

Document downloaded from:

<http://hdl.handle.net/10251/50785>

This paper must be cited as:

Glenk, K.; Martin-Ortega, J.; Pulido-Velazquez, M.; Potts, J. (2014). Inferring Attribute Non-attendance from Discrete Choice Experiments: Implications for Benefit Transfer. *Environmental and Resource Economics*. 60(4):497-520. doi:10.1007/s10640-014-9777-9.



The final publication is available at

<http://dx.doi.org/10.1007/s10640-014-9777-9>

Copyright Springer Verlag (Germany)

Inferring attribute non-attendance from discrete choice experiments: implications for benefit transfer

Klaus Glenk · Julia Martin-Ortega · Manuel Pulido-Velazquez · Jacqueline Potts

Abstract

Typical convergent validity tests of benefit transfer based on stated preference data assume that willingness to pay (WTP) estimates have been accurately measured, and that differences in WTP arise from differences in observable and unobservable characteristics between the study and the policy sites. In this paper, we conduct a convergent validity test assuming equality of underlying preferences, but allow for the possibility that transfer errors arise from differences in the way that respondents process information in the preference elicitation tasks. Using data from an identical survey instrument applied to the population of two river basins in Spain, we obtain marginal and total WTP estimates for ecological improvements of water bodies and the corresponding transfer errors across sites. Results of equality constraint latent class (ECLC) models that infer attribute non-attendance (AN-A) are compared to results from mixed logit (MXL) models in WTP space. We found large absolute and relative differences in marginal and total WTP between sites for the MXL models, and significantly reduced transfer errors for the ECLC model. This paper therefore provides further evidence that AN-A can significantly affect environmental values derived from attribute-based stated preference methods and is the first to investigate the implications for benefit transfer.

Keywords: attribute non-attendance, benefit function transfer; convergent validity; discrete choice modelling;; Water Framework Directive

1. Introduction

The discrete choice experiment (DCE) literature on attribute non-attendance (hereafter AN-A) describes a phenomenon where each respondent may focus only on a particular subset of attributes while making his or her choices. Such behaviour is in conflict with the assumption of fully compensatory preferences typically made when using modelling stated choice data. Fully compensatory behaviour implies that respondents with complete knowledge of all the information presented in the choice task make trade-offs between *all* attributes and their levels when choosing the preferred alternative. However, a growing body of literature has demonstrated that it is quite common for respondents to ignore one or several attributes when making their choices (Hensher et al. 2005, 2012; Scarpa et al. 2009; Campbell et al. 2011).

AN-A behaviour can have a variety of underlying reasons including, for example, irrelevance of an attribute to a particular choice situation or coping strategies to reduce the complexity of the choice task and hence the cognitive cost associated with choosing (Alemu et al. 2013). Irrespective of the particular source of AN-A behaviour, the violation of the continuity axiom associated with AN-A rules out the possibility of estimating marginal rates of substitution between certain attributes, including implicit prices, at the level of the individual respondent (Scarpa et al. 2009). At the sample level, not accounting for AN-A can result in biased estimates of preference parameters and willingness to pay (WTP), with potentially serious implications on the further use of WTP estimates in policy analysis. While a substantial body of literature exists on the incidence of AN-A in DCEs and the resulting consequences for the estimation of marginal WTP (MWTP) and welfare measures (Hensher et al. 2005; Puckett and Hensher 2008; Campbell et al. 2008; Scarpa et al. 2009; Hensher and Greene 2010; Scarpa et al. 2010; Kragt 2013; Scarpa et al. 2013), the implications of AN-A have not yet been investigated in the context of DCE applications to benefit transfer.

The benefit transfer method, which estimates the value of a change of an environmental good at a target 'policy' site using information from a study conducted at another 'study' site, appeals to policy makers because of its great potential to reduce costs and time needed for the estimation of non-market values (Navrud and Ready, 2007). Policy makers' demand for 'quick and cheap' approaches to

deriving non-market values for policy analysis has been mirrored by an increasing body of literature on the benefit transfer method in general (Wilson and Hoehn 2006; Bateman et al. 2011; Kaul et al. 2013) and on applications using DCEs in particular (Morrison and Bennett 2004; Hanley et al. 2006; Colombo et al. 2007; Johnston and Duke 2009, 2010; Baskaran et al. 2010; Martin-Ortega et al. 2012). Despite significant research progress in recent years, it is still being debated how to increase the accuracy and reliability of benefit transfer, and which guiding criteria should be applied to generate an acceptable level of transfer errors (Bateman et al. 2011; Martin-Ortega et al. 2012; Kaul et al. 2013).

Most of the research efforts to improve the reliability of the benefit transfer method thus far aim to mitigate the *generalisation error* which arises from differences between the study and the policy sites that are expected to relate to the differences in the parameter values of the utility functions (Rosenberger and Stanley 2006). In contrast, *measurement error*¹ is the error endogenous to the research process that can potentially affect the accurate measurement of MWTP and welfare estimates for use in benefit transfer. Empirical assessments of theoretical models require analyst judgment on preference elicitation methods and data analysis (Hanemann 2000), and measurement error occurs if decisions on the approach to empirical analysis made by analysts affect the transferability of values. Some of this error is inevitably ‘passed on’ from original value estimation and may even be amplified by benefit transfer if care is not taken to minimize such effects (Wilson and Hoehn 2006). Measurement error has received comparatively less attention in the benefit transfer literature². So far, measurement error in benefit transfer related to methodological factors has largely been dealt with in meta-analysis-based benefit transfer (e.g., Johnston et al. 2006; Moeltner et al. 2007), but less so in

¹ For completeness, we note that Rosenberg and Stanley (2006) identify a third category of transfer error, the *publication bias* arising from economic journals giving more attention to methodological innovations than to reporting ‘conventional’ valuation studies.

² Both generalisation and measurement error can be confounded. Certain types of measurement error related to, for example, design principles for the development of valuation scenarios are implicitly addressed when aiming at reducing the generalisation error (Bateman et al. 2011).

value and function transfers. This paper is specifically concerned with measurement error arising from the choice of the econometric specification used for analysis, which may be denoted as *econometric specification error*. An example investigating the implications of econometric specification error is Colombo et al. (2007), who find that model choice matters in a function transfer context and conclude that the inclusion of preference heterogeneity reduces the magnitude of transfer errors.

In much of the benefit transfer research based on stated preference data, tests of convergent validity, (testing for WTP differences between study site estimates and ‘actual’ values at a policy site and observing the magnitude of ‘transfer errors’) have certainly been the main methodological tool to assess the validity of the transfer (Boyle et al. 2010; Kaul et al. 2013). These tests rest on the assumption that both study site and policy site WTP estimates have been accurately measured; or that at least any measurement errors, including those related to the econometric specification, are systematically affecting both study and policy site data (Boyle et al. 2010). Hence, any WTP differences between sites are assumed to arise from observable and unobservable differences in site characteristics. In this paper, we conduct a convergent validity test assuming equality of underlying preferences, but allow for the possibility that transfer errors arise from differences in the way respondents processed information in the preference elicitation tasks. In particular, we conduct a function transfer testing for differences in MWTP and compensating surplus (CS) between study site and policy site after adjusting for differences in observable socio-economic characteristics and *additionally* account for AN-A behaviour that may be present and differ between sites

Differences in AN-A behaviour can be a result of the research process (for example, due to differences in survey administration and enumerators), reflecting hitherto unobserved differences between study and policy site (for example, it may reflect cultural differences between two sites that impact on response behaviour in surveys), or both. Whatever the source of differences in AN-A behaviour, if the study site respondents exhibit selective attribute attendance, estimated WTP values of the study site might be biased and transferred to the policy site. If AN-A behaviour differs between study and policy sites, there is a risk that conclusions drawn from tests of convergent validity are flawed. Scarpa et al. (2009) suggest that researchers involved in previous DCE studies might want to

explore whether the conclusions drawn from previous data analysis are robust when AN-A is accounted for. This question may as well be extended to the analysis of DCE data for use in benefit transfer.

Therefore, the aims of this paper are to i) investigate whether transfer errors in benefit transfer of non-market values can (at least partially) be attributed to AN-A behaviour; and ii) to analyse the potential of accounting for AN-A behaviour in choice models to reduce transfer error. To the best of our knowledge, these questions have not yet been explored in the literature. In order to do so, we use data from an identical choice experiment questionnaire that had been previously applied to two sites. We estimate and compare MWTP indicators, CS and transfer errors using results of mixed logit (MXL) models in WTP space and equality constrained latent class (ECLC) models that explicitly take AN-A into account.

The DCE data used in this paper was collected in the policy context of the Water Framework Directive (WFD), which prescribes the improvement of water bodies across Europe to ‘good ecological status’. The non-market benefits of ecological improvements in water bodies can be used for an assessment of economic efficiency and disproportionality in the implementation of the WFD (Brouwer 2008; Vinten et al. 2012). The use of benefit transfer is of particular interest in this context, due to the need of estimating environmental benefits of water status improvements across Europe on different spatial scales in a limited amount of time, for which it is unrealistic to expect the use of primary valuation studies only.

2. Method

To test the effects of considering AN-A behaviour when estimating benefit transfer errors, we use data collected from DCEs in two Spanish river basins: the Guadalquivir River Basin (GRB) in the South of Spain, and the Serpis River Basin (SRB) in the East of the country; where main problems in complying with the WFD relate to water scarcity. Valuation scenarios were characterized by different

levels of the river's water flow causing varying degrees of deviation from natural ecological conditions with respect to species diversity and habitat conditions (ENV); different frequencies of household water restrictions (RES); and a cost operationalised as an increase in the annual water bill (COST).

The choice models that do not incorporate AN-A behaviour are estimated in WTP space following Train and Weeks (2005) and Scarpa et al. (2008). The modelling approach is based on random utility theory, with a utility function U for respondent n and alternative i in choice task t characterised by price p (COST) and non-price attributes \mathbf{x} (ENV, RES) of the experimental design, and a random error term ε :

$$U_{nit} = -\alpha_n p_{nit} + \boldsymbol{\beta}_n' \mathbf{x}_{nit} + \varepsilon_{nit} \quad (1)$$

where α and $\boldsymbol{\beta}$ are parameters to be estimated and ε is assumed to be identically and independently distributed (*iid*) and related to the choice probability with a Gumbel distribution with respondent specific error variance $\text{Var}(\varepsilon_{ni}) = \mu_n^2(\pi^2/6)$, with μ_n being a respondent specific scale factor.

Train and Weeks (2005) show that equation (1) can be divided by μ_n to derive a scale-free and behaviourally equivalent utility function with a new error term that is constant across individuals:

$$U_{nit} = -(\alpha_n/\mu_n)p_{nit} + (\boldsymbol{\beta}_n/\mu_n)' \mathbf{x}_{nit} + \varepsilon_{nit} \quad (2)$$

where ε_{nit} is *iid* with constant error variance $\pi^2/6$. Substituting $\gamma_n = \alpha_n/\mu_n$ and $\mathbf{c}_n = \boldsymbol{\beta}_n/\mu_n$ in equation (2) as the parameters to be estimated provides what Train and Weeks (2005) refer to as the model in preference space. Exploiting the fact that MWTP is $\mathbf{w}_n = \mathbf{c}_n/\gamma_n$, the utility function in WTP space can be written as:

$$U_{nit} = -\gamma_n p_{nit} + (\gamma_n \mathbf{w}_n)' \mathbf{x}_{nit} + \varepsilon_{nit}. \quad (3)$$

Denote the sequence of choices over T_n choice tasks for respondent n as \mathbf{y}_n , i.e. $\mathbf{y}_n = \langle i_{n1}, i_{n2}, \dots, i_{nT_n} \rangle$.

In a mixed logit (MXL) model, heterogeneity across respondents is introduced by allowing γ_n and \mathbf{w}_n to deviate from the population means following a random distribution. Conveniently, as opposed to models in preference space, the distribution of MWTP can be estimated directly. In a MXL model, the

unconditional or mixed logit choice probability of respondent n 's sequence of choices is the integral of the logit formula over all possible values of γ_n and \mathbf{w}_n :

$$\Pr(\mathbf{y}_n | \gamma_n, \mathbf{w}_n) = \int \prod_{t=1}^{T_n} \frac{\exp(-\gamma_n p_{nit} + (\gamma_n \mathbf{w}_n)' \mathbf{x}_{nit})}{\sum_{j=1}^J \exp(-\gamma_n p_{njt} + (\gamma_n \mathbf{w}_n)' \mathbf{x}_{njt})} f(\boldsymbol{\eta}_i | \boldsymbol{\Omega}) d\boldsymbol{\eta}_i \quad (4)$$

where $f(\boldsymbol{\eta}_i | \boldsymbol{\Omega})$ is the joint density of the parameter vector for price and K non-price attributes [$\gamma_n, w_{n1}, w_{n2}, \dots, w_{nK}$], $\boldsymbol{\eta}_i$ is the vector comprised of the random parameters and $\boldsymbol{\Omega}$ denotes the parameters of these distributions (e.g. the mean and variance). This integral does not have a closed form and thus requires approximation through simulation (Train, 2003), in our case using 2,000 Halton draws. The price attribute parameter is assumed to follow a lognormal distribution, the MWTP parameters of the non-price attributes are assumed to follow a normal distribution, and the MWTP parameter for an alternative specific constant (ASC) associated with the status quo is fixed (i.e. non-random). Socio-economic characteristics and perceptual variables related to respondents' attitudes and beliefs regarding water scarcity and its impact are included as interactions with non-price attributes and the ASC. Conveniently, these interactions directly capture effects of individual-specific characteristics on MWTP. The resulting parameter estimates are used as a transfer function that adjusts MWTP to observed characteristics at the policy site. The MXL model has been used in several recent benefit transfer convergent validity tests based on DCEs (Hanley et al. 2006; Colombo et al. 2007; Johnston and Duke 2009, 2010; Baskaran et al. 2010). To our knowledge, this is the first application of MXL models in WTP space in a benefit function transfer context.

To incorporate AN-A behaviour into the estimation of preferences from stated choice data, we use the equality constrained latent class model (ECLC) proposed by Scarpa et al. (2009) and applied in several other recent studies (e.g., Campbell et al. 2011; Hensher and Greene 2010; Hensher et al. 2012). Instead of using latent classes to represent a discrete mixture of people's preferences, the ECLC models heterogeneity in attribute processing strategies, in our case AN-A behaviour. Each latent class in the ECLC model represents a different pattern of attribute attendance. The probability of respondent n choosing alternative i , conditional on belonging to class c out of K classes reflecting a certain AN-A pattern, can be written as:

$$P_n(i|c) = \frac{\exp(\beta'_c x_i)}{\sum_{j=1}^J \exp(\beta'_c x_j)} ; c \in \{1 \dots K\}. \quad (5)$$

The unconditional probability is then given by:

$$P_n(i) = \sum_{c=1}^K \frac{\exp \theta_c}{\sum_{k=1}^K \exp \theta_k} \frac{\exp(\beta'_c x_i)}{\sum_{j=1}^J \exp(\beta'_c x_j)} ; \sum_{k=1}^K \theta_k = 0. \quad (6)$$

where θ is a set of K class-specific constants identified by ensuring that they sum to zero. The model allows for fully compensatory preferences where all attributes are considered; any combination of ignoring one or more attributes; and non-attendance to all attributes in the choice set, reflecting random choice between alternatives (Scarpa et al. 2009). These AN-A patterns are accommodated by imposing restrictions on the utility coefficients in each latent class. A zero utility weight is assigned to attribute coefficients that are not considered, while all coefficients to be estimated are constrained to be equal across classes. That is, preferences are assumed to be homogeneous across the sample, as in CL models. There is a maximum of 2^k permutations of AN-A, and hence, latent classes in the model, where k is the number of attributes that are either considered or ignored by a respondent. A key outcome of the ECLC model are the class probabilities for the 2^k AN-A patterns, which can be used to infer the probability that an attribute has been ignored.

While it is implicit that the parameters are constrained to be equal across classes in the ECLC, it is less obvious if this should also apply to ASCs. Previous literature applying the ECLC is mostly silent about the treatment of ASCs, although it can impact on results. An exception is Hensher et al. (2012), which estimates ASCs for the one class where all attributes are attended to, but omit constants (constraining ASC coefficients to zero) for all other classes where one or more attribute coefficients are assigned zero values. In environmental DCM applications, an ASC is usually added to capture differences in utility between the status quo alternative (or an ‘opt-out’ alternative) and the other choice options. It may be possible to assign a behavioural interpretation to the status quo option, for example related to protest attitudes (Meyerhoff and Liebe 2009), if the attribute levels used to describe the status quo also appear in the designed part of the choice experiment. If this is not the case, however, it is not possible to estimate the difference in utility between the baseline level of the

attributes (used in the status quo option) and the reference level (the lowest level appearing in the designed part of the choice experiment). In this situation, the ASC parameter captures the differences in utility between the baseline level and the reference levels for all the attributes, together with any utility from choosing the status quo that is unrelated to attribute information. In an ECLC context, this means that it is counterintuitive to constrain the ASC to be equal across classes, as non-attendance to one or more of the attributes will almost certainly influence the ASC parameter. Constraining the ASC to be equal hence imposes unnecessary restrictions and may bias inferred AN-A patterns, preference estimates and MWTP estimates. Therefore, we do not impose ASC equality across classes in our case.

Moreover, there may be a class of respondents who regard the options in the designed part of a choice experiment as broadly similar, since all involve some environmental improvement along with an increase in cost, whereas the status quo option involves no environmental improvement and no extra cost. Such respondents may make a choice between the status quo and the other options, but choose one of the other two options at random, giving rise to a non-zero ASC but zero coefficients for the other attributes. Therefore, we additionally include a class where all attribute coefficients *and* the ASC are constrained to be zero, to reflect completely random choice among alternatives. In sum, we consequently estimate 2^{k+1} classes, 2^k classes for all permutations of AN-A patterns without constraining the ASC to be equal across classes, and an additional one that reflects random choice among all the alternatives, totalling to 9 classes in the case study application reported in this paper.

To enable a transfer function, socio-economic characteristics and perceptual variables are included as interactions with the ASC and the attributes. Note that the interaction terms only enter the utility function in classes where the associated attributes are attended to, and that the parameters of the interaction terms are constrained to be equal across those classes. Interactions with the ASC follow the same pattern.

To analyse the impacts of considering AN-A behaviour in benefit transfer, MWTP estimates based on MXL model results using data from GRB and SRB serve as reference models for comparison to

ECLC estimates. MWTP estimates based on MXL models are calculated as unconditional mean parameter estimates.

MWTP calculation from ECLC models requires assumptions to be made about respondents with a high probability of non-attendance to cost. We calculate the weighted average WTP over those classes where cost is attended to, assuming that WTP is zero for those classes where cost is attended to but the attribute is not attended to, with the probability of class membership conditional on cost being attended to as weights. This is a compromise between two approaches that have recently been applied in the literature. Campbell et al. (2011) calculate the weighted average MWTP over those classes where both the attribute and cost are attended to, where the weights are the probability of class membership conditional on both the attribute and cost being attended to. Kragt (2013) suggests that an alternative approach is to assume that MWTP is zero for those respondents who do not make full trade-offs³.

Calculating WTP from ECLC models generally implies accepting that a proportion of respondents is not sensitive to changes in the cost attribute (and hence income), which it is at odds with economic theory, and is ‘probably not a realistic representation of [peoples’] preferences with respect to money’ (Scarpa et al. 2009, 166). Hence, the inferred ‘non-attendance’ to changes in cost may be due to the fact that people care for money, but behaved in the choice experiment as if their marginal utility was (close) to zero. As investigated by Puckett and Hensher (2008), and emphasised by Hensher et al. (2012), it may also be due to the fact that some respondents may not perceive the between-alternative difference in cost large enough to influence their choice – at least for a subset of the choice tasks. This would suggest that the inability to measure marginal utility of income in every case is related to the experimental design, which induces the – from an analyst’s perspective – undesirable behaviour of not trading cost when choosing among alternatives. When observing impacts of AN-A on the value of travel time savings (VTTS) using an ECLC model, Hensher et al. (2012, 243) are ‘inclined to support the mean VTTS [derived from classes] where we can observe a time-cost trade-off’, arguing that

³ We calculated MWTP based on the approaches applied in Campbell et al. (2011) and Kragt (2013). Results differed slightly in magnitude but not qualitatively to the MWTP estimates reported in this paper.

applying this value to the whole sample is appealing from the perspective of behavioural interpretation.

Alternatively, inferred ‘non-attendance’ to changes in cost may reflect an inability to capture a non-zero marginal utility of income using the ECLC. This could be due to not accounting for preference heterogeneity with respect to cost, resulting in the ECLC erroneously assigning probabilities for zero marginal utility of income when in fact it is merely very low (Campbell et al. 2012; Hess et al. 2013). Some of those respondents who are identified as having a high probability of not attending to cost may actually have a very high MWTP for all attributes. Because it is not possible within the usual model framework to assess whether this is actually the case, it seems preferable to assume that their MWTP is similar to that of those respondents who do attend to cost rather than that it is zero.

Compensating surplus (CS) is calculated for the six improvement options with respect to water use restrictions (RES) and river water flows (ENV) following Hanemann (1984) as:

$$CS = -\frac{1}{\alpha} [\ln \sum \exp V^1 - \ln \sum \exp V^0] \quad (7)$$

where CS is the compensating surplus welfare measure, α is the marginal utility of income (represented by the coefficient of the monetary attribute β_s in the choice experiment) and V^0 and V^1 represent the value of the indirect utility functions before and after the change under consideration.

Using the MWTP and CS estimates for the GRB and SRB, we calculate relative transfer errors (%) for transfers in both directions (i.e. GRB to SRB and *vice versa*) using the equation:

$$Relative\ Transfer\ Error = \left| \frac{Predicted\ MWTP/CS_{policy\ site} - Estimated\ MWTP/CS_{policy\ site}}{Estimated\ MWTP/CS_{policy\ site}} \right| \times 100\% \quad (8)$$

We adjust MWTP and CS estimates for the policy site by inserting observed values of socio-economic characteristics and perceptual variables into the utility function derived for the study site to derive estimates of $Predicted\ MWTP/CS_{policy\ site}$. Absolute transfer errors (i.e., $Estimated\ MWTP/CS_{study\ site} - Estimated\ MWTP/CS_{policy\ site}$) and relative transfer errors are then compared for the MXL models in relation to the ECLC model to assess the impact of taking AN-A into account.

3. Data

We employ data of DCEs collected from the population of the two Spanish river basins (GRB and SRB). Respondents were asked to select the most preferred alternative in a set of different scenarios of water allocation to the environment and frequencies of household water restrictions. Table 1 presents the attributes and levels of the choice experiment - see Martin-Ortega et al. (2011) for details on the valuation scenario.

TABLE 1 AROUND HERE

The design comprises 24 choice cards organized in 6 blocks representing a fractional factorial design. Each respondent hence faced four choice occasions, in which they were asked to choose between the current situation with no rise in the water bill, and option A and B with different levels of improvement of the two attributes in exchange for an annual increase in the water bill. **¡Error! No se encuentra el origen de la referencia.** shows an example of a typical choice card.

FIGURE 1 AROUND HERE

The environmental attribute was coded as a set of dummy variables omitting the moderate level of environmental quality as the reference category in the models⁴, while the frequency of water restrictions was entered as a continuous variable taking values from 1 to 4 years with restrictions during summer within the next ten years. Apart from the choice experiment, the survey collected

⁴ The design does not allow estimation of the utility impact of a move from ‘poor’ to ‘moderate’ for environmental status related to river flows (ENV). Because the ‘poor’ level used to describe the status quo does not also appear in the remaining (designed) alternatives, it is not possible to estimate the difference in utility between ‘poor’ and ‘moderate’, which serves as the reference level (see section 2 for further details). This was done to serve the original purpose of the study, but it is unfortunate for the present paper. Strictly speaking, therefore, when we refer to non-attendance to ENV, this involves non-attendance to any improvement in ENV over ‘moderate’, since any choice of a non-status quo alternative implies a move from ‘poor’ to ‘moderate’.

information on basic socio-economic characteristics and respondents' perceptions on and attitudes towards aspects related to water scarcity and the environment.

The same questionnaire, with the same experimental design, was administered between June and September 2008 to two samples of the population of the GRB (354 respondents) and SRB (394 respondents) derived from a quota-based sampling procedure with quotas based on age and sex, and an additional fixed sample size for respondents residing in either urban or rural areas. The surveys were designed by a joint team of researchers and applied simultaneously following the same survey protocol. The only difference in the survey process between the two sites concerned the enumerators. They were trained by the same team of researchers, but were residents of the respective regions where the survey was carried out.

Preceding the survey, existing case study status reports were examined, and river basin experts were consulted in a series of scientific workshops in the original study. In addition, a series of focus groups with a total of about 30 people in each case study were organised to discuss the attributes. The main survey was preceded by three rounds of pre-testing and a small pilot study, involving approximately one hundred individuals. The pre-tests focused specifically on i) the refinement of attributes and levels; ii) the clarity and credibility of the policy scenarios and baseline conditions; and iii) the wording and understanding of survey questions.

While being considerably different in size (the GRB being larger than the SRB), these two basins share common features relevant for the valuation scenario used, because they face similar challenges to the implementation of the WFD. In both river basins, the target of achieving good ecological status is compromised by structural water scarcity and competing demands for water use. Both have a Mediterranean climate with heterogeneous precipitation distribution. Agricultural abstraction is the most important consumptive water use in both cases (between 70 and 80%)⁵. Both regions have suffered from several periods of severe drought in the past 40 years, the latest occurring in 2004/5 in

⁵ The second and third largest consumptive water uses are in both cases domestic water supply and industrial abstraction and use. The relative magnitude of domestic water consumption is higher in the SRB (25%) compared to the GRB (11.5%).

the GRB (CHG, 2007) and in 2007 in the SRB (CHJ, 2007), resulting in restrictions in the agricultural and industrial use of water in both cases. Households in the GRB have also experienced water use restrictions due to limited water availability (CHG, 2007).

In both basins, climate change is expected to result in lower water yield and more frequent drought events as a consequence of both a rise in temperatures and a reduction in rainfall. For 2030, simulations considering a temperature increase of 1°C and a reduction of 5% in mean rainfall project a decrease of mean water yield of almost 12% in both river basins, which is above the national Spanish average of 8% (Iglesias et al. 2005). Overall, the similarities between both river basins in terms of environmental challenges and ecological site characteristics make this a typical example where a benefit transfer between sites may be considered to be appropriate.

Table 2 reports summary statistics of socio-economic and perceptual variables elicited in the survey. In terms of representativeness, the samples generally compare favourably to population statistics of the two regions. Some notable differences are a higher share of upper secondary education in the samples (statistically significant for SRB), greater household size and slight differences in household income (statistically significant for both regions), while the samples match age and gender distributions of the population very well.

TABLE 2 AROUND HERE

With regards to perceptual variables relevant to the valuation context, all variables are significantly different across the two sites, with a higher percentage of the GRB sample reporting past experience with water use restrictions, and a greater tendency of the GRB sample respondents to perceive water scarcity as problem, to believe that scenarios of future water scarcity are credible, and to consider the environment a priority when allocating water in case of scarcity (relative to agriculture and industry). The average annual water bill as recalled by respondents, on the other hand, is higher in the SRB.

Several socio-economic and perceptual/attitudinal variables are included in the utility function and thus form the basis for a benefit function transfer. *Age, income, gender*, the belief that water *scarcity* is a problem in the region and a subjective assessment of whether it is *credible* that there will be a

future increase in water scarcity are interacted with the ASC. In the WTP space models, *income* is interacted with the environmental attributes (RES, ENV), while income is interacted with COST in the ECLC models. Additionally, the effect of whether a respondent has *suffered* from water scarcity conditions in the past on preferences for reducing the frequency of household water restrictions (RES), and the effect of considering the environment as a *priority* for allocating water in case of scarcity on preferences for ecological status improvements are added to the utility function.

Protest answers were distinguished from legitimate zero bids by means of follow up questions about the reasons why respondents were not willing to pay. Less than 5% of the responses were classified as ‘protesters’ and were excluded from the analysis of the choice data⁶, following standard procedure in stated preferences literature (Dziegielewska and Mendelsohn, 2007).

4. Results

Table 3 reports the model results for the GRB and SRB sites. In all models, all environmental attributes effects carry the expected sign after considering the interaction terms with socio-economic and perceptual variables. That is, a one year increase in the frequency of water use restrictions in the next ten years (RES) is on average associated with a utility loss, while improvements in the ecological status related to river water flow relative to ‘moderate’ (ENV2 for improvement to good, and ENV3 for improvement to very good) are associated with a utility gain. The monetary attribute (COST) is negative, implying that respondents are price sensitive TABLE 3 AROUND HERE

Several of the interaction terms between attributes and socio-economic and perceptual covariates are statistically significant. The interaction term of the cost attribute with *income* is positive for the ECLC models, indicating that wealthier respondents are less sensitive to increases in the water bill, but the coefficient is only significant for the SRB sample. In the WTP space MXL models for both regions, income has a positive effect on MWTP for reducing the frequency of water use restrictions (RES) and

⁶ The omission of protest responses has no effect on the model results and the conclusions drawn in this paper.

for improving the ecological status of the rivers (ENV). *Suffering* from water scarcity conditions in the past impacts positively on MWTP to reduce the frequency of water use restrictions in the GRB sample. This effect is, however, only significantly different from zero in the MXL model. The same effect is negative but not significantly different from zero for the SRB for both MXL and ECLC models. Considering the environment as a *priority* for allocating water impacts positively on MWTP for improving ecological status, and is significant for improvements to both good (ENV2) and very good (ENV3) levels in the MXL model for SRB, and significant for ENV3 and the SRB in the ECLC model. Several of the interaction terms with the ASC are significant, but both sign and level of statistical significance differ markedly between models and regions.

The MXL models for both GRB and SRB exhibit significant random taste heterogeneity. The inclusion of random parameters in the MXL models significantly improved model fit in terms of log-likelihood compared to the CL models, especially for the SRB sample. For both sites, and consistently with previous findings in the literature (see Hensher et al. 2012 and articles cited therein), model fit improves significantly when considering AN-A behaviour using the ECLC compared to MXL models⁷. This can be interpreted as an indication of the presence of AN-A in the data (Scarpa et al. 2009). Across all three types of models, MWTP of improving the ecological status related to water flow (ENV) is increasing for increasing levels of improvement (i.e. $MWTP_{ENV3} > MWTP_{ENV2}$), showing respondent's sensitivity to scope as theoretically expected (Carson 1997).

Table 4 shows the non-attendance patterns as inferred by the ECLC models for the GRB and SRB data. Generally, the probability of AN-A is considerable for both data sets, but within the range of AN-A probabilities reported in the literature (e.g., Scarpa et al. 2009). While AN-A to the frequency of water use restrictions (RES) is of similar magnitude (+/- 60%) in the GRB and SRB data, inferred non-attendance to improvements in ecological status (ENV2 and ENV3) is more pronounced in the

⁷ Test statistics of a BenAkiva and Swait (1986) test are ~0, i.e. the probability of erroneously choosing MXL as the true model when the ECLC is the true model is effectively zero for both GRB and SRB. For convenience the reader may also wish to compare values of the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) listed in Table 3.

GRB sample (66%) than in the SRB sample (45%). Non-attendance to increases in the water bill (COST), as inferred by the ECLC, is at 58% more than twice as high for the SRB model compared to the GRB model. Given that MWTP is estimated as the ratio of the marginal utility of non-cost attributes and the marginal utility of cost (or income minus cost, to be precise), we would, *ceteris paribus*, expect *AN-A to cost* to result in an upward bias if not considered by the modelling approach, and *AN-A to environmental attributes* to result in a downward bias. Given the AN-A patterns for GRB and SRB found here, we would therefore expect that, *relative to each other*, MWTP was downward biased in the MXL model of the GRB data, while MWTP as derived from the MXL model is biased upwards for the SRB data.

TABLE 4 AROUND HERE

This can be confirmed by inspecting the MWTP estimates reported in Table 5. MXL-based MWTP estimates are between three to eight times higher for the SRB than for the GRB. CS estimates based on MXL results are about five times larger for the SRB compared to the GRB⁸, as shown in Table 6.

MWTP estimates derived from ECLC models are considerably lower for the SRB. Comparing the ECLC MWTP estimates of the SRB sample with those derived from the GRB data, it becomes apparent that as a result of accounting for AN-A, mean MWTP for all attributes is higher in magnitude for the GRB, reversing the findings based on the MXL models. Mean CS estimates based on ECLC models are, however, still higher in magnitude for the SRB⁹. The impact of adjusting for differences in socio-economic and perceptual covariates is relatively modest for both MXL and ECLC models.

TABLES 5 AND 6 AROUND HERE

⁸ The applied tests (Poe et al. 2005) confirm that differences in MWTP and CS are statistically significant at the 1% level.

⁹ Based on results of Poe et al. (2005) tests, the null hypothesis that mean MWTP distributions of RES, ENV2 and ENV3 are equal for samples can only be rejected for RES, but the null hypothesis of equal CS distributions cannot be rejected for any of the six CS estimates.

Of key interest in a benefit transfer context testing for convergent validity are estimates of relative and absolute transfer errors in both directions, i.e. considering the GRB as the policy site and the SRB as the study site and *vice versa*, shown in Table 7. Relative transfer errors when using MWTP and CS estimates derived from MXL models are between 69% and 87% if values are transferred from GRB to SRB, and >300% (MXL) for the majority of MWTP and CS estimates if the transfer is conducted in the opposite direction (SRB to GRB). Relative transfer errors derived from the ECLC models are considerably lower in all cases for a transfer from SRB to GRB. This also applies to the reverse transfer from GRB to SRB with the exception of RES, which has a slightly larger relative transfer error. Considering absolute differences in MWTP and CS, all estimates based on the ECLC models are *much* lower compared to estimates derived from MXL models.

TABLE 7 AROUND HERE

5. Discussion and conclusions

We have found that, for our case studies, transfer errors based on mixed logit (MXL) choice models are considerable, both in relative and absolute terms. While being large, relative transfer errors of the order of magnitude of 100% or even several hundreds per cent are not unusual and have been reported in the literature. For example, Loomis et al. (1995) find transfer errors of up to 475% in multi-site lake recreation; Van den Berg et al. (2001) report relative transfer errors of 239% for groundwater quality improvements; Morrison and Bennett (2000) estimate relative transfer errors of 191% in a case study on wetland valuation; Colombo et al. (2008) find transfer errors based on compensating surplus (CS) estimates of changes in landscape attributes to range between 25% and 2,234%; Bateman et al. (2011) report transfer errors ranging from 37 to 137% for international water quality transfers. Beyond the magnitude of transfer errors estimated from MXL models, however, we find that it is not straightforward to relate differences in both marginal and total willingness to pay (WTP) estimates to observed differences between the sites. Mean marginal WTP (MWTP) estimates per household are significantly lower for the Guadalquivir River Basin (GRB), while its population expressed a

significantly higher concern for current and future developments related to water scarcity and the environmental impacts of reduced water flows in rivers. Additionally, key socio-economic drivers that have been applied in functional benefit transfer such as income do not markedly differ between both sites – mean income is even slightly but statistically significantly lower in the Serpis River Basin (SRB) sample. Given the assumption of equality of underlying preferences in convergent validity tests, these results would suggest the existence of unobserved factors that have a large impact on MWTP. A reduction of the frequency of water use restrictions by three years out of ten and an improvement of ecological status related to water flow from moderate to very good would be valued at 144% (MXL) of the average annual water bill in the SRB. When faced with such benefit estimates, some analysts and policy makers may therefore not consider them of a realistic magnitude. Scarpa et al. (2009) make a similar observation regarding WTP estimates derived from a conditional logit model valuing landscape attributes in Ireland.

In this paper, we have explored an alternative approach to address this clearly unsatisfactory situation. We test if accounting for heterogeneity in attribute processing – and attribute non-attendance (AN-A) in particular – can help to improve the convergent validity of discrete choice experiments (DCEs) collected at two different sites. We infer AN-A behaviour analytically using an equality-constrained latent class (ECLC) model. Results show that the large transfer errors across the two sites are partly due to the existence of non-attendance patterns that are dissimilar across the two sites. After accounting for AN-A behaviour using an ECLC model, both absolute and relative transfer errors are significantly reduced.

The results of this study suggest that econometric specification errors impact on conclusions drawn from convergent validity tests in benefit transfer. This insight has received little attention compared to the ongoing search for ways to reduce the generalisation error, for example by introducing new sets of explanatory variables for use in transfers of preference functions (e.g., Norton et al. 2012). As demonstrated in this paper, the consequences of ignoring differences in respondent behaviour to stated choice tasks across different sites may result in erroneous assessments of convergent validity. AN-A may not be the only source of error arising from the data collection and analysis process. This casts

doubt on the use of convergent validity tests as the ‘workhorse’ to developing benefit transfer guidelines for stated preference data. At worst, the presence of an econometric specification error implies that such tests provide only random results that, in addition to publication bias towards studies that demonstrate methodological innovations and other concerns about convergent validity tests themselves (Kaul et al. 2013), are of little use for improving our understanding of when benefit transfer actually works. However, we do not believe that the situation is as extreme: as a recent review by Kaul et al. (2013) demonstrates, some generally applicable conclusions can be drawn from about two decades of convergent validity testing. The same authors suggest that there is a need to investigate why some value comparisons across sites have small and others have large errors. In this respect the findings of this paper could motivate a re-assessment of convergent validity tests after taking AN-A behaviour into account in the modelling process, especially in cases where transfer errors have been reported to be large. Ultimately, this could contribute to the development of reliable guidance criteria for benefit transfer.

Before concluding, we critically reflect on the methodological approach that we have employed. The reasons underlying AN-A behaviour are generally difficult to disentangle, and possibly not fully identifiable by the researcher, although first attempts have been made (Alemu et al. 2013). The body of empirical evidence suggests that individuals use a number of different information processing strategies when completing a series of stated choice tasks. These strategies may reflect actual preferences or various decision heuristics (Hensher and Greene 2009) related to choice task complexity and the experimental design (Puckett and Hensher 2008; Hensher et al. 2012), including differential allocation of attention across attributes depending upon both the expected marginal benefits and marginal costs of further information processing (Cameron and De Shazo 2011).

In the case study presented here, only the enumerators differed in the approach to collecting data from the two case study sites. While the study was carefully administered in both river basins, including an extensive training of enumerators, we cannot deny the possibility that interviewer effects affected the differences in AN-A patterns across sites. However, the reasons for these differences may also be related to differences in attention to the survey instrument, or intrinsic social and cultural differences

between the populations of the two sites, which are very difficult, if not impossible, to predict *a priori*. All of these issues apply to most stated preference studies conducted across different regions.

Regarding the modelling procedure, different approaches to account for AN-A in stated choice modelling either based on analytical approaches (e.g., Scarpa et al. 2009; Hensher et al. 2012) or based on respondents stated behaviour (Hensher et al. 2005; Campbell et al. 2008; Scarpa et al. 2010) have been explored in the literature. The ability of the different approaches to reliably assess AN-A may be questioned (Alemu et al. 2013; Scarpa et al. 2013), but the collective body of evidence suggests that AN-A is a real phenomenon (Hensher et al. 2012) with potentially severe implications on key behavioural interpretations of results. The ECLC approach to inferring AN-A behaviour taken in this paper does not allow for preference heterogeneity to be considered, while Campbell et al. (2012) and Hess et al (2013) suggest that it can be important. Generally, it is very difficult to disentangle AN-A and preference heterogeneity, in particular low sensitivity to changes in attributes, analytically in the model structure. However, in the context of benefit transfer based on DCEs, the use of a combined latent class-mixed logit model or the stochastic attribute selection approach (STAS) described in Scarpa et al. (2009) certainly deserves further investigation.

Our study was presented as a benefit function transfer, adjusting for differences in socio-economic characteristics and perceptual variables between sites. We acknowledge that the choice of covariates and their inclusion into the utility function can affect the magnitude of transfer errors to some degree. We tested several specifications and sets of covariates, which in our case did not significantly affect the general results. Beyond applications of benefit function transfer using discrete choice experiments, the impact of the econometric specification error related to considering information processing on WTP values derived via meta-analysis also deserves further investigation.

We included ASCs for the estimation of CS. There are arguments both for and against their inclusion. ASCs may represent protest motives (status quo bias) or yeah-saying behaviour. In a choice experiment setting with two or more policy alternatives and a status quo, naïve yeah-saying would imply choosing randomly from the two policy alternatives while consistently ignoring the status quo alternative. Ignoring the status quo alternative may, however, be a decision that reflects genuine

preferences for environmental quality improvements not captured by the attributes that should be accounted for in welfare estimation. That is, if a respondent decides to choose only from the policy alternatives, *and* is considering at least cost to arrive at her/his final choice, it can be argued that ASCs should be included. In contrast to the ECLC models that incorporate AN-A behaviour, the MXL models do not allow this distinction, and the resulting contribution of the ASCs to the CS estimates is particularly large, especially for the SRB sample. The CS estimates based on the MXL models should therefore be treated with caution. Finally, if ASCs are included, it is implicitly assumed that the two river basins do not differ in unobserved aspects (i.e., in aspects that are not covered by the attributes that are related to water scarcity and ecological status). If this is the case, Morrison et al. (2002) suggest that it may be prudent to only use MWTP estimates for benefit transfer. While the two river basins share similar characteristics regarding the challenges they face, we cannot deny the possibility that unobserved differences between sites may exist, for example related to differences in patterns of water use by the respective populations. Irrespective of the above issues, our main finding that econometric specification error related to AN-A behaviour can matter in function benefit transfer based on discrete choice experiments is not affected by whether or not we include ASCs for welfare estimation.

The policy implications of not accounting for the econometric specification error due to AN-A in benefit transfer based on DCEs clearly depend on the valuation context and the magnitude of the bias in WTP estimates, and are therefore case study specific. Generally, if biased WTP estimates enter a cost-benefit analysis, this can have important consequences on the economic efficiency of the related policy decision. The study presented in this paper is relevant to the policy context of the WFD. A WFD-specific element is the possibility of justifying derogation of investments aimed at improving the ecological status of water bodies based on economic grounds. In particular, the WFD provides Member States with the option to derogate investments based on an assessment of disproportionality of costs. Such an assessment requires the comparison of marginal costs and benefits of water status improvements across all river basins within a relatively short period of time (Martin-Ortega 2012). Benefit transfer using information from a limited number of existing primary valuation studies is

expected to play a key role in this analysis. Large transfer errors on the assessment of benefits could lead to decisions reflecting inefficient resource allocation (i.e. justifying derogations when the benefits actually exceed costs, or supporting interventions that would actually be disproportionately costly). In our study, a transfer based on the MXL model results in statistically significantly different WTP estimates at the policy sites. Once AN-A is taken into account, differences in both MWTP and CS between the two river basins are considerably reduced. Therefore, econometric specification error related to differences in AN-A behaviour between the two river basins may significantly affect outcomes of analyses of disproportionality of public investments in water quality improvements.

Acknowledgements

We like to thank two anonymous reviewers and the editor, Christian Vossler, for their helpful comments and suggestions on earlier versions of the paper. The research work was supported by the Scottish Government's Rural and Environment Science and Analytical Services Division (RESAS). The data used in this article were collected as part of the Collaboration Agreement between the University of Córdoba (Spain) and the Spanish Ministry of the Environment for the *Development of Water Demand Analysis and Assessment of Environmental and Resource Benefits of the Water Framework Directive*, and the AquaMoney project of the EU VI Framework Programme (SSPI-022723, *Development and Testing of Guidelines for the Assessment of Environmental and Resource Costs and Benefits of the Water Framework Directive*, www.aquamoney.org).

References

Alemu MH, Mørkbak MR, Olsen SB, Jensen CL (2013) Attending to the reasons for attribute non-attendance in Choice Experiments. *Environ Resource Econ* 54(3):333–359

Baskaran R, Cullen R, Colombo S (2010) Testing Different Types of Benefit Transfer in Valuation of Ecosystem Services: New Zealand Winegrowing Case Studies. *Ecol Econ* 69(5):1010–1022

Bateman IJ, Brouwer R, Ferrini S et al (2011) Making benefit transfers work: deriving and testing principles for value transfers for similar and dissimilar sites using a case study of the non-market benefits of water quality improvements across Europe. *Environ Resource Econ* 50(3):365–387

Ben-Akiva ME, Swait JD (1986) The Akaike Likelihood Ratio Index. *Transport Sci* 20:133–136

Boyle KJ, Kuminoff NV, Parmeter CF, Pope JC (2010) The Benefit-Transfer Challenges. *Annu Rev Resource Econ* 2(1):161–182

Brouwer R (2008) The role of stated preference methods in the Water Framework Directive to assess disproportionate costs. *J Environ Plann Manag* 51:597–614

Brouwer R, Martin-Ortega, J, Dekker T, Raggi M, Sardonini L, Viaggi D, Andreu J, Pulido-Velázquez M, Kontogianni A, Machleras A, Skourtos M, Rolfe J, Windle J (2012) Improving value transfer through socio-economic adjustments in a multi-country choice experiment of water conservation alternatives. Paper presented at the 19th Annual Conference of the European Association of Environmental and Resource Economists, Charles University, Prague, 27-30 June

Cameron and De Shazo 2011

Campbell D, Hutchinson WG, Scarpa R (2008) Incorporating Discontinuous Preferences into the Analysis of Discrete Choice Experiments. *Environ Resource Econ* 41:401–417

Campbell D, Hensher DA, Scarpa R (2011) Non-attendance to attributes in environmental choice analysis: A latent class specification. *J Environ Plann Manag* 54:1061–1076

Campbell D, Hensher DA, Scarpa R (2012) Cost thresholds, cut-offs and sensitivities in stated choice analysis: Identification and implications. *Resour Energy Econ* 34:396–411

Carson RT (1997) Contingent Valuation and Tests of Insensitivity to Scope In Determining the Value of Non-Marketed Goods: In: Kopp RJ, Pommerhene W, Schwartz N (eds) *Economic, Psychological, and Policy Relevant Aspects of Contingent Valuation Methods*. Kluwer, Boston

CHG (2007) Plan Especial de actuación en situaciones de alerta y eventual sequía de la Cuenca Hidrográfica del Guadalquivir. Confederación Hidrográfica del Guadalquivir. Ministerio de Medio Ambiente, España

CHJ (2007) Plan Especial de actuación en situaciones de alerta y eventual sequía de la Cuenca Hidrográfica del Júcar. Confederación Hidrográfica del Júcar. Ministerio de Medio Ambiente, España

Colombo S, Calatrava-Requena J, Hanley N (2007) Testing Choice Experiment for Benefit Transfer with Preference Heterogeneity. *Am J Agr Econ* 89:135–151

Colombo S, Hanley N (2008) How Can We Reduce the Errors from Benefit transfer? An Investigation Using the Choice Experiment Method. *Land Econ* 84:128–147

Dziegielewska D, Mendelsohn R (2007) Does No Mean No? A Protester Methodology. *Environ Resource Econ* 38:71–87

Hanemann WM (2000) Adaptation and its measurement. *Climatic Change* 45:571–581

Hanley N, Colombo S, Tinch D et al (2006) Estimating the benefits of water quality improvements under the Water Framework Directive: are benefit transferable? *Eur Rev Agr Econ* 33:391–413

Hensher DA, Rose JM, Greene WH (2005) The Implications of Willingness to Pay of Respondents Ignoring Specific Attributes. *Transportation* 32:203–222

Hensher DA, Greene W (2010) Non-attendance and dual processing of common-metric attributes in choice analysis: a latent class specification. *Empir Econ* 39:413–426

Hensher DA, Rose J, Greene W (2012) Inferring attribute non-attendance from stated choice data: implications for willingness to pay estimates and a warning for stated choice experiment design, *Transportation* 39:235–245

Hess S (2005) Advanced discrete choice models with applications to transport demand. PhD dissertation, Imperial College London

Hess S, Stathopoulos A, Campbell D et al (2013) It's not that I don't care, I just don't care very much: confounding between attribute non-attendance and taste heterogeneity. *Transportation* 40(3):583–607

- Iglesias A, Estrela T, Gallart F (2005) Impactos sobre los recursos hídricos. In: Moreno J (ed) Evaluación preliminar de los impactos en España por el efecto del cambio climático. Ministerio de Medio Ambiente, Madrid, pp 303–352
- Johnston RJ, Besedin EY, Ranson MH (2006) Characterizing the effects of valuation methodology in function-based benefit transfer. *Ecol Econ* 60:407–419
- Johnston RJ, Duke JM (2009) Willingness to Pay for Land Preservation across States and Jurisdictional Scale: Implications for Benefit Transfer. *Land Econ* 85:217–237
- Johnston RJ, Duke JM (2010) Socioeconomic adjustments and choice experiment benefit function transfer: evaluating the common wisdom. *Resour Energy Econ* 32:421–438
- Kaul S, Boyle KJ, Kuminoff NV, Parmeter CF, Pope JC (2013) What can we learn from benefit transfer errors? Evidence from 20 years of research on convergent validity. *J Environ Econ Manag* 66: 90–104
- Kragt ME (2013) Stated and Inferred Attribute Attendance Models: A Comparison with Environmental Choice Experiments. *J Agr Econ* 64:719–736
- Krinsky I, Robb A (1986) On approximating the statistical properties of elasticities. *Rev Econ Stat* 68:715–719
- Loomis JB, Roach B, Ward F, Ready R (1995) Testing the transferability of recreation demand models across regions: a study of Corps of Engineers reservoirs. *Water Resour Res* 31:721–730
- Martin-Ortega, J. (2012). Economic prescriptions and policy applications in the implementation of the European Water Framework Directive. *Environmental Science & Policy*, 24, 83-91.
- Martin-Ortega J, Giannocaro G, Berbel J (2011) Environmental and resource costs under water scarcity conditions: an estimation in the context of the European Water Framework Directive. *Water Resour Manag* 25:1615–1633
- Martin-Ortega J, Brouwer R, Ojea E, Berbel J (2012) Benefit transfer of water quality improvements and spatial heterogeneity of preferences. *J Environ Manag* 106:22–29

- McFadden D (1974) Conditional logit analysis of qualitative choice behavior. In: Zarembka P (ed) *Frontiers in Econometrics*. Academic Press, New York, pp 105–142
- McFadden D, Train K (2000) Mixed MNL Models for discrete response. *J Appl Econometrics* 15:447–470
- Meyerhoff J, Liebe U (2009) Status Quo Effect in Choice Experiments: Empirical Evidence on Attitudes and Choice Task Complexity. *Land Econ* 85:515–528
- Moeltner K, Boyle KJ, Paterson RW (2007) Meta-analysis and benefit transfer for resource valuation-addressing classical challenges with Bayesian modeling. *J Environ Econ Manag* 53:250–269
- Morrison M, Bennett J (2000) Choice modelling, non-use values and benefit transfers. *Econ Anal Pol* 30:13–32
- Morrison M, Bennett J, Blamey R, Louviere J (2002) Choice modeling and tests of benefit transfer. *Am J Agr Econ* 84(1):161–170
- Morrison M, Bennett J (2004) Valuing New South Wales rivers for use in benefit transfer. *Aust J Agr Resource Econ* 48:591–612
- Navrud S, Ready R (2007) Review of methods for value transfer. In: Navrud S, Ready R (eds) *Environmental value transfer: Issues and methods*. Springer, Dordrecht, The Netherlands
- Norton D, Hynes S, Hanley N (2012) Accounting for cultural dimensions in estimating the value of coastal zone ecosystem services using international benefit transfer. Paper presented at the 19th Annual Conference of the European Association of Environmental and Resource Economists, Charles University, Prague, 27-30 June
- Poe G, Severance-Lossin E, Welsh M (2005) Simple Computational Methods for Measuring the Difference of Empirical Distributions. *Am J Agr Econ* 87:353–365
- Puckett SM, Hensher DA (2008) The role of attribute processing strategies in estimating the preferences of road freight stakeholders. *Transport Res E Logist Transport Rev* 44:379–395
- Revelt D, Train K (1998) Mixed Logit with repeated choices: households' choices of appliance efficiency level. *Rev Econ Stat* 80:647–657

Rosenberger RS, Stanley TD (2006) Measurement, generalization, and publication: Sources of error in benefit transfers and their management. *Ecol Econ* 60:372–378

Scarpa R, Gillbride TJ, Campbell D, Hensher DA (2009) Modelling attribute non-attendance in choice experiments for rural landscape valuation. *Eur Rev Agr Econ* 36:151–174

Scarpa R, Thiene M, Hensher DA (2010) Monitoring Choice Task Attribute Attendance in Nonmarket Valuation of Multiple Park Management Services: Does It Matter? *Land Econ* 86:817–839

Scarpa R, Zanolli R, Bruschi V, Naspetti S (2013) Inferred and stated attribute non-attendance in food choice experiments. *Am J Agr Econ* 95:165–180

Train K, (1998) Recreation demand models with taste differences over people. *Land Econ* 74:185–194

Train (2003)

Train and Weeks (2005)

Van den Berg TP, Poe GL, Powell JR (2001) Assessing the accuracy of benefit transfers: evidence from a multi-site contingent valuation study of groundwater quality. In: Bergstrom JC, Boyle KJ, Poe GL (eds) *The Economic Value of Water Quality*. Edward Elgar, Massachusetts

Vinten AJA, Martin-Ortega J, Glenk K et al (2012) Application of the WFD cost proportionality principle to diffuse pollution mitigation: a case study for Scottish Lochs. *J Environ Manag* 97:28–37

Wilson MA, Hoehn JB (2006) Valuing environmental goods and services using benefit transfer: state-of-the art and science. *Ecol Econ* 60:335–342

Figure 1. Example of a typical choice card







SITUATION A	SITUATION B	NO MEASURES ARE TAKEN
<p>1 out of 10 years</p> 	<p>3 out of 10 years</p> 	<p>4 out of 10 years</p> 
<p>moderate</p> 	<p>good</p> 	<p>poor</p> 
<p>80 € per year</p>	<p>40 € per year</p>	<p>0 € per year</p>

Table 1. Description of the choice experiment attributes and levels

Attributes	Description
<i>Attribute 1: Levels of ecological status related to water flow (ENV)</i>	
Poor (baseline)	This is the expected future situation of low water levels and environmental quality status. It entails a large deviation from the natural situation due to the increased water scarcity conditions. Many fish species will disappear and riverbanks will lose most of their vegetation. As a result most birds and insects will disappear as well.
Moderate	Less than natural water levels and environmental quality status. This is still a substantial deviation from the natural situation. A limited number of fish species are present. Riverbanks have some vegetation supporting a limited number and variety of birds and other wildlife.
Good	Water levels and environmental quality status are close to their natural levels. There is a small deviation from the natural situation. Under these conditions riverbanks have a lower than natural vegetation cover. As a result the breeding and nesting conditions for a number of species are limited.
Very good	Water levels and the environmental quality status are at their natural state where there is almost no pressure from human activities. Conditions for wildlife are optimal under these circumstances.
<i>Attribute 2: Frequency of household water restrictions (RES)</i>	
Restrictions in 4 out of the next 10 years (baseline)	Restrictions on watering gardens, filling swimming pools, using washing machines and other 'secondary water uses'. Water shortfall lasting 6 to 7 hours during up to 20 days in dry summers.
Restrictions in 3 out of the next 10 years	
Restrictions in 2 out of the next 10 years	
Restrictions in 1 out of the next 10 years	
<i>Monetary attribute (COST)</i>	Increase in the household water bill of 0 (baseline) 20, 40, 60, 80, 100 and 120 euros per year during the next 10 years

Table 2. Summary statistics of case study descriptive variables

Variable	GRB			SRB			GRB vs SRB p-value
	Sample	Target Population	p-value Sample vs Target population	Sample	Target Population	p-value Sample vs Target population	
Socio-economic characteristics							
Age groups (%)							
18-34 years	31.4	32.5	0.95	30.5	30.3	1	0.095
35-59 years	51.1	48.8		44.9	44.8		
60 or more years	17.5	18.7		24.6	25.0		
Age (mean)	43.6	-	-	46.2	-	-	< 0.01**
Gender (male=0; female=1)	0.5	0.5	0.88	0.5	0.5	0.88	0.83
Education (%)							
No studies	12.7	17.1	0.067	13.5	36.8	<0.01 ^a	0.95
Secondary studies	58.7	69.9		62.1	51.7		
Upper Secondary studies	25.7	13.0		24.4	11.5		
Household size	3.5	2.8	0.051	3.0	2.7	0.68	< 0.01**
Annual household <i>income</i> (10k €)	2.3	2.1	< 0.01 ^a	2.2	2.5	< 0.01 ^a	< 0.01**
Recalled Annual water bill (€)	168	-	-	213	-	-	< 0.05**
Perceptual variables							
Water <i>scarcity</i> is a problem in the region (0=not a problem at all; 4=a big problem)	3.1	-	-	2.5	-	-	< 0.01**
Respondent has <i>suffered</i> from water use restrictions in the past (1=yes;0=no)	0.6	-	-	0.2	-	-	< 0.01**
Environment should have <i>priority</i> when allocating water in case of scarcity (after households) (1=yes;0=no)	0.4	-	-	0.3	-	-	< 0.01**
It is <i>credible</i> that water scarcity in the region will increase (0=not credible at all; 4=totally credible)	2.8	-	-	2.3	-	-	< 0.01 ^{a**}

Note: Variables in italics are included in the utility function of the choice models; Reported summary statistics are mean values; ^a Sample is not representative with respect to characteristic based on a Chi-square test at the 5% significance level; ** Statistically significant differences between SRB sample and GRB sample at the 5% significance level following a Chi-square test

Table 3. Model results

	CL GRB		CL SRB		MXL GRB ^a		MXL SRB ^a		ECLCM GRB		ECLCM SRB	
RES	-0.118	*	-0.448	***	-0.024		-0.108	***	-1.673	***	-1.857	***
	(0.065)		(0.047)		(0.034)		(0.008)		(0.366)		(0.276)	
RES_suffered	-0.087		0.176	***	-0.023	**	0.004		-0.386		0.135	
	(0.057)		(0.067)		(0.01)		(0.007)		(0.229)		(0.298)	
RES_income	-		-		-0.010		-0.027	***	-		-	
					(0.01)		(0.004)					
ENV2	0.115		0.673	***	-0.025		0.277	***	3.925	***	3.227	***
	(0.126)		(0.107)		(0.054)		(0.027)		(0.826)		(0.481)	
ENV2_priority	0.482	***	0.513	***	0.029		0.126	***	0.565		0.417	
	(0.161)		(0.157)		(0.025)		(0.019)		(0.655)		(0.468)	
ENV2_income	-		-		0.04	*	0.07	***	-		-	
					(0.023)		(0.008)					
ENV3	0.381	***	1.063	***	-0.122		0.378	***	5.513	***	5.641	***
	(0.136)		(0.113)		(0.076)		(0.033)		(0.945)		(0.819)	
ENV3_priority	0.172		0.962	***	0.01		0.374	***	0.331		1.354	*
	(0.175)		(0.161)		(0.037)		(0.033)		(0.841)		(0.700)	
ENV3_income	-		-		0.103	***	0.116	***	-		-	
					(0.031)		(0.010)					
COST	-0.042	***	-0.016	***	-2.19	***	2.02	***	-0.138	***	-0.246	***
	(0.007)		(0.004)		(0.115)		(0.175)		(0.022)		(0.032)	
COST_income	-0.000		0.001				-		0.006		0.03	***
	(0.003)		(0.001)						(0.008)		(0.008)	
ASC	-1.422	**	2.093	***	-0.452	***	-1.82	***	-		-	
	(0.665)		(0.433)		(0.102)		(0.113)					
ASC_age ^b	0.013		0.037		-0.003		0.293	***	-0.093		-0.126	
	(0.047)		(0.042)		(0.009)		(0.015)		(0.099)		(0.230)	
ASC_income	0.034		-0.338	***	0.074	***	-0.225	***	0.340		-0.916	**
	(0.216)		(0.125)		(0.03)		(0.026)		(0.397)		(0.429)	
ASC_gender	-0.119		0.059		-0.053	**	0.764	***	-0.157		2.390	***
	(0.125)		(0.137)		(0.025)		(0.050)		(0.279)		(0.834)	
ASC_scarcity	-0.052		-0.318	***	0.014		-0.412	***	0.045		-4.228	***
	(0.058)		(0.054)		(0.01)		(0.025)		(0.125)		(0.637)	
ASC_credible	-0.077		-0.398	***	0.032	*	0.003		-0.386	*	-3.298	***
	(0.113)		(0.062)		(0.017)		(0.017)		(0.229)		(0.532)	

Table 3 ctd. Model results

	CL GRB	CL SRB	MXL GRB ^a	MXL SRB ^a	ECLCM GRB	ECLCM SRB
ASC class 1	-	-	-	-	7.721 *** (1.939)	1.286 (2.468)
ASC class 2	-	-	-	-	-1.945 (1.601)	-19.614 *** (6.986)
ASC class 3	-	-	-	-	-3.472 ** (1.673)	13.504 *** (3.087)
ASC class 4	-	-	-	-	3.116 (5.158)	5.786 (6.875)
ASC class 5	-	-	-	-	-5.391 *** (1.388)	24.555 *** (4.177)
ASC class 6	-	-	-	-	-0.973 (4.887)	12.916 *** (2.970)
ASC class 7	-	-	-	-	6.034 * (3.556)	5.328 (4.479)
ASC class 8	-	-	-	-	-0.689 (1.231)	11.394 *** (2.486)
ASC class 9 ^c	-	-	-	-	0	0
RES STDV	-	-	0.023 *** (0.007)	0.159 *** (0.007)	-	-
ENV2 STDV	-	-	0.076 *** (0.018)	0.177 *** (0.013)	-	-
ENV3 STDV	-	-	0.105 *** (0.026)	0.06 *** (0.010)	-	-
COST STDV	-	-	1.32 *** (0.193)	2.54 *** (0.205)	-	-
$\sigma_{A,B}^d$			2.38 *** (0.283)	11.3 *** (1.90)		
LL	-1219.8	-1387.0	-1096.1	-985.5	-1069.7	-928.7
Observations	1348	1544	1348	1544	1348	1544
AIC	2467.6	2802.0	2234.2	2013.0	2197.4	1915.4
BIC	2483.4	2818.6	2257.9	2038	2230.2	1949.9
Adjusted Pseudo R ²	0.17	0.17	0.25	0.41	0.26	0.44

Note: ^a Cost scaled by 1/100 (multiply by 100 to arrive at WTP in €); ^b Age scaled by 1/10; ^c The ASC is constrained to be equal to zero in class 9; ^d variance of the additional error component shared between alternatives A and B; MXL models estimated using 2,000 Halton draws; *, **, *** statistically significant at the 1%, 5% and 10% levels

Table 4. Attribute Non-Attendance patterns in the Guadalquivir River Basin (GRB) and Serpis River Basin (SRB) data based on equality constraint latent class (ECLC) model results

GRB	Class membership probability GRB	Class membership probability SRB	Attendance pattern specified in ECLC model ^a			
			RES	ENV2	ENV3	COST
Class 1	0.102	0.056	1	1	1	1
Class 2	0.178	0.143	0	1	1	1
Class 3	0.253	0.097	1	0	0	1
Class 4	0.016	0.21	1	1	1	0
Class 5	0.221	0.129	0	0	0	1
Class 6	0.043	0.147	0	1	1	0
Class 7	0.002	0.087	1	0	0	0
Class 8	0.18	0.114	0	0	0	0
Class 9 ^b	0.008	0.017	0	0	0	0
Probability of non-attendance GRB			0.628	0.660	0.660	0.246
Probability of non-attendance SRB			0.55	0.445	0.445	0.575

Note: ^a Attendance (1=attended; 0 = not attended); ^b As opposed to class 8, class 9 restricts the ASC to be zero, reflecting random choice among all three alternatives

Table 5. Mean marginal willingness to pay estimates

	At mean of covariates of study site				At mean of covariates of policy site (Function transfer)			
	MXL GRB	MXL SRB	ECLCM GRB	ECLCM SRB	MXL GRB	MXL SRB	ECLCM GRB	ECLCM SRB
RES ^a	6.00 (4.39;7.67)	16.47 (15.23;17.82)	7.14 (4.83;9.42)	3.66 (2.83;4.62)	5.10 (3.23;6.95)	16.49 (15.20;17.81)	6.61 (4.37;8.8)	3.59 (2.81;4.48)
ENV2 ^b	7.79 (4.79;10.76)	47.36 (44.09;50.72)	12.63 (7.71;17.70)	8.69 (7.45;10.04)	7.28 (4.25;10.28)	48.88 (45.57;52.21)	12.44 (7.56;17.41)	8.88 (7.62;10.25)
ENV3 ^b	11.45 (7.18;15.76)	76.16 (71.13;81.08)	17.06 (11.45;22.79)	15.72 (13.72;17.76)	10.70 (6.32;15.15)	80.09 (75.07;85.13)	16.92 (11.55;22.46)	16.19 (14.2;18.27)

Note: all values in € per household and year; 95% confidence bounds for mean marginal WTP estimates reported in parentheses were derived using the Krinsky and Robb (1986) method with 2,000 draws; ^aRES: Marginal WTP for a reduction of the frequency of water use restrictions of one year out of ten; ^bENV: marginal WTP for an improvement in environmental quality to good (ENV2) and very good (ENV3) status

Table 6. Compensating surplus estimates

	Change in frequency of water use restrictions ^a	Change in environmental quality of water bodies ^b	Compensating surplus ^c			
	RES	ENV	MXL GRB	MXL SRB	ECLCM GRB	ECLCM SRB
Scenario 1	-1	good	54.33 (49.48;59.03)	249.82 (175.69;322.2)	35.18 (31.92;38.52)	59.62 (48.17;71.56)
Scenario 2	-1	very good	55.55 (50.45;60.86)	274.93 (201.58;346.88)	39.65 (36.22;43.10)	66.94 (55.28;79.02)
Scenario 3	-2	good	58 (53.37;62.44)	265.56 (190.22;338.36)	41.79 (38.13;45.51)	63.22 (51.77;75.38)
Scenario 4	-2	very good	59.22 (54;64.45)	290.66 (216.58;363.88)	46.26 (42.55;50.13)	70.53 (58.78;82.76)
Scenario 5	-3	good	61.67 (56.44;66.94)	281.3 (205.3;355.35)	48.4 (43.18;53.58)	66.81 (55.29;79.12)

Scenario 6	-3	very good	62.89 (56.95;68.49)	306.4 (232.37;380.83)	52.87 (47.77;58.13)	74.12 (62.34;86.54)
------------	----	-----------	------------------------	--------------------------	------------------------	------------------------

Note: ^a number of years out of ten (baseline: four years); ^b baseline level: poor; ^c € per household and year; 95% confidence bounds for WTP estimates reported in parentheses were derived using the Krinsky and Robb (1986) method with 2,000 draws

Table 7. Mean transfer errors

	MXL		ECLCM	
	GRB → SRB	SRB → GRB	GRB → SRB	SRB → GRB
<i>Absolute transfer errors (€)^a</i>				
RES	-11.39	11.39	3.01	-3.01
ENV2	-41.60	41.60	3.57	-3.57
ENV3	-69.39	69.39	0.73	-0.73
Scenario 1	-182.69	182.69	-24.44	24.44
Scenario 2	-210.48	210.48	-27.28	27.28
Scenario 3	-194.08	194.08	-21.43	21.43
Scenario 4	-221.86	221.86	-24.27	24.27
Scenario 5	-205.46	205.46	-18.42	18.42
Scenario 6	-233.25	233.25	-21.26	21.26
<i>Relative transfer errors (%)^b</i>				
RES	69.23	232.86	82.57	44.98
ENV2	85.19	602.65	39.39	27.25
ENV3	86.72	682.37	9.03	8.68
Scenario 1	78.81	373.17	40.40	69.90
Scenario 2	80.03	402.26	40.31	69.06
Scenario 3	78.16	358.85	33.37	51.50
Scenario 4	79.38	386.26	34.01	52.57
Scenario 5	77.59	347.26	27.09	38.29
Scenario 6	78.80	373.10	28.32	40.35

Note: GRB →SRB indicates transfers from the GRB (study site) to the SRB (policy site); SRB→GRB indicates transfers from the SRB (study site) to the GRB (policy site)

