Received: 22 June 2022 | Accepted: 11 January 2024

DOI: 10.1111/1467-8489.12557

The Australian Journal of Agricultural and Resource Economics | AARES

ORIGINAL ARTICLE

A nonparametric random effects model for the valuation of forest recreation services: An application to forest sites in Tuscany, Italy

Andrea Pellegrini[1] | **Ginevra Virginia Lombardi**[2] | **Riccardo Scarpa**[3,4,5] | **John M. Rose**[6]

[1]Sustainable Mobility and Accessibility, Institute of Transport and Logistic Studies, Business School, University of Sydney, Sydney, New South Wales, Australia

[2]Department of Economics and Management, University of Florence, Firenze, Italy

[3]Durham University Business School, Durham University, Durham, UK

[4]Department of Economics, Waikato Management School, University of Waikato, Hamilton, New Zealand

[5]Department of Business Economics, University of Verona, Verona, Italy

[6]Sustainable Transport Futures, Institute of Transport and Logistic Studies, Business School, University of Sydney, Sydney, New South Wales, Australia

Correspondence
Andrea Pellegrini, Sustainable Mobility and Accessibility, Institute of Transport and Logistic Studies, Business School, University of Sydney, Sydney, NSW, Australia.
Email: andrea.pellegrini@sydney.edu.au

Abstract

This study assesses individuals' preferences for the use of forest sites for recreational purposes by means of the logit-mixed logit (LML) model. The appeal of the LML is that the analyst does not need to assume any specific functional form for the mixing distributions of random preferences. The empirical analysis generates a data-driven nonparametric representation of individuals' preference heterogeneity. We apply this approach to data collected using an unlabelled discrete choice experiment (DCE), consisting of three recreational options, two of which are in two hypothetical forest sites. Forest destinations are described by means of six attributes: forest type, signposting, hiking time, access to rivers or lakes, wildlife watch hides for visitors and cost of access. The empirical findings reveal that the signpost for each trail is the attribute for which respondents are on average willing to pay the most (6.565€). Further evidence suggests that respondents have strong positive preferences for those forest sites that offer amenities such as wildlife watching hides and access to rivers or lakes. Finally, the histograms derived from the semi-parametric LML estimation reveal multimodality of random taste amongst respondents for different hypothetical forest sites.

KEYWORDS

correlation patterns, discrete choice experiment, forest recreation sites, multimodal and asymmetric mixing distributions, nonparametric distributions, willingness to pay

JEL CLASSIFICATION

Q23, Q26, C14

1 | INTRODUCTION

The steady growth in demand for land for agricultural production and firewood products has profoundly altered the structure and distribution of forests worldwide (Keenan et al., 2015; Tavárez & Elbakidze, 2019). The UN Food and Agriculture Organization has estimated that from 1990 to 2015, one-third of the earth's land covered by trees has vanished (around 129 million hectares—almost equal to the area of South Africa), resulting in the extinction of numerous plant and animal species (FAO, 2015; MacDicken, 2015). In a recent study, based on remote sensing evidence (Potapov et al., 2022) with a 30 mt resolution grid, it was found that most of the net forest average aboveground biomass carbon loss between 2000 and 2020 has come from tall forest conversion. This type of loss is particularly damaging to biodiversity across tree canopy layers. In addition to the loss of diversity of fauna and flora, deforestation and forest degradation have notably contributed to the ongoing climate crisis, with an annual average of 4.8 billion tonnes of carbon dioxide released into the atmosphere over just the biennium 2015–2017 (Austin et al., 2020). In recent years, national governments have implemented a series of policies and programs specifically designed to mitigate the impact of forest destruction on the environment, spanning from afforestation (e.g., tree planting projects) to novel forest management approaches (Barry et al., 2014; Oldfield et al., 2013).

In densely populated areas of economically developed countries, such as Tuscany in central Italy, which is the subject of this study, woodland management must ensure the provision of multifunctional services, amongst which one of the most valuable has long been recognised to be outdoor recreation. Forested areas here have long been shaped by the presence of humans and have lost much of their naturalness. Woodland management decisions are nowadays crucially important to maintain the presence of viable communities in marginal hilly and mountain areas, which is where most of the woodlands are located. Settlements in these areas were previously threatened by mass movement of locals to urban areas, and this trend has recently been exacerbated by the demographic crisis and by the disappearance of the older generations that were custodians of these forested landscapes. The consequent fast disappearance of these communities contributes to the vulnerability to dramatic weather events of the more prosperous community settlements of the lower flooding planes. Recent flood events in Italy in May 2023 have caused an estimated 3 billion Euros of damage, which could have been in part attenuated if the water absorption properties of the topsoil in high mountain areas had been still in place. As adaptation to climate change impacts becomes increasingly important, the role of adaptive forest management in mountain areas should increase (Yousefpour et al., 2017).

Although some of the policy actions taken so far have proven to be effective in slowing down the current ecological and hydrogeological disruption, others have struggled to attain the intended goals. A possible explanation for such mixed results is that forest conservation interventions related to tourism and recreational activities are often implemented regardless of their economic and societal values (Austin et al., 2020; Elomina & Pülzl, 2021; Giergiczny et al., 2021; Scarpa, Chilton, & Hutchinson, 2000; Scarpa, Chilton, Hutchinson, & Buongiorno, 2000; Scarpa, Hutchinson, Chilton, & Buongiorno, 2000). Yet, forest recreation can be deemed a pivotal ecosystem service insofar as it provides a wide range of intangible benefits and cultural values to society, including social interaction, spiritual renewal and physical wellness (e.g., bird watching and hiking; see, for an overview, Berlinhn & Gómez-Baggethun, 2021; Boncinelli et al., 2015; Brack, 2002; Brown et al., 2016; Dou et al., 2017; Ignatyeva et al., 2020; Jim & Chen, 2009; Lankia et al., 2015; MEA, 2005; Queiroz et al., 2015; Weller & Elasser, 2018). Further, recreational uses of forests are also seen as educational opportunities from which it is possible to increase the awareness of the general public towards the existing environmental issues (Immerzeel et al., 2022; Larson et al., 2016). Hence, it is paramount for policymakers and stakeholders to account for the impact that recreational activities have on ecological outcomes when planning managerial and preservation strategies (Ferraro et al., 2011).

Within the environmental and resource economics literature, many studies have valued the recreational benefits related to environmental nonmarket goods and services by analysing stated preference (SP) survey data collected via discrete choice experiments (DCEs; see, e.g., Adamowicz, 2004; Boxall & Adamowicz, 2002; Doherty et al., 2013; Juutinen et al., 2017; Legg et al., 2020; Pelletier et al., 2022; Yao et al., 2019). In contrast to revealed preference (RP) data, SP data consist of hypothetical consumption decisions made by economic agents within experimentally designed scenarios that are constructed by the analyst, involving multiple alternatives (Adamowicz et al., 1994, 1998; Bazzani et al., 2018; Boxall et al., 2009; Boxall & Macnab, 2000; Hanley et al., 1998; Juutinen et al., 2011, 2012; Louviere et al., 2000; MacDonald et al., 2019; Morrison & Bennett, 2004; Pellegrini et al., 2022; Scheufele & Bennett, 2013).

Since the early 2000s, there have been several SP studies on the value of forest recreation that used contingent valuation across several woodland sites with different forest attributes (Hutchinson et al., 2001; Scarpa, Chilton, & Hutchinson, 2000; Scarpa, Chilton, Hutchinson, & Buongiorno, 2000; Scarpa, Hutchinson, Chilton, & Buongiorno, 2000). Later on, the advantages of choice experiments induced a shift to this multi-attribute SP method. Christie et al. (2007) explored the propensity of four forest visitor segments—cycling, horse riding, nature watching and general forest visitors—to participate in recreational activities in forests and woodland in Great Britain, so as to value a series of improvements to the recreational facilities. Brey et al. (2007) estimated that the annual willingness to pay (WTP) of Spanish rural residents for enabling picking mushrooms in the new forests stood at approximately12.82€. Nielsen et al. (2007) calculated that WTP values of Danish people for forests with varied tree heights in the stand and two tree heights in stand were around 856 DKK and 205 DKK, respectively. Upton et al. (2012) found that gaining access and facilities was the attribute associated with the highest WTP value (89.94€) within a study on afforestation preferences (see also, Vecchiato and Tempesta (2013) who examined the impact of an afforestation project in Italy). Yao et al. (2014) studied the value of iconic species in New Zealand commercial forests, demonstrating how even these forests can host biodiversity valuable to recreationists. Further examples of DCEs applied to forest-based recreation contexts include Abildtrup et al. (2013) who embedded a DCE within an online survey to acquire data on visitor preferences for recreational use of forests in Lorraine. They employed choice scenarios that displayed three forest options, two of which were depicted with five attributes: dominant tree species, hiking paths, facilities, access to water and distance from home. The authors reported that excursionists accrued positive utility from visiting forests with access to lakes or rivers and forests located near their dwellings. Giergiczny et al. (2015) administered a DCE to a sample of 1000 Poles to gather information on their preferences for 12 structural forest attributes (e.g., tourist infrastructure, forest type, stand age and residue), which were adapted from the work of Edwards et al. (2012) (we note, though, that Edwards et al. interviewed solely landscape and forest experts). Tu and Abildtrup (2015) investigated the impact of experience on the likelihood of visiting forest sites. The results reported by Sagebiel et al. (2017) showed that excursionists held positive preferences not only for afforestation but also for mixed landscape (75 per cent woodland and 25 per cent meadows) as opposed to a single typology of forest. Tavárez and Elbakidze (2019) evaluated residents' preferences for urban forests in Puerto Rico, concluding that sampled respondents were WTP \$29 and \$26 for improved trails and stands with binoculars stood, respectively. Giergiczny et al. (2021) combined SP and RP sets of data to explore the intentions to visit forest areas in 10 different European countries (Austria, Belarus, Czech Republic, Denmark, France, Germany, Poland, Slovakia, Scotland and Switzerland). A more complete bibliography of forest ecosystem valuation in German-speaking countries can be found in Elsasser et al. (2016). Some studies also address the supply side of ecosystem service provision: the willingness to accept payments by forest owners to enter provision contracts (e.g., Vedel et al., 2015).

We note that, when preference heterogeneity is addressed, limited allowance is made in all the above studies for investigating flexible distributional assumptions on the preference parameters. The main impetus of this paper is to assess what this flexibility means in an otherwise rather standard application. More broadly, the aim of our study was to quantify in monetary terms preferences held by visitors and excursionists towards recreational uses of forests located in the Tuscany mountains in Italy. The local (regional) consequences of national forest policies are often ignored. The emphasis on location is hence useful to evaluate the effects on forest access (Nielsen et al., 2016). And from this viewpoint, the region of Tuscany represents an ideal case study for woodland visitation as it provides a wide set of opportunities to recreationists: 50 per cent of the region's area is covered by woodland and more than 90 per cent of the territory is labelled as rural (Fagarazzi et al., 2021).

At the core of our study lies an unlabelled DCE, with each choice task consisting of three discrete choice alternatives, two of which are characterised by six recreation-relevant attributes such as type of forest, signposting, hiking time, access to rivers or lakes, wildlife watch hides and cost of access. The choice data so collected are analysed by means of a logit-mixed logit (LML) model, first introduced by Train (2016). The LML belongs to the class of random effects discrete choice models that explicitly acknowledge the presence of flexible forms of heterogeneity in consumers' preferences (Bansal et al., 2017, 2019; Daziano, 2020; Mäntymaa et al., 2018; Morey et al., 2008; Pellegrini et al., 2023; Scarpa et al., 2007; Scarpa & Thiene, 2005; Tabasi et al., 2023). Unlike the very popular mixed multinomial logit (MMNL) model, the LML requires no prior assumption as to the functional form of the mixing taste distributions across visitors. This is because, when the sample size is above a certain threshold (Scarpa et al., 2020), it allows researchers to accurately retrieve and identify the underlying shapes of taste distributions in a manner that is nearly completely data-driven. This results in two main advantages. First, it yields a great degree of flexibility in capturing random taste variation amongst decision-makers. Second, it limits potential misspecifications that may arise from erroneous distributional assumptions made by the researcher during the analysis. Importantly, LML addresses explicitly the issues of asymmetry and multimodality of taste distributions, which in our case study can provide important management information to woodland managers. For example, the most common parametric distributions used in MNML impose unimodality, often combined with symmetry and unbounded intervals of taste variation. However, the existence of more than one modal value in taste distribution can provide woodland managers with more than one focal value for the provision of an attribute, whilst bounded intervals of variation avoid densities over unrealistic preference values. With this information, visitors with different taste intensities can be supplied with desirable/undesirable woodland attributes. Further allowing for asymmetry provides even more flexibility. Whilst the LML estimator has been used in food choice preference analysis (Caputo et al., 2018), in time preference analysis for resource extraction (West et al., 2021), it is still relatively new in nonmarket valuation of environmental goods.

The paper is structured as follows. Sections 2 and 3 outline the data collection process and the main features of the LML model, respectively. Section 4 highlights the empirical findings with the aid of histograms and provides some policy implications, whilst Section 5 concludes the paper.

2 | DATA

2.1 | Survey structure

The design of the DCE resulted from the records of several meetings conducted with local authorities and stakeholders, as well as from running a series of focus group discussions with excursionists who usually engaged in outdoor activities. After providing their consent to participate in the survey, respondents were asked to answer a series of attitudinal questions designed to understand the role they expect the forest to play in society, as well as the importance they assign to forest ecosystem services. Next, they were asked to provide information as to their travel habits. For example, how often they visit forest sites over a year, what means of transport they typically use, how far they travel, as well as the typical travel party composition and size. Since the survey was conducted during the COVID-19 pandemic, respondents were invited to envisage a recreational destination choice scenario with no travel or access restrictions in place to combat the spread of the virus. Section 3 of the questionnaire showed to respondents the DCE along with an example of a choice task.

In introducing the DCE, a careful explanation was provided illustrating the details of the choice alternatives, attributes and attribute levels. Respondents were asked to assume having embarked on a daily recreational trip to a specific location, with multiple recreational services. In the DCE, these were described by two unlabelled forest sites (Forest A and Forest B) and a third option representing non-forest-related activities. The latter option was included to explicitly acknowledge the possibility of individuals spending time at the recreation site without necessarily having to gain access to one of the two forests listed in the DCE. The two unlabelled forest alternatives were presented with a battery of six attributes each.[1] The attributes and their corresponding levels are reported in Table 1, whilst Table 2 displays one example of a choice scenario used in the questionnaire.

The first attribute in the DCE corresponded to the *type of forest* that the recreation site offered to its visitors, consisting of three levels: tall trees with the same height, tall trees with varying heights and copse. The latter is taken as the baseline and represents a form of woodland periodically harvested for timber that regrows from the stumps from natural regeneration. It

TABLE 1 Attributes and levels.

Attributes	Levels
Type of forest	Tall trees with the same height
	Tall trees with varying heights
	Copse
Signposting	No signpost
	One signpost for all trails
	Signpost for each trail
Hiking time	1 h
	1–2 h
	2–3 h
	3–4 h
Access to rivers or lakes	Yes
	No
Wildlife watch hides	Yes
	No
Cost of access	0€, 4€, 8€ and 12€

[1]A Bayesian D-efficient experimental designed with one thousand Sobol draws was employed to distribute the attribute levels over 48 different choice tasks (see e.g., Rose & Bliemer, 2012). The resulting design was next blocked into eight blocks of six choice tasks each via an algorithm that minimized the maximum absolute value correlation between the blocking column and the final design attributes.

TABLE 2 Example choice task.

	Forest A	Forest B	None of them
Type of forest	Copse	Tall trees with the same height	
Signposting	One signpost for all trails	Signpost for each trail	I would not choose any of the forest locations.
Hiking time	3–4 h	1 h	I would engage in other leisure activities at the recreation site.
Access to river or lakes	Yes	No	
Wildlife watch hides	No	Yes	
Cost of access	4€	8€	

typically has multi-aged trees with crowns of different heights, and with space below crown is rather crowded and often unsuitable for easy walking. These levels emerged from discussions with local policymakers, as well as from collecting information on the forest landscape characterising the Apennine Mountains in Italy.

The second attribute, *signpost*, had three levels: The base level was associated with trails without any signposting at all. The second level referred to the presence of a single sign for all woodland trails, and the final level was set as a system providing separate colour-coded or number-coded signposts for each separate woodland trail (see Appendix S1 for further details).

The third attribute, *hiking time*, is related to the expected walk duration, expressed in hours, necessary to complete each woodland trail itinerary. Four levels were used to depict this attribute: 1 h (the baseline), 1–2, 2–3 h and 3–4 h. The fourth attribute, *access to rivers or lakes*, denoted whether visitors have the opportunity to gain access to rivers or lakes along the woodland trail. The fifth attribute, *wildlife watch hides*, indicated the possible provision of hides where visitors can watch wildlife populating the forest site in their natural environment and without being spotted by the animals. The last attribute, *cost of access*, represented the per person entrance fee to the woodland trail, with levels 0€, 4€, 8€ and12€ (1 Euro = 1.65 AUD = 1.10 USD in early 2024).

To facilitate the comprehension of the DCE by respondents, the meaning of the attributes and their levels were explained prior to the respondents' undertaking their first-choice task. Respondents were also given the opportunity to re-examine the description of the choice alternative attributes in real time. This was done by coding a routine in Java Script that produced a pop-up *explanation window* every time the respondent hovered the mouse over the table cells. Doing so enabled the user to complete the DCE without spending additional time and cognitive effort in memorising and recalling the features of each choice alternative, thus ensuring a high-quality data collection. Finally, respondents had to answer the standard sociodemographic and economic questions regarding their age, gender, level of education and income.

2.2 | Empirical sample

A high-quality profile provider was hired to conduct the data collection process with the survey being administrated to 1400 respondents via the online platform Qualtrics. The expected length of the survey was approximately 15 min. Of the 1400 survey participants, 309 (22 per cent) were excluded from the analysis due to response times either being excessively fast (lower than 10 min) or excessively long (longer than 38 min[2]), leaving 1091 suitably completed questionnaires. Of these, 636 (58.30 per cent) were respondents sampled from Tuscany, whilst the remaining 435 (41.70 per cent) were recruited from the neighbouring region of Emilia-Romagna. The average age of the respondents was 44 years of age, with 601 (55.00 per cent) being female, and 576 (52.80 per cent) declaring to be married. Around two in five respondents reported achieving at least a bachelor's degree whilst 71 (6.51 per cent) indicated to have achieved a high school certificate. The average number of annual visits to forest sites was 14. The vast majority of trips were taken by private car, with an average return trip distance of 57.19 km. The average travel party composition was two adults with an average of 1.84 children. Six in 10 visits to forest areas reported that excursionists had lunch onsite, with packed lunch being the most favourite option.

3 | METHODS

The primary purpose of this paper was not to compare different econometric models, but rather to explore preferences for various forest recreational services. However, as a benchmark, we estimate two distinct discrete choice models, a conventional MMNL model and a novel LML model, representing a variation of the first LML introduced by Train (2016). Our emphasis is placed on interpreting the LML model results, with the MMNL model provided as a basis of comparison. Given the ubiquity of the MMNL model, we do not describe its operation here. Rather, we refer to the reader to other texts (e.g., Hensher et al., 2015; Train, 2009). We therefore confine ourselves here to describing the much less commonly used LML model.

Unlike standard random taste choice models based on MMNL, the LML requires no prior assumptions with respect to the functional form that the random parameters should take, resulting in a nonparametric and more data-driven representation of preference heterogeneity across decision agents. To understand the model, we begin by introducing the analytical notion, after which we outline the backbone of the empirical structure that lies at the heart of the employed methodology.

An individual n in a choice situation t makes a single decision from a finite set of three mutually exclusive alternatives ($j = 1, 2, 3$), with the last option representing visiting none of the forest options accessible onsite (see the previous section for further details). Under this specification, the economic agent n is assumed to act as a utility maximiser meaning that he/she chooses the discrete alternative that returns the highest utility in expectation. Let the utility function, U_{ntj}, be defined as follows:

$$U_{ntj} = V_{ntj} + \epsilon_{ntj}. \tag{1}$$

From Equation (1) above, total utility from visiting each forest alternative is defined as the sum of two elements, the systematic component of the utility V_{ntj} and the stochastic error term ϵ_{ntj}. The latter is assumed to be independently and identically (IID) extreme value type 1 (EV1) distributed across alternatives j and respondents n. In order for the analyst to accommodate fixed and random coefficients, V_{ntj} can be further decomposed as follows:

$$V_{ntj} = \sum_{k=1}^{K} \beta_{nk} x_{ntjk} + \sum_{i=1}^{I} \omega_i q_{ntji}, \tag{2}$$

where x_{ntjk} and q_{ntji} represent the attributes associated with the j alternative, and β_{nk} and ω_i consist of weights that capture the effect of x_{ntjk} and q_{ntji} on the utility, respectively. Further, ω_i is treated

[2]Specifically, we calculated the minimum and the maximum survey completion time from a pilot study involving 80 respondents. The pilot study was undertaken at an early stage of the data collection process. By doing so, we could evaluate the duration of each section of the survey and collect feedback with respect to the perceived complexity of the DCE.

as fixed taste parameters across agents, whilst β_{nk} is taste parameters varying across agents according to an unknown distribution. The alternative specific constants (henceforth ASCs) can be incorporated into q_{ntji} by adding up to $J-1$ column vectors of ones in the case of non-random constants. Within this context of application, the $J-1$ ASCs are assumed to not vary across the sample (i.e., fixed parameters) and hence are included in q_{ntji}. Next, the utility function formalised in Equation (1) can be re-written as the WTP space specification (Scarpa, Thiene, & Train, 2008; Train & Weeks, 2005) as below:

$$U_{ntj} = \mu_n \left(\sum_{k=1}^{K} \beta_{nk} x_{ntjk} + \sum_{i=1}^{I} \omega_i q_{ntji} - p_{ntj} \right) + \epsilon_{ntj}, \tag{3}$$

where p_{ntj} is price and μ_n is a random positive scalar.

The LML model encompasses two probability expressions. The first expression is the probability that the individual n selects the alternative j in a choice situation t (i.e., this represents the standard probability function). The second expression corresponds to the probability that the random coefficients associated with the first probability expression belong to respondent n. The mechanism by which the model operates differs from that of a standard random effects model insofar as the coefficients related to the choice probabilities are not estimated. Rather, with respect to the first probability (i.e., the choice probability), the researcher specifies the support for each random coefficient. The parameter support for the k^{th} random coefficient is such that $\beta_{nk} \in \left[a_k, b_k\right]$. The choice probability that individuals choose a certain choice alternative is then given by:

$$P_{ntj}^{r} = \frac{\exp\left[\mu_n^r \left(\sum_{k=1}^{K} \beta_{nk}^r x_{ntjk} + \sum_{i=1}^{\mathrm{I}} \omega_i q_{ntji} - p_{ntj} \right)\right]}{\sum_{i \in J_{nt}} \exp\left[\mu_n^r \left(\sum_{k=1}^{K} \beta_{nk}^r x_{ntik} + \sum_{i=1}^{I} \omega_i q_{ntji} - p_{nti} \right)\right]}. \tag{4}$$

The probability in Equation (4) is calculated by taking R draws for each random coefficient from within the support ranges previously set by the analyst. Let β_{nk}^r represents the r^{th} random discrete draw extracted from within the parameter support associated with β_{nk}. Each random draw is taken from a finite and discrete base vector, with the latter being defined by a grid of points. Specifically, the analyst first defines the number of grid points, L, after which L integer values are randomly drawn from within the interval $[1, L]$. Each integer value drawn, D, is next constrained to lay between zero and one by applying the following transformation function: $g_l = \left(D_l - 1\right) / (L-1)$, where $g_L \in [0, 1]$.

At this point, each grid point is rescaled in such a way as to reproduce the parameter ranges previously set up by the analyst. This is done as follows:

$$\beta_k^l = a_k + \left(b_k - a_k\right) g_l. \tag{5}$$

In the optimisation problem, the draws are randomly taken from the rescaled finite set of coefficients obtained from Equation (5). In order to compute the second probability function, each draw is mapped onto multiple points in parameter space such that $\beta_{nk}^r \rightarrow \left\{z_{nk1}^r, \dots, z_{nkH}^r\right\}$. The $\left\{z_{nk1}^r, \dots, z_{nkH}^r\right\}$ are subsequently used to determine the probability mass function for each vector of random parameter as follows:

$$S_n^r = \frac{\exp\left(\sum_{k=1}^{K} \sum_{h=1}^{H} \alpha_{kh} z_{nkh}^r \right)}{\sum_{r=1}^{R} \exp\left(\sum_{k=1}^{K} \sum_{h=1}^{H} \alpha_{kh} z_{nkh}^r \right)}, \tag{6}$$

where α_{kh} are parameters to be estimated.

In order to map the random coefficients β^r_{nk} to z^r_{nkh}, we make use of a Spline function with three knots, paired with the cross-products of the first-order polynomial terms. The adoption of the cross-products is crucial for capturing potential correlation patterns across the utility coefficients included in the modelling specification (for additional details, the reader is redirected to Pages 42 to 44 in Train, 2016). The overall number of parameters to be estimated, H, is given by the following formula: $H = \mathrm{NV} \times (\#\mathrm{knots} + 1) + 2 \times (\mathrm{NV} - 1) + \left[\frac{(\mathrm{NV}-1)\times(\mathrm{NV}-2)}{2}\right] + \mathrm{NFX}$, where NV refers to the number of random variables, whereas NFX corresponds to the number of fixed variables.

The unconditional choice probability can then be formulated as follows:

$$O_{ntj} = \sum_{r=1}^{R} P^r_{ntj} S^r_n. \tag{7}$$

If multiple observations for the same respondent are present, then Equation (7) can be reformulated as follows:

$$O^*_n = \prod_{t=1}^{T} O_{ntj}^{y_{ntj}}, \tag{8}$$

where y_{ntj} consists of an indicator variable, which takes the value of one if alternative j is selected by respondent n in choice situation t, and zero otherwise.

By simply applying the logarithm function to Equation (8), we can write the final log-likelihood function of the model overall sampled respondents as below:

$$\mathrm{LL} = \sum_{n=1}^{N} \ln\left(O^*_n\right). \tag{9}$$

It is worth noting here that α_{kh} are the only parameters to be estimated within the LML model. Despite α_{kh} representing points within a finite and discrete parameter space, they can be used to derive interesting information as to individuals' preference heterogeneity. First, we can obtain the probability densities for each random coefficient β_{nk} in the form of histograms by applying the following equation:

$$b^r_{nk} = \left\lfloor B_k \frac{(\beta^r_{nk} - a_k)}{(b_k - a_k)} \right\rceil. \tag{10}$$

In Equation (10), B_k corresponds to the number of bins that the analyst chooses to display the histogram for parameter k, whilst $\lfloor . \rceil$ is a function that rounds the value to the nearest integer (note that $b^r_{nk} = 0$ is assigned to the first bin). Histograms are useful visual aids to the researcher as they allow for a pictorial visualisation of the distribution shape of the random utility coefficients (see the next section). Second, the model parameter estimates, α_{kh}, can be employed to compute the population moments related to each mixing distribution. For the parameter k, the first moment of the mixing distribution (i.e., the mean) is given by:

$$\overline{\beta_k} = \frac{\sum_{r=1}^{R} \beta^r_{nk} S^r_n}{\sum_{n=1}^{N} \sum_{r=1}^{R} \beta^r_{nk} V^r_n}. \tag{11}$$

As can be noted from Equation (11), the first moment of the distribution essentially stems from weighing the probability over the draws. The second central moment of the mixing distribution for the parameter k, the square root of which yields the standard deviation, is

$$\mathrm{var}\left(\beta_k\right) = \frac{\sum_{r=1}^{R}\left(\beta_{nk}^{r} - \overline{\beta_k}\right)^2 S_n^r}{\sum_{n=1}^{N}\sum_{r=1}^{R} V_n^r}. \tag{12}$$

Finally, a bootstrapping technique is necessary to approximate the standard errors associated with the mean and the standard deviation of the mixing distributions.

4 | RESULTS AND DISCUSSION

This section illustrates the results of our analysis pertaining to two empirical models, namely the MMNL model and the LML model. For the MMNL model, 2000 Halton draws were used in model estimation, with the log-likelihood function specified so as to account for the pseudo-panel nature of the data. For the estimation of the LML model, 2000 draws were randomly taken from 1000 grid points, with a three knots spline function utilised to describe random taste variation across agents. In order to account for possible correlation patterns, we coupled the spline vector-valued function with the cross-products of the second-order polynomial terms. As for the modelling specification, the utility function was assumed to be additively linear in attributes and parameters, with two ASCs linked to the unlabelled forest site options (i.e., Forest A and Forest B). Type of forest, signposting, hiking time, access to river or lakes and wildlife watch hides were dummy-coded with respect to their respective baselines, whereas cost of access was employed as a continuous variable.

To operationalise the LML model, the analyst establishes the parameter supports for each random utility coefficient adopted in the estimation process, as discussed in Section 3. To do this, we first estimated a MMNL model in WTP space with 2000 Halton draws, after which we set the intervals of variation as the estimated means plus/minus two times the corresponding estimated standard deviations (see, for further details, Train, 2016). Next, we identify the final specification of the LML by trialling different support points. The process used to determine the support values for the various coefficients involved examining the coefficient densities obtained from the estimated model as represented by the coefficient histograms. Where a coefficient has a near-zero density in one or both tails of the distribution, the support range for that coefficient is decreased and the model re-estimated. Similarly, where there exists a significant mass at the support point of a coefficient's density function, the support range is extended in that direction. Table 3 reports the parameter supports associated with the model specification that yields the best goodness of fit. The reader will note that whilst all the attributes used to

TABLE 3 Parameter supports.

	Minimum	Maximum
Tall trees with the same height	−5.204	8.499
Tall trees with varying heights	−5.063	9.938
One signpost for all trails	−6.292	11.214
Signpost for each trail	−7.679	20.115
1–2 h	−7.316	9.140
2–3 h	−11.275	10.396
3–4 h	−15.071	11.325
Access to rivers or lakes	−5.062	10.677
Wildlife watch hides	−5.984	11.867
Cost of access	0.000	0.919

describe the two discrete choice alternatives are randomly distributed over the sample, the ASCs are treated as fixed parameters.

For the MMNL model, the cost/scale parameter is assumed to be lognormally distributed, whilst the remaining random parameters are multivariate normally distributed. Similar to the LML model, the constants are maintained fixed across respondents.

The model results for the MMNL and LML models are presented in Table 4. For each model, the first column shows the estimated means of the random taste coefficients whilst the second column shows the estimated standard deviations. Recall that a positive mean value indicates that individuals are on average willing to pay more for forest sites with that feature/service with respect to the baseline attribute level, whereas a negative mean suggests that attribute is valued comparatively less. With respect to the constant terms, both models suggest that respondents have a positive WTP for the two hypothetical alternatives, suggesting that the sample has a positive sentiment towards visiting forests for the purpose of recreation, *all else being equal*. The WTP outcomes for the constants are slightly higher for the MMNL model than for the LML model.

With respect to the forest type attribute, based on the MMNL model, visitors are willing to pay 1.02€ and 2.324€ more for forests that are composed by tall trees all of the same height and tall trees of various heights, respectively (copse with natural regeneration represents the chosen base category). From the LML model, we derive slightly higher WTP estimates (1.33€ and 2.51€, respectively). The associated standard deviation estimates from both models are found to be statistically significant, implying a sizeable degree of preference heterogeneity amongst respondents in relation to these woodland attributes. The standard deviation estimates for tall trees are almost identical across models, whilst the MMNL model suggests slightly more heterogeneity in preferences for trees of varying height relative to the estimate obtained from the LML model.

The mean estimates for the signpost attribute reveal that respondents hold strong positive preferences for forests endowed with adequately signposted trails *vis-à-vis* forests lacking any such signposts. These results in estimates of WTP of 4. 64€ and 6.47€ for the MMNL model and 5.17€ and 7.00€ for the LML model for forests with one signpost for all trails and one signpost for each trail, respectively. Both models return evidence of significant preference heterogeneity for having signs on all trails relative to having a single signpost for all trails.

For the duration of the hike attribute, we note a number of differences across the MMNL and LML models. In both cases, on average, sample respondents prefer hikes of 1- to 2-h duration; however, the average WTP based on the LML model is almost twice that obtained from the MMNL model (1.13€ compared with 0. 70€). For hikes longer than 2 h, both models indicate respondents derive a lower utility than for hikes of 1-h duration. However, for hikes of 2–3 h, the mean WTP to avoid such a hike relative to a 1-h hike is slightly larger for the MMNL model than for the LML model (1.19€ compared with 0.80€). Overall, both models suggest forest visitors are keen to engage in hiking trips of 1–2 h, but with a significant and large preference variation within the sampled population.

Both models support the hypothesis that respondents on average have a positive WTP to gain access to rivers or lakes (2.95€ for the MMNL model and 3.20€ for the LML model), with similar levels of heterogeneity being present. Likewise, respondents have a positive WTP on average for access to hides from which to watch wildlife (2.39€ and 2.56€, respectively, for the MMNL and LML models). Note that the cost of access attribute (i.e., the scale parameter) has no practical interpretation for models estimated using a WTP utility specification, as cost is perfectly confounded with scale in such models.

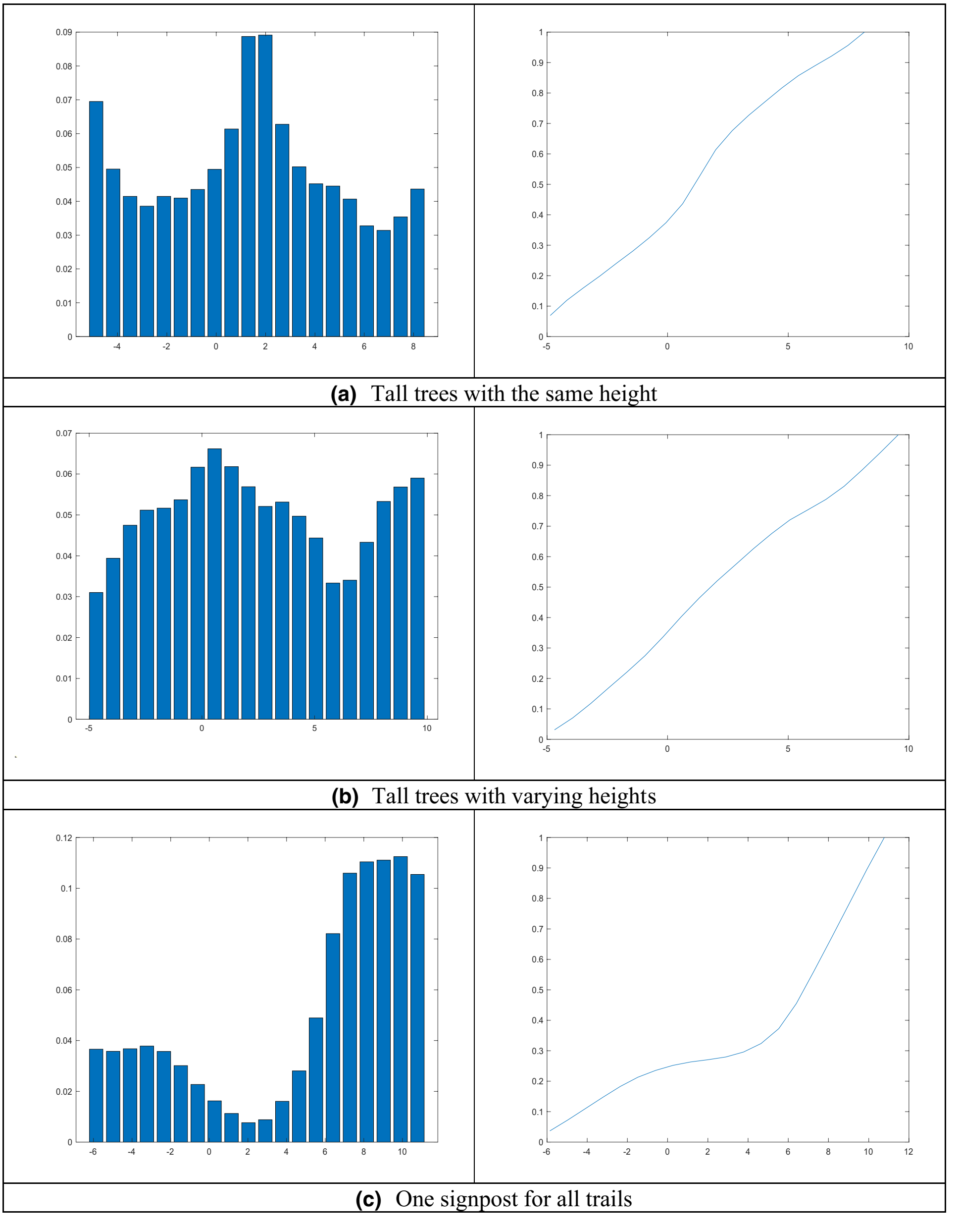
To explore the implications of the nonparametric distribution estimates obtained from the LML model, we examine both the probability density function (PDF) and cumulative density functions (CDF) of each of the random parameters derived from the model. First, a series

TABLE 4 Model results.

	MMNL model				LML model			
	Mean		Standard deviation		Mean		Standard deviation	
	Par.	(*t*-rat.)	Par.	(*t*-rat.)	Par.	(*t*-rat.)	Par.	(*t*-rat.)
ASC1	6.003	(14.66)	—	—	4.789	(9.17)	—	—
ASC2	6.583	(16.16)	—	—	5.405	(10.34)	—	—
Copse (natural regenerated woodland)	Base							
Tall trees with the same height (same age)	1.016	(3.56)	3.563	(8.56)	1.331	(3.52)	3.733	(13.18)
Tall trees with varying heights (multi age)	2.324	(7.87)	4.833	(12.69)	2.512	(6.75)	4.210	(16.09)
No signpost	Base							
One signpost for all trails	4.636	(12.92)	5.961	(14.30)	5.168	(14.12)	5.292	(19.19)
Signpost for each trail	6.565	(16.98)	6.463	(14.45)	7.000	(15.01)	6.626	(16.52)
1 h	Base							
1–2 h	0.695	(1.84)	4.340	(9.31)	1.132	(2.30)	4.798	(11.85)
2–3 h	−1.192	(−3.31)	6.000	(13.33)	−0.799	(1.79)	5.395	(15.25)
3–4 h	−2.223	(−5.41)	7.426	(12.58)	−2.155	(4.57)	6.620	(16.74)
Access to rivers or lakes (dummy = 1)	2.946	(10.99)	3.852	(11.45)	3.199	(9.39)	3.830	(13.20)
Wildlife watch hides (dummy = 1)	2.389	(7.89)	4.824	(6.56)	2.563	(6.56)	5.042	(15.67)
Cost of access (scale)	0.643	(−16.55)	1.313	(8.26)	0.307	(13.36)	0.206	(10.64)
LL (0)	−7191.516				−7191.516			
LL (β)	−5611.607				−5666.886			
p^2	0.220				0.212			
Adj. p^2	0.213				0.201			
AIC	11339.214				11521.773			
Number of parameters	58				96			
Sample size	1091							
Number of choice tasks per respondent	6							
Number of observations	6546							
Number of grid points	—				1000			
Number of draws	2000				2000			

of Kolmogorov–Smirnov tests were conducted to establish the confidence level with which any of the obtained preference distributions could be normally distributed. The resulting p-values for the null of normality were found to be smaller than 0.05, and hence, we conclude that none of the estimated distributions approximate a normal distribution. This contrasts with the MMNL model results, which typically and by design impose normality on the WTP distributions a priori. Given the similarity in population moments of the random parameter outputs reported in Table 4, the fact that the distributions obtained from the LML model are not normally distributed warrants further examinations.

To do so, we graph both the PDF and CDFs of each random coefficient in Figure 1. With respect to the woodland-type dummy tall trees of the same height, the resulting distribution

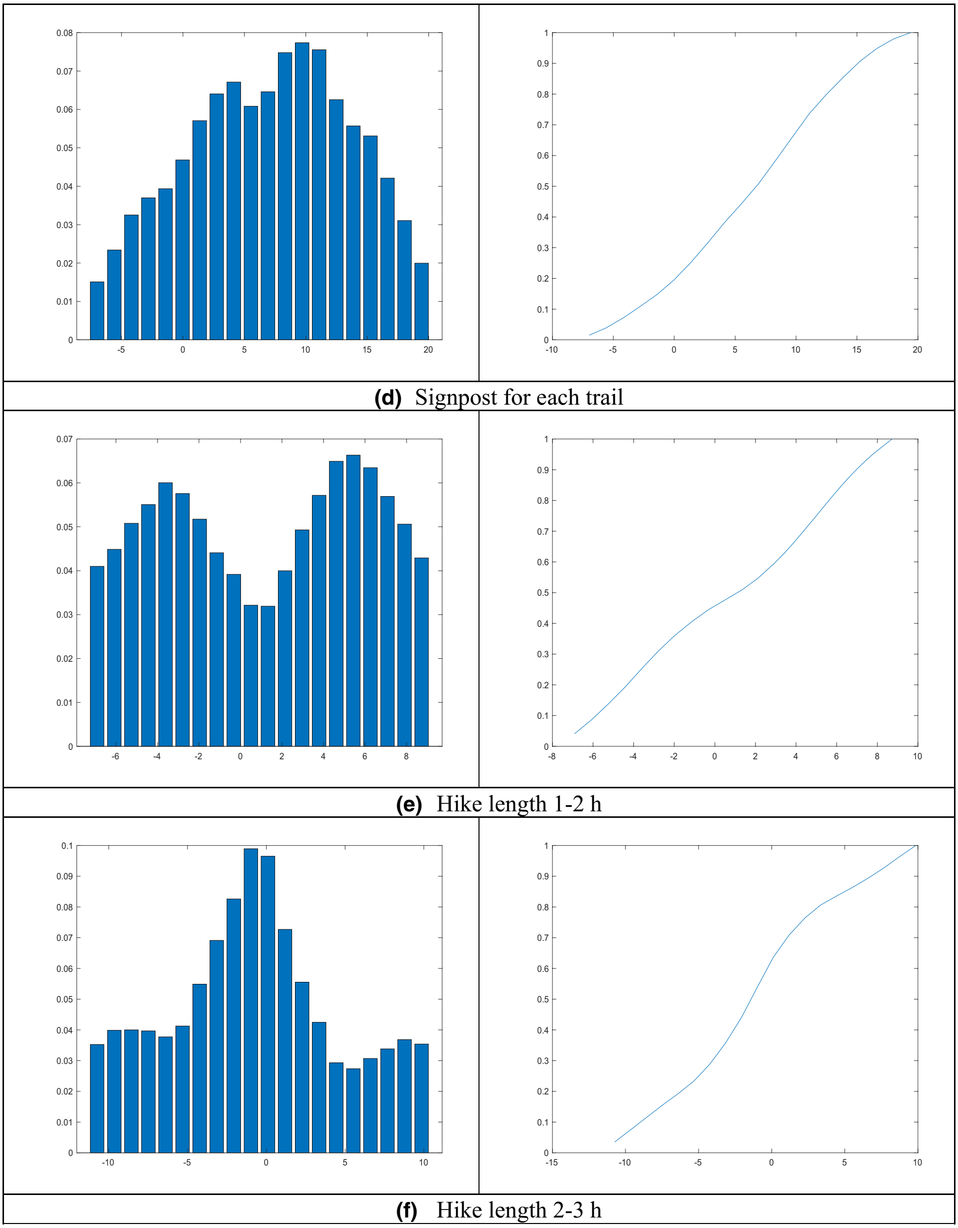


(a) Tall trees with the same height

(b) Tall trees with varying heights

(c) One signpost for all trails

FIGURE 1 Random parameter probability densities from nonparametric logit-mixed logit model. [Colour figure can be viewed at wileyonlinelibrary.com]



(d) Signpost for each trail

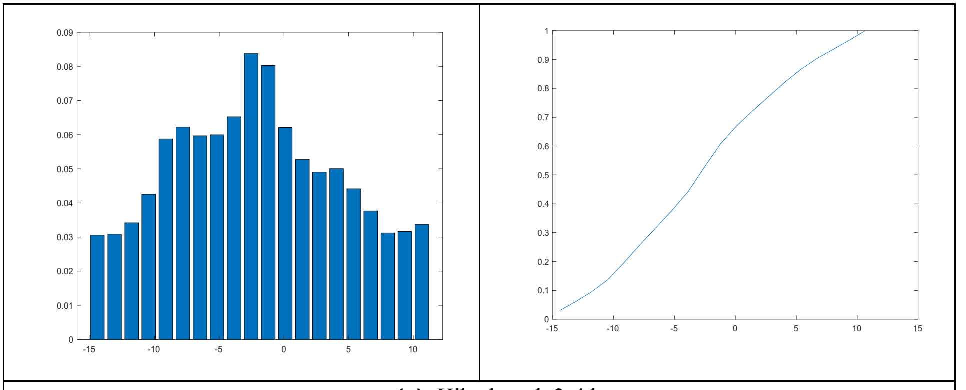
(e) Hike length 1-2 h

(f) Hike length 2-3 h

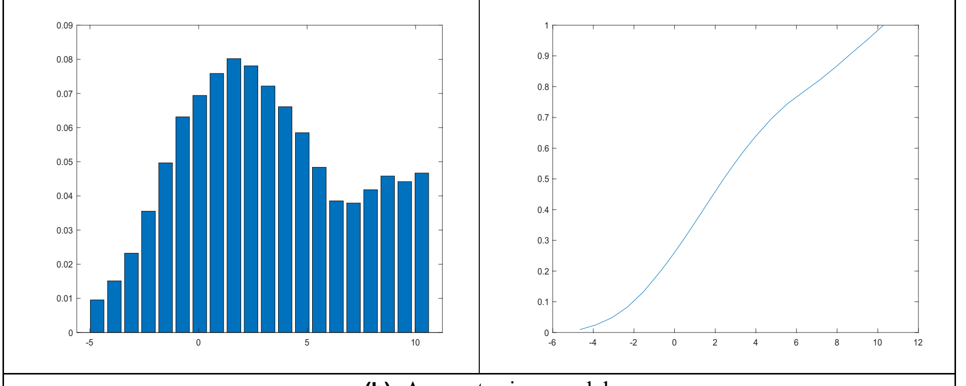
FIGURE 1 (Continued)

is trimodal in nature, as shown in Figure 1a. Whilst the largest mass exists at the central point of the distribution (between 1.31€ and 1.98€), other peaks occur at both tails of the distribution.[3] Given the range of the support values assumed, it is unsurprising that there exists a probability mass over the negative domain of the density function, which represents 37.44 per cent of the cumulative probability. This suggests that whilst the estimated mean

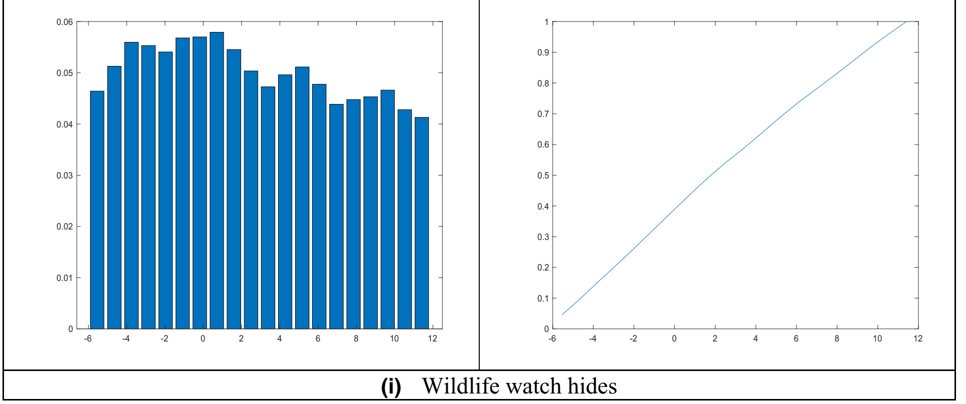
[3]The density near the supports is still rather small. Attempts were made to reduce these localised masses, as described in the section on how the support values were determined, however doing so led to areas of the distribution having zero density. Similar situations arose for other distributions.



(g) Hike length 3-4 h



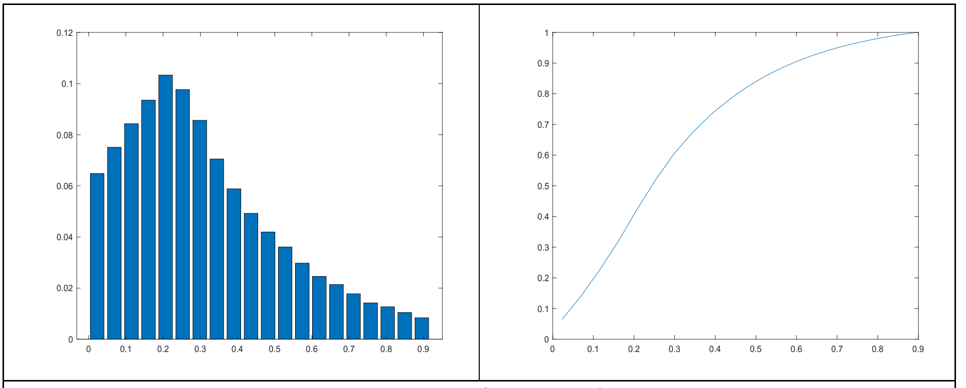
(h) Access to rivers or lakes



(i) Wildlife watch hides

FIGURE 1 (Continued)

parameter is positive (Table 4), a relevant proportion of the sampled population holds negative preferences for forests with tall trees with the same height, and some strongly prefer copse to this form of woodland. In addition to allowing for negative WTP outcomes for this attribute, the nonparametric nature of the distribution allows for skewness. In this case, 56.42 per cent of the distribution falls above the mean estimate. For the MMNL, based on the estimated mean and standard deviations, 38.77 per cent of the taste parameters for the woodland-type dummy tall trees of the same height falls within the negative range (Table 5), whilst being symmetrical around the mean, as about 50 per cent of the density falls either side of the mean/median.



(j) Cost of access (scale)

FIGURE 1 (Continued)

TABLE 5 Random coefficient percentage of density within different domains.

	MMNL model			LML model		
	−ve	**+ve**	**% above mean**	**−ve**	**+ve**	**% above mean**
Tall trees with the same height (same age)	38.77%	61.23%	50.00%	37.44%	62.56%	56.42%
Tall trees with varying heights (multi-age)	31.53%	68.47%	50.00%	33.61%	66.39%	47.90%
One signpost for all trails	21.84%	78.16%	50.00%	23.55%	76.45%	67.64%
Signpost for each trail	15.49%	84.51%	50.00%	19.42%	80.58%	55.68%
1–2 h	43.64%	56.36%	50.00%	44.44%	55.56%	52.35%
2–3 h	57.87%	42.13%	50.00%	53.93%	46.07%	46.07%
3–4 h	61.77%	38.23%	50.00%	60.78%	39.22%	55.62%
Access to rivers or lakes (dummy = 1)	22.22%	77.78%	50.00%	19.63%	80.37%	50.02%
Wildlife watch hides (dummy = 1)	31.03%	68.97%	50.00%	37.69%	62.31%	51.07%
Price/Scale	0.00%	100.00%	—	0.00%	100.00%	—

A similar distribution pattern is observed for the random coefficient of the second dummy variable pertaining to tall trees of different heights; 33.61 per cent of the density falls into the negative support of the distribution (Figure 1b). This compares with 31.53 per cent of the density for this coefficient falling within the negative domain of the distribution based on the MMNL model. The distribution is bimodal in shape, with the largest mass surrounding 0.56€ and the second mass at the right-hand side of the distribution. It is also positively skewed, with 47.90 per cent of the density falling above the mean.

Preferences for one signpost for all trails appear to be bimodal distributed, with two masses situated at both extremes of the support, with 11.08 per cent of the density lying in between (Figure 1c). Based on the LML model, 23.55 per cent of the density for this attribute falls within the negative range, a similar fraction of 21.84 per cent is predicted by the MMNL model. The two models differ significantly in terms of how much of the density falls above and below the mean, however. For the LML model, 67.64 per cent of the density for this random coefficient attribute sits above the mean, suggesting that the distribution of WTP is highly negatively skewed, with high density in the 7–11€ interval. This result highlights the strength of using the LML model over a standard MMNL model. From a policy perspective, the MMNL model implies that for this attribute, 50 per cent of the population has a WTP greater than the mean, implying that 50 per cent of the population would be expected to obtain benefits equal to or greater than the mean value (4.64€ in this instance). Based on the LML model however, the mean WTP for this attribute is 5.17€, whilst 67.64 per cent of WTP values fall above this value. This suggests that according to the LML estimates, a greater proportion of the population would benefit compared with the MMNL model if one were to implement a one signpost policy across all forests that currently have the baseline. Indeed, the median for the distribution is 7.28€, meaning that a public policy based on a referendum outcome would be supported if it cost any amount below the median value. This result further highlights the dangers of working with population moments from random coefficients models in which taste distributions are assumed a priori, rather than obtained from the data, as doing so may result in incorrect policy outcomes.

The density distribution associated with the random coefficient for one signpost for each trail dummy variable (Figure 1d) appears to have a single mode around the interval 9.69€–11.08€. Overall, 19.42 per cent of WTP values fall within the negative domain of the distribution, compared with 15.49 per cent based on the MMNL model results. The distribution is slightly negatively skewed, with 55.68 per cent of the density located above the mean. From a policy standpoint, the adoption of signposting of forest trails, despite being broadly supported, has the potential to alienate a tourism segment that would enjoy outdoor recreational activities in environments that are uncongested by visitors.

For the hiking time attribute, the random coefficient for the 1–2 h of dummy has two major density peaks located at either side of zero, respectively, around −4.44€ and 6.26€ (Figure 1e). The mass within the positive domain of the distribution accounts for 55.56 per cent of the distribution, similar to the MMNL estimates. It is worth noting, however, that when comparing the population moments for the coefficient based on the two models, the average WTP value obtained from the MMNL model is almost half that of the value obtained from the LML model. This result is due to the low-density mass for this coefficient around the centre of the distribution, whereas from the MMNL model estimates in this interval have the highest density. In effect, this indicates some form of aggregation bias occurring with respect to WTP for this attribute within the MMNL model, where respondents are observed to mostly have preferences at the extremes, but when averaged, preferences based on a symmetrical distribution will coalesce to the midpoint of the distribution, which is not representative of the preferences of most of the population. A distribution with a highest mass around zero WTP has very different implications than one with substantive densities at both sides of zero. A similar effect was reported in earlier DCE studies using flexible multimodal distributions based on Legendre polynomials (Scarpa, Thiene, & Marangon, 2008; Scarpa, Thiene, & Train, 2008). Again, this highlights the dangers of working solely with population moments of a distribution such as those reported in Table 4, without considering the actual underlying shape of the distribution.

With respect to hiking trails of 2–3 h (Figure 1f) and 3–4 h (Figure 1g), both distributions appear to be quite symmetrical, with the former having a peak at −0.98€ and the later at −2.53€. In both cases, a significant proportion of the mass falls within the negative domain of the distribution (Table 5). Compared with the WTP distribution for 3–4 h of hike length, the one for the 2–3 h of hike length appears to be Leptokurtic.

The density functions for access to river or lakes and wildlife watching location attributes are presented in Figure 1h,i, respectively. This WTP density appears to be quite low in the negative part of the distribution, then peak at 1.63€, after which it flattens out after reaching 5.56€. In contrast, the one for the provision of wildlife watching locations appears to be somewhat uniform in shape across the entire range.

Cumulatively, the graphical analysis of the flexible taste distributions produced by our analysis provides a much richer picture than the standard mean–variance representation of the normal distributions of MMNL analyses. This can clearly be seen when examining the density profiles of the random coefficients associated with having one signpost for one trail or one signpost for all trails. In terms of the former, the bimodal nature of the distribution demonstrates the risk of working with symmetrical parametric distributions, and in particular relying on the population moments of distributions to derive policy implications and advice. This is further highlighted with respect to the latter attribute, where the WTP distribution is highly skewed, such that the mean and median of the distribution are very different, again resulting in potential policy misinterpretation when the analyses do not account for flexible preference distributions. In addition to exploring the densities of the random coefficients, further insights can be obtained from examining the correlations between random parameters.

Table 6 outlines the estimated correlation parameters for the MMNL model (lower triangular sub-matrix) and the LML model (upper triangular sub-matrix). The first thing to note is that magnitudes of the correlations are larger for the MMNL model than for the LML model. Six of the MMNL model have positive values higher than 60 per cent, versus only two of those from the LML model. Out of 36 possible correlation terms, 17 are found to be statistically significant in the MMNL model at the 0.01 per cent level, with an additional correlation found to be statistically significant at the 0.05 significance level. Instead, in the LML model, only 13 correlations are found to be statistically significant at the 0.01 per cent level, with no correlations being statistically significant at the 0.05 level of significance. Of these, only seven correlation terms are found to be jointly significant across the two model forms. Seven correlation coefficients are of different signs across the two models; however, all seven of these are not statistically significant in either model. Finally, the patterns of correlation differ between the two models.

For the MMNL model, the random parameters associated with the type of forest dummy for tall trees are positively correlated with that involving trees of varying heights, as well as with the coefficients for both signpost dummy variables and the wildlife hide variable. The coefficient for tall trees with varying heights is also correlated with both signpost dummy variables, as well as with all hiking time dummies, but not with the wildlife hide dummy. In contrast, both coefficients for forest type are correlated only with the single signpost dummy variable in the LML model. As such, both models corroborate the hypothesis that individuals who prefer non-copse-type forests also prefer having at least one signpost describing the area, whereas the MMNL model also suggests that respondents who prefer forests populated with taller trees also prefer having signposts on each trail. Moreover, the MMNL model implies that those who prefer forests consisting of tall trees of the same height also prefer areas with animal hides from which to observe wildlife without being noticed.

Both models report that visitors have strong positive preferences towards forests with a single signpost and those with multiple signposts located along different tracks. The LML model, however, suggests that preferences for signposting are positively correlated with hiking tracks that take longer than an hour, something that is not found by the MMNL estimates.

For both models however, visitors who prefer having signposts also prefer forests with access to rivers or lakes. For the LML model, preferences towards having at least one signpost for all trials are positively correlated with the provision of wildlife watch hides. Of particular interest for both models is the positive correlation between having access to rivers or lakes and having wildlife watch hides in place. From a managerial perspective, both models suggest that forests with tall trees should have at least one signpost for all trails and that forests with access to rivers or lakes should also be properly signposted. Further, forests with access to rivers or lakes should also provide animal hides for watching local wildlife.

TABLE 6 Correlation matrix for random coefficients[a].

	Tall trees with the same height	Tall trees with varying heights	One signpost for all trails	Signpost for each trail	1–2 h	2–3 h	3–4 h	Access to rivers or lakes	Wildlife watch hides
Tall trees with the same height	1.000	0.104	0.305***	0.126	0.011	−0.238*	−0.165	0.211	0.030
Tall trees with varying heights	0.738***	1.000	0.299***	0.064	0.231	0.033	0.100	0.024	−0.002
One signpost for all trails	0.701***	0.460***	1.000	0.652***	0.384***	0.330***	0.092	0.450***	0.308***
Signpost for each trail	0.661***	0.274***	0.902***	1.000	0.294***	0.321***	−0.045	0.250***	−0.111
1–2 h	0.067	0.595***	0.202*	0.002	1.000	0.247	0.461	−0.025	−0.200
2–3 h	−0.080	0.325***	0.031	−0.069	0.860***	1.000	0.685***	0.317***	−0.083
3–4 h	−0.161*	0.372***	−0.149*	−0.289***	0.797***	0.915***	1.000	−0.187*	−0.021
Access to rivers or lakes	0.201*	−0.073	0.588***	0.543***	0.154	0.120**	−0.126	1.000	0.280***
Wildlife watch hides	0.378***	0.033	0.212	0.154	−0.175	0.015	−0.057	0.291***	1.000

[a]Lower triangular MMNL model—Upper triangular LML.

***Significance at the 1% level; ** significance at the 5% level; * significance at the 10% level.

5 | CONCLUSIONS

We started from the observation that better taste heterogeneity for the outdoor recreation features of forests can better guide woodland management in high population density regions of economically developed countries. We hence argued that preference models with flexible data-driven preference distributions, such as the LML model, could be of help in this context, if adequately adjusted to the case at hand. Yet, insufficient applications have appeared in this literature to date.

This study aimed to fill this gap, and to elicit individuals' preferences for recreational visits to forest in Tuscany via DCE and conducted the analysis of the data by implementing an adequately adapted version of the flexible LML model (Train, 2016). This flexible model differs from the traditional MMNL model in the way it captures random taste variation across respondents. These data are collected with an online DCE with choice tasks consisting of three alternative destinations, two of which are woodland sites described by a list of six attributes such as type of forest, signage, hiking time, access to rivers or lakes, wildlife watching spots and cost of access. The third alternative in the choice task allowed respondents to partake in outdoor activities onsite, without accessing any of the two woodlands. The empirical results suggest that sampled respondents hold strong preferences for forest attributes, with tall trees with different heights, signposting, wildlife watch hides, access to rivers or lakes and short hiking trails (up to 1–2 h).

The population moment estimates obtained from both the MMNL and LML models are quite similar, which at a superficial level, suggests that one could favour the MMNL model based on parsimony and ease of estimation. However, our results highlight a potential issue with how the results of MMNL models can fail to identify policy-salient features of preference heterogeneity. By defaulting on the common assumption of normal distributions, many researchers limit their result discussion to the unconditional mean estimates of the random parameters from the model. Take the 4.63€ WTP estimate for the mean of the one signpost dummy resulting from the MMNL model. This implies that on average, respondents are willing to pay 4.63€ more to access a forest with a single signpost for hiking trails relative to a forest with no signposting. If management was to implement a 4.63€ entry fee to such a forest, this would imply that 50 per cent of visitors would either obtain a gain in utility (that is they have a higher willingness to pay than 4.63€) or be indifferent relative to a forest with no signposting. However, 50 per cent of visitors would obtain a disutility from such an action. Based on the LML model however, the mean WTP for the same attribute is 5.17€, and 67.64 per cent of the population have a WTP value above this mean value, indicating that a greater proportion of the population would obtain a utility gain from having signposting even if a 5.17€ fee were to be imposed.

In this study, the ability to visualise estimated density distributions with histograms enabled us to uncover interesting nuances of preference distributions for woodland features, which would otherwise be overlooked in standard parametric random effects models, such as the MMNL model with normally distributed random tastes. We exhort fellow researchers to refrain from drawing conclusions based solely on the first and second central moments of the random taste distributions, to avoid making misleading policy evaluations, which might be critically driven by untested assumptions about symmetry and unimodality.

We would be amiss not to mention that our study suffers from at least three main limitations. First, the identification of the LML model that returns the best goodness of fit may require a laborious testing process prior to determining the final parameter ranges. Second, the LML model is data-intensive as emerged from the simulation exercise conducted by Scarpa et al. (2020). However, our study does use a sample size that exceeds the recommended thresholds from that simulation study. Finally, the estimated model parameters of the LML are not of direct interest to the analyst and hence necessitate conversion to provide policy-relevant outcomes.

ACKNOWLEDGEMENTS

Open access publishing facilitated by The University of Sydney, as part of the Wiley - The University of Sydney agreement via the Council of Australian University Librarians.

FUNDING INFORMATION

No external funding to report.

DATA AVAILABILITY STATEMENT

Data sharing not applicable.

ORCID

Andrea Pellegrini https://orcid.org/0000-0003-1049-1395

REFERENCES

Abildtrup, J., Garcia, S., Olsen, S.B. & Stenger, A. (2013) Spatial preference heterogeneity in forest recreation. *Ecological Economics*, 92, 67–77.

Adamowicz, W. (2004) What's it worth? An examination of historical trends and future directions in environmental valuation. *The Australian Journal of Agricultural and Resource Economics*, 48, 419–443.

Adamowicz, W., Boxall, P., Williams, M. & Louviere, J. (1998) Stated preference approaches for measuring passive use values: choice experiments and contingent valuation. *American Journal of Agricultural Economics*, 80, 64–75.

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.

Austin, K.G., Baker, J.S., Sohngen, B.L., Wade, C.M., Daigneault, A., Ohrel, S.B. et al. (2020) The economic costs of planting, preserving and managing the world's forests to mitigate climate change. *Nature Communications*, 11, 5946.

Bansal, P., Daziano, R.A. & Achtnicht, M. (2017) Extending the logit-mixed logit model for a combination of random and fixed parameters. *Journal of Choice Modelling*, 27, 88–96.

Bansal, P., Hurtubia, R., Tirachini, A. & Daziano, R.A. (2019) Flexible estimates of heterogeneity in crowding valuation in the New York City subway. *Journal of Choice Modelling*, 31, 124–140.

Barry, L.E., Yao, R.T., Harrison, D.R., Paragahaweva, Y.H. & Pannel, D.J. (2014) Enhancing ecosystem services through afforestation: how policy can help. *Land Use Policy*, 39, 135–145.

Bazzani, C., Palma, M.A. & Nayga, R.M., Jr. (2018) On the use of flexible mixing distributions in WTP space: an induced value choice experiment. *The Australian Journal of Agricultural and Resource Economics*, 62, 185–198.

Berlinhn, E.C. & Gómez-Baggethun, E. (2021) Ecosystem services from urban forests: the case of Oslomarka, Norway. *Ecosystem Services*, 51, 101358.

Boncinelli, F., Riccioli, F. & Marone, E. (2015) Do forests help to keep my body mass index low? *Forest Policy and Economics*, 54, 11–17.

Boxall, P. & Adamowicz, W. (2002) Understanding heterogeneous preferences in random utility models: a latent class approach. *Environmental and Resource Economics*, 23, 421–446.

Boxall, P., Adamowicz, W. & Moon, A. (2009) Complexity in choice experiments: choice of the status quo alternative and implications for welfare measurement. *The Australian Journal of Agricultural and Resource Economics*, 53, 503–519.

Boxall, P.C. & Macnab, B. (2000) Exploring the preferences of wildlife recreationists for features of boreal Forest management: a choice experiment approach. *Canadian Journal of Forest Research*, 30, 1931–1941.

Brack, C.L. (2002) Pollution mitigation and carbon sequestration by an urban forest. *Environmental Pollution*, 116, 195–200.

Brey, R., Riera, P. & Mogas, J. (2007) Estimation of forest values using choice modelling: an application to Spanish forests. *Ecological Economics*, 64, 305–312.

Brown, G., Pullar, D. & Hausner, V.H. (2016) An empirical evaluation of spatial value transfer methods for identifying cultural ecosystem services. *Ecological Indicators*, 60, 1–11.

Caputo, V., Scarpa, R., Nayga, R.M., Jr. & Ortega, D.L. (2018) Are preferences for food quality attributes really normally distributed? An analysis using flexible mixing distributions. *Journal of Choice Modelling*, 28, 10–27.

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.
Daziano, R.A. (2020) Flexible customer willingness to pay for bundled smart home energy products and services. *Resource and Energy Economics*, 61, 101175.
Doherty, E., Campbell, D., Hynes, S. & van Rensburg, T.M. (2013) Examining labelling effects within discrete choice experiments: an application to recreational site choice. *Journal of Environmental Management*, 125, 94–104.
Dou, Y., Zhen, L., de Groot, R., Du, B. & Yu, X. (2017) Assessing the importance of cultural ecosystem services in urban areas of Beijing municipality. *Ecosystem Services*, 24, 79–90.
Edwards, D., Jay, M., Jensen, F.S., Lucas, B., Marzano, M., Montagné, C. et al. (2012) Public preferences for structural attributes of forests: towards a pan-European perspective. *Forest Policy and Economics*, 19, 12–19.
Elomina, J. & Pülzl, H. (2021) How are forests framed? An analysis of EU forest policy. *Forest Policy and Economics*, 127, 102448.
Elsasser, P., Meyerhoff, J. & Weller, P. (2016). *An updated bibliography and database on forest ecosystem service valuation studies in Austria, Germany and Switzerland*. Available from: https://www.econstor.eu/handle/10419/148396
Fagarazzi, C., Sergiacomi, C., Stefanini, F. & Marone, E. (2021) A model or the economic evaluation of cultural ecosystem services: the recreational hunting function in the Agroforestry Territories of Tuscany (Italy). *Sustainability*, 13, 11229.
FAO. (2015) *Global Forest Resources Assessment 2015*. FAO Forestry Paper No. UN Food and Agriculture Organization, Rome.
Ferraro, P.J., Lawlor, K., Mullan, K.L. & Pattanayak, S.K. (2011) Forest figures: ecosystem services valuation and policy evaluation in developing countries. *Review of Environmental Economics and Policy*, 6, 20–44.
Giergiczny, M., Czajkowski, M., Żylicz, M.T. & Angelstam, P. (2015) Choice experiment assessment of public preferences for forest structural attributes. *Ecological Economics*, 119, 8–23.
Giergiczny, M., Jacobsen, J., Glenk, K., Abildtrup, J., Czajkowski, M., Faccioli, M. et al. (2021) *Shaping the future of temperate forests in Europe: why outdoor recreation matters*. https://doi.org/10.21203/rs.3.rs-841881/v1
Hanley, N., Wright, R.E. & Adamowicz, W. (1998) Using choice experiments to value the environment: design issues, current experience and future prospects. *Environmental and Resource Economics*, 11, 413–428.
Hensher, D.A., Rose, J.M. & Greene, W.H. (2015). *Applied choice analysis*, 2nd edition, Cambridge: Cambridge University Press.
Hutchinson, G., Scarpa, R., Chilton, S.M. & Mc Callion, T. (2001) Parametric and non-parametric estimates of WTP for forest recreation in Northern Ireland: a multi-site analysis using discrete choice contingent valuation with follow-ups. *Journal of Agricultural Economics*, 52, 104–122.
Ignatyeva, M., Yurak, V. & Logvinenko, O. (2020) A new look at the natural capital concept: approaches, structure, and evaluation procedure. *Sustainability*, 12, 9236.
Immerzeel, B., Vermaat, J.E., Juutinen, A., Pouta, E. & Artell, J. (2022) Appreciation of Nordic landscapes and how the bioeconomy might change that: results from a discrete choice experiment. *Land Use Policy*, 113, 105909.
Jim, C.Y. & Chen, W.Y. (2009) Ecosystem services and valuation of urban forests in China. *Cities*, 26, 187–194.
Juutinen, A., Kosenius, A.-K., Ovaskainen, V., Tolvanen, A. & Tyrväinen, L. (2017) Heterogeneous preferences for recreation-oriented management in commercial forests: the role of citizens' socioeconomic characteristics and recreational profiles. *Journal of Environmental Planning and Management*, 60, 399–418.
Juutinen, A., Mitani, Y., Mäntymaa, E., Shoji, Y., Siikamäki, P. & Svento, R. (2011) Combining ecological and recreational aspects in National Park Management: a choice experiment application. *Ecological Economics*, 70, 1231–1239.
Juutinen, A., Svento, R., Mitani, Y., Mäntymaa, E., Shoji, Y. & Siikamäki, P. (2012) Modelling observed and unobserved heterogeneity in choice experiments. *Environmental Economics*, 3, 57–65.
Keenan, R.J., Reams, G.A., Achard, F., de Fretis, J.V., Grainger, A. & Lindquist, E. (2015) Dynamics of global forest area: results from the FAO Global Forest Resources Assessment 2015. *Forest Ecology and Management*, 352, 9–20.
Lankia, T., Kopperoinen, L., Pouta, E. & Neuvonen, M. (2015) Valuing recreational ecosystem service flow in Finland. *Journal of Outdoor Recreation and Tourism*, 10, 14–28.
Larson, L.R., Keith, S.J., Fernandez, M., Hallo, J.C., Shafer, C.S. & Jennings, V. (2016) Ecosystem services and urban greenways: what's the public's perspective? *Ecosystem Services*, 22, 111–116.
Legg, P., MacDonald, D.H., Bark, R.H., Tocock, M., Tinch, D. & Rose, J. (2020) Cultural values, deep mining operations and the use of surplus groundwater for towns, landscapes and jobs. *Ecological Economics*, 178, 106808.
Louviere, J., Hensher, D. & Swait, J. (2000) *Stated choice methods: analysis and application*. Cambridge: Cambridge University Press.
MacDicken, K.G. (2015) Global Forest Resources Assessment 2015: what, why and how? *Forest Ecology and Management*, 352, 3–8.

MacDonald, D.H., Rose, J., Johnston, R.J., Bark, R.H. & Pritchard, J. (2019) Managing groundwater in mining region: an opportunity to compare best-worst and referendum data. *Agricultural and Resource Economics*, 63, 897–921.

Mäntymaa, E., Ovaskainen, V., Juutinen, A. & Tyrväinen, L. (2018) Integrating nature-based tourism and forestry in private lands under heterogeneous visitor preferences for forest attributes. *Journal of Environmental Planning and Management*, 61, 724–746.

MEA Millenium Ecosystem Assessment (MA). (2005) *Ecosystems and human well-being: synthesis*. Washington, DC: Island Press.

Morey, E., Thiene, M., De Salvo, M. & Signorello, G. (2008) Using attitudinal data to identify latent classes that vary in their preference for landscape preservation. *Ecological Economics*, 68, 536–546.

Morrison, M. & Bennett, J. (2004) Valuing New South Wales rivers for use in benefit transfer. *The Australian Journal of Agricultural and Resource Economics*, 48, 591–611.

Nielsen, A.B., Olsen, S.B. & Lundhede, T. (2007) An economic valuation of the recreational benefits associated with nature-based forest management practices. *Landscape and Urban Planning*, 80, 63–71.

Nielsen, A.S.E., Lundhede, T.H. & Jacobsen, J.B. (2016) Local consequences of national policies—a spatial analysis of preferences for forest access reduction. *Forest Policy and Economics*, 73, 68–77.

Oldfield, E.E., Warren, R.J., Felson, A.J. & Bradford, M. (2013) FORUM: challenges and future directions in urban afforestation. *Journal of Applied Ecology*, 50, 1169–1177.

Pellegrini, A., Borriello, A. & Rose, J. (2023) Assessing the willingness of Australian households for adopting home charging stations for electric vehicles. *Transportation Research Part C: Emerging Technology*, 148, 104034.

Pellegrini, A., Rose, J. & Scarpa, R. (2022) Multiple herbicide use in cropland: a discrete continuous model for stated choice data. *Land Economics*, 98, 355–375.

Pelletier, M.C., Tocock, M., MacDonald, D.H., Rose, J. & Sullivan, C.A. (2022) Does information matter in the value of a wetland? *Journal of Environmental Planning and Management*, 65, 1323–1348.

Potapov, P., Hansen, M.C., Pickens, A., Hernandez-Serna, A., Tyukavina, A., Turubanova, S. et al. (2022) The global 2000–2020 land cover and land use change dataset derived from the Landsat archive: first results. *Frontiers in Remote Sensing*, 3, 856903.

Queiroz, C., Meacham, M., Richter, K., Norstrom, A.V., Andersson, E., Norberg, J. et al. (2015) Mapping bundles of ecosystem services reveals distinct types of multifunctionality within a Swedish landscape. *Ambio*, 2015, 89–101.

Rose, J.M. & Bliemer, M.C.J. (2012). Sample optimality in the design of stated choice experiments. *Travel behavior research in the evolving world*. India: IATBR, pp. 119–145.

Sagebiel, J., Glenk, K. & Meyerhoff, J. (2017) Spatially explicit demand for afforestation. *Forest Policy and Economics*, 78, 190–199.

Scarpa, R., Chilton, S.M. & Hutchinson, W.G. (2000) Benefits from forest recreation: flexible functional forms for WTP distributions. *Journal of Forest Economics*, 6, 41–54.

Scarpa, R., Chilton, S.M., Hutchinson, W.G. & Buongiorno, J. (2000) Valuing the recreational benefits from the creation of nature reserves in Irish forests. *Ecological Economics*, 33, 237–250.

Scarpa, R., Franceschinis, C. & Thiene, M. (2020) Logit Mixed Logit under asymmetry and multimodality of WTP: a Monte Carlo evaluation. *American Journal of Agricultural Economics*, 103, 643–662. Available from: https://doi.org/10.1111/ajae.12122

Scarpa, R., Hutchinson, W.G., Chilton, S.M. & Buongiorno, J. (2000) Importance of forest attributes in the willingness to pay for recreation: a contingent valuation study of Irish forests. *Forest Policy and Economics*, 1, 315–329.

Scarpa, R. & Thiene, M. (2005) Destination choice models for rock climbing in the north-eastern Alps: a latent-class approach based on intensity of preferences. *Land Economics*, 85, 426–444.

Scarpa, R., Thiene, M. & Marangon, F. (2008) Using flexible taste distributions to value collective reputation for environmentally friendly production methods. *Canadian Journal of Agricultural Economics*, 56, 145–162.

Scarpa, R., Thiene, M. & Tempesta, T. (2007) Latent class count models of total visitation demand: days out hiking in the eastern Alps. *Environmental and Resource Economics*, 38, 447–460.

Scarpa, R., Thiene, M. & Train, K. (2008) Utility in willingness to pay space: a tool to address cofounding random scale effects in destination choice to the Alps. *American Journal of Agricultural Economics*, 90, 994–1010.

Scheufele, G. & Bennett, J. (2013) Effects of alternative elicitation formats in discrete choice experiments. *The Australian Journal of Agricultural and Resource Economics*, 57, 214–233.

Tabasi, M., Rose, J.M., Pellegrini, A. & Rashidi, T.H. (2023) An empirical investigation of the distribution of travellers' willingness-to-pay: a comparison between a parametric and nonparametric approach. *Transport Policy*, 146, 312–321.

Tavárez, H. & Elbakidze, L. (2019) Valuing recreational enhancements in the San Patricio Urban Forest of Puerto Rico: a choice experiment approach. *Forest Policy and Economics*, 109, 102004.

Train, K. (2009). *Discrete choice methods with simulation*, 2nd edition, Cambridge: Cambridge University Press.

Train, K. (2016) Mixed logit with a flexible mixing distribution. *Journal of Choice Modelling*, 19, 40–53.

Train, K. & Weeks, M. (2005) Discrete choice models in preference space and willingness to pay space. In: Scarpa, R. & Alberini, A. (Eds.) *Applications of simulation methods in environmental and resource economics. The economics of non-market goods and resources*, Vol. 6. Dordrecht: Springer.

Tu, G. & Abildtrup, J. (2015) The effect of experience on choosing where to go: an application to a choice experiment on forest recreation. *Journal of Environmental Planning and Management*, 59, 2064–2078.

Upton, V., Dhubháin, Á.N. & Bullock, C. (2012) Preferences and values for afforestation: the effects of location and respondent understanding on forest attributes in a labelled choice experiment. *Forest Policy and Economics*, 23, 17–27.

Vecchiato, D. & Tempesta, T. (2013) Valuing the benefits of an afforestation project in a peri-urban area with choice experiments. *Forest Policy and Economics*, 26, 111–120.

Vedel, S.Z., Jacobsen, J.B. & Thorsen, B.J. (2015) Forest owners' willingness to accept contracts for ecosystem service provision is sensitive to additionality. *Ecological Economics*, 113, 15–24.

Weller, P. & Elasser, P. (2018) Preferences for forest structural attributes in Germany—evidence from a choice experiment. *Forest Policy and Economics*, 93, 1–9.

West, G.H., Snell, H., Kovacs, K.F. & Nayga, R.M., Jr. (2021) Flexible estimation of groundwater service values and time preferences. *Journal of the Association of Environmental and Resource Economists*, 8, 825–861.

Yao, R.T., Scarpa, R., Harrison, D.R. & Burns, R.J. (2019) Does the economic benefit of biodiversity enhancement exceed the cost conversation in planted forests. *Ecosystem Services*, 38, 100954.

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

Yousefpour, R., Temperli, C., Jacobsen, J.B., Thorsen, B.J., Meilby, H., Lexer, M.J. et al. (2017) A framework for modeling adaptive forest management and decision making under climate change. *Ecology and Society*, 22. Available from: https://www.jstor.org/stable/26799027

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Pellegrini, A., Lombardi, G.V., Scarpa, R. & Rose, J.M. (2024) A nonparametric random effects model for the valuation of forest recreation services: An application to forest sites in Tuscany, Italy. *Australian Journal of Agricultural and Resource Economics*, 68, 229–252. Available from: https://doi.org/10.1111/1467-8489.12557