression models.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Unit</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VR</td>
<td>Visit rate under current condition</td>
<td>Trips/year</td>
<td>17.27</td>
<td>34.63</td>
</tr>
<tr>
<td></td>
<td>Hypothetical visit rate under scenario 1 (−1 m)</td>
<td>Trips/year</td>
<td>16.32</td>
<td>33.70</td>
</tr>
<tr>
<td></td>
<td>Hypothetical visit rate under scenario 2 (−3 and 5 m)</td>
<td>Trips/year</td>
<td>12.28</td>
<td>30.37</td>
</tr>
<tr>
<td></td>
<td>Hypothetical visit rate under scenario 3 (−5 m)</td>
<td>Trips/year</td>
<td>9.65</td>
<td>25.94</td>
</tr>
<tr>
<td>WTP</td>
<td>Willingness to pay to avoid scenario 1 (−1 m)</td>
<td>€/trip day</td>
<td>0.45 (0.73)</td>
<td>0.99 (1.18)</td>
</tr>
<tr>
<td></td>
<td>Willingness to pay to avoid scenario 2 (−3 and 5 m)</td>
<td>€/trip day</td>
<td>0.72 (1.17)</td>
<td>1.23 (1.40)</td>
</tr>
<tr>
<td></td>
<td>Willingness to pay to avoid scenario 3 (−5 m)</td>
<td>€/trip day</td>
<td>1.03 (1.69)</td>
<td>1.46 (1.55)</td>
</tr>
<tr>
<td><strong>Explanatory variables</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Order_up</td>
<td>Presentation of scenarios in ascending (−1, −3.5, and −5 m) or descending order (−5, −3.5, and −1 m)</td>
<td>1 = −1, −3.5, −5, 0 = −5, −3.5, −1</td>
<td>0.61</td>
<td></td>
</tr>
<tr>
<td>D_Sc−1m</td>
<td>Dummy indicating contingent behaviour data under scenario 1</td>
<td>1 = yes</td>
<td>0 = no</td>
<td></td>
</tr>
<tr>
<td>D_Sc−3.5m</td>
<td>Dummy indicating contingent behaviour/WTP data under scenario 2</td>
<td>1 = yes</td>
<td>0 = no</td>
<td></td>
</tr>
<tr>
<td>D_Sc−5m</td>
<td>Dummy indicating contingent behaviour/WTP data under scenario 3</td>
<td>1 = yes</td>
<td>0 = no</td>
<td></td>
</tr>
<tr>
<td>Cost</td>
<td>Travel costs (round trip kilometres commuter tax allowance)</td>
<td>€/trip</td>
<td>48.92</td>
<td>78.62</td>
</tr>
<tr>
<td>Inc</td>
<td>Net household income</td>
<td>€/month</td>
<td>1902.93</td>
<td>933.36</td>
</tr>
<tr>
<td>Age</td>
<td>Age of the respondent</td>
<td>Years</td>
<td>45.83</td>
<td>15.21</td>
</tr>
<tr>
<td>Swim</td>
<td>Activity at reservoir is swimming</td>
<td>1 = yes</td>
<td>0 = no</td>
<td>0.66</td>
</tr>
<tr>
<td>Hiking</td>
<td>Activity at reservoir is hiking</td>
<td>1 = yes</td>
<td>0 = no</td>
<td>0.54</td>
</tr>
<tr>
<td>Subs</td>
<td>Respondent would visit another lake/reservoir if water levels were reduced</td>
<td>1 = yes</td>
<td>0 = no</td>
<td>0.06</td>
</tr>
</tbody>
</table>

* Results from the sensitivity analysis (73 zero WTP bids excluded, n=115).
4. Modelling Approach

This section describes the selection of a model that suits the data and at the same time facilitates comparability of TCM–CB and CV data.

4.1. Model Specification: TCM–CB

Since the research focuses on water level changes at a specific recreational site, a single site demand model seemed appropriate for the analysis of TCM–CB data. Among single demand models, count data models are most commonly applied (e.g. Chakraborty and Keith, 2000; Eiswerth et al., 2000; Martínez-Martínez-Espiñeira and Amoako-Tuffour, 2008; Whitehead et al., 2000). Count data specifications are seen as a logical extension to accommodate the particular properties of trip data. Contrary to classical regression analysis, count data models have features that are better suited to analyse actual and hypothetical trip data to the reservoir, because the response variable is discrete with a distribution that places probability mass at integer and non-negative values only. Furthermore, count data distributions are characterised by a concentration of values on few small discrete values, skewness to the left and heteroskedasticity with variance increasing with the mean.

The count variable $Y$ (number of trips) in our TCM–CB model is an integer and non-negative. One approach to model such count data is to assume that the dependent variable is drawn from a Poisson distribution

$$f(y|x) = \frac{e^{-\lambda} \lambda^y}{y!},$$

where the Poisson parameter $\lambda$ is conditioned on the observed covariates, $x$, such that

$$\ln(\lambda_x) = \beta^t x_t.$$  

However, a problem with Poisson specifications is that the distribution constrains the mean and the variance of $Y$ to be equal, a property known as equidispersion. Due to this fact Poisson specifications are often not suitable for trip data, as the number of trips is often overdispersed, that is, variance is larger than mean, because a few visitors take many trips while most visitors take few trips. An alternative is to use a negative binomial regression model, where the count variable is assumed to be generated by a Poisson-like process, except that the variation is greater than in a true Poisson distribution (Englin and Shonkwiler, 1995). This model relaxes the assumption that variance equals mean and thus permits more flexible modelling of variance. In order to not impose this restriction a priori, a negative binomial specification was chosen, which nests the Poisson distribution as a special case. The negative binomial model extends the Poisson model by adding an additional parameter, $\alpha$, that reflects the unobserved heterogeneity that the Poisson distribution fails to capture. This can be written as

$$\ln(\lambda_x) + \ln(\alpha_x) = \beta^t x_t + \epsilon_i.$$  

The probability of observing $y$ conditional on both $x$ and $\alpha$ has the same structure as the Poisson distribution. To evaluate this distribution conditional only on $x$, a distribution for $\alpha_\psi = \exp(\epsilon_i)$ must be specified. If a gamma density is assumed, the conditional distribution, $f(y|x,\alpha)$, is negative binomial with mean $\lambda$ and variance $\lambda(1+\psi\lambda)$, where $\psi$ is the parameter of the gamma distribution.

TCM and CB data were combined in a panel data format since this allows the application of either pooled or panel data models. Generally, panel approaches with fixed or random effects estimators have advantages over simpler pooled approaches as they take account of differences in individual behaviour (see Englin and Cameron, 1996). However, due to our interest in separate contingent behaviour effects on ‘visit rate’ in the model, it was decided to estimate a pooled data model (Eiswerth et al., 2000 and Grijalva et al., 2002). The pooled TCM–CB count data model assumes a semi-log function and can be written as

$$\ln(y_t) = \beta_0 + \beta_1 \text{Cost} + \beta_2 \text{Inc} + \beta_3 \text{Age} + \beta_4 D_{5c-1m} + \beta_5 D_{5c-3.5m} + \beta_6 D_{5c-5m} + \beta_7 \text{Swim} + \beta_8 \text{Hiking} + \beta_9 \text{Subs} + \beta_10 \text{Order}_{up} + \epsilon_i,$$

where $\beta_0$ is the unknown coefficient to be estimated.

A common feature associated with TCM data collected on site is truncation and endogenous stratification of the data. Truncation is caused naturally since all the respondents are visitors and thus have taken at least one trip to the reservoir (Shaw, 1988; Creel and Loomis, 1990; Ven and Adamowicz, 1993; Englin and Shonkwiler, 1995). Similarly, CB data on expected trips do not include zero trips when the envisaged environmental change is positive. Endogenous stratification occurs when the likelihood of being sampled is related to the number of trips taken, thus frequent visitors are usually overrepresented in visitor samples (Shaw, 1988; Englin and Shonkwiler, 1995). Both features lead to an upward bias in visit rate. In order to address this problem Englin and Shonkwiler (1995) derived and operationalised a density function of the negative binominal distribution truncated at zero and adjusted for endogenous stratification for

\[2\] A comparison of pooled and panel (random effects negative binomial regression) models showed that the less restrictive pooled model delivered more plausible results (e.g. in the panel model the coefficient cost is not significant and leads to inflated values of consumer surplus).
overdispersed data. However, our study differs from the majority of the existing studies, as it is concerned about a negative environmental change (lowering reservoir water levels) and therefore includes zero trips in the CB data. Approaches to correct for truncation and endogenous stratification exist merely for data without zero values, that is expected trips under scenarios of environmental improvements (see Englin and Shonkwiler, 1995 and Yen and Adamowicz, 1993) and not for data including zero trips. Given this gap and considering our focus on a methodological comparison between TCM–CB and CV a standard negative binomial specification was applied. Hence, the estimators of our model are not corrected for truncation and endogenous stratification and will not be used to make inferences to the underlying population.

4.2. Model Specification: CV

Since WTP data was elicited using a payment card approach, an ordinary linear regression model was employed rather than logit or probit models that are commonly used for the analysis of dichotomous choice data. In order to ensure comparison with the TCM–CB model, the data for the CV model was combined in a panel data format and estimated using the maximum likelihood estimation technique instead of the preferred algebraic solution with ordinary least squares. The motivation for this choice was that we wanted to ensure comparability of the TCM–CB and the CV models in terms of log-likelihood values. The value of the log-likelihood function indicates the fit of each model for the sampled data. The CV model differs from the TCM–CB model in that CV is incapable of generating estimates for the status quo. Visitors’ WTP was entered to the equation in a logarithmic form. The semi-logarithmic function of the CVM may be written as

$$\ln(WTP_i) = \beta_0 + \beta_1 Inc + \beta_2 Age + \beta_3 D_{Sc-3.5m} + \beta_4 D_{Sc-5m} + \beta_5 Swim + \beta_6 Hiking + \beta_7 Subs + \beta_8 Order_{up} + \epsilon_i,$$

where \(\beta_i\) is the unknown coefficient to be estimated.

5. Results

Regression models for the TCM–CB and CV data form the basis for an in depth investigation of the validity of the two valuation methods. To begin with, we explore theoretical validity and order effects by means of regression analysis. This is followed by a comparison of mean estimates to examine convergent validity.

The results for the maximum likelihood estimators of the TCM–CB and CV models are summarised in Table 2 with parameter estimates and their corresponding z-scores in column 2 and 3. The regression disturbances are clustered at the level of the household, in order to correct for the non-independence of repeated household observations in the panel data format. This means that the measures of statistical significance are robust to this dataset design. Using robust standard errors the Wald chi square based on log pseudo-likelihoods is calculated instead of a likelihood ratio chi-square. The overdispersion coefficient \(\alpha\) confirms that in the TCM–CB overdispersion is a problem. Hence, to assume a Poisson distribution would have been overly restrictive and not suitable for the sampled data.

### Table 2

<table>
<thead>
<tr>
<th>Variable</th>
<th>TCM–CB</th>
<th>CV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>0.7647 (1.63)</td>
<td>−1.7401 (−2.27)</td>
</tr>
<tr>
<td>Cost</td>
<td>−0.0047 (−4.24)</td>
<td>b</td>
</tr>
<tr>
<td>Inc</td>
<td>0.0349 (5.47)</td>
<td>b</td>
</tr>
<tr>
<td>Age</td>
<td>−0.1161 (−1.70)</td>
<td>b</td>
</tr>
<tr>
<td>D (_{Sc-1m})</td>
<td>−0.5291 (−5.28)</td>
<td>b</td>
</tr>
<tr>
<td>D (_{Sc-3m})</td>
<td>0.7463 (3.36)</td>
<td>b</td>
</tr>
<tr>
<td>Swim</td>
<td>0.4789 (2.28)</td>
<td>b</td>
</tr>
<tr>
<td>Hiking</td>
<td>−0.8359 (−4.40)</td>
<td>b</td>
</tr>
<tr>
<td>Subs</td>
<td>0.4437 (2.18)</td>
<td>b</td>
</tr>
<tr>
<td>Costs(<em>{Sc-1m}) &amp; Costs(</em>{Sc-3.5m}) &amp; Costs(_{Sc-5m})</td>
<td>&amp;</td>
<td></td>
</tr>
<tr>
<td>Wald chi squared</td>
<td>231.70 (dif10)</td>
<td>164.59 (dif8)</td>
</tr>
<tr>
<td>Log likelihood</td>
<td>−2340.52</td>
<td>−1287.28</td>
</tr>
<tr>
<td>AIC</td>
<td>4705.05</td>
<td>2594.55</td>
</tr>
<tr>
<td>(\alpha)</td>
<td>1.2841</td>
<td></td>
</tr>
<tr>
<td>Number of observations</td>
<td>748</td>
<td>561</td>
</tr>
</tbody>
</table>

*Significance at 5% or lower p level.
| Significance at 1% or lower p level.
| Significance at 10% or lower p level.

5.1. Theoretical Validity

The results presented in Table 2 illustrate that the theoretical validity of our two measures of welfare is high. TCM–CB and CV estimations are very robust, because most parameters are very similar, have the theoretically expected sign to predict ‘visit rate’ and WTP, and have significant influence on the dependent variables.

In particular, the negative coefficient on travel cost (Cost) in the TCM–CB model is consistent with economic theory and significant at the 1% level. The variable Income is found to be highly significant, but has contradictory signs between the two models. While one would expect ‘visit rate’ and WTP to rise with higher income, the negative influence of income in the TCM–CB model is intuitively surprising, but in line with existing travel cost studies (Creel and Loomis, 1990; Liston-Heyes and Heyes, 1999; Loomis, 2003). It is likely that locals with low income and short travel distance visit the reservoir more frequently, while people with higher income travel to other, more exotic, destinations. This would mean that visits to the reservoir are an inferior good. A further coefficient requiring explanation is the coefficient of the variable Age. The positive sign in the TCM–CB model suggests that the number of trips increases with age, whereas the negative sign in the CV model implies WTP to become smaller with age. A likely explanation for these parameter sign reversals might be differences in the utility function. This finding is in line with Hensher et al. (1999) who found inequality for some parameters between revealed and stated preferences measuring the same construct. The variable Subs has a significant negative sign in both models, indicating that visitors who reported that they would visit a substitute lake if water levels were decreased, would visit the reservoir less frequently under the scenarios of water level decrease.

With respect to the general fit of the models, Akaike’s information criterion suggests that the CV model (2594.55) fits the data better than the TCM–CB model (4705.05), although such a comparison should be done with caution since the two dependent variables in the two models differ.

Further evidence for theoretical validity is revealed by the dummy variables D\(_{Sc-1m}\), D\(_{Sc-3.5m}\) and D\(_{Sc-5m}\) indicating that the size of water level change under the three scenarios has a significant influence on both ‘visit rate’ and WTP. Thus, the models confirm the descriptive statistics reported in Fig. 2. Since the parameter estimates increase with the size of water level change there is evidence that the welfare effects elicited with the two valuation methods are sensitive to size. These findings enable us to support H\(_{Cost}\)(Sc−1m ≤ Costs−3.5m ≤ Costs−5m) and show that both valuation methods are equally strong with respect to size effects.

---

* Please cite this article as: Lienhoop, N., Ansmann, T., Valuing water level changes in reservoirs using two stated preference approaches: An exploration of validity, Ecol. Econ. (2011), doi:10.1016/j.ecolecon.2011.02.014

---

3 Both models (maximum likelihood and OLS) were estimated and revealed identical results.

4 It is specified in the model that the observations are independent across households, but not necessarily within the observations on households. This affects the estimated standard errors and variance–covariance matrix of the estimators, but not the estimated coefficients. (StataCorp, 2005a, pp. 43–47, StataCorp 2005b, pp. 275–280).
5.2. Order Effects

Our initial finding in the descriptive statistics about order effects on ‘visit rate’ and WTP is confirmed in both models (Table 2). The variable Orderš reveals that the order in which the scenarios of water level change is presented affects individual preferences. The direction of the parameter estimates suggests that the welfare effects are smaller if scenarios are presented in an ascending order (−1 m,−3.5 m, and −5 m) and higher if they are presented in a descending order (−5 m,−3.5 m, and −1 m). These findings show that both welfare measures are context-dependent, and fail our hypothesis that welfare measures are equally independent of the order in which they are presented (H0: CostsSc = CostsSc = CostsSc = CostsSc = CostsSc = CostsSc = CostsSc).

5.3. Estimating Recreational Costs and Convergent Validity

In order to estimate the recreational costs associated with water level reductions in monetary terms, it is assumed that trip demand in the TCM–CB model follows a negative binomial distribution (see Eq. (1)). Marshallian consumer surplus (CS) is obtained by integrating the demand curve from the initial price to the choke price, where the choke price reflects the cost at which an individual demands zero trips. Based on Eq. (1) the total CS per visit is simply the inverse of the coefficient of the travel cost variable. That is,

\[
CS = -\frac{1}{\beta_1}. \tag{6}
\]

To calculate CS per year and household, a specific interpretation of the error term is required. Since we assume that the specification errors dominate measurement error in the analysis, we multiply the expression by the observed number of household trips (Y) (Bockstaels et al. 1989, pp. 94–95). Table 3 shows the observed and predicted number of trips under the current situation as well as for the three scenarios of water level change.

The marginal effect of changes in water levels under the scenarios on CS per year is calculated by multiplying Eq. (6) by the decrease in trips indicated by the respondent as a response to the hypothetical scenario, that is the stated number of trips (YScn) subtracted from the observed number of trips (YSc)

\[
ME = \left(\frac{Y_{Sc} - Y_{Scn}}{\beta_1}\right) \tag{7}
\]

where \( n = -1, -3.5, -5 \). A marginal effects value indicates how much CS per household and year would be reduced under the three scenarios, and thus reflects the marginal recreational cost of water level decreases (Table 4). Hicksian welfare measures were not appropriate for our analysis because the reservoir turned out to be an inferior good (see income coefficient in the TCM–CB model in Table 2). Moreover, the income effect in the TCM–CB model is very small, and a comparison of Marshallian CS and Hicksian equivalent variation revealed that the differences between the two welfare measures are very small in this study. The marginal recreational costs derived from CV are directly obtained from the survey data. Stated WTP per day is multiplied by the mean duration of the visit (1.29 days) at the reservoir and the mean number of trips per year (Table 4). Estimates from the sensitivity analysis excluding 73 zero bids are presented in brackets. Confidence intervals were estimated using bootstrap, because the maximum likelihood estimators of the welfare measures from the TCM–CB model are functions of a random variable, and hence their asymptotic distribution is unknown.

The findings indicate that the recreational value of the reservoir would be significantly reduced if water levels were to be decreased by −1, −3.5, and −5 m. The most surprising result is the significant difference between the estimates of marginal recreational costs between the two valuation methods, with CB generating higher recreational costs by a factor ranging between 20 and 71 depending on the scenario. Moreover, the confidence intervals of the estimated models do not overlap, so that the hypothesis expecting equal marginal recreational costs (H0: CostsCB = CostsCV) cannot be supported and thus convergent validity between TCM–CB and CV cannot be confirmed. These significant differences remain when 73 zero bids are omitted.

6. Discussion

TCM–CB is a relatively new method in the recreational valuation context. Since it is able to value marginal changes similar to the longer established CV method, the main interest of this paper was to assess the performance of TCM–CB in terms of its validity in comparison to CV. So far, research effort has been spent on comparisons between TCM and CV, where the focus has been on welfare estimates of the status quo, rather than marginal changes, due to the incapability of TCM to value environmental quality/quantity changes (Gillig et al., 2003). Since the estimates of marginal costs and benefits provide valuable ex ante guidance when changes to the environmental quality/quantity or management regimes at recreational sites are considered, our investigation into the capability of the two methods to measure marginal welfare effects is highly policy relevant. Overall, the findings reveal that TCM–CB and CV perform equally well with respect

<table>
<thead>
<tr>
<th>Table 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recreational value of the status quo and marginal recreational costs of the scenarios of water level decrease (€/household/year).</td>
</tr>
<tr>
<td>TCM–CB</td>
</tr>
<tr>
<td>Status quo</td>
</tr>
<tr>
<td>Status quo</td>
</tr>
<tr>
<td>Lower CI 95%</td>
</tr>
<tr>
<td>Point</td>
</tr>
<tr>
<td>Upper CI 95%</td>
</tr>
<tr>
<td>a This corresponds to 79.95€/person/trip.</td>
</tr>
<tr>
<td>b This corresponds to −4.41€/person/trip.</td>
</tr>
<tr>
<td>c Results from the sensitivity analysis in brackets (73 zero WTP bids excluded, n = 115).</td>
</tr>
</tbody>
</table>

---

Please cite this article as: Lienhoop, N., Ansmann, T., Valuing water level changes in reservoirs using two stated preference approaches: An exploration of validity, Ecol. Econ. (2011), doi:10.1016/j.ecolecon.2011.02.014

---
to our theoretical validity tests, and differ considerably in their estimates of marginal recreational costs. We also observe that both methods face similar internal difficulties with respect to the stability of values when the order of a set of valuation questions is changed.

6.1. Theoretical Validity

The assessment of theoretical validity in terms of regression analysis and sensitivity to size suggests that both methods generate highly valid and robust welfare estimates in this study. Although these findings do not prove that the methods generally generate valid estimates, they confirm that meaningful answers were elicited for hypothetical ‘visit rate’ and WTP for a well-defined and familiar public good, and that both survey design and administration seemed to enable respondents to form valid responses. The self-administered survey may have supported the elicitation of meaningful answers to the valuation questions, as it gave the respondents sufficient time to think about their preferences. While self-administered and mail surveys are criticised for the lack of opportunity to clarify issues or ask for more information (Dillman, 1978; Mitchell and Carson, 1989), it seems that these survey modes do not affect validity when the good under investigation is familiar and comprises use values. We can therefore not make any statement about the theoretical validity of the two methods in general, as we do not know how respondents would perform in a personal interview context, where they have less time to think. An interesting finding is that sign reversals were observed for two out of ten variables and this may indicate differences in the utility function between the two methods. While these rather marginal differences should not jeopardise the comparison of the two methods, our finding shows that caution is required when different sources of preference data are combined (e.g. joint TCM–CV models).

6.2. Convergent Validity

The most surprising finding of our study is the result of our convergent validity test, showing that the marginal recreational costs are considerably higher in CB than in CV. Since the confidence intervals do not match, convergent validity cannot be confirmed. Although tests of convergent validity are ambiguous as it is impossible to say, which of the two valuation methods is more accurate (Mitchell and Carson, 1989), these differences should be taken seriously, as they may certainly influence the outcome of reservoir management decisions. In our view, explanations for the difference can be found in the survey design, but also in the extent to which the two methods reflect lexicoographic preferences: First, the fact that marginal recreational costs estimated with CV are relatively small may be related to the respondents’ perceived property rights and the choice of welfare measure. In Germany, property rights for reservoir management are held by the respective federal states. However during the survey it turned out that the respondents had contrary perceptions regarding the allocation of property rights: most visitors strongly felt that they were entitled to high water levels in the reservoir. From a theoretical point of view it was therefore correct to estimate equivalent surplus (WTP to avoid), as this welfare measure assumes that individuals do not hold the property rights for the status quo. However, since visitors perceived to have the right to the status quo, it would have been more realistic to ask them for their willingness to accept (WTA) in compensation to put up with a decrease in water levels. According to Bromley (1995) and Halstead et al. (2002) the use of WTP instead of WTA in such a situation can severely underestimate the value, because losses are valued more than gains. Our finding suggests that the choice of welfare measure is not trivial and that the recommendation to use ‘WTP to avoid’ instead of WTA when losses are valued (e.g. NOAA, 1993) can lead to distorted welfare estimates. Our sensitivity analysis (excluding 73 potential protest zeros) resulted in higher marginal recreational costs, but still lie far below the costs estimated with CB. Hence there may be further causes for the failure of convergent validity. Second, numerous respondents expressed that they held lexicoographic preferences for high water levels when they were first approached by the interviewers on site. Lexicoographic motives tend to be present among individuals who have ethical concerns about nature or feel they have the right to the status quo and therefore refuse to trade off changes in the provision of environmental goods or services with changes in income (Harris et al., 1989; Spash and Hanley, 1995). We suppose that the high estimates of recreational costs estimated with TCM–CB are results of lexicoographic behaviour which is reflected by a considerable reduction of trips under the scenarios of water level change. In the CV context, it seems as if the hypothetical market alleviated lexicoographic preferences, as respondents were given the opportunity to avoid decreases in water level by supporting alternative flood protection measures. In contrast, CB did not offer respondents an alternative to water level changes and lexicoographic preferences were directly expressed in terms of hypothetical trips, leading to a serious drop in ‘visit rate’ and high recreational costs. Although less than 6% of the respondents stated that they have other lakes in mind, it might still be that the availability of such substitutes may have caused the reductions in ‘visit rate’. So far, there is hardly any research looking at the motives behind hypothetical ‘visit rates’. Given the unexpectedly high estimates generated with TCM–CB, further research is required to better understand how respondents decide on hypothetical trips and how these are affected by the presence of substitutes.

6.3. Order Effects

The analysis of order effects shows that both valuation methods are equally prone to the order in which water level scenarios are presented in the questionnaire. This finding is interesting since it shows that order effects are not an exclusive anomaly in CV, but also occur when preferences are revealed in terms of travel behaviour. According to psychologists order effects in CV are triggered due to the absence of pre-defined preferences, because preferences are constructed depending on the context in which they are presented and because they have no experience in valuing, even familiar, goods in monetary terms (e.g. Payne et al., 1999). The fact that order effects also occur in CB underlines that preferences may not be fully established in the respondents’ minds and that order effects are not caused by the demanding translation of preferences into monetary units. While this finding relaxes the criticism of CV that monetary valuation is cognitively too demanding, it shows that preferences elicitation for public goods is not a straightforward task, regardless of the valuation method used. Other possible causes, such as learning effects and fatigue may also play a role and need further investigation.

Similar to Clark and Fiesen (2008) the order does not affect the degree of sensitivity to size in our CV exercise, since the marginal costs for water level reductions are similar regardless of order. Thus, the initial valuation seems to provide an anchor to subsequent valuations (see also Kahneman et al., 1982). An obvious avenue for future research would therefore be to investigate, under which circumstances order effects occur, e.g. the survey method and the type of good (familiar/unfamiliar). Furthermore, it would be worthwhile to further explore the reasons for anchoring with follow-up discussions or think-aloud techniques and to use split sample experiments investigating the (in)dependence of several scenarios presented to the respondents. Certainly, an initial recommendation to abate the influence on order on mean welfare estimates when more than one scenario is valued would be to design surveys in a way that all possible orders are presented in split samples.

6.4. Policy Implications

With respect to the policy implications for flood control, our findings indicate that a decrease in reservoir water levels reduces the
recreational value of reservoirs in the Ore Mountains. While a water level decrease of 1 m does not have a serious effect on the recreational value, more severe water level decreases (e.g., −3.5 m and −5 m) would create considerable recreational costs. Thus, we can conclude that the recreational costs of management rules considering water level decreases exceeding 1 m are expected to be severe. These findings also facilitate the weighting of the respective interests of reservoir management for flood control and recreational users. We refrain from aggregating mean recreational costs over the total number of annual visitors or residents to guide cost-benefit decisions for two reasons: first, the overrepresentation of frequent visitors in the on-site sample (endogenous stratification) was not corrected and non-users were not taken into account, as approaches to correct for these biases are merely available for TCM–CB data without zero trips. An interesting and relevant future research task would be to identify ways to counter truncation and endogenous stratification when data includes zero bids, as this is always the case when trip data is applied for the valuation of environmental deterioration. Second, given the differences in mean estimates between the two valuation methods – and bearing in mind that both methods achieved a high degree of theoretical validity – we are not able to judge which estimates are more adequate.

Acknowledgement

We thank Melanie Heyde and Daniela Siedschlag for valuable support during the data collection phase. We are also very grateful to Erik Gavel and three anonymous referees who provided helpful comments on earlier drafts of this manuscript. We acknowledge funding from the German Ministry of Education and Research for the GLOWA-Elbe Project.

References


