Structurally-consistent estimation of use and non-use values for landscape-wide environmental change

Abstract

We address the problem of estimating the use and non-use value derived from a landscape-wide programme of environmental change. Working in the random utility framework, we develop a structural model that describes both demand for recreational trips to the landscape's quality-differentiated natural areas and preferences over different landscape-wide patterns of environmental quality elicited in a choice experiment. The structural coherence of the model ensures that the parameters of the preference function can be simultaneously estimated from the combination of revealed and stated preference data. We explore the properties of the model in a Monte Carlo experiment and then apply it to a study of preferences for changes in the ecological quality of rivers in northern England. This implementation reveals plausible estimates of the use and non-use parameters of the model and provides insights into the distance decay in those two different forms of value.

Keywords: random utility models; non-use utility; structural modelling; water quality; recreation; travel cost model; choice experiment

1. Introduction

The problem addressed by this paper concerns the appraisal of programmes or policies whose environmental impacts are not constrained to a particular site but which have widespread yet spatially differentiated impacts across a landscape. Our motivating example concerns a programme designed to improve the ecological status of a region's rivers[1], a change that delivers benefit flows not only through improved recreational experiences at riverside sites but also, potentially, through increased non-use value.

Valuing a programme that delivers a simultaneous change in the quality of natural areas across a landscape presents a significantly more complex challenge than the single-site appraisal problem. With respect to recreational use, perhaps, the most significant of those challenges is in understanding how individuals assess the value of a landscape offering a multiplicity of recreational opportunities. In the recreational demand literature, that problem is most frequently approached through adoption of the discrete choice travel cost (TC) method a revealed preference (RP) approach that derives estimates of value from examination of recreational choice behaviour. Discrete choice TC modelling is theoretically underpinned by the random utility model (RUM) (McFadden, 1973) which formally defines the structure of individuals' preferences and the process through which individuals choose which of the set of quality-differentiated recreational destinations to visit. Armed with estimates of the preference function derived from a clear structural model of behaviour, TC analysts can explore the potential welfare consequences that might arise should environmental quality change in a variety of different ways across any number of sites in a landscape.

While the discrete-choice TC model has many advantages it also has short-comings. One obvious short-coming is that the technique is unable to estimate value derived from non-use. Moreover, reliance on observed behaviour may result in problems of identification, for example, when the range of current qualities fails to cover that to be evaluated under the proposed

[1] Other similar programmes include agri-environment schemes that deliver regional changes in agricultural landscapes, air quality regulations that tighten controls over regional air pollution sources and hazardous waste clean-up programs that rehabilitate an array of contaminated locations.

programme. An alternative method that addresses both those short-comings is provided by stated preference (SP) methods of valuation.

In theory, an SP approach could generate data that would allow identification of both use and non-use values for any desired pattern of environmental quality change. Realising that outcome is, of course, more difficult. One problem concerns how to convey to respondents the complex spatial reality of the landscape within which hypothetical changes in environmental quality occur. As we discuss further in Section 2, here we champion the use of *visual spatial choice experiment*s (VSCEs), an SP elicitation method in which respondents are asked to choose between policy options presented to them in the form of maps displaying the environmental quality delivered by a policy at each location in the landscape. This use of spatial visualisation provides a mechanism whereby complex differences in the spatial organisation of supply under different policy options can be conveyed to respondents in a simple and accessible manner.

A second complexity, concerns analysis of responses to SP exercises like the VSCE. The particular difficulty here is that the preferences underpinning those responses reflect value in use and non-use and are shaped by the substitution possibilities afforded by the diversity of natural areas present in a landscape. As we describe in Section 2, to date, the analysis of SP data has failed to convincingly address these complexities. Analysts have tended to forego the formality of structural modelling relying instead on highly reduced-form specifications which confound use and non-use value and address issues of spatial location and substitution through estimation of 'distance-decay' parameters and broad indices of substitute availability and quality.

The core contribution of this paper is to propose an estimation strategy for SP data that directly addresses these issues. In Section 3 we describe a formulation for the preference function that captures both use and non-use values. Moreover we show how within the RUM framework, that specification can be developed into a coherent structural model of choice behaviour that not only describes respondents' observed recreational activity but also their choices in a VSCE exercise. Indeed, one significant advantage of our approach is that it results in econometric specifications for both observed recreational behaviour and VSCE responses that are derived from the same structural model. This common derivation has the added benefit of allowing RP and SP data to simultaneously inform estimation of the same structural parameters.

As we describe in Section 4, deriving our econometric specification from a structural model, results in a preference function that is highly nonlinear in the structural parameters. Moreover, unlike most previous applications of the RUM method, our structurally derived model results in error terms that are sums of independent logistic variates. Given these complexities, in Section 5, we present the results of a Monte Carlo experiment in which we investigate the properties of this econometric specification. In particular, we examine a variety of data generating processes in which individuals place different relative weights on use and non-use and examine under what conditions the model can successfully recover the preference parameters. Finally, in Section 6, we estimate the model in the context of a real data set collected in a large scale survey carried out across Northern England. This implementation reveals intuitively plausible estimates of the use and non-use parameters of the model and provides interesting insights into the distance decay in those two different forms of value.

As we describe in detail in the next section, our paper makes contributions to a number of literatures. Primarily, it makes a contribution to the field of structural econometric modelling in environmental economics (Timmins and Schlenker, 2009). As far as the authors are aware, we provide the first attempt to underpin the analysis of non-market valuation data for landscapes of quality-differentiated sites with a coherent structural description of use and non-use value. In that way, our approach builds on the structurally-coherent model developed by Eom and Larson (2006) to explore use and non-use values in the single site setting. Moreover, in providing a new approach to the simultaneous analysis of TC and CE data, our work makes a novel contribution to the literature on the combined analysis of RP and SP data (Whitehead et al., 2008). Our paper is also of interest to the field of discrete choice modelling, presenting as it does a RUM specification not previously explored in the literature. Finally, our research contributes to various literatures within SP research by showing how issues of distance decay, substitute location and quality as well as use and non-use value can be coherently addressed through utility-theoretic modelling of preferences.

2. Literature Review

There is a long history of applying SP methods to the problem of valuing spatially-explicit environmental quality change. Early applications used the contingent valuation method and focused on valuing environmental quality change at some particular location (Davis and Knetsch, 1966; Oster, 1977). The same problem has since been examined through application of the choice experiment (CE) method, an approach that allows analysts to derive a richer description of values for different dimensions of quality change at the focus location (e.g. Boxall et al., 1996; Hanley et al., 1998). More recently, SP methods have been applied in efforts to value more complex patterns of landscape-wide environmental change (e.g. Brouwer et al., 2010; Meyerhoff et al., 2014)

In stark contrast to the structural coherence of the RUM model underpinning discrete choice TC modelling, analysis of responses to SP studies has almost universally adopted extreme reduced-form descriptions of preferences. In those studies, preferences are described as a simple, often linear, value function relating willingness to pay (WTP) to the level of environmental quality change and a variety of qualifiers. While many analysts include amongst those qualifiers measures of the distance from a respondent's home to the site of the environmental change (e.g. Bateman et al., 2006; Concu, 2007; Hanley et al., 2003; Sutherland and Walsh, 1985), few use the trip expenditure measure that is so central to the structural model of recreational-value formation used in TC analysis[2]. One possible justification for using a distance rather than travel cost measure is that SP surveys capture value flows derived both in use and non-use. While, all else equal, use values must fall with increasing distance as a consequence of the rising costs of access, the same expectation is not self-evident for non-use values (Bateman et al., 2006). Indeed, the empirical evidence for distance decay in non-use values is mixed and based almost entirely on studies that have compared values expressed by those that currently use the resource (users) with those that currently do not (non-users). Hanley et al. (2003), for example, find that values for users decay more rapidly than those for non-users while Bateman et al. (2006) find no distance decay for non-users in one study and comparable levels of distance decay for users and

[2] Though some authors have used travel time, a close correlate of travel cost (e.g. Jørgensen et al., 2013; Taylor and Longo, 2010)

non-users in another. Of course, values of users and non-users should not be conflated with use values and non-use values; for a start, it would be unreasonable to assume that current users of a resource do not also hold non-use values. Indeed, Cummings and Harrison (1995) argue that a short-coming of SP methods is that they fail to provide an operationally meaningful mechanism by which use and non-use values can be separately identified.

A central contribution of our paper is to present a method of analysis that overcomes the Cummings and Harrison (1995) critique. As we describe shortly, our method derives from a clear structural representation of preferences for use and non-use that progresses to a method of estimation that allows for separate identification of distance decay effects from use and non-use. In line with the structural model, use value is assumed to decay with the costs of access while we use a flexible functional form to investigate whether and how non-use value decays with distance.

Even less structurally convincing than the treatment of distance is the way in which SP studies have addressed the issue of substitutes. Indeed, many SP studies simply ignore the issue altogether (amongst many others; Birol et al., 2006; Doherty et al., 2014; Hanley et al., 2003; Stithou et al., 2013). Where attempts have been made to control for substitute availability those controls tend to have been included in the model specification in ways that bear little resemblance to any formal model of recreational demand behaviour. Often that means including some aggregate measure of the quantity or density of environmental assets in the vicinity of a respondent's home (e.g. De Valck et al., 2017; Pate and Loomis, 1997; Yao et al., 2014) or the proximity of the nearest substitute (e.g. Caudill et al., 2011; Söderberg and Barton, 2014).

Of course, the influence of substitutes on values is not only determined by the proximity of alternative sites but also by their quality. While the discrete choice TC model explicitly incorporates the quality of each substitute into its description of recreational demand behaviour, the question of substitute quality has, until recently, received little attention in the SP literature. Part of the problem has been in finding methods to present respondents with descriptions of complex spatial patterns of quality-differentiated substitute locations. Recently, a growing number of studies have sought to address those difficulties by depicting the context of substitutes and their qualities on colour-coded and annotated maps (e.g. Brouwer et al., 2010; Kataria et al., 2012; Schaafsma and Brouwer, 2013; Söderberg and Barton, 2014)[3]. The particular map-based SP elicitation method we explore in this paper is similar to that used by Horne et al. (2005) and Meyerhoff et al. (2014). In those studies, respondents are presented with a selection of maps each illustrating a different spatial pattern of quality change and each associated with some particular cost. Respondents are asked to identify which costly pattern of quality change is their most preferred. We describe this form of CE as a *visual spatial choice experiment* (VSCE).

Map-based presentations like the VCSE allow SP studies to elicit preferences for complex landscape-wide patterns of environmental change, explicitly presenting respondents with information on the quality and location of substitutes upon which they may condition their responses. How those responses should be modelled to properly reflect respondents' decision processes, however, remains an open question in the SP literature. Perhaps the most complete representation of substitute quality and location in that literature is that provided by Meyerhoff et al. (2014) who adopt a reduced-form specification in which distance decay parameters are estimated specific to each level of quality change at each site. While the Meyerhoff et al. (2014) specification represents a considerable improvement on previous analyses, it captures only the main effects of each site's quality and proximity. The interaction terms, which identify substitution effects and determine how the value of one site is influenced by the quality and proximity of all other sites, are ignored. Moreover, the reliance on a specification that delivers site-specific parameter estimates inhibits effective transfer of the value estimates outside the study area to locations exhibiting different spatial patterns of quality-differentiated substitutes.

The central contribution of this paper is to build an econometric specification for the analysis of VSCE data that is derived from a coherent structural model of preferences for landscape-wide environmental quality change. In that pursuit, our starting point is the RUM used in discrete-choice TC studies to describe the preferences that drive recreational choices over substitute, quality-differentiated sites. Accordingly, our work has parallels to the contingent

e3 A number of these map-based SP studies stem from the same EU-funded project (Aquamoney) which sought evidence on the value of landscape-wide water quality improvement in order to inform implementation of the EU's Water Framework Directive. Indeed, the empirical case study we describe subsequently as part of that same effort.

behaviour literature (e.g. Adamowicz et al., 1994; Christie et al., 2007; Englin and Cameron, 1996; Whitehead et al., 2000) which combines observed recreational behaviour with SP data recording how respondents report they would behave if the qualities or availability of recreation sites were to change. Since the SP and RP data are assumed to be driven by the same choice process, the structural modelling analyses used in TC studies can be applied directly to the combined data.

Of course, contingent behaviour studies are limited to identifying information on preferences for recreational use. The method we outline is for application to VSCE data that contains expressions of preferences reflecting not only use values, but also values from non-use. Accordingly, our model of responses to VSCE exercises has to incorporate a structural description for non-use utility into the preference specification. Following the lead of Mäler et al. (1994) and Freeman (2003) we take the operational definition of use value to be 'values revealed by market behaviour' and of non-use value to be a 'value _not_ revealed by market behaviour'. Those definitions imply separability of the utility function (for review of concepts of separability see Deaton and Muellbauer, 1980). In our case, we have a sub-utility function in which the environmental quality of sites is combined with travel in order to deliver recreational use values, while a second sub-utility function delivers non-use value simply through environmental quality.

In recovering the structural parameters of a preference function defining both use and non-use values for environmental quality, our work is comparable to that of Eom and Larson (2006). Those authors explore the decomposition of total value into use and non-use value in the context of valuing a single recreational site. Indeed, they begin with a particular specification of the Marshallian trip demand function for that site and integrate back to reveal the form of the quasi-expenditure function, interpreting the constant of integration as the source of non-use value. In a sense, the approach we develop in this paper can be considered as the discrete choice counterpart to the continuous demand model developed by Eom and Larson (2006) and one that moves away from the focus on a single site to consider the complexities of environmental valuation in the context of a landscape of quality-differentiated sites[4].

[4] Note that Eom and Larson (2006) acknowledge that the river basin which forms the subject of their investigation, consists of a variety of different sites, visits to six of which are elicited in their survey. Rather than estimating a system of demand equations for these substitute sites, Eom and Larson (2006)

In estimating the parameters of their model Eom and Larson (2006) combine RP data on recreational visits to a site with SP data recording how greatly respondents value improvements in environmental quality at that site. In the field of non-market valuation, efforts to combine RP and SP data in model estimation stretching back at least as far as Cameron (1992) and Adamowicz et al. (1994). Like Eom and Larson (2006), we contribute to that literature by showing how the derivation of an econometric specification from a common structural model of behaviour ensures that both RP and SP data contribute to identification of the same structural parameters.

3. The Structural Model

The data we wish to interrogate reports on the preferences of a sample of individuals indexed $i = 1,2, \dots, N$, living in a region endowed with an assortment of natural areas, indexed $j = 1,2, \dots, J$. The welfare that an individual realises from those natural areas during time period t arises as a result of the qualities that those areas exhibit; our particular interest being their environmental qualities. Qualities differ across natural areas and may differ across time periods, though are assumed to remain constant for the duration of any one period. The qualities of natural areas can also differ across possible states of the world, $s = 0,1, \dots, S$. The reality of the current state of the world is indicated $s = 0$ and the S alternative states of the world are those constructed to describe the qualities of natural areas for the purposes of a non-market valuation exercise. The qualities of natural area j, under scenario s is given by the vector $\boldsymbol{q}_{j,t,s}$, though for simplicity of notation we assume that those qualities remain constant across

Natural areas can be used for outdoor recreation, though to enjoy the recreational experience offered by natural area j, an individual must make the round trip to that location. We indicate the consumption levels of those trips by the vector $\boldsymbol{x}_{i,t,s} = \left(x_{1,i,t,s}, x_{2,i,t,s}, \dots, x_{J,i,t,s}\right)$ and take those to be goods whose purchase prices (comprising the costs of travel and the opportunity cost of travel time) are identified by the price vector $\boldsymbol{p}_i = \left(p_{i,1}, p_{i,2}, \dots, p_{i,J}\right)$.

Individuals might also gain utility from natural areas without having to purchase any

choose to focus on estimating the value for a "typical site", defining the level of demand for that typical site as the number of visits a respondent takes to their most frequently visited site.

complementary market goods; perhaps through the pleasure they derive simply from knowing that such natural areas exist or from the knowledge that others may benefit from their existence.[5] Again we assume that utility derived in this way arises as a consequence of the qualities of a natural area.[6] Following evidence from the SP literature (Bateman et al., 2006; Schaafsma et al., 2012) we allow for the possibility that the non-use value derived from a natural area with particular qualities may differ as distance from an individual's home increases. Those distances are identified by the vector $\boldsymbol{d}_i = \left(d_{i,1}, d_{i,2}, \dots, d_{i,M}\right)$.

Our assumptions lead us to the direct utility function;

$$U\left(U^{use}\left(\boldsymbol{x}_{i,t,s}, \boldsymbol{q}_{1,t,s}, \dots, \boldsymbol{q}_{J,t,s}\right), U^{non-use}\left(\boldsymbol{q}_{1,t,s}, \dots, \boldsymbol{q}_{J,t,s}, \boldsymbol{d}_i\right), z_{i,t,s}\right) \quad (\forall\, i, t, s) \tag{1}$$

where z is a numeraire good with unit price. Observe from (1) that as a result of our definition of use and non-use value (see Section 2) the utility function is separable into value derived from use of natural areas, $U^{use}(\cdot)$, value derived from non-use, $U^{non-use}(\cdot)$, and value derived from consumption of other goods, $z_{i,t,s}$, and that utility is increasing in all three of those arguments. Moreover, we assume that $U^{use}(\cdot)$ exhibits weak complementarity such that the utility from use derived from the qualities of a natural area falls to zero when consumption of trips to that site is zero; that is to say, $\partial U^{use}/\partial \boldsymbol{q}_{j,t,s} = \boldsymbol{0}$ when $x_{j,t,s} = 0$. We also assume that preferences are strongly separable over time.

If the choice period is reduced to a length of time such that in each period, t, an individual can make at most one recreational trip, then an individual's consumption decision amounts to solving the discrete choice problem given by (Phaneuf and von Haefen, 2009);

$$\max_{x_{i,t,s}, z_{i,t,s}} \; U\left(U^{use}\left(\boldsymbol{x}_{i,t,s}, \boldsymbol{q}_{1,t,s}, \dots, \boldsymbol{q}_{J,t,s}\right), U^{non-use}\left(\boldsymbol{q}_{1,t,s}, \dots, \boldsymbol{q}_{J,t,s}, \boldsymbol{d}_i\right), z_{i,t,s}\right) \tag{2}$$

[5] For simplicity of notation, we assume that all natural areas are accessible. Individuals may gain non-use utility from sites even if they are not accessible for recreation, a fact that analysts might exploit in attempts to identify the separate contribution of environmental quality to utility in use and non-use.

[6] The qualities which deliver value in non-use could potentially differ from those offering value in use. Our notation assumes, therefore, that the vector $\boldsymbol{q}_{j,t,s}$ is a comprehensive list of utility-relevant quality attributes, but that the contribution which a particular quality element makes to value in use or non-use may be zero.

$$s.t. \qquad y_{i,t} = \boldsymbol{p}_i'\boldsymbol{x}_{i,t,s} + z$$
$$x_{j,t,s} \in \{0,1\} \qquad\qquad (\forall\, i,t,s)$$
$$x_{j,t,s}x_{k,t,s} = 0 \; (\forall\, k \neq j)$$

The conditional indirect utility function that arises from (2) takes the form;

$$u_{i,t,s|j} = u\big(u^{use}(\boldsymbol{q}_{j,t,s}), u^{non-use}(\boldsymbol{q}_{1,t,s}, \dots, \boldsymbol{q}_{J,t,s}, \boldsymbol{d}_i\,), y_{i,t} - p_{i,j}\,\big) \qquad (\forall\, i,t,s) \tag{3}$$

Observe that our assumptions regarding weak complementarity, imply that an individual only derives use utility from the qualities of the natural area that they choose to visit in choice period t. In contrast, during that period individuals derive non-use utility from the qualities of all natural areas. Our model of recreational behaviour is completed through the rational choice rule;

$$\text{choose to visit } j \text{ in period } t \text{ if:} \quad u_{i,t,s|j} > \{u_{i,t,s|k}\}_{\forall\, k \neq j} \qquad (\forall\, i,t,s) \tag{4}$$

Over the course of a year we assume that individuals face $t = 1,2,\dots,T$ recreational choice periods of equal length and that in each period individuals follow (4) in determining their recreational choice behaviour. Accordingly our model follows the tradition of repeated discrete choice models as per Morey et al. (1993).

The research we describe subsequently, involves a VSCE exercise in which respondents are asked to consider alternative states of the world in which the qualities of the natural areas differ from those experienced in the current state of the world ($s = 0$). The quality changes described in each alternative state of the world ($s = 1, 2, \dots, S$) cannot be achieved without cost, a cost to which individuals must contribute through a hypothetical coercive annual charge C_s. Since the year is divided into T equally-sized choice periods indexed $t = 1, 2, \dots, T$ we assume that the annual payment can be equivalently expressed as a series of per period payments; $c_s = C_s/T$. Preferences for these different states of the world have the same fundamental structure, though the conditional indirect utility function (3) must be modified to include the hypothetical payment;

$$u_{i,t,s|j} = u\big(u^{use}(\boldsymbol{q}_{j,t,s}), u^{non-use}(\boldsymbol{q}_{1,t,s}, \dots, \boldsymbol{q}_{M,t,s}, \boldsymbol{d}_i\,), y_{i,t} - p_{i,j} - c_s\,\big) \qquad (\forall\, i,t,s) \tag{5}$$

which reduces to (3) in the current state of the world since $c_0 = 0$;

In a typical hypothetical choice task, individuals are presented with a set of scenarios, $\mathbb{s}$, drawn from the S scenarios constructed for the VSCE. Respondents are asked to indicate which scenario is their most preferred. According to our model, to make that choice, respondents must first solve the site visitation problem (4) for each time period such that their declared preference over hypothetical scenarios should be made according to the choice rule;

$$\text{choose } s \text{ if:} \quad \sum_{t=1}^{T} \max_{j}\left(u_{i,t,s|j}\right) > \left\{\sum_{t=1}^{T} \max_{j}\left(u_{i,t,r|j}\right)\right\}_{\forall\, r\neq s} \qquad (s, r \in \mathbb{s}) \tag{6}$$

where the summation over the T time periods in a year follows from our assumption of inter-temporal additive separability of the utility function.

4. The Econometric Model

We develop our econometric model by first specifying a functional form for the conditional indirect utility function (4). Our separability assumptions are compatible with the additive form;

$$\begin{aligned} u_{i,t,s|j} &= v_{i,t,s|j}^{use} + v_{i,t,s}^{non-use} + v_{i,t,s|j}^{other} + \varepsilon_{i,j,t,s} \\ &= v_{i,t,s|j} + \varepsilon_{i,j,t,s} \qquad\qquad (j = 1, 2, \dots, J+1 \text{ and } \forall i, t, s) \end{aligned} \tag{7}$$

where $\varepsilon_{i,j,t,s}$ is an econometric error term introduced to capture the divergence between our model of conditional indirect utility $\left(v_{i,t,s|j}^{use} + v_{i,t,s}^{non-use} + v_{i,t,s|j}^{other}\right)$ and the individual's experienced utility $\left(u_{i,t,s|j}\right)$. Moreover, we treat $\varepsilon_{i,j,t,s}$ as a compound error comprising an element reflecting the numerous unmodelled influences on use utility, $\epsilon_{i,j,t,s}^{use}$, and an element reflecting the numerous unmodelled influences on non-use utility, $\epsilon_{i,t,s}^{non-use}$, such that;

$$\varepsilon_{i,j,t,s} = \epsilon_{i,j,t,s}^{use} + \epsilon_{i,t,s}^{non-use} \qquad (j = 1, 2, \dots, J+1 \text{ and } \forall i, t, s) \tag{8}$$

Notice that since individuals derive non-use utility from the J environmental areas independent of their recreation activity, the non-use error component is not dependent on their choice of which site to visit.

Moreover we specify;

$$v_{i,t,s|j}^{use} = \alpha_{j,i,t} + \boldsymbol{q}_{j,t,s}\boldsymbol{\beta}_i \qquad (j = 1, 2, \dots, J \text{ and } \forall\, i, t, s) \tag{9}$$

where $\alpha_{j,i,t}$ is a site-specific utility element and $\boldsymbol{\beta}_i$ is the vector of coefficients describing the marginal use utilities of site qualities. Of course, in any choice period an individual may choose not to make a recreational trip to a natural area. We give that option the index $J+1$, and specify the use utility from choosing that option as;

$$v_{i,t|J+1}^{use} = \alpha_{J+1,i,t} \qquad (\forall i,t) \tag{10}$$

Observe that since this option does not involve visiting one of the J natural areas, the use utility associated with choosing this option does not change across scenarios. We gather the parameters of the use element of individual i's utility into the vector $\boldsymbol{\theta}_{\boldsymbol{i}}^{\boldsymbol{use}} = \left[\alpha_{1,i,t} \dots \alpha_{J+1,i,t}\ \boldsymbol{\beta}_i\right]$;

Our model of the non-use utility element of the preference function is given by;

$$v_{i,t,s}^{non-use} = \sum_{j=1}^{J} d_{i,j}{}^{\lambda_i} \left(a_{j,i,t} + \boldsymbol{q}_{j,t,s}\boldsymbol{b}_{\boldsymbol{i}}\right) \qquad (\forall\ i,t,s) \tag{11}$$

where $d_{i,j}$ is the distance from individual i's home to area j, $a_{j,i,t}$ is an area-specific element contributing to non-use utility, $\boldsymbol{b}_i$ is the vector of coefficients on site qualities and λ_i is a parameter that establishes the rate of distance decay in non-use utility. Notice from (11) that non-use utility is specified as a distance-weighted sum across the non-use utility provided by each individual natural area. The use of summation imposes the assumption that no substitution or complementarity relationships exist between sites in delivering non-use value. The power function used to describe that distance weighting, nests a number of plausible specifications: for example, $\lambda_i = 0$ suggests that non-use utility does not decline with distance, while $\lambda_i = -1$ suggests that the non-use utility declines inversely with distance. Again we use the notation $\boldsymbol{\theta}_{\boldsymbol{i}}^{\boldsymbol{non-use}} = \left[a_{1,i,t} \dots a_{J,i,t}\ \boldsymbol{b}_i\ \lambda_i\right]$ to denote parameters of the non-use element of utility.

Finally we assume a simple linear form for utility from other consumption, such that conditional on travelling to j

$$v_{i,t,s|j}^{other} = \gamma_i\left(y_{t,i} - p_{j,i} - c_s\right) \qquad (j = 1, 2, \dots, J+1 \text{ and } \forall\ i,t,s) \tag{12}$$

We imagine a dataset, like that of the empirical exercise we describe subsequently, in which a sample of respondents provide both RP and SP data. The RP data details the visits each respondent made to the different natural areas over the course of the last year. The SP data is collected from a series of hypothetical choice tasks that, as described earlier, ask respondents to choose between quality-differentiated states of the world. Our objective is to build an econometric model that is derived from the coherent behavioural model described in equations (4) and (6) such that the parameters of the structural equations in equations (8), (9), (10) and (11) can be estimated simultaneously from both RP and SP data.

Our econometric model proceeds through building a likelihood in the manner of the standard RUM. As a result of the error term $\varepsilon_{i,j,t,s}$, probabilistic behavioural equations replace the deterministic choices envisaged by (4) and (6). As such, our econometric model of the probability of observing individual i choosing to visit site j in period t can be written as;

$$\begin{aligned} P_{i,j,t,0}(\boldsymbol{\theta}_{\boldsymbol{i}}^{\boldsymbol{use}},\gamma_i) &= Prob\big[u_{i,t,0|j} > u_{i,t,0|k} \quad \forall\, j \neq k\,\big] \\ &= Prob\big[v_{i,t,0|j}^{use} + v_{i,t,0}^{non-use} + v_{i,t,0|j}^{other} + \epsilon_{i,j,t,0}^{use} + \epsilon_{i,t,0}^{non-use} \\ &> v_{i,t,0|k}^{use} + v_{i,t,0}^{non-use} + v_{i,t,0|k}^{other} + \epsilon_{i,k,t,0}^{use} + \epsilon_{i,t,0}^{non-use} \quad \forall\, k \neq j\,\big] \\ &= Prob\big[v_{i,t,0|k}^{use} + v_{i,t,0|k}^{other} - v_{i,t,0|j}^{use} - v_{i,t,0|j}^{other} > \epsilon_{i,j,t,0}^{use} - \epsilon_{i,k,t,0}^{use} \quad \forall\, k \neq j\,\big] \end{aligned} \tag{13}$$

Since individuals derive the same level of non-use value independent of their choice of which area to visit, the non-use element of modelled utility nets out of line 3 of equation (13).[7] It follows, that the parameters determining values through non-use cannot be estimated from discrete-choice data on recreational behaviour. In a similar vein, the differencing of errors, ensures that elements that relate to unmodelled influences on non-use utility also net out of the errors in equation (13).

Probabilities for responses to the VCSE can be developed in a similar manner. In particular, in a VSCE with M exercises indexed $m = 1,2,\dots,M$, the probability that individual i chooses option s from the choice set $\mathbb{s}_m$ amounts to;

[7] Since income, $y_{t,i}$, remains constant across choice options, the term $\gamma_i y_{t,i}$ from equation (12) also drops out of equation (13) as is true of all RUM applications assuming constant marginal utility of income.

$$
\begin{aligned}
P_{i,s,m}(\boldsymbol{\theta}_{\boldsymbol{i}}^{\boldsymbol{use}}, \boldsymbol{\theta}_{\boldsymbol{i}}^{\boldsymbol{non-use}}, \gamma_i) &= Prob\left[\sum_{t=1}^{T} \max_j\left(u_{i,t,s|j}\right) > \left\{\sum_{t=1}^{T} \max_j\left(u_{i,t,r|j}\right)\right\}_{\forall\, r\neq s}\right] \\
&= Prob\left[\sum_{t=1}^{T} \max_j\left(v_{i,t,s|j} + \varepsilon_{i,j,t,s}\right) > \left\{\sum_{t=1}^{T} \max_j\left(v_{i,t,r|j} + \varepsilon_{i,j,t,r}\right)\right\}_{\forall\, r\neq s}\right] \\
&= Prob\left[\sum_{t=1}^{T} \max_j\left(v_{i,t,s|j}^{use} + v_{i,t,s|j}^{other} + \epsilon_{i,j,t,s}^{use}\right) + v_{i,t,s}^{non-use} + \epsilon_{i,t,s}^{non-use}\right. \\
&\qquad \left. > \left\{\sum_{t=1}^{T} \max_j\left(v_{i,t,r|j}^{use} + v_{i,t,r|j}^{other} + \epsilon_{i,j,t,r}^{use}\right) + v_{i,t,r}^{non-use} + \epsilon_{i,t,r}^{non-use}\right\}_{\forall\, r\neq s}\right]
\end{aligned}
\tag{14}
$$

$$(s, r \in \mathbb{s}_m)$$

where the modelled and non-modelled elements of non-use utility can be taken out of the maximisation problem in the third equation since their magnitudes are, by definition, independent of the choice of recreation activity.

The nature of the probabilities in (13) and (14) are determined by the assumptions the analyst makes regarding the distribution of the error terms. For our purposes, we make the initial assumption both the use error components $\left(\epsilon_{i,j,t,s}^{use}\right)$ and the non-use error components $\left(\epsilon_{i,t,s}^{non-use}\right)$ are IID random variates though potentially following different distributions. Moreover we assume independence between use and non-use error components.

Under those assumptions, the probability of RP choices (13), defines a standard discrete choice TC model in which probabilities can be expressed in terms of differences in IID errors.

A somewhat more difficult econometric challenge is posed by the probability of SP choices (14) which involve sums of maxima across random variables. One of the central contributions of this research is to show how making the assumption that the use utility error terms $\left(\epsilon_{i,j,t,s}^{use}\right)$ are drawn from the family of generalised extreme value (GEV) distributions (McFadden, 1978) enables us to derive closed form solutions to (13) and (14) that facilitate joint estimation of the parameters of use and non-use utility.

To illustrate assume first that the use-utility error components are independent draws from a Type I Extreme Value distribution with location parameter zero and scale parameter σ^{RP} $\left(IID\ EV(0,\sigma^{RP})\right)$. Furthermore let us apply the standard normalisation and set σ^{RP} to a value of 1. Under that assumption, (13) can be solved to give an expression for the probability of observing a particular recreational choice that takes the familiar multinomial logit (MNL) form;

$$P_{i,j,t,0}(\boldsymbol{\theta_i^{use}},\gamma_i) = \frac{e^{v_{i,t,0|j}^{use}+v_{i,t,0|j}^{other}}}{\sum_{k=0}^{J+1} e^{v_{i,t,0|k}^{use}+v_{i,t,0|k}^{other}}} \qquad (\forall i,j,t) \tag{15}$$

To derive an equivalent expression for the probabilities of choice in the VSCE exercise (14) is a more complex challenge. First we need to deal with the expression $\max_j\left(v_{i,t,s|j}^{use} + v_{i,t,s|j}^{other} + \epsilon_{i,j,t,s}^{use}\right)$, which describes the use utility a respondent expects to derive by solving the site-visitation problem and choosing which natural area to visit in time period t under state of the world s. Notice that from the analyst's point of view the presence of the error component $\epsilon_{i,j,t,s}^{use}$ results in this maximum use utility being a random variate.

Of course, in a VCSE respondents choose between states of the world but do not provide details of that anticipated recreational behaviour. Accordingly, it is not possible to simply replace the maximisation expression with the utility of the particular site solving that maximisation problem. One way to proceed, follows from the observation that the set of arguments to the visitation problem are, by assumption, independent Type I Extreme Value variates with equal variance. It follows from properties of that distribution that an individual's maximum use utility in state of the world s must also be an extreme value variate; [8]

$$\max_{j\in 1,..,J+1} v_{i,t,s|j}^{use} + v_{i,t,s|j}^{other} + \epsilon_{i,j,t,s}^{use} \sim EV\left(\ln\sum_{j=1}^{J+1} e^{v_{i,t,s|j}^{use}+v_{i,t,s|j}^{other}}, 1\right) \quad (\forall i,t,s) \tag{16}$$

Accordingly, our specification allows us to write the utility enjoyed by individual i in period t in

[8] Other GEV distributions such as those resulting in the nested and cross-nested logit models have similar properties. To avoid burdening the reader with the notation needed for the expressions describing the maximum of iid variates under those alternative GEV distributions, we do not present those here.

state of the world s as;

$$u_{i,t,s} = \ln \sum_{j=1}^{J+1} e^{v_{i,t,s|j}^{use} + v_{i,t,s|j}^{other}} + \epsilon_{i,t,s}^{use} + v_{i,t,s}^{non-use} + \epsilon_{i,t,s}^{non-use} \qquad (\forall i, t, s) \tag{17}$$

where, as a consequence of (16), the error term $\epsilon_{i,t,s}^{use}$ is a standard Type I Extreme Value variate.

The final error assumption we need to make concerns the distribution of the compound error $\epsilon_{i,t,s}^{use} + \epsilon_{i,t,s}^{non-use}$. Again for the sake of tractability we assume that this is distributed as a Type1 Extreme Value variate with mean zero and scale σ_s^{SP}. Since we have no reason to suspect that the error scales differ across scenarios, we impose the normalisation $\sigma_s^{SP} = \sigma^{SP}$ for all $s = 1, 2, \dots, S$. It follows that (17) can be rewritten as;

$$u_{i,t,s} = \frac{1}{\sigma^{SP}} \ln \sum_{j=1}^{J+1} e^{v_{i,t,s|j}^{use} + v_{i,t,s|j}^{other}} + v_{i,t,s}^{non-use} / \sigma^{SP} + e_{i,t,s} \qquad (\forall i, t, s) \tag{18}$$

where $e_{i,t,s} = \left(\epsilon_{i,t,s}^{use} + \epsilon_{i,t,s}^{non-use} \right) / \sigma^{SP}$ and, by construction, is distributed as an IID standard Type I Extreme Value variate

Of course, the VSCE scenarios are framed as choices made over the duration of one year such that the final step in deriving the econometric specification for the utility derived from a particular choice experiment scenario is to sum over all periods;

$$\begin{aligned} u_{i,s} &= \sum_{t=1}^{T} \frac{1}{\sigma^{SP}} \ln \sum_{j=1}^{J+1} e^{\left(v_{i,t,s|j}^{use} + v_{i,t,s|j}^{other} \right)} + \sum_{t=1}^{T} v_{i,t,s}^{non-use} / \sigma^{SP} + \sum_{t=1}^{T} \varepsilon_{i.t.s} \\ &= v_{i,s} + \sum_{t=1}^{T} \varepsilon_{i,t,s} \qquad (\forall i, s) \end{aligned} \tag{19}$$

In the VSCE we describe subsequently individuals are presented with a series of tasks, $m = 1, 2, \dots, M$ each of which asks them to state a preference over two particular scenarios, s and r, such that the choice set $\mathbb{s}_m$ has only two members. Accordingly, replacing (19) into (14) reveals the probability of observing individual i choosing option s in choice task m, to be;

$$
\begin{aligned}
P_{i,s,m}(\boldsymbol{\theta}_{\boldsymbol{i}}^{\boldsymbol{use}}, \boldsymbol{\theta}_{\boldsymbol{i}}^{\boldsymbol{non-use}}, \gamma_i, \sigma^{SP}) &= Prob\big[u_{i,s} > u_{i,r}\big] \\
&= Prob\left[v_{i,s} + \sum_{t=1}^{T} \varepsilon_{i,t,s} > v_{i,r} + \sum_{t=1}^{T} \varepsilon_{i,t,r}\right] \\
&= Prob\left[v_{i,s} - v_{i,r} > \sum_{t=1}^{T} \big(\varepsilon_{i,t,r} - \varepsilon_{i,t,s}\big)\right] \\
&= Prob\left[v_{i,s} - v_{i,r} > \sum_{t=1}^{T} e_{i,t,m}\right] \qquad (\,\forall i, m \text{ and } s, r \in \mathbb{s}_m)
\end{aligned} \tag{20}
$$

where, from a property of the Type I Extreme Value distribution, $e_{i,t} \sim Logistic(0,1)$. Observe that in differencing the utilities across the two scenarios any additive elements that are constant across scenarios are removed. For that reason, the data provides no means of identifying the area-specific non-use utility elements $a_{m,i,t}$.

To evaluate the probability in (20) we use a result from George and Mudholkar (1983) that shows how, as a convolution of standard logistic variates, the distribution of $\sum_{t=1}^{T} e_{i,t,m}$ can be very closely approximated by Student's t distribution. In particular;

$$
Prob\left[z > \sum_{t=1}^{T} e_{i,t,m}\right] \sim t_{5T+4}\left(0, \pi\left(\frac{15T+12}{5T^2+2T}\right)^{-\frac{1}{2}}\right) \qquad (\forall i, m) \tag{21}
$$

where $t_{5T+4}(\cdot)$ is the cumulative density function of Student's t distribution with $5T + 4$ degrees of freedom.

To complete our econometric specification, we note that our independence assumptions allow us to write the likelihood of observing individual i's recreational visit and SP choices as;

$$
L_i(\boldsymbol{\theta}_i) = \prod_{t}\prod_{j} P_{i,j,t}(\boldsymbol{\theta}_i^{use}, \gamma_i)^{Y_{i,j,t}} \prod_{m}\prod_{s\in\mathbb{s}_m} P_{i,s,m}(\boldsymbol{\theta}_i^{use}, \boldsymbol{\theta}_i^{non-use}, \gamma_i)^{Y_{i,s,m}} \tag{22}
$$

Where $\boldsymbol{\theta}_i = [\boldsymbol{\theta}_i^{use} \quad \boldsymbol{\theta}_i^{non-use} \quad \gamma_i \quad \sigma^{SP}]$ is a vector gathering together all the parameters of the behavioural model. $Y_{i,j,t}$ records visit choices such that $Y_{i,j,t} = 1$ if individual i chose to visit site j

in choice period t and $Y_{i,j,t} = 0$ otherwise. And, $Y_{i,s,m}$ records SP choices where $Y_{i,s,m} = 1$ if individual i chose s from the set of scenarios presented to them in choice task m and $Y_{i,s,m} = 0$ otherwise.

Since our data are not sufficiently rich to allow estimation of a parameter vector $(\boldsymbol{\theta}_i)$ for each respondent, in our empirical application we adopt a random parameters specification (Revelt and Train, 1998; Train, 1998). Accordingly, we assume that each respondent's preference parameters are drawn from the population distribution of preference parameters, $f(\boldsymbol{\theta}|\boldsymbol{\Omega})$ which is specified up to some unknown set of parameters $\boldsymbol{\Omega}$ which must also be estimated from the data. The log likelihood for estimation is given by;

$$\ln L(\boldsymbol{\theta}, \boldsymbol{\Omega}) = \sum_{i=1}^{N} \ln \int L_i(\boldsymbol{\theta})\, f(\boldsymbol{\theta}|\boldsymbol{\Omega})\, d\boldsymbol{\theta} \tag{23}$$

Optimising (23) over the parameters of the model provides maximum likelihood estimates of both use and non-use parameters of the preference function.

5. Monte Carlo Analysis

The estimation strategy described in the previous section offers the possibility of recovering both preference parameters relating to use utility and those relating to non-use utility by combining information from RP data with that from SP data. Of course, using just RP data, the use-value parameters of the utility function could be estimated using (15). A key concern in understanding the limitations of the proposed estimation strategy is discovering the circumstances under which the SP data can provide information that (i) contributes to the improved estimation of the use utility parameters and (ii) allows for estimation of the non-use utility parameters. We investigate those questions through a series of Monte Carlo experiments.

To facilitate the large number of estimations needed to complete a Monte Carlo analysis, we simplify the utility function specification. In particular, we assume that in each period there is only one alternative to taking a recreational trip to a river site which gives each individual the same utility α_{J+1}. Likewise we assume that the site-specific utility element for each river site is zero for all individuals in each period that is $\alpha_{j,i,t} = 0\ \forall j, i, t$. Finally we assume that the other parameters of the model are the same for each simulated individual (that is; $\boldsymbol{\beta}_i = \boldsymbol{\beta}_0, \boldsymbol{b}_i = \boldsymbol{b}_0, \gamma_i = \gamma_0, \lambda_i = \lambda_0,\ \forall i$). Indeed, the utility function used in the Monte Carlo analysis is;

$$u_{i,t,s|j} = \boldsymbol{q}_{j,s}\boldsymbol{\beta}_0 + \sum_{k=1}^{J} d_{i,k}{}^{\lambda_0}\, \boldsymbol{q}_{k,s}\boldsymbol{b}_0 + \gamma_0\left(y_i - p_{k,i} - c_s\right) + \varepsilon_{i,j,t,s} \qquad (j = 1, \dots, J \text{ and } \forall i, t, s) \tag{24}$$

$$u_{i,t,s|J+1} = \alpha_{J+1} + \sum_{k=1}^{J} d_{i,k}{}^{\lambda_0}\, \boldsymbol{q}_{k,s}\boldsymbol{b}_0 + \gamma_0\left(y_i - c_s\right) + \varepsilon_{i,J+1,t,s} \qquad (\forall i, t, s)$$

Observe from (22) that natural area qualities $\left(\boldsymbol{q}_{j,s}\right)$ and income (y_i) are also taken to be constant across periods.

As shown in Figure 1, to execute the Monte Carlo analysis we construct a hypothetical landscape that mimics that of the empirical problem we analyse subsequently. Our hypothetical landscape is a unit square traversed by three rivers within which we randomly locate the residences of a sample of 500 individuals (Panel A of Figure 1). For the purposes of evaluating non-use utility, each river is divided up into 27 stretches of equal length (Panel B of Figure 1). Likewise, 120 recreational sites are located at random along the rivers. In the simulation, river site attributes are limited to ecological status which ranges from bad through to excellent and these are ascribed to sites using a random walk procedure that ensures spatial correlation in water quality (Panel C of Figure 1).

[INSERT FIGURE 1 ABOUT HERE]

We assume that there are 50 choice occasions in one year ($T = 50$) and simulate utilities for each recreational option in each period using (22) with particular choices of parameters, a travel cost calculated from the straight line distance from residence to site and random draws of the error term. An RP dataset is simulated by taking the recreational choice in each period to be the option that provides the highest utility. That choice may change from period to period for one simulated individual since new error terms are drawn each time utilities are evaluated.

To simulate an SP dataset, we created a very simple choice experiment design in which the rivers are divided up into nine lengths of equal length. As shown in Panel B of Figure 6 each of these lengths comprises nine river stretches. A scenario was created by randomly attributing a water quality to each river length and associating with that scenario a randomly determined cost. One hundred and twenty such scenarios were generated, paired-off so as to create choice tasks and then randomly ascribed into five blocks of twelve tasks. Each simulated individual was assigned to one of the five blocks and their choices in the choice experiment determined by selecting the option providing the highest utility from realisations of the utility function in (24).

Our first experiment explores the contribution that the SP data can make to the identification of use utility parameters. Recall from (21) that those parameters enter our expression for the utility of a choice option in the VSCE within the log-sum-exp term, $\ln\left(\sum_{j=1}^{J} e^{\left(v_{i,t,s|j}^{use}+v_{i,t,s|j}^{other}\right)} + e^{\left(\alpha_{J+1}+v_{i,t,s|J+1}^{other}\right)}\right)$. For each choice option the qualities of the recreation sites change and those changes are manifested as changes in the $v_{i,t,s|j}^{use}$elements of the log-sum-exp term. Of course that log-sum-exp term also includes an element reflecting the utility of the outside option, α_{J+1}, and this element remains constant across all choice options.

Now consider what happens to the variability manifested in the magnitude of the log-sum-exp term as a result of changes in environmental qualities across different VSCE choice options. In particular, when the utility of the outside option is relatively small, we might expect to see relatively large responses in the log-sum-exp term and hence considerable information from which to aid identification of the use-utility parameters. In contrast, as a result of the log operation, when the utility of the outside option is relatively large, the change in magnitude of the log-exp-sum from the same quality changes will be relatively small. Indeed as the size of outside good utility increases the variation in the log-sum-exp term resulting from different VSCE choice options diminishes and the less information that choice experiment can provide regarding the values of the use-utility parameters.

Our first Monte Carlo analysis investigates this identification issue by comparing the distribution of estimates of the site quality parameters of use utility ($\boldsymbol{\beta}_0$) over three experiments where the outside good utility parameter $\left(\alpha_{J+1}\right)$ is fixed at the values of 1, 3, and 5 respectively.

While the SP data might contribute to the identification of the site quality parameters, it is still the case that those parameters can be estimated directly form the RP data. Accordingly, for the purposes of our Monte Carlo experiment we generate a more demanding test by introducing an additional parameter into the use utility specification that can only be estimated from the SP data. In particular, we imagine that individuals responding to the VSCE may over-estimate their propensity to take trips to environmental areas. Accordingly, we introduce the parameter α_{J+1}^{SP} and set its value to -0.5. In the analysis of the simulated responses to VSCE choice tasks, α_{J+1}^{SP} is added to α_{J+1} thereby reducing the utility ascribed to choosing the outside option in determining choices in the VSCE.

[INSERT TABLE I ABOUT HERE]

The results of that Monte Carlo experiment are presented in Table 1. Notice that drawing on information provided by the RP data estimates of the river quality parameters for use utility (β_1, β_2 and β_3) are estimated without bias. Notice, however, that as the size of the utility of the outside good option increases the standard deviation of the Monte Carlo estimates of those parameters also increases. That pattern meets our expectations thatSP data provides less information on use utility parameters as the size of the outside option utility increases. Figure 2 plots out the distribution of β_1 estimates from each Monte Carlo experiment providing visual affirmation of the reduced informational content of the data at larger values for the outside option utility.

[INSERT FIGURE 2 ABOUT HERE]

With respect to the hypothetical bias parameter α_{J+1}^{SP} we see a similar, though more extreme response to increasing the utility offered by the outside good option. When that utility element is small the hypothetical bias parameter can be successfully recovered solely from the SP data. At the medium level of outside option utility, the parameter is still estimated without appreciable bias, though with a greater degree of variability. Indeed, at the high level of outside option utility the parameter is estimated with such imprecision that the mean value of the parameter over 1000 Monte replicates (1.884) is a long way from the true value (-0.5). Figure 3 plots out the distribution of α_{J+1}^{SP} estimates for each Monte Carlo experiment. The Figure clearly demonstrates the increasing difficulty of identifying the hypothetical bias parameter as an increasingly large utility for the outside good diminishes the informational content of the SP data.

[INSERT FIGURE 3 ABOUT HERE]

Our second key concern with the combined-data estimation strategy concerns the model's ability to return estimates of non-use utility parameters. Put crudely, the estimation strategy for disentangling non-use values is to observe the degree to which choices in the VSCE differ from the choices that would be expected if determined solely by use value. Of course, the degree to which choices will be influenced by non-use considerations will depend on the relative size of the non-use component of utility to the use component of utility. If the non-use component is relatively large, then we would imagine that the SP data from the choice experiment would give good identification of the non-use parameters. In contrast, if the non-use component is relatively small, then identification may be difficult.

Three different sets of parameters were chosen for the Monte Carlo analysis which differed only in the size of the preference parameters determining the non-use value derived from water qualities. To determine the size of those parameters we selected one scenario at random and evaluated the average gain in welfare that would be realised by the simulated individuals if all rivers were improved up to the excellent water quality level.[9] As shown in Table I, we chose values for the non-use parameters that resulted in the non-use element of this average welfare gain being twice as large as the use element ("Large"), roughly equal to the use element ("Equal") and half the size of the use element ("Small").

[INSERT TABLE II ABOUT HERE]

A summary of the outcomes from the Monte Carlo analysis using 1,000 simulations is presented in Table III. Observe first that, for all three treatments, the use parameters of the utility function are estimated with almost no bias and with high precision. As expected, the RP data provides a good source of identification for those parameters.

[INSERT TABLE III ABOUT HERE]

In the case of the 'Large' treatment where non-use is a major component of utility, the estimator fares very well; the non-use parameters are estimated with almost no bias and a

[9] The randomly chosen scenario had 2 river lengths of excellent quality, four of good quality, three of poor quality but no bad quality river lengths.

reasonable level of precision. Moving from the 'Large' to the 'Equal' to the 'Small' treatments, however, introduces increasing bias into the non-use parameters. The most significant bias is in the distance-decay parameter (λ) whose mean value from the simulations significantly increases in absolute value. To better understand that finding, Figure 4 provides a density plot of the λ values estimated in each Monte Carlo treatment.

[INSERT FIGURE 4 ABOUT HERE]

It is apparent from Figure 4, that for all three treatments the distributions are centred on the true value of -1. What is also apparent from Figure 4 is the fact that as the relative size of the non-use utility element declines the precision with which the distance-decay parameter is estimated also declines. Indeed, a more detailed examination of the data shows that while in the 'Large' treatment the distance-decay parameters estimated in the simulations fall within the range -0.770 to -1.163, the 'Equal' and 'Small' treatments are characterised by a significant extension to the left hand tail of the distribution. Indeed in the 'Equal' treatment 88% of the parameter estimates lie in the range -0.368 to -1.325 the remaining 12% span the range -12.539 to -180.630. Likewise in the 'Small' treatment 19% of the estimates fall in the range -10.150 to -2,651.6. It appears that when the contribution of non-use is relatively small, the estimator has difficulties identifying the non-use parameters for some realisations of the data. In those cases, we can assume that the non-use elements of utility make little difference to the choices made in the choice experiment. Accordingly, the estimator tends towards a distance-decay parameter that, in effect, discounts non-use utility away to practically zero.

Interestingly, the difficulties in estimating the non-use components of utility do not appear to carry over into the estimation of welfare impacts. Observe from the second to last row of Table III that the average welfare gain from improving all rivers to excellent quality from our selected baseline scenario is estimated with relatively little bias. Figure 5, confirms that finding showing that the distributions of welfare estimates from the simulations are centred around their true values with a similar level of precision being realised in all three treatments.

[INSERT FIGURE 5 ABOUT HERE]

Now imagine that rather than the structural model combining RP and SP, an analyst had decided to use an approach that employed only the data from a VSCE. To maintain relative comparability, suppose that the approach adopted was to estimate a utility function that assumed the value derived from a choice experiment scenario could be approximated as the weighted sum of river qualities in that scenario according to;

$$u_{i,t,s} = \sum_{m=1}^{M} {d_{i,m}}^{\lambda_0}\, \boldsymbol{x}_{m,s}\boldsymbol{b_0} + \gamma_0(I_i - p_s/T) + \varepsilon_{i,j,t,s} \qquad (\forall i,t,s) \tag{25}$$

Notice that (23) is simply our specification of the non-use elements of the utility function. In applying (23) to data in which choices are made according to both use and non-use considerations, however, we might expect the water quality parameters, $\boldsymbol{b_0}$, to pick up the combined effect of water quality on both use and non-use. Likewise, the distance-decay parameter, λ_0, will pick up not only the effect of distance on non-use utility but also the effect of travel costs on use utility. The question we wish to answer is whether (23) provides a sufficiently close approximation to the full structural model (22) that the added complexity of estimating the full structural model might be considered an unnecessary luxury. Table IV summarizes a Monte Carlo experiment used to explore that question.

[INSERT TABLE IV ABOUT HERE]

The Monte Carlo analyses reported in Table IV use exactly the same simulations as those underpinning Table III. Observe that for all three treatments, the parameters on the quality variables and the distance decay show the bias that would be expected if the parameters were picking up elements relating to use utility as well as to non-use utility. More importantly observe the welfare estimates shown in the final row of Table IV. For each treatment these welfare estimates are biased upwards with the size of bias in terms of standard deviations from the true value increasing from 1.31 for the ‘Large’ treatment, to 1.89 for the ‘Equal’ treatment, to 2.29 for the ‘Small’ treatment. We conclude that relying on approximating reduced-form equations may result in significant errors in the calculation of welfare effects. Our Monte Carlo analysis lends support to the idea that, if possible, econometric models should be carefully constructed to reflect the underlying structure of the data-generating process.

6. Empirical Case Study

The data which motivated the research reported in this paper came from a large scale valuation exercise carried out in northern England as part of the EU Aquamoney project. The focus of the exercise was to establish the value of improvements in the ecological status of rivers. The study area and the rivers upon which the survey focussed are shown in Figure 6.

[INSERT FIGURE 6 ABOUT HERE]

Using computer-aided survey methods data was collected from a large sample of residents in the region. Using interactive maps, respondents first located their home residence and then the river sites they had visited over the last 12 months. Amongst other things, data was collected on how frequently they had visited each of those sites as well as the number of outdoor recreational trips they had taken to non-river locations. In the second part of the survey, respondents participated in a VSCE. To construct the choice experiment the major rivers in the region were divided into nine river lengths of equal extent. To construct a scenario to describe a future possible state of the world, each river length was ascribed one of four levels of ecological status; bad, poor, good or excellent, with those statuses being represented visually on a map of the rivers by the colours red, yellow, green and blue respectively. The meaning of the four ecological statuses was explained to respondents in detail using the procedure described in Hime et al. (2009). Each scenario was associated with a cost and a choice task constructed by pairing two scenarios, as illustrated in Figure 7. Using a fractional factorial design, 60 choice tasks were constructed which were divided into five blocks of 12 tasks. Each respondent, therefore, provided answers to 12 choice tasks.

[INSERT FIGURE 7 ABOUT HERE]

An important source of identification in estimating the parameters of the utility function is to observe the choices of individuals at different distances from river locations. To that end, surveying locations were chosen that evenly sampled the spatial extent of the study area. Between 40 and 100 interviews were conducted in each sampling location giving a total sample size of 1,805. Some of those observations were incomplete or lacked crucial information such that the final dataset consisted of recreational choice data for 1,794 individuals with 1,708 of those individuals also providing a complete set of responses to the choice experiment.

The recreational river sites available in the study region were identified using a GIS to find locations where the river could be accessed either by walking or driving and confirming those locations using aerial photographs. In total, 531 recreational sites were identified along the study rivers ($J = 531$). Their locations are shown in Figure 1. Information on the environmental characteristics of the recreational sites was identified in the GIS using Ordnance Survey's MasterMap and the Centre for Ecology and Hydrology's LandCover Map 2007. These provided details of the predominant land use around each of the recreational sites, which were grouped into four broad categories including woodland, farmland, grassland, and urban. The current water quality at each of the recreational sites was calculated from Environment Agency long-term water quality monitoring data and categorised on the four-point ecological status scale.

Of the 531 recreational sites, 286 had been visited by the sample of respondents. Of those respondents 18% made no recreational trips to a river site in the previous year, 27% made 1 to 5 trips to a river site, 12% made 6 to 10 trips, with the remaining 33% making more than 10 trips a year. Since some respondents visited a river site every day in the year, the recreational choice period was established as one day giving $T = 365$. Travel costs to recreational sites $\left(tc_{i,j}\right)$ were calculated in the GIS with the cost of time valued at a quarter of the imputed wage rate.

In order to evaluate non-use utility the 9 river lengths defined for the purposes of the choice experiment were each further divided into 9 river stretches giving a total of 81 river stretches ($M = 81$) each of which was a little under 3km in length. Distance to each river stretch from each respondent's home $\left(d_{i,m}\right)$ was measured in the GIS

7. Results

For the purposes of estimating the parameters of the model outlined in Sections 3 and 4 with real world data, we make a number of simplifying assumptions. First, our data do not record changes in river qualities over time, accordingly we assume that those qualities remain constant over the period of one year. Accordingly, our vector of river qualities in the use utility element of the utility function, $\boldsymbol{x}_{j,s}$, consists of a set of dummy variables capturing ecological status (with bad status being the baseline), a set of dummy variables capturing the predominant land use at the site (with farmland being the baseline) and a variable measuring population density in the local area. Likewise, our vector of river qualities in the non-use utility element of the utility function, $\boldsymbol{x}_{m,s}$, consists only of a set of dummy variables indicating the ecological status of the river stretch. Notice that all other features of rivers are assumed to stay constant across scenarios and hence cancel out by differencing in the estimating equations.

In addition, we capture heterogeneity in tastes across the sample by adopting a random parameters specification. In particular, we assume that the marginal utility of money parameter, the distance-decay parameter and the utility of the no trip option are draws from a normal distribution; specifically, $\gamma_i \sim N(\gamma, \sigma_\gamma^2)$, $\lambda_i \sim N(\lambda, \sigma_\lambda^2)$ and $\alpha_{J+1,i} \sim N\left(\alpha_{J+1}, \sigma_{\alpha_{J+1}}^2\right)$. In contrast, we constrain the utility of the other outdoor recreation trip and the taste parameters on river quality attributes for both use and non-use to be constant across individuals; that is to say, $\alpha_{J+1,i} = \alpha_{J+1}$ $(\forall i)$, $\boldsymbol{\beta}_i = \boldsymbol{\beta}$ $(\forall i)$ and $\boldsymbol{b}_i = \boldsymbol{b}$ $(\forall i)$. Similarly, we constrain the parameters on the site-specific element of use utility to be constant across individuals, but allow for unobserved differences in quality across sites by allowing those elements to be draws from a normal distribution; $\alpha_{j,i,t} = \alpha_j \sim N(\alpha, \sigma_\alpha^2)$ $(\forall j, i, t)$.

Since the model contains random parameters, we use simulated maximum likelihood to recover parameter estimates. Those estimates are reported in Table V along with those of a standard discrete choice TC model estimated from just the RP data using an identical specification of the utility function. To permit direct comparison of the parameters across those models we impose the same scaling normalisation on both models, $\sigma_{RP} = 1$, and allow a separate parameter, σ_{SP}, to capture differences in the scale of the error terms between the RP and SP choices.

[INSERT TABLE V ABOUT HERE]

Notice that the outside good utility parameter is large, a finding that our Monte Carlo analysis suggests will limit the contribution that the SP data can make to the estimation of use-utility parameters. Indeed, the estimated parameters support that contention; we observe that the use-utility parameters from the combined data model closely resemble those from the ‘travel cost’ model estimated from just the RP data. Moreover, we attempted to estimate a model with a hypothetical bias parameter but as per the Monte Carlo experiment found that parameter to be unidentified in this case.

Second, notice that at its mean the non-use distance decay parameter takes a value of −1.06. Again drawing on the evidence provided in the Monte Carlo experiment the relatively modest size of that parameter and its high level of significance provide some reassurance that the non-use parameters are identified from the SP data.

The parameters of the models are all plausibly signed and in the main statistically significant at the usual levels of confidence. The notable exceptions are the parameters on poor and good ecological status in the use element of the utility function which are found to be insignificantly different from the baseline case of bad ecological status. On the other hand, the parameter on excellent ecological status indicates that individuals obtain a significant utility dividend from using rivers that are at the highest ecological status. It appears that in their use of rivers for recreational activities, individuals only differentiate sites on the basis of whether or not they offer that highest level of ecological status.

In contrast, consider the parameters on the ecological status of rivers for the non-use element of utility. Each of these parameters are statistically significant with a very high level of confidence and progress in a natural order with excellent status being preferred to good status being preferred to poor status being preferred to bad status. One conclusion that might be drawn from these findings is that individuals gain non-use utility from any improvement in the ecological status of a river, but in using those rivers the only change that significantly impacts on their utility is the achievement of the very highest level of ecological status.

Finally, the distance-decay parameter on non-use utility is significantly different from zero indicating that the non-use utility an individual enjoys from a river declines with the distance that river is from their home. Indeed, the parameter value of −1.06 suggests the rate of decline to be approximately equal to the inverse of distance.

Taking the estimated parameters from the combined model, it is possible to carry out a welfare analysis that differentiates between changes in use and non-use value. To do that, we explore the average welfare gains that would be realised in our sample if all rivers in the region were improved from their current ecological status to excellent ecological status. We find that the average annual welfare gain from that change is £12.04 (std. dev. £7.47), of which £10.45 (std. dev. £5.92) is derived in non-use utility and £1.59 (std. dev. £1.75) in use utility. Accordingly, our data suggest that the values of improvements in ecological status are relatively small in magnitude and come mainly from changes in non-use utility.

8. Concluding Remarks

The central contribution of this paper is to build an econometric specification for the analysis of VSCE data that is derived from a coherent structural model of preferences for landscape-wide environmental quality change. The functional form of the econometric specification of the preference function describing choice behaviour in a VSCE is highly nonlinear (see equation 16) which stands in stark contrast to standard practice in the field of choice modelling. Indeed, analysts have come to increasingly depend on linear specifications of preference functions that are amenable to estimation using the mixed logit model. Often that modelling choice is justified through appeal to the results of McFadden and Train (2000) who show that any RUM can be approximated to any degree of accuracy by a mixed logit with appropriate choice of variables and mixing distribution. That reliance on linear preference functions has recently been questioned by (Andersen et al., 2012) and this paper lends weight to that criticism. In particular, our Monte Carlo analysis shows how a reduced form model (admittedly without random coefficients) fails to accurately predict welfare changes. Indeed, our research suggests that there may be good reasons to be suspicious of welfare calculations emanating from models using reduced-form specifications of the preference function especially when, as in the case studied in this paper, there is good reason to believe that the true preference function is highly non-linear.

In the context of our study, another important justification for basing our econometric specification on a structurally coherent model of preferences results from our use of both RP and SP data in estimation. Evidently, any theoretically consistent attempt at joint estimation must clearly identify which parameters of the preference function are informed on by the two different data sources (Eom and Larson, 2006). In our case, the behavioural data reflect just on the parameters determining use value and the hypothetical choice data on those determining both use and non-use value. Those differences fall naturally out of our derivation of the econometric models for the two different forms of data from the same structural description of preferences.

A final justification for the importance of structural modelling results from the requirement for benefits transfer. The use of reduced-form specifications that confound use and non-use values or inadequately describes substitution relationships inhibit effective transfer of the value estimates outside the study area to locations exhibiting different spatial patterns of quality-differentiated substitutes.

With regards to the findings of the empirical exercise, a number of results stand out. First, while it has long been established that utility from the use of a natural resource declines with distance from an individual's home our research provides evidence to show that the same is true of utility from non-use. Indeed, we find that non-use values for the ecological quality of rivers decline at a rate approximately equal to the inverse of distance. Our research, therefore, supports the speculation of Bateman et al. (2006) that there may be a cultural identity or 'ownership' dimension to non-use values that precipitates distance decay in those values. Those speculations were based on the empirical findings of distance decay in the expressions of value made by non-users of a resource (Bateman et al., 2005; Hanley et al., 2003). As far as we are aware, our empirical findings are the first to identify distance decay in non-use values themselves.

Our empirical application also reveals that value flows from river quality attributes differ in use and non-use. In particular, our models suggest that incremental improvements in the ecological status of rivers are associated with increasing non-use utilities, but only achievement of the highest ecological status provides meaningful increases in use utility. Those findings may reflect the fact that relatively little of the recreational activity associated with rivers in our study region involves physical contact with the water. One might speculate that respondents derive non-use value from the knowledge that increments in river quality deliver improvements in ecological functioning even if those increments result in little meaningful change in their experience of visiting a river until the highest level of ecological quality is achieved.

Finally, our empirical estimates suggest that non-use utility may be a significant component of the welfare gains that arise from improving the ecological status of rivers.

Accordingly, ignoring non-use values may significantly understate the welfare gains that might arise from landscape-wide programmes of river quality improvement.

References

Adamowicz, W., Louviere, J., Williams, M., 1994. Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities. Journal of Environmental Economics and Management 26, 271-292.

Andersen, S., Harrison, G.W., Hole, A.R., Lau, M., Rutström, E.E., 2012. Non-linear mixed logit. Theory and Decision 73, 77-96.

Bateman, I.J., Cooper, P., Georgiou, S., Navrud, S., Poe, G.L., Ready, R.C., Riera, P., Ryan, M., Vossler, C.A., 2005. Economic valuation of policies for managing acidity in remote mountain lakes: Examining validity through scope sensitivity testing. Aquatic Sciences 67, 274-291.

Bateman, I.J., Day, B.H., Georgiou, S., Lake, I., 2006. The aggregation of environmental benefit values: Welfare measures, distance decay and total WTP. Ecological Economics 60, 450-460.

Birol, E., Karousakis, K., Koundouri, P., 2006. Using a choice experiment to account for preference heterogeneity in wetland attributes: The case of Cheimaditida wetland in Greece. Ecological Economics 60, 145-156.

Boxall, P.C., Adamowicz, W.L., Swait, J., Williams, M., Louviere, J., 1996. A comparison of stated preference methods for environmental valuation. Ecological Economics 18, 243-253.

Brouwer, R., Martin-Ortega, J., Berbel, J., 2010. Spatial preference heterogeneity: a choice experiment. Land Economics 86, 552-568.

Cameron, T.A., 1992. Combining Contingent Valuation and Travel Cost Data for the Valuation of Nonmarket Goods. Land Economics 68, 302-317.

Caudill, S.B., Groothuis, P.A., Whitehead, J.C., 2011. The Development and Estimation of a Latent Choice Multinomial Logit Model with Application to Contingent Valuation. American Journal of Agricultural Economics 93, 983-992.

Christie, M., Hanley, N., Hynes, S., 2007. Valuing enhancements to forest recreation using choice experiment and contingent behaviour methods. Journal of Forest Economics 13, 75-102.

Concu, G.B., 2007. Investigating distance effects on environmental values: a choice modelling approach*. Australian Journal of Agricultural and Resource Economics 51, 175-194.

Cummings, R.G., Harrison, G.W., 1995. The measurement and decomposition of nonuse values: A critical review. Environmental and Resource Economics 5, 225-247.

Davis, R.K., Knetsch, J., 1966. Comparisons of Methods for Recreation Evaluation, in: Kneese, A., Smith, S. (Eds.), Water Research. Johns Hopkins University Press, Baltimore, MD.

De Valck, J., Broekx, S., Liekens, I., Aertsens, J., Vranken, L., 2017. Testing the Influence of Substitute Sites in Nature Valuation by Using Spatial Discounting Factors. Environmental and Resource Economics 66, 17-43.

Deaton, A., Muellbauer, J., 1980. Economics and consumer behavior. Cambridge university press.

Doherty, E., Murphy, G., Hynes, S., Buckley, C., 2014. Valuing ecosystem services across water bodies: Results from a discrete choice experiment. Ecosystem Services 7, 89-97.

Englin, J., Cameron, T.A., 1996. Augmenting travel cost models with contingent behavior data.

Environmental and Resource Economics 7, 133-147.

Eom, Y.-S., Larson, D.M., 2006. Improving environmental valuation estimates through consistent use of revealed and stated preference information. Journal of Environmental Economics and Management 52, 501-516.

Freeman, A.M., 2003. The measurement of environmental and resource values: Theory and methods. Resources for the Future, Washington, DC.

George, E., Mudholkar, G., 1983. On the convolution of logistic random variables. Metrika 30, 1-13.

Hanley, N., Schläpfer, F., Spurgeon, J., 2003. Aggregating the benefits of environmental improvements: distance-decay functions for use and non-use values. Journal of Environmental Management 68, 297-304.

Hanley, N., Wright, R.E., Adamowicz, V., 1998. Using Choice Experiments to Value the Environment. Environmental and Resource Economics 11, 413-428.

Hime, S., Bateman, I.J., Posen, P., Hutchins, M., 2009. A transferable water quality ladder for conveying use and ecological information within public surveys. CSERGE working paper EDM.

Horne, P., Boxall, P.C., Adamowicz, W.L., 2005. Multiple-use management of forest recreation sites: a spatially explicit choice experiment. Forest Ecology and Management 207, 189-199.

Jørgensen, S.L., Olsen, S.B., Ladenburg, J., Martinsen, L., Svenningsen, S.R., Hasler, B., 2013. Spatially induced disparities in users' and non-users' WTP for water quality improvements—Testing the effect of multiple substitutes and distance decay. Ecological Economics 92, 58-66.

Kataria, M., Bateman, I., Christensen, T., Dubgaard, A., Hasler, B., Hime, S., Ladenburg, J., Levin, G., Martinsen, L., Nissen, C., 2012. Scenario realism and welfare estimates in choice experiments – A non-market valuation study on the European water framework directive. Journal of Environmental Management 94, 25-33.

Mäler, K.-G., Gren, I.-M., Folke, C., 1994. Multiple use of environmental resources: A household production function approach to valuing natural capital, in: Janson, A.M., Hammer, M., Folke, C., Costanza, R. (Eds.), Investing in Natural Capital: The Ecological Economics Approach to Sustainability. Island Press, Washington, D.C. and Covelo, CA, pp. 233-249.

McFadden, D., 1973. Conditional Logit Analysis of Qualitative Choice Behavior, in: Zarembka, P. (Ed.), Frontiers of Econometrics. Academic Press, New York.

McFadden, D., 1978. Modelling the choice of residential location, in: Karlqvist, A., Lundqvist, L., Snickars, F., Weibull, J. (Eds.), Spatial Interaction Theory and Planning Models. Amsterdam, Amsterdam, pp. 75-96.

McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. Journal of Applied Econometrics 15, 447-470.

Meyerhoff, J., Boeri, M., Hartje, V., 2014. The value of water quality improvements in the region Berlin–Brandenburg as a function of distance and state residency. Water Resources and Economics 5, 49-66.

Morey, E.R., Rowe, R.D., Watson, M., 1993. A Repeated Nested-Logit Model of Atlantic Salmon Fishing. American Journal of Agricultural Economics 75, 578-592.

Oster, S., 1977. Survey results on the benefits of water pollution abatement in the Merrimack River Basin. Water Resources Research 13, 882-884.

Pate, J., Loomis, J., 1997. The effect of distance on willingness to pay values: a case study of wetlands and salmon in California. Ecological Economics 20, 199-207.

Phaneuf, D.J., von Haefen, R.H., 2009. Chapter 12 Estimating the Demand for Quality with Discrete Choice Models, Quantifying Consumer Preferences. Emerald Group Publishing Limited, pp. 317-347.

Revelt, D., Train, K., 1998. Mixed logit with repeated choices: households' choices of appliance efficiency level. Review of economics and statistics 80, 647-657.

Schaafsma, M., Brouwer, R., 2013. Testing geographical framing and substitution effects in spatial choice experiments. Journal of Choice Modelling 8, 32-48.

Schaafsma, M., Brouwer, R., Rose, J., 2012. Directional heterogeneity in WTP models for environmental valuation. Ecological Economics 79, 21-31.

Söderberg, M., Barton, D.N., 2014. Marginal WTP and Distance Decay: The Role of ‘Protest’ and ‘True Zero’ Responses in the Economic Valuation of Recreational Water Quality. Environmental and Resource Economics 59, 389-405.

Stithou, M., Hynes, S., Hanley, N., Campbell, D., 2013. Estimating the Value of Achieving “Good Ecological Status” in the Boyne River Catchment in Ireland Using Choice Experiments. The Economic and Social Review; Vol 43, No 3, Autumn (2012).

Sutherland, R.J., Walsh, R.G., 1985. Effect of Distance on the Preservation Value of Water Quality. Land Economics 61, 281-291.

Taylor, T., Longo, A., 2010. Valuing algal bloom in the Black Sea Coast of Bulgaria: A choice experiments approach. Journal of Environmental Management 91, 1963-1971.

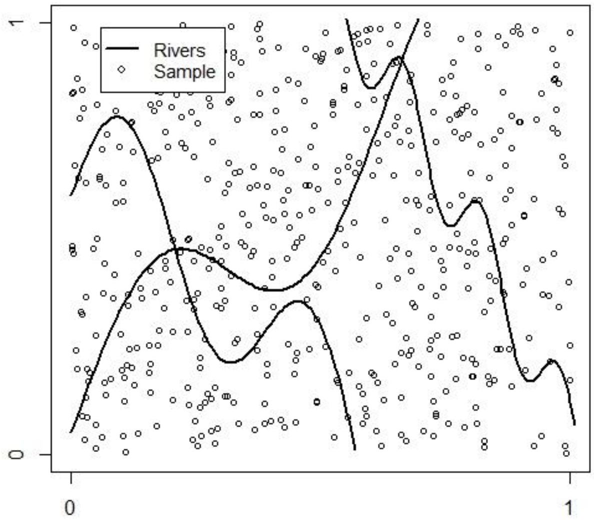
Timmins, C., Schlenker, W., 2009. Reduced-form versus structural modeling in environmental and resource economics. Annual Review of Resource Economics 1, 351-380.

Train, K.E., 1998. Recreation Demand Models with Taste Differences over People. Land Economics 74, 230-239.

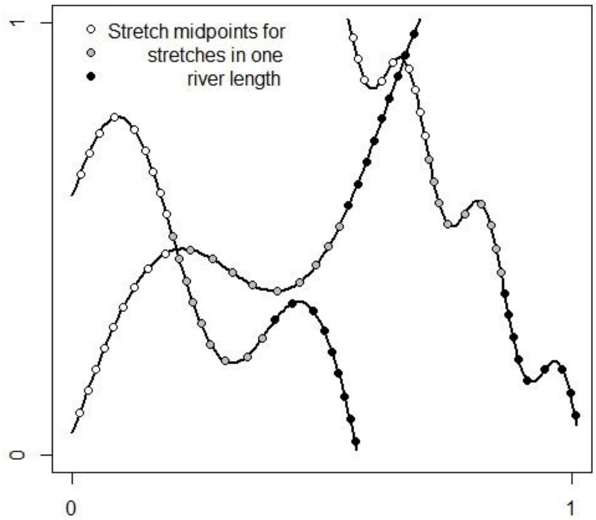
Whitehead, J.C., Haab, T.C., Huang, J.-C., 2000. Measuring recreation benefits of quality improvements with revealed and stated behavior data. Resource and Energy Economics 22, 339-354.

Whitehead, J.C., Pattanayak, S.K., Van Houtven, G.L., Gelso, B.R., 2008. Combining revealed and stated preference data to estimate the nonmarket value of ecological services: an assessment of the state of the science. Journal of Economic Surveys 22, 872-908.

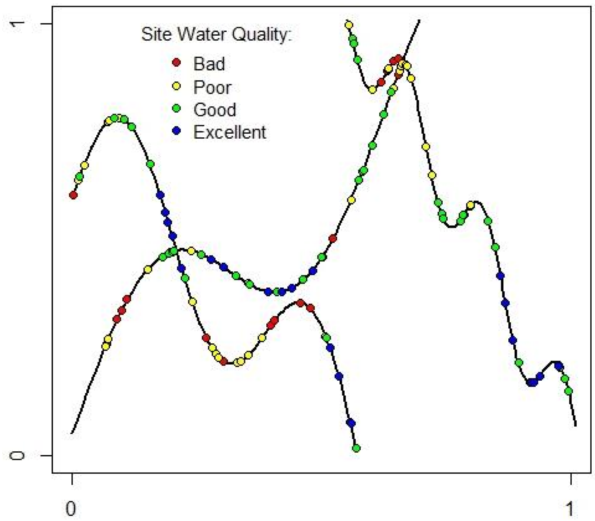
Yao, R.T., Scarpa, R., Turner, J.A., Barnard, T.D., Rose, J.M., Palma, J.H.N., Harrison, D.R., 2014. Valuing biodiversity enhancement in New Zealand's planted forests: Socioeconomic and spatial determinants of willingness-to-pay. Ecological Economics 98, 90-101.



Panel A: Rivers & Sample Locations



Panel B: Midpoints of Stretches



Panel C: Recreational Sites

Figure 1: Rivers, Stretches and Sites in the Monte Carlo Analysis

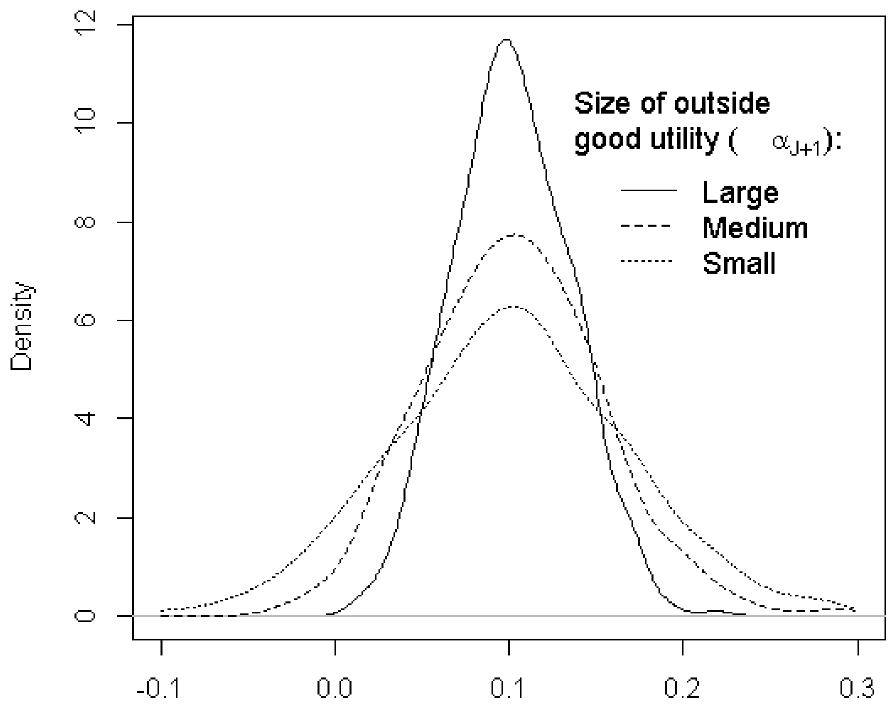


Figure 3: Distribution of poor water quality parameter $(\boldsymbol{\beta_1^0})$ from Monte Carlo experiments differing in the size of the outside option utility parameter $(\boldsymbol{\alpha})$

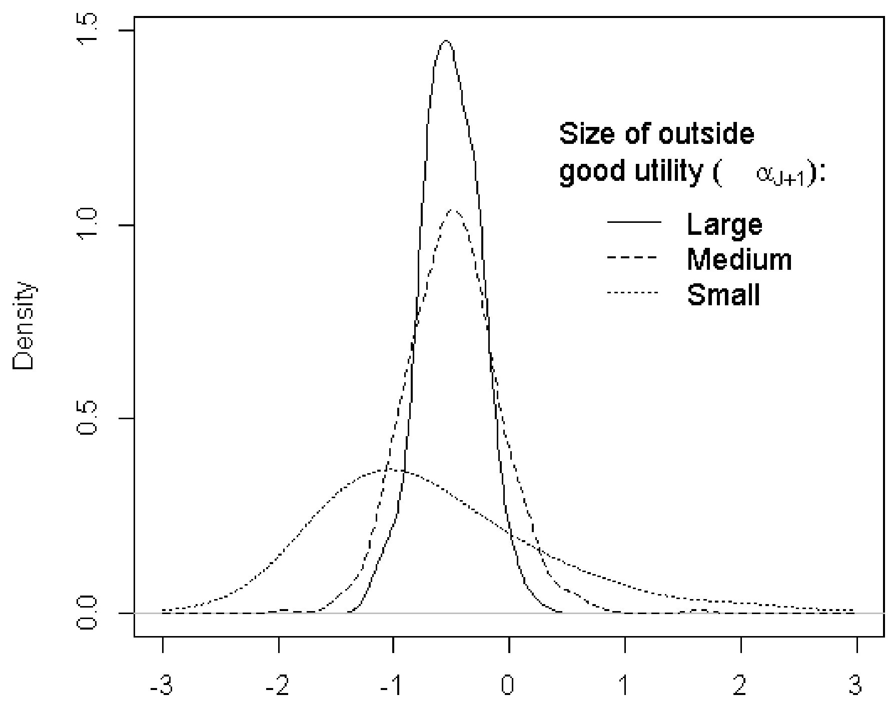


Figure 2: Distribution of hypothetical bias parameter $(\boldsymbol{\alpha^{SP}})$ from Monte Carlo experiments differing in the size of the outside option utility parameter $(\boldsymbol{\alpha})$

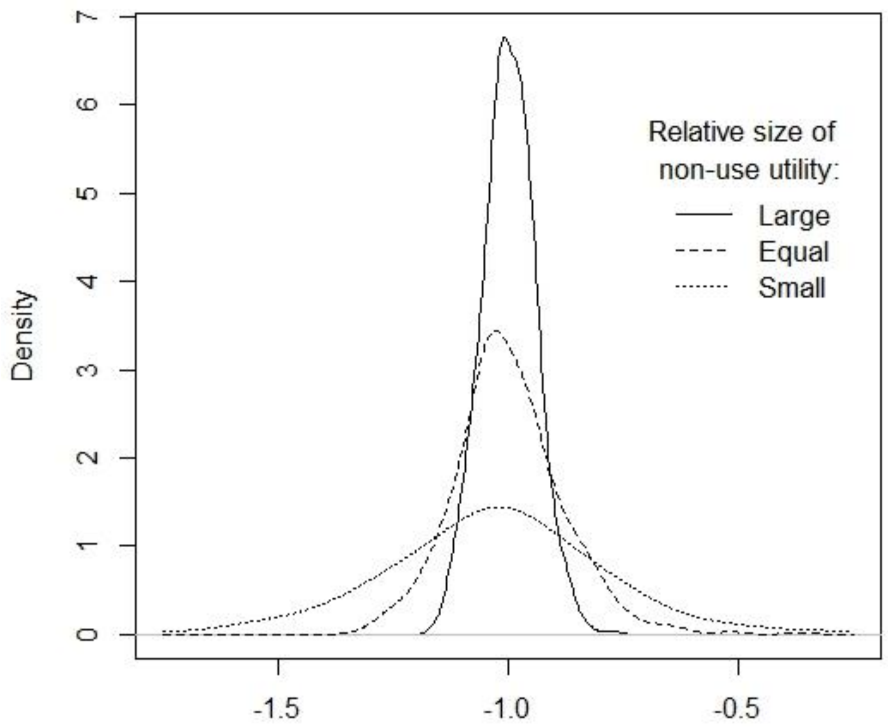


Figure 4: Distribution of distance-decay parameter (λ) from Monte Carlo experiments differing in the relative size of the non-use utility elements

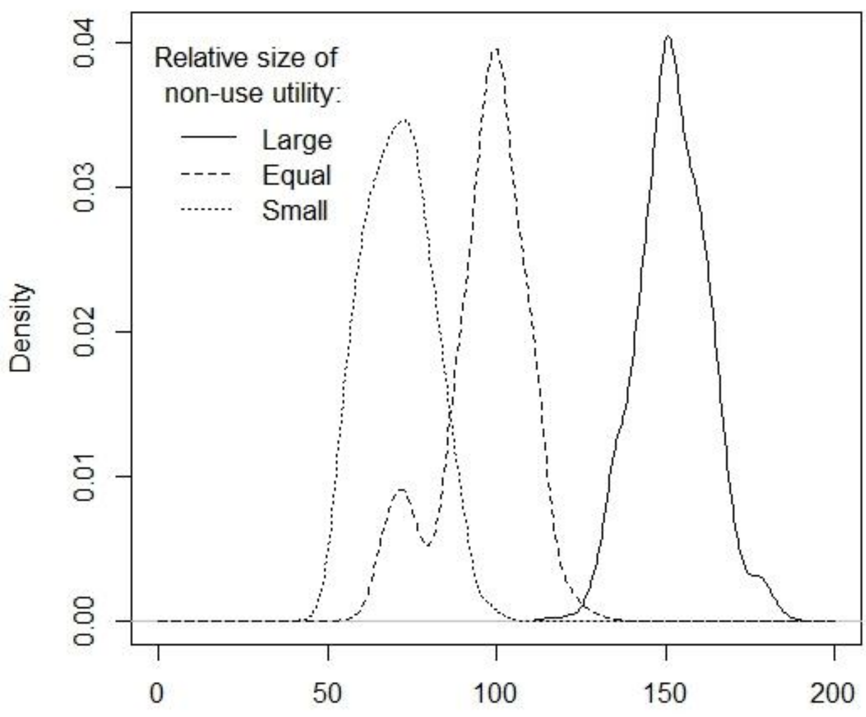


Figure 5: Distribution of welfare gains from Monte Carlo experiments differing in the relative size of the non-use utility elements

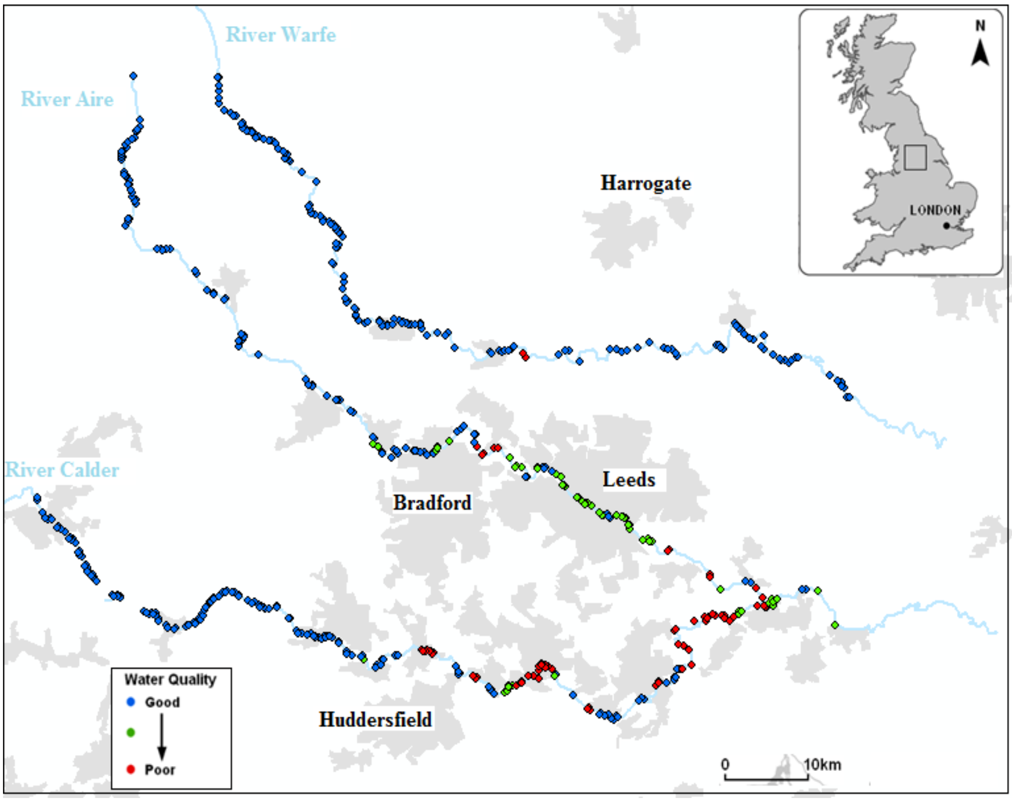


Figure 6: Study area, location of recreational river sites and their current water quality

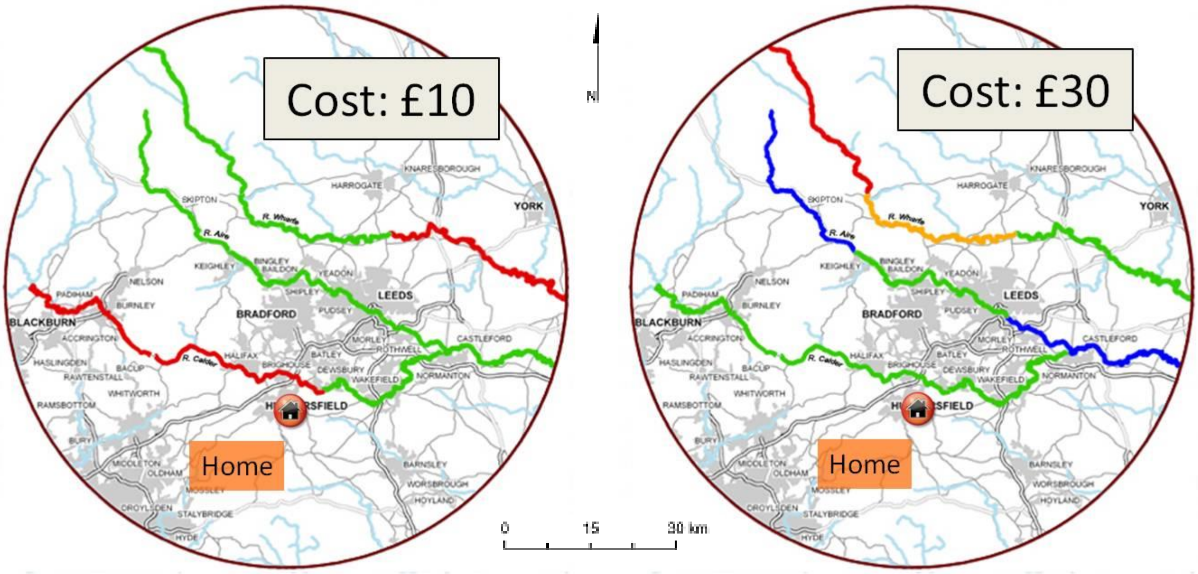


Figure 7: A typical choice task

Table I: Summary of Monte Carlo simulations for different sizes of the utility for the outside option

Parameters	Outside Good Utility: Small			Outside Good Utility: Medium			Outside Good Utility: Large		
	True	Mean (sd)	median	True	Mean (sd)	median	True	Mean (sd)	median
No Trip (α)	1	1.004 (0.068)	1.002	3	2.998 (0.181)	2.992	5	4.992 (0.319)	4.996
No Trip (α^{SP})	-0.5	-0.507 (0.257)	-0.514	-0.5	-0.484 (0.400)	-0.501	-0.5	1.884 (5.849)	-0.605
Poor Quality Use(β_1)	0.1	0.102 (0.034)	0.101	0.1	0.103 (0.051)	0.102	0.1	0.100 (0.068)	0.102
Good Quality Use (β_2)	0.3	0.304 (0.036)	0.303	0.3	0.302 (0.053)	0.305	0.3	0.300 (0.072)	0.301
Excellent Quality Use (β_3)	0.5	0.502 (0.043)	0.500	0.5	0.502 (0.063)	0.500	0.5	0.498 (0.078)	0.493
Cost (γ)	-0.1	-0.100 (0.004)	-0.100	-0.1	-0.100 (0.006)	-0.100	-0.1	-0.100 (0.006)	-0.100
Scale (σ_{RP})	2	1.999 (0.089)	1.997	2	1.997 (0.113)	1.995	2	1.996 (0.123)	2.003
Poor Quality Non-Use (b_1)	1	1.065 (0.831)	1.003	1	1.064 (0.662)	0.973	1	1.048 (0.643)	0.997
Good Quality Non-Use (b_2)	3	3.141 (1.272)	3.003	3	3.136 (1.248)	3.065	3	3.074 (1.111)	3.016
Excellent Quality Non-Use (b_3)	5	5.297 (1.901)	5.172	5	5.224 (1.975)	5.066	5	5.064 (1.692)	5.018
Distance Decay (λ)	-1	-0.997 (0.126)	-1.007	-1	-0.992 (0.144)	-1.011	-1	-1.381 (3.001)	-1.007
Mean Iterations to Convergence:		107.00			112.30			158.46	

Table II: Monte Carlo Simulation Treatments

MC Simulation Treatment Name	Non-Use Parameters			Average Welfare Change from Improvement of all Rivers to Excellent Quality		
	Poor Quality (b_1)	Good Quality (b_2)	Excellent Quality (b_3)	Non-Use Element	Use Element	Ratio of Use to Non-Use Element
Large	2	6	10	104.3	46.9	2.22
Equal	1	3	5	52.1	46.9	1.11
Small	.5	1.5	2.5	26.1	46.9	0.56

Table III: Summary of Monte Carlo simulations for different relative sizes of the non-use component of utility

Parameters	Non-Use Component: Large			Non-Use Component: Equal			Non-Use Component: Small		
	True	**Mean (sd)**	**median**	**True**	**Mean (sd)**	**median**	**True**	**Mean (sd)**	**median**
No Trip (α)	1	1.001 (0.072)	0.999	1	0.997 (0.068)	0.995	1	0.996 (0.070)	0.994
Poor Quality Use(β_1)	0.1	0.101 (0.036)	0.099	0.1	0.107 (0.036)	0.106	0.1	0.104 (0.034)	0.105
Good Quality Use (β_2)	0.3	0.301 (0.037)	0.301	0.3	0.311 (0.044)	0.307	0.3	0.308 (0.039)	0.307
Excellent Quality Use (β_3)	0.5	0.501 (0.044)	0.500	0.5	0.514 (0.055)	0.507	0.5	0.510 (0.047)	0.506
Cost (γ)	-0.1	-0.100 (0.005)	-0.100	-0.1	-0.099 (0.006)	-0.099	-0.1	-0.099 (0.005)	-0.099
Scale (σ_{RP})	2	2.004 (0.100)	2.004	2	1.975 (0.113)	1.981	2	1.980 (0.101)	1.981
Poor Quality Non-Use (b_1)	2	2.051 (0.733)	2.012	1	1.070 (1.129)	0.924	0.5	0.700 (1.449)	0.485
Good Quality Non-Use (b_2)	6	6.098 (1.223)	5.933	3	3.195 (1.420)	3.062	1.5	1.830 (1.765)	1.635
Excellent Quality Non-Use (b_3)	10	10.198 (1.871)	10.119	5	5.316 (2.000)	5.135	2.5	2.993 (1.977)	2.826
Distance Decay (λ)	-1	-0.999 (0.058)	-1.000	-1	-6.634 (20.52)	-1.025	-1	-20.68 (100.24)	-1.084
Welfare	151.22	152.22 (10.48)	151.86	99.08	96.60 (12.64)	98.34	72.99	70.71 (10.11)	70.47
Mean Iterations to Convergence:		97.62			137.10			132.56	

Table IV: Summary of Monte Carlo simulations for mis-specified model applied to stated preference data

Parameters	Non-Use: Large		Non-Use: Equal		Non-Use: Small	
	True	**Mean (sd)**	**True**	**Mean (sd)**	**True**	**Mean (sd)**
Cost (γ)	-0.1	-0.101 (0.006)	-0.1	-0.102 (0.006)	-0.1	-0.1023 (0.006)
Poor Quality (b_1)	2	4.953 (0.785)	1	3.541 (0.763)	0.5	2.809 (0,736)
Good Quality (b_2)	6	15.152 (1.337)	3	11.049 (1.175)	1.5	8.941 (1.037)
Excellent Quality (b_3)	10	26.435 (2.131)	5	19.666 (1.791)	2.5	16.211 (1.572)
Distance Decay (λ)	-1	-1.169 (0.028)	-1	-1.195 (0.033)	-1	-1.214 (0.037)
Welfare	151.22	166.08 (11.35)	99.08	115.38 (8.623)	72.99	90.30 (7.538)

Table V: Parameter estimates

Parameter	Revealed Preference Model		Revealed & Stated Preference Model	
	Estimate (std.err.)	***p*-value**	**Estimate (std.err.)**	***p*-value**
Use & Non-Use Utility Parameters				
Cost $\left(\gamma_i \sim N(\gamma, \sigma_\gamma^2)\right)$				
• Location of Distribution (γ)	-0.338 (0.018)	<0.001	-0.312 (0.022)	<0.001
• Scale of Distribution (σ_γ)	0.191 (0.011)	<0.001	0.170 (0.010)	<0.001
Use Utility Parameters				
Recreational Trip Type:				
No Trip $\left(\alpha_{J+1,i} \sim N\left(\alpha_{J+1}, \sigma_{\alpha_{J+1}}^2\right)\right)$				
• Location of Distribution (α_{J+1})	8.655 (0.452)	<0.001	8.580 (0.398)	<0.001
• Scale of Distribution $\left(\sigma_{\alpha_{J+1}}\right)$	2.325 (0.124)	<0.001	2.236 (0.139)	<0.001
Other Outdoor Trip (α_{J+2})	6.206 (0.456)	<0.001	6.211 (0.419)	<0.001
River Trip $\left(\alpha_j \sim N(\alpha, \sigma_\alpha^2)\right)$				
• Location of Distribution (α)	0.000	Baseline	0.000	Baseline
• Scale of Distribution (σ_α)	3.665 (0.092)	<0.001	3.419 (0.101)	<0.001
River Site Qualities:				
Ecological Status: Bad (β_0)	0.000	Baseline	0.000	
Ecological Status: Poor (β_1)	-0.395 (0.395)	0.317	-0.137 (0.330)	0.678
Ecological Status: Good (β_2)	-0.127 (0.500)	0.800	-0.061 (0.324)	0.852
Ecological Status: Excellent (β_3)	0.879 (0.350)	0.012	0.622 (0.277)	0.025
Land Use: Farmland (β_4)	0.000	Baseline	0.000	
Land Use: Urban (β_5)	1.070 (0.251)	<0.001	1.017 (0.262)	<0.001
Land Use: Grassland (β_6)	0.890 (0.251)	<0.001	0.857 (0.246)	0.001

Land Use: Woodland (β_7)	1.160 (0.261)	<0.001	1.190 (0.269)	<0.001
Population Density (β_8)	-0.371 (0.106)	<0.001	-0.350 (0.105)	0.001
Non-Use Utility Parameters				
River Site Qualities:				
Ecological Status: Bad (b_0)	-		0.000	baseline
Ecological Status: Poor (b_1)	-		3.078 (0.460)	<0.001
Ecological Status: Good (b_2)	-		6.754 (0.848)	<0.001
Ecological Status: Excellent (b_3)	-		8.120 (0.994)	<0.001
Distance Decay: $\left(\lambda_i \sim N\left(\lambda, \sigma_\lambda^2\right)\right)$				
• Location of Distribution (λ)	-		-1.060 (0.041)	<0.001
• Scale of Distribution (σ_λ)	.		-0.379 (0.024)	<0.001
Relative Scale of CE (σ_{SP})	.		0.519 (0.062)	<0.001
Log Likelihood	-271,700.0		-283,007.0	
N	1794		1794	