Utility in Willingness to Pay Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps

Riccardo Scarpa, Mara Thiene, and Kenneth Train

We compare two approaches for estimating the distribution of consumers' willingness to pay (*WTP*) in discrete choice models. The usual procedure is to estimate the distribution of the utility coefficients and then derive the distribution of *WTP*, which is the ratio of coefficients. The alternative is to estimate the distribution of *WTP* directly. We apply both approaches to data on site choice in the Alps. We find that the alternative approach fits the data better, reduces the incidence of exceedingly large estimated *WTP* values, and provides the analyst with greater control in specifying and testing the distribution of *WTP*.

Key words: destination site choice, mixed logit, nonmarket valuation, outdoor recreation, random parameters, random willingness to pay, travel cost.

Nonmarket values of qualitative changes in sites for outdoor recreation are often investigated by estimating random utility models (RUMs) of site selection (Bockstael, Hanemann, and Kling 1987; Morey, Rowe, and Watson 1993). Most recent applications address the issue of unobserved taste heterogeneity by using continuous (Train 1998, 1999) or finite mixing (Provencher, Baerenklau, and Bishop 2002; Scarpa and Thiene 2005) of individual taste distributions by means of panel mixed logit models. Such approaches are shown to produce more informative and realistic estimates of nonmarket values than models without taste heterogeneity and are now part of the state of practice in the profession. However, models with conveniently tractable distributions for taste coefficients, such as the normal and the log-normal, often obtain estimates that imply counter-intuitive distributions of marginal willingness to pay (*WTP*). This is due to the fact that the analytical expression for *WTP* involves a ratio where the denominator is the cost coefficient. Values of the denominator that are close to zero (which are possible under most standard distributions such as the lognormal) cause the ratio to be exceedingly large, such that the derived distribution of *WTP* obtains an untenably long upper tail. The mean and variance of the skewed distribution are both raised artificially by these implausibly large values.

One solution is to assume that the cost coefficient is constant and not random (e.g., Revelt and Train 1998; Goett, Train, and Hudson 2000; Layton and Brown 2000; Morey and Rossmann 2003). This restriction allows the distributions of *WTP* to be calculated easily from the distributions of the nonprice coefficients, since the two distributions take the same form. For example, if the coefficient of an attribute is distributed normally, then *WTP* for that attribute, which is the attribute's coefficient divided by the price coefficient, is also normally distributed. The mean and standard deviation of *WTP* is simply the mean and standard deviation of the attribute coefficient scaled by the inverse of the (fixed) price coefficient. The fixed cost coefficient restriction also facilitates estimation. For example, Ruud (1996) suggests that a model specification with all random coefficients can be empirically unidentified, especially in data sets with few observed choices for each decision-maker (short panels). However, this restriction is counter-intuitive as there are very good theoretical and common-sense

Riccardo Scarpa is professor in the Department of Economics, Waikato Management School, University of Waikato, Hamilton, New Zealand. Mara Thiene is assistant professor in the Dipartimento Territorio e Sistemi Agroforestali, University of Padua, Italy. Kenneth Train is adjunct professor in the Department of Economics, University of California-Berkeley, U.S.A.

Senior authorship is equally shared. The authors would like to acknowledge useful comments from three anonymous reviewers, the associated editor Stephen Swallow and Timothy Gilbride. They are also grateful to Kelvin Balcombe for making available Gauss routines to graphically assess convergence of the Gibbs sampler. The usual disclaimer on the remaining errors applies.

DOI: 10.1111/j.1467-8276.2008.01155.x

reasons as to why response to costs should vary across respondents according to factors that can be independent of observed socioeconomic covariates.

Train and Weeks (2005) note on the topic that assuming a fixed price coefficient implies that the standard deviation of unobserved utility (i.e., the scale parameter) is the same for all observations. On the other hand, it is important to recognize that the scale parameter can, and in many situations clearly does, vary randomly over observations. Estimation practices that ignore such source of variation may lead to erroneous interpretation and policy conclusions. For example, in the context of destination choice modeling, if the travel cost coefficient is constrained to be fixed when in fact scale varies over observations, then the variation in scale will be erroneously attributed to variation in *WTP* for site attributes.

Another solution is to re-parameterize the model such that the parameters are the (marginal) *WTP* for each attribute rather than the utility coefficient of each attribute. That is, instead of the usual approach of parameterizing the model in "preference space" (i.e., coefficients in the utility), the model is parameterized in "*WTP* space." This alternative procedure has recently been utilized to represent taste heterogeneity by Train and Weeks (2005) and Sonnier, Ainslie, and Otter (2007). However, the idea of specifying utility in the *WTP* space is not new. For example, readers familiar with the analysis of discrete-choice contingent valuation data may recall the so-called *variation* function or *expenditure* function approach suggested by Cameron and James (1987) and Cameron (1988), which—as discussed in some more detail by McConnell (1995)—in some cases boils down to a simple re-parameterization of the RUM model proposed by Hanemann (1984, 1989).

Train and Weeks (2005) and Sonnier, Ainslie, and Otter (2007) extended the approach by Cameron and James to multinomial choice models with random tastes, where distributional assumptions and restrictions can be placed on the *WTP*'s. They point out that the two approaches are formally equivalent because any distribution of coefficients translates into some derivable distribution of *WTP*'s, and vice versa. However, the appeal of the approach is that it allows the analyst to specify and estimate the distributions of *WTP* directly, rather than deriving them indirectly from distributions of coefficients in the utility function. To researchers in nonmarket valuation this is an important advantage.

Comparisons of estimates obtained from the two parameterizations on an identical data set have already been investigated using hierarchical Bayes (HB) estimation on stated preference (SP) data. Train and Sonnier (2005) compared the estimates of the two approaches and the implied *WTP* for attributes related to cars with different fuel (fossil, hybrid, and electric). Sonnier, Ainslie, and Otter (2007) investigate the same issues in the context of stated preference data for car brand choice and photographic cameras. Both studies found that the specifications in the preference space fit their in-sample data better but produce distributions of *WTP* with fatter tails than specifications in the *WTP* space. The studies differed for out-of-sample fit, as Train and Weeks found that their model in preference space fit out-of-sample better than their model in *WTP* space, while Sonnier, Ainslie, and Otter found the reverse.[1]

We apply these concepts to revealed-preference (RP) data, the first such application to our knowledge. In order to ensure that results are not dependent on the estimation method, we estimate our models by both HB and maximum simulated likelihood (MSL). To our knowledge, this is the first application of MSL to random coefficient models in *WTP* space. We find, like in both previous studies, that models in *WTP* space imply *WTP* distributions with lower incidence of extreme values than models in preference space, a feature that some analysts may associate with enhanced model plausibility. Further, we show with an example that assuming bounded distributions of coefficients in the preference space does not avoid the "fat tail" problem of *WTP* distributions when the cost coefficient is assumed to be random over a support approaching zero. Unlike previous studies, though, we find that the models in *WTP* space also fit the in-sample data better than the models in preference space. This improved fit arises with both HB and MSL estimation. Our findings indicate

[1] Importantly for the practice of RUM estimation, Train and Weeks emphasize how assuming independence across utility coefficients in the presence of a scale parameter which varies across visitors implies dependence (correlation) across implied *WTP* distributions, and vice versa. This issue may escape the attention of analysts, and it is worth bearing in mind for its consequences in interpretation of results, because in general neither marginal *WTP*s for attributes nor their taste intensities are independently distributed, and hence correlation matrices should be estimated whenever the data allow it, regardless of the choice of utility specification.

that in our data there is no trade-off between goodness of fit and density over extreme values: the model in *WTP* space outperforms on both criteria.

Specification

In this section, we start with the conventional specification of utility in the preference space, and describe the implications for correlation of utility coefficients and implied *WTP*s. We then re-parameterize the model in *WTP* space and discuss the implications. Throughout, the notation and language is adapted for our application to Alpine site choice.

Day trippers are indexed by n, destination sites by j, and choice situations by t. To ease the illustration, we specify utility as separable in price, p, and a vector of nonprice attributes, $\boldsymbol{x}$:

$$(1) \qquad U_{njt} = -\alpha_n p_{njt} + \boldsymbol{\theta}_n' \mathbf{x}_{njt} + \epsilon_{njt}$$

where scalar α_n and vector $\boldsymbol{\theta}_n$ vary randomly over day visitors and ϵ_{njt} is Gumbel distributed. The variance of ϵ_{njt} is visitor-specific: $\text{Var}(\epsilon_{njt}) = \mu_n^2(\pi^2/6)$, where μ_n is the scale parameter for day visitor n. Since utility is ordinal one can divide equation (1) by the scale parameter to obtain its scale-free equivalent. This division does not affect behavior and yet it results in a new error term that has the same variance for all decision makers:

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

where ε_{njt} is i.i.d. type-one extreme value, with constant variance $\pi^2/6$. The utility coefficients are defined as $\lambda_n = (\alpha_n/\mu_n)$ and $\mathbf{c}_n = (\boldsymbol{\theta}_n/\mu_n)$, such that utility may be written:

$$(3) \qquad U_{njt} = -\lambda_n p_{njt} + \mathbf{c}_n' \mathbf{x}_{njt} + \varepsilon_{njt}.$$

Note that if μ_n varies randomly, then the utility coefficients are correlated, since μ_n enters the denominator of each coefficient. Specifying the utility coefficients to be independent implicitly constrains μ_n to be constant. If the scale parameter varies and α_n and $\boldsymbol{\theta}_n$ are fixed, then the utility coefficients vary with *perfect* correlation. If the utility coefficients have correlation less than unity, then α_n and $\boldsymbol{\theta}_n$ are necessarily varying in addition to, or instead of, the scale parameter. Finally, even if μ_n does not vary over visitors (e.g., the standard deviation in unobserved factors over sites and trips is the same for all visitors), utility coefficients can be correlated simply due to correlations among tastes for various attributes.

The specification in equation (3) parameterizes utility in "preference space." The implied *WTP* for a site attribute is the ratio of the attribute's coefficient to the price coefficient: $\mathbf{w}_n = \mathbf{c}_n/\lambda_n = \boldsymbol{\theta}_n/\alpha_n$. Using this definition, utility can be rewritten as:

$$(4) \qquad U_{njt} = -\lambda_n p_{njt} + (\lambda_n \mathbf{w}_n)' \mathbf{x}_{njt} + \varepsilon_{njt}$$

which we name "utility in *WTP* space," while Sonnier, Ainslie, and Otter (2007) called it the "surplus model." In a context in which scale can vary over people—such as in our alpine destination choice—this specification is very useful for distinguishing *WTP* variation (i.e., the distributional features of $\mathbf{w}_n$) from variation in scale. To what extent this distinction affects the derived welfare estimates remains an empirical question, and one of the objectives of our investigation. We note that, although any coefficient can be used as the base that incorporates scale, the reason to focus on the travel cost coefficient in this case is that the scale-free terms can be directly interpreted as *WTP*s, which are easy to rationalize. This utility specification is distinctive for another reason as it gives a nonlinear-in-the-parameter utility function, which poses some computational challenges in the context of MSL estimation (and is probably the reason MSL has not been previously used for models in *WTP* space). In contrast, nonlinearity is readily accommodated in HB estimation.

The utility expressions are behaviorally equivalent and any distribution of λ_n and $\mathbf{c}_n$ in (3) implies a distribution of λ_n and $\mathbf{w}_n$ in (4), and vice versa. The general practice in nonmarket valuation and elsewhere has been to specify distributions in preference space, estimate the parameters of those distributions, and derive the distributions of *WTP* from these estimated distributions in preference space (Train 1998). While fully general in theory, this practice is usually limited in implementation by the use of computationally convenient distributions for utility coefficients. However, empirically tractable distributions for coefficients do not necessarily imply convenient distributions for *WTP*, and vice versa. For example, if the travel cost coefficient is distributed log-normal and the coefficients of site attributes are normal, then *WTP* is the ratio of a normal term to a log-normal term. Similarly, in (4), normal distributions for *WTP* and a log-normal for the (negative of) travel cost coefficient imply that the utility coefficients are the product of a log-normal variate and a normal one: $\lambda_n \times w_n$.

A similar asymmetry exists for the placement of restrictions on patterns of correlations (such as independence). In the travel cost site selection literature it is fairly common for researchers to specify uncorrelated utility coefficients. However, this restriction implies that scale is constant, as stated above, and moreover that *WTP* is correlated in quite a particular way via the common variation in the price coefficient. Researchers might not be aware of such implications of their choice of specification, as few papers discuss its consequences. Symmetrically, specifications assuming uncorrelated *WTP* imply a pattern of correlation in utility coefficients that is difficult to implement in preference space. We know of only one other application of travel cost RUMs that assumes a random scale parameter, but in that case the authors do not explicitly address correlation across *WTP* estimates (Breffle and Morey 2000).

The issue becomes: does the use of convenient distributions and restrictions in preference space or *WTP* space result in more accurate models? The answer is necessarily situationally dependent, since the true distributions differ across applications. However, some insight into this issue can be obtained by comparing alternative specifications on a given data set under alternative estimators. Description of our data is the topic of the next section.

Survey and Site Attribute Data

The data for our analysis were collected through a survey of 858 members of the local (Veneto Region) chapter of the CAI (Italian Alpine Club). Respondents provided information on the mountain visits that they took during the year 1999. Importantly for this application, respondents were asked the total number of trips they took to each of 18 sites in the last twelve months. A total of 9,221 trips were reported, for which table 1 gives summary statistics. The most visited sites are Piccole Dolomiti, Asiago, Lessini-Baldo, which are located in the pre-Alps, and Civetta, Pale S.Martino and Tre Cime, all of which are in the Dolomites. Unsurprisingly the most frequently attended sites are those closest to the urban centers located in the plains.

The interviewers contacted the CAI members at club meetings taking place in the municipalities of the Veneto region. The various parts of the questionnaire were explained to the group, and then the members filled out the questionnaire on their own. Respondents were asked questions about their mountaineering abilities and experience (i.e., when they started mountain recreation, whether they attended mountaineering training courses, and the kind of activities they usually undertook at the sites, etc.). Respondents also provided socioeconomic information about themselves and their households.

Table 1. Site-Specific Data

	Descriptive Statistics of Trips				Site Attributes				
Destination Sites	Mean	Std. Dev.	Visits	Percentage	Degree of Difficulty	Ferratas	Easy Trails	Shelters	Hard Trails
1. Vette Feltrine	0.7	1.5	642	7.0	3	3	0.61	25	0.07
2. P. Dolomiti-Pasubio	2.1	4	1,808	19.6	1	4	0.54	13	0.17
3. Alpago-Cansiglio	0.5	1.7	414	4.5	3	4	0.86	10	0.08
4. Asiago	1.5	2.8	1,318	14.3	1	0	1	13	0
5. Grappa	0.9	2.1	757	8.2	1	1	0.99	5	0.01
6. Baldo-Lessini	1.2	3.6	1,045	11.3	1	2	0.76	18	0.02
7. Antelao	0.3	0.7	244	2.6	3	0	0.68	6	0.08
8. Pelmo	0.3	0.6	243	2.6	3	0	0.66	9	0.04
9. Cortina	0.3	0.8	220	2.4	2	22	0.53	32	0.11
10. Duranno-Cima Preti	0.1	0.3	44	0.5	3	0	0.33	4	0.09
11. Sorapis	0.1	0.5	128	1.4	3	4	0.36	9	0.23
12. Agner-Pale S.Lucano	0.1	0.5	112	1.2	3	2	0.51	7	0.14
13. Tamer-Bosconero	0.2	0.6	188	2.0	3	0	0.3	6	0.06
14. Marmarole	0.2	0.7	161	1.7	2	1	0.51	9	0.07
15. Tre Cime-Cadini	0.6	1.2	547	5.9	2	4	0.6	9	0.08
16. Civetta-Moiazza	0.7	1.3	561	6.1	2	4	0.34	16	0.11
17. Pale S.Martino	0.7	1.3	564	6.1	2	11	0.46	14	0.14
18. Marmolada	0.3	0.7	225	2.4	3	2	0.21	13	0.25

Round-trip distance from the respondent's residence to each of the destinations in the choice set was calculated using the software package "Strade d'Italia e d'Europa." These data were used to estimate the "travel cost" to each destination. Distance was converted into monetary values using a rate of €0.35 per km, which was the car running cost at the time. Each reported trip was a "day out" (i.e., a trip lasting one day and not involving overnight stay), as is customary for this generic form of local outdoor recreation. The eighteen mountain destinations differ substantially from both a morphological and mountaineering point of view, but they can provide both specialist and nonspecialist outdoor recreation, and so are all destinations for local visitors planning "a day out."

Two broad geographically determined groups of destinations can be distinguished. Destinations 1–6 (table 1) belong to the Prealps, which are mountains with gentler slopes and lower peaks separating the plane from the proper Alps. Because of their distinct nature, the Prealps are often chosen for different recreational objectives than the Alps. Destinations 7–18 are in the Northeastern Alps, in the mountain chain of the Dolomites, which is an extended rocky area mostly made of dolomite rocks. This rare and distinguished rock type is geologically well-defined as it originates from coral reefs. Mountains made of this rock are scenically quite attractive as they tend to show orange–pink reflections at sunset.

Some of the recreational attributes describe the land-use of the sites, while others provide specific information about hiking conditions at each destination. "Degree of difficulty" is a score from 1 to 3 describing the degree of technical difficulty of the trailing itineraries that are available at each destination. This score takes into account not only the total length of the trails network but also the average degree of adversity of the mountain environment at each destination. "Ferrata" is the number of trails equipped with safety ropes, to which visitors can secure themselves in the ascent toward hard-to-reach vantage points. "Alpine shelters" is the number of equipped alpine shelters accessible in the destination area.

The recreational attractiveness of a destination is also measured as percent of total trail length that is "easily" walkable ("% of easy trails"). These trails require lower-than-average physical effort and are selected on the basis of a composite set of measurements, such as width, incline and accessibility. At the other extreme of the spectrum, we use the percent of total trail length that is "hard" walkable ("% of hard trails"), requiring higher-than-average physical effort and fitness. Finally, because the "Prealps" offer an experience distinctively different from the Dolomites, an alternative-specific constant is included for the "Prealps" to capture this difference.

Method

Revelt and Train (1998) derived the mixed logit specification in the context of repeated choices by individuals with continuous taste distributions, the so-called panel mixed logit. In our alpine destination choice context, visitor n faces a choice among J destination alternatives in each of T_n trips taken over an outdoor season. J in our case is 18 while we have a maximum of $T_n = 40$ which represents a reasonable maximum number of days out over a year. We have an unbalanced panel since the number of trips varies across individuals, hence the subscript n.

To ensure that the range of variation of the travel cost coefficient is negative we define $\lambda_n = -\exp(\upsilon_n)$, where υ_n can be considered the latent random factor underlying such coefficient.[2] Let $\boldsymbol{\beta}_n$ denote the random terms entering utility, which are υ_n and $\mathbf{c}_n$ for the model in preference space (equation (3)) and υ_n and $\mathbf{w}_n$ for the model in *WTP* space (equation (4)). Similarly, let utility be written $U_{njt} = V_{njt}(\boldsymbol{\beta}_n) + \varepsilon_{njt}$, with $V_{njt}(\boldsymbol{\beta}_n)$ being defined by either equation (3) or (4), depending on the parameterization.

Visitor n chooses destination i in period t if $U_{nit} > U_{njt}$ $\forall j \neq i$. Denote the visitor's chosen destination in choice occasion t as y_{nt}, the visitor's sequence of choices over the T_n choice occasions as $\mathbf{y}_n = \langle y_{n1}, \ldots, y_{nT_n} \rangle$. Conditional on $\boldsymbol{\beta}_n$, the probability of visitor n's sequence of choices is the product of standard logit formulas:

$$(5) \qquad L(\mathbf{y}_n \mid \boldsymbol{\beta}_n) = \prod_{t=1}^{t=T_n} \frac{e^{V_{ny_{nt}t}(\boldsymbol{\beta}_n)}}{\sum_j e^{V_{njt}(\boldsymbol{\beta}_n)}}.$$

The unconditional probability is the integral of $L(\mathbf{y}_n \mid \boldsymbol{\beta}_n)$ over all values of $\boldsymbol{\beta}_n$ weighted by its density:

[2] We note that in our estimations without random "travel cost" coefficient the point estimate is negative, as expected.

$$P_n(\mathbf{y}_n) = \int L(\mathbf{y}_n \mid \boldsymbol{\beta}_n) g(\boldsymbol{\beta}_n)\, d\boldsymbol{\beta}_n \quad (6)$$

where $g(\cdot)$ is the density of $\boldsymbol{\beta}_n$ which depends on parameters to be estimated. This unconditional probability is called the mixed logit choice probability, since it is a product of logits mixed over a density of random factors reflecting tastes.

Mixed Logit Estimation via Hierarchical Bayes

Because MSL estimation of mixed logit models is well-documented (e.g., Train 2003), in this section we mostly focus on HB estimation. For the MSL estimation we just mention that to deal with nonlinearity of V_{nit} we used BIOGEME (Bierlaire 2002, 2003) and the algorithm CFSQP Lawrence, Zhou, and Tits (1997) so as to avoid the problem of local optima. All MSL estimates were obtained using 100 quasi-random draws via Latin-hypercube sampling (Hess, Train, and Polak 2006).

The Bayesian procedure for estimating the model with normally distributed coefficients was developed by Allenby (1997) and implemented by Sawtooth Software (1999). This estimation method was also applied by Rigby and Burton (2006) to derive transforms that address mass distribution at zero (indifference to attributes) of utility coefficients (not *WTP* coefficients) for choice over GM food products in the United Kingdom. Related methods for probit models were developed by Albert and Chib (1993), McCulloch and Rossi (1994), Allenby and Rossi (1999). Layton and Levine (2005) made a contribution in the context of sequential learning from previous applications. A review of applications to marketing methods is found in Rossi, Allenby, and McCulloch (2005).

We specify the density of $\boldsymbol{\beta}_n$ to be normal with mean b and variance $\boldsymbol{\Omega}$, denoted $g(\boldsymbol{\beta}_n \mid \mathbf{b}, \boldsymbol{\Omega})$. Although terminology differs over authors and fields,[3] we call b and $\boldsymbol{\Omega}$ "population parameters" since they describe the distribution of visitor-level $\boldsymbol{\beta}_n$'s in the population. With this usage, the distribution $g(\boldsymbol{\beta}_n \mid \mathbf{b}, \boldsymbol{\Omega})$ is interpreted as the actual distribution of tastes for the recreational attributes of destination sites in the population of the regional branch of the Italian Alpine Club, from which we drew the sample. Note that, given the expression above for the price coefficient, the specification of normal $\boldsymbol{\beta}_n$ implies that the price coefficient is log-normally distributed.

In Bayesian analysis, a prior distribution is specified for the parameters. We lack previous information on the type of visitors in our sample[4] and therefore specify the prior on $\mathbf{b}$ to be a diffuse normal, denoted $N(\mathbf{b} \mid 0, \boldsymbol{\Theta})$, which has zero mean and a sufficiently large variance $\boldsymbol{\Theta}$ such that the density is essentially flat from a computational perspective.[5] A normal prior on $\mathbf{b}$ has a computational advantage since it provides a conditional posterior on $\mathbf{b}$ (i.e., conditional on $\boldsymbol{\beta}_n \forall n$ and $\boldsymbol{\Omega}$) that is also normal and hence easy to draw from, while the large variance ensures that the prior has minimal (effectively no) influence on the posterior, reflecting the absence of a-priori knowledge, especially in the presence of large samples, such as in our case. The standard diffuse prior on $\boldsymbol{\Omega}$ is inverted Wishart with low degrees of freedom. This specification is also computationally advantageous as it provides a conditional posterior on $\boldsymbol{\Omega}$ that is also Inverted Wishart and hence easy to draw from. The conditional posterior on $\boldsymbol{\beta}_n \forall n$, given $\mathbf{b}$ and $\boldsymbol{\Omega}$, is

$$\Lambda(\boldsymbol{\beta}_n \mid, \mathbf{b}, \boldsymbol{\Omega}) \propto \prod_n L(\mathbf{y}_n \mid \boldsymbol{\beta}_n) \cdot g(\boldsymbol{\beta}_n \mid \mathbf{b}, \boldsymbol{\Omega}). \quad (7)$$

Information about the posterior is obtained by taking draws from the posterior and calculating relevant statistics, such as moments, over these draws. Draws from the joint posterior are obtained by Gibbs sampling (Casella and George 1992). In particular, a draw is taken from the conditional posterior of each parameter, given the previous draw of the other parameters. The sequence of draws from the conditional posteriors converges, after a sufficient number of iterations (called "burn-in"), to draws from the joint posterior. Technical information about the algorithm can be found in Train and Sonnier (2005) and Train (2003).[6]

[3] In Bayesian applications $\mathbf{b}$ and $\boldsymbol{\Omega}$ tend to be called hyperparameters, with the $\boldsymbol{\beta}_n$'s themselves being the parameters of interest. Sometimes, however, the $\boldsymbol{\beta}_n$'s are called nuisance parameters, to reflect the concept that they are incorporated into the analysis to facilitate estimation of $\mathbf{b}$ and $\boldsymbol{\Omega}$.

[4] The only other study we know of on the region is Scarpa and Thiene (2005) and it focused on rock-climbers and not generic day-out visitors.

[5] The final results are not sensitive to the choice of prior as when we started the search algorithm from different prior values we obtained very similar results. This is unsurprising given the large sample size available.

[6] For the HB models in preference space, we used the GAUSS code that is available on K. Train's website at http://elsa.berkeley.edu/ train/software.html. We adapted this code appropriately for the HB models in *WTP* space.

It is worth reminding the reader not familiar with Bayesian estimation that the Bernstein–von Mises theorem states that, under quite unrestrictive conditions, the mean of the Bayesian posterior of a parameter is a classical estimator that is asymptotically equivalent to the maximum likelihood estimator of the parameter. Similarly, the variance of the posterior distribution is the asymptotic variance of this estimator. See Train (2003) for an extended explanation with citations. Hence, the results obtained by Bayesian procedures can be interpreted from a purely classical perspective. In the tables below, results are presented in the way that is standard for classical estimation, giving the estimate and standard error for each parameter. These statistics are the mean and standard deviation, respectively, of the draws from the posterior for each parameter.

Estimation Results

We discuss the preference space estimation results first and those in *WTP* space afterwards.

Preference Space

The estimates for models in the preference space (i.e., equation (3)) are reported in table 2. The left column reports the HB estimates, while the right column the MSL ones. Estimates with uncorrelated coefficients are reported in the top part and estimates with correlated coefficients are reported in the bottom part, while the estimates of the Cholesky matrix associated with the MSL correlated model are available in Scarpa, Thiene, and Train (2008). Allowing for full correlation amongst coefficients increases the log-likelihood simulated at the posterior means from −20,773.59 to −20,383.65 in the HB case. Similarly in the MSL case, the value of the simulated log-likelihood improves from −20,469.86 to −20,147.91.

In interpreting the figures in tables 2, recall that the coefficient for "travel cost" is log-normally distributed, such that the estimated mean and standard deviation are the mean of the latent normally distributed random factor underlying the travel cost coefficient. The other coefficients were all normally distributed, such that their means and standard deviations are estimated directly. The estimated mean and standard deviation together determine the proportion of the population implied to have coefficients of each sign.

The estimated means have the same signs and orders of magnitude across models (with and without correlation) and estimators (HB and MSL). The signs are plausible considering that the population of reference are the members of the Italian Alpine Club selecting days out in the Alps. A negative mean is observed for the "Degree of technical difficulty." To tackle technically difficult sites

Table 2. Hierarchical Bayes (HB) and Maximum Simulated Likelihood (MSL) Estimates for Preference Space Models

Prefer. Parameters $\ln(\hat{\lambda})$ and $\hat{\mathbf{c}}$	Statistics of HB post. Distribution				Statistics of MSL Estimates			
	Mean	Std. err.	Std Dev.	Std. err.	Mean	Std. err.	Std. Dev.	Std. err.
$\ln(\hat{\lambda})$	−1.29	0.04	0.73	0.25	−1.41	0.06	0.71	0.06
Degree of difficulty	−0.76	0.04	0.72	0.24	−0.51	0.04	0.48	0.06
Ferrata	−0.12	0.01	0.09	0.03	−0.07	0.01	0.02	0.01
% of easy trails	0.02	0.002	0.06	0.001	0.01	0.001	0.01	0.002
Alpine shelters	0.11	0.005	0.08	0.001	0.07	0.01	0.03	0.01
% of hard trails	0.09	0.01	0.10	0.03	0.05	0.005	0.07	0.005
Prealps ASC	−1.54	0.10	1.28	0.46	−0.98	0.11	0.98	0.09
Uncorrelated: $\ln \mathcal{L}^*$ at means of post. dist. −20,774					$\ln \mathcal{L}^*$ at convergence −20,470			
$\ln(\hat{\lambda})$	−1.22	0.05	0.88	0.28	−1.43	0.07	0.92	(∗)
Degree of difficulty	−1.16	0.07	1.17	0.39	−0.67	0.12	0.73	
Ferrata	−0.19	0.01	0.23	0.06	−0.10	0.01	0.11	
% of easy trails	0.04	0.004	0.11	0.03	0.01	0.002	0.01	
Alpine shelters	0.15	0.01	0.18	0.05	0.09	0.01	0.08	
% of hard trails	0.14	0.01	0.19	0.06	0.07	0.01	1.95	
Prealps ASC	−2.74	0.16	2.84	0.94	−1.62	0.25	0.07	
With correlation: $\ln \mathcal{L}^*$ at means of post. dist. −20,384					$\ln \mathcal{L}^*$ at convergence −20,148			

Note: (∗) Std. err. for the elements of the Cholesky matrix are reported in Scarpa, Thiene, and Train (2008).

requires rigorous training and experience, and it is expected that in general visitors are not attracted by technically challenging destinations. The negative mean for the number of "Ferrata" seems reasonable when one bears in mind that the number of "Ferrata" is mostly a consequence of strategic access for the military, established during the World War I period against invading Austrians, and not necessarily designed to facilitate tourist access to such vantage points.

Destinations with many "Alpine shelters" tend to be liked more than those with few. "Alpine shelters" are often themselves the destinations of days out in the Alps and offer opportunities to encounter other visitors and eat local specialities, as well as providing shelter for unexpected bad weather. Everything else equal, one would be more inclined to plan a day out to a destination with shelters.

Sites with higher percent of easily walkable trails ("% easy Trails") and hard walkable trails ("% hard trails") are, on the average, both liked by visitors from the Alpine Club, but with large estimated taste variation. Trail-walking is still one of the most popular activities in the Alps because it is cheap and attracts visitors of all ages and abilities. These results indicate that visitors like destinations with easy as well as more challenging trails and that there is considerable heterogeneity in visitors' response to trails' features. For example, we note that the MSL estimate imply that nearly 50% of the population do not like hard trails. Perhaps the nature of trails helps in sorting the composition of the visiting party or the purpose of the recreational visit.

Using the estimates for the means of the latent normal variables and their variance-covariance matrices, one can simulate the implied distribution of *WTP* in the population of visitors. The means, medians and standard deviations are given in table 3. The implied distribution of *WTP* is highly skewed, as evidenced by the absolute values of the mean *WTP* being considerably larger than those of the median for all attributes. Importantly, the estimates imply a fairly large proportion of visitors have *WTP* values that might appear to some as excessively large when compared to the sample distribution of round-trip travel costs ($\bar{x} =$ €9.41, $s^2 =$ €4.29, 95th quantile = €16.73). Among these are the "Degree of difficulty" of excursions, the number of "Ferrata" and the "% of hard trails." For example, the MSL model in preference space implies that 10% of visitors are, on the margin, *WTP* over €20 to avoid 1 extra level of difficulty per choice occasion, 5% are *WTP* more than €3 to avoid an extra "Ferrata" at destination, and 10% are willing to pay over €30 to have the network of trails classified as difficult increased by 10%. Similar results are typically found in specification searches for applications in which the price coefficient is allowed to vary across agents, and indeed it often motivates the assumptions of a fixed "travel cost" coefficient, which, despite the low conceptual plausibility, are currently prevailing in the published literature.

The correlation matrices across *WTP*s obtained by simulating the population distribution of the utility parameters according to these estimates are reported in the lower

Table 3. Statistics of Simulated *WTP*s from Models in Preference Space in €

	From HB Estimates			From MSL Estimates		
Site Attributes	Median	Mean	Std. Dev.	Median	Mean	Std. Dev.
			Uncorrelated models			
Degree of difficulty	−2.35	−3.62	5.44	−1.65	−2.99	6.49
Ferrata	−0.39	−0.58	0.77	−0.21	−0.40	1.22
% of easy trails	0.06	0.11	0.35	0.03	0.06	0.79
Alpine shelters	0.34	0.51	0.65	0.22	0.42	1.42
% of hard trails	0.28	0.44	0.73	0.16	0.32	2.18
Prealps ASC	−4.83	−7.34	10.07	−3.31	−5.75	10.10
			Correlated models			
Degree of difficulty	−3.04	−4.52	8.77	−2.08	−3.12	7.10
Ferrata	−0.48	−0.67	1.65	−0.31	−0.29	1.08
% of easy trails	0.09	0.18	0.84	0.05	0.09	0.16
Alpine shelters	0.36	0.40	1.29	0.26	0.21	0.83
% of hard trails	0.35	0.53	1.44	0.21	0.34	0.70
Prealps ASC	−6.93	−7.87	19.52	−4.72	−2.67	20.77

Table 4. *WTP* Correlations

Site Attributes	HB Estimates					
Degree of difficulty	1	0.60	−0.35	−0.40	−0.59	0.73
Ferrata	0.43	1	−0.30	−0.80	−0.42	0.61
% of easy trail	−0.13	−0.12	1	0.04	0.68	−0.51
Alpine shelters	−0.20	−0.48	0.04	1	0.27	−0.40
% of hard trail	−0.32	−0.27	0.34	0.14	1	−0.46
Prealps ASC	0.63	0.48	−0.14	−0.38	−0.21	1
	MSL Estimates					
Degree of difficulty	1	0.80	−0.80	−0.66	−0.73	0.71
Ferrata	0.57	1	−0.52	−0.93	−0.46	0.83
% of easy trail	0.16	−0.07	1	0.41	0.68	−0.64
Alpine shelters	−0.38	−0.97	0.11	1	0.33	−0.75
% of hard trail	−0.21	−0.02	−0.04	0.02	1	−0.32
Prealps ASC	0.63	0.70	−0.01	−0.67	0.31	1

Note: Upper triangular from *WTP* space and lower triangular from preference space.

triangular part of table 4, with the top part of the table showing the HB estimates, and the bottom part showing the MSL ones. These estimates mostly concord in signs across estimators with only 2 out 15 correlations being different. A large positive correlation is found between *WTP* for the number of "Ferrata" at destination and the "Degree of difficulty," which is behaviorally very plausible, and similarly plausible is the strong negative correlation between *WTP* for "Alpine shelters" and the "Degree of difficulty." "Alpine shelters" are particularly valuable to hikers embarking on routes leading to peaks and vantage points, many of which are accessible via "Ferrata." On the other hand, those who enjoy sites with a high "Degree of difficulty" are likely to be less dependent on the proximity of shelters for protection from sudden change of weather as they would probably be better equipped.

WTP Space

A salient feature of the *WTP* space model is that estimated parameters are also the parameters of the implied *WTP* distributions. In table 5 we report HB and MSL estimates of models parameterized in *WTP* space, i.e.,

Table 5. Hierarchical Bayes (HB) and Maximum Simulated Likelihood (MSL) Estimates for *WTP* Space Models in €

WTP Parameters	Statistics of HB post. Distribution				Statistics of MSL Estimates			
$\ln(\hat{\lambda})$ and $\hat{\mathbf{w}}$	Mean	Std. err.	Std. Dev.	Std. err.	Mean	Std. err.	Std. Dev.	St. err.
$\ln(\hat{\lambda})$	−1.41	0.04	0.74	0.24	−1.22	0.06	0.67	0.05
Degree of difficulty	−2.80	0.16	2.24	0.83	−1.99	0.20	2.19	0.33
Ferrata	−0.37	0.02	0.21	0.08	−0.31	0.03	0.06	0.04
% of easy trails	0.07	0.01	0.09	0.03	0.07	0.01	0.03	0.01
Alpine shelters	0.35	0.01	0.17	0.06	0.32	0.02	0.12	0.02
% of hard trails	0.30	0.02	0.23	0.08	0.28	0.03	0.16	0.01
Prealps ASC	−4.54	0.32	4.60	1.72	−4.39	0.46	3.97	0.39
Uncorrelated: $\ln \mathcal{L}^*$ at means of post. dist. −20,471					$\ln \mathcal{L}^*$ at convergence −20,420			
$\ln(\hat{\lambda})$	−1.81	0.05	0.74	0.25	−1.16	0.04	0.04	(∗)
Degree of difficulty	−5.59	0.34	5.87	2.25	−2.85	0.16	2.98	
Ferrata	−0.60	0.05	0.74	0.28	−0.37	0.02	0.37	
% of easy trails	0.16	0.02	0.27	0.09	0.10	0.01	0.08	
Alpine shelters	0.53	0.04	0.51	0.19	0.36	0.02	0.23	
% of hard trails	0.56	0.05	0.70	0.26	0.37	0.02	0.38	
Prealps ASC	−7.37	0.78	13.78	5.13	−5.76	0.36	6.57	
With correlation: $\ln \mathcal{L}^*$ at means of post. dist. −20,326					$\ln \mathcal{L}^*$ at convergence −20,068			

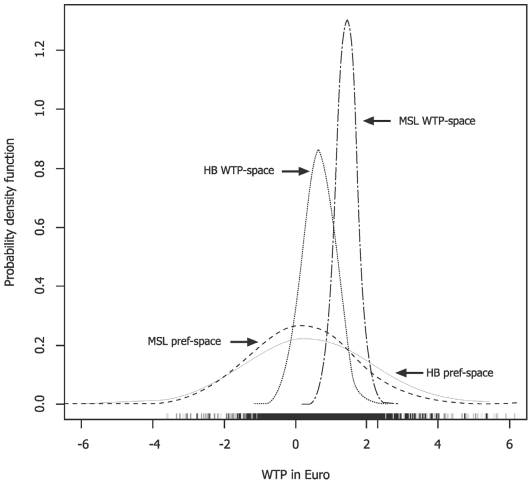
Note: (∗) Std. err. for the elements of the Cholesky matrix are reported in Scarpa, Thiene, and Train (2008).

according to equation (4). Estimates for the model without correlation are reported in the top part of the table, while the estimates with full correlation are in the bottom part of the table. The simulated log-likelihood is higher for the models in *WTP* space than in preference space: −20,470.89 versus −20,773.59 for uncorrelated terms, and −20,325.55 compared to −20,383.65 for models with correlated terms using HB. This result, which differs from the findings of Train and Weeks (2005) and Sonnier, Ainslie, and Otter (2007) on SP data, indicates that it is possible for models in *WTP* space to statistically outperform models in preference space. A similar improvement is found for the MSL estimates reported in the right column of the table. The associated estimates for the Cholesky matrix are available in Scarpa, Thiene, and Train (2008).

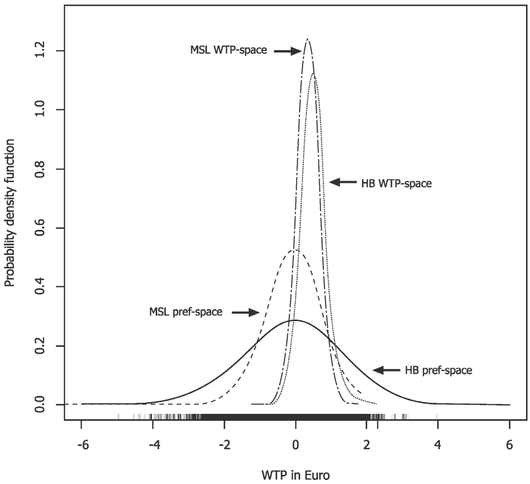
The MSL estimates imply smaller *WTP* variation than the HB ones for all attributes, but means have identical signs and very similar magnitudes. Models with correlation also uniformly imply smaller *WTP* variation in the population, with exclusion of the “Prealps ASC” in the MSL model. Examining the upper triangular sections of table 4 we note that estimated correlation match perfectly in sign between the HB (upper part of the table) and MSL estimates.

The estimated standard deviations of *WTP* are uniformly lower for the models in *WTP* space than the models in preference space. For example, in the HB models with correlated terms, the standard deviation of *WTP* for “Alpine shelters” is 1.29 for the model in preference space and 0.51 for the model in *WTP* space. However, the estimated means are not consistently higher or lower under either parameterization: with correlated terms, the HB model in *WTP* space gives a higher mean than the model in preference space for three attributes and a lower mean for the other three. The share of extreme values for *WTP* is much smaller with the models in *WTP* space than with the models in preference space. For example, the correlated preference space HB model implies that 5% of the population is willing to pay at least €1.41 for 1% increase in easy trails.[7] In contrast, the correlated model in *WTP* space implies a more plausible €0.60.

[7] The distribution of $\widehat{WTP}$ was simulated with 100,000 draws from the distributions evaluated at the estimated location and scale parameters.



(a) Models with no correlation.



(b) Models with correlation.

Figure 1. Distributions of *WTP* for one additional “Alpine shelter”

This point is visually described in figure 1 obtained with the SM package in R (Bowman and Azzalini 1997). Here we plot the kernel smoothing with cross-validated bandwidth of a simulation of 100,000 draws from each model’s *WTP* for an extra alpine shelter at destination. The densities implied by the models in *WTP* space are much “tighter” than those implied by the models in preference space. As a result, some analysts could find the *WTP* distributions implied by *WTP* space models more plausible.

Comparison

The two estimation methods produced very similar results, but the estimation via MSL was more difficult and took much longer than HB. When the MSL estimation was started at the convergence values of the HB procedures, the *WTP* space model with correlation took four times as long to run as the equivalent model using HB. With other starting values, it took even longer. For these reasons, HB estimation has an advantage. The potential disadvantage of HB estimation is in determining whether the sampler has actually converged. We evaluated convergence of the HB sampler both visually and formally using the test suggested by Geweke (1992) and Koop (2003). The tests were passed for all coefficients at conventional significance levels of 10%.

The correlated models always show significantly better fit than the uncorrelated ones. These models are also more reasonable a priori, since one would expect correlations among both *WTP* and utility coefficients. In particular, the scale parameter induces correlation among utility coefficients, such that allowing correlation in models in preference space is recommended even if the analyst believes that *WTP* is not correlated, as discussed more extensively in Train and Weeks (2005).

The "long tails" of the *WTP* distributions implied by the models in preference-space might be thought to be a direct consequence of the distributional assumptions invoked for the utility coefficients. In particular, they might be thought to be linked to the long tail of the lognormal distribution used to model the variation of the travel cost coefficient and to the normal distribution, which is known to span the entire real line. If this were the case such effect would disappear if the assumed distributions were bounded over a plausible range of values for each utility coefficient. In order to empirically investigate this issue we used the S_b distribution, which was introduced in this literature by Train and Sonnier (2005) and used previously by Rigby and Burton (2006), which is defined over an upper and lower bound.[8] We estimate two models with fully correlated taste intensities using the HB estimator—one in preference space and the other in *WTP* space where the S_b distribution was assumed for the price coefficient λ and the coefficients for Alpine shelters. All other taste intensities were assumed normally distributed.[9] In both specifications the $\lambda \sim S_b[0, 2]$, while the distribution for Alpine shelters was $c_n^{\text{shelter}} \sim S_b[0, 2]$ in preference space and $w_n^{\text{shelter}} \sim S_b[0, 1.5]$ in *WTP* space. The different upper bound reflects differences in the intervals associated with the highest density observed in estimation results with unbounded distributions illustrated in figure 1.

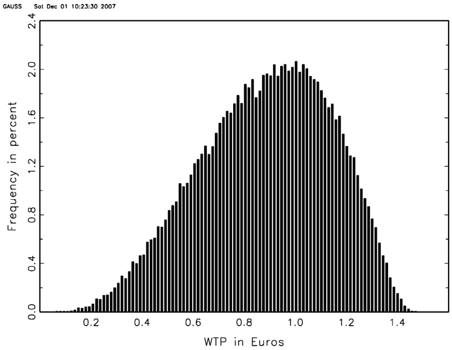
The results are best represented graphically. Figure 2 reports the histograms describing the implied *WTP* distributions for the two specifications. The distribution from the model in *WTP* space (in the top panel) is contained in the interval €[0,1.5], which are the bounds of the S_b distribution that was specified for it. However, the use of bounded distributions in preference space does not have the same effect: the implied distribution of *WTP* (in the bottom two panels) shows a very long tail. In the top two panels, the units on the x-axis are the same and cover the interval €[0,1.5]. However, the preference space *WTP* distribution extends far beyond 1.5; the bottom panel shows the extension of the histogram for this distribution over the interval €[1.5,16]. These results illustrate that assuming bounded distributions for utility coefficients does not eliminate the long tails of the *WTP* distributions. More importantly, it does not imply boundedness of such distributions. This is because the skewness arises from the *WTP* formula being a ratio of two random coefficients. When the denominator is close to zero, the ratio becomes arbitrarily large, even when the numerator is bounded. The ability to invoke and specify distributional assumptions directly in *WTP*-space is an important practical advantage afforded to analysts. It avoids the analyst achieving such control by assuming a conceptually undesirable fixed "travel cost" coefficient or bounding the price coefficient to be some arbitrary distance from zero.

Policy Implications and Conclusions

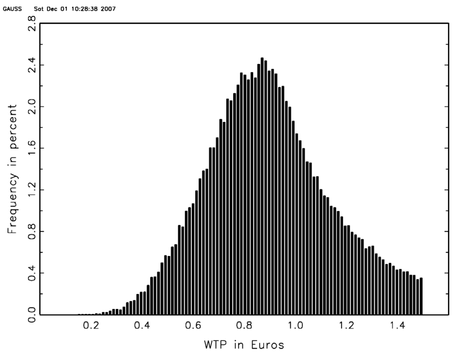
This study investigated destination choices of an inherently diverse population of visitors to alpine destinations in the northeast of Italy: those members of the local (Veneto) chapter of the alpine club visiting the Alps for day-out trips. Using a panel data set of 858 respondents who took a total of 9,221 trips, we estimate *WTP* distributions for key site attributes using models parameterized in preference space and in *WTP* space. Because parameters enter nonlinearly in the model in *WTP* space and the number of parameters

[8] If $s \sim S_b$ then $s = a + (b - a)(1 + e^{-x})^{-1}$ and $x \sim N(\mu, \sigma^2)$ while a and b are the lower and upper bound of the range.

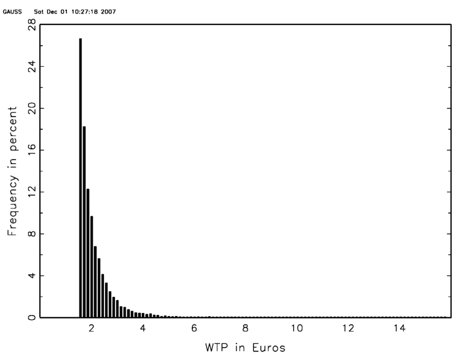
[9] The simulated log-likelihood at the posterior distribution was −20,706 for the preference space model and −20,177 for the *WTP* space model, hence showing that this bounded distributions imply a slight worsening of statistical fit. The entire set of model estimates are reported in Scarpa, Thiene, and Train (2008).



(a) From WTP-space model with $\hat{w}_{shelter} \sim S_b$.



(b) From preference-space model with $\hat{\beta}_{shelter} \sim S_b$.



(c) From preference-space model with $\hat{\beta}_{shelter} \sim S_b$.

Figure 2. Histogram of *WTP* for one additional "Alpine shelter"

is large when correlations are allowed, previous studies used hierarchical Bayes estimation procedures, which are computationally faster than maximum simulated likelihood for models of this form. We contrasted HB and MSL estimates and found them to produce similar results, with the latter implying smaller variation of taste and hence of values. However, we note that MSL estimates are significantly more time consuming to derive.

Our results confirm previous findings obtained by Train and Weeks (2005) and Sonnier, Ainslie, and Otter (2007) that the models in *WTP* space provide estimates of *WTP* distributions that have lower densities associated with extreme *WTP* values than the models in preference space. As a result some analysts might judge *WTP* space models more reasonable. However, unlike these previous studies, we find that the specification in *WTP* space statistically outperforms that in preference space. This means that practitioners need not face a trade-off between plausibility of *WTP* estimates and model fit to the data, as was previously suggested. We also illustrate how the undesirable skewness of *WTP* distributions derived from preference space specifications based on a random travel cost coefficient models is not eliminated by assuming bounded distributions. *WTP* space specifications therefore emerge as a natural choice when the analyst wants to directly control the distributions of marginal *WTP*.

Although the main objective of the paper is methodological, the estimation results from the MSL model in *WTP* space with correlated terms—which fits the data best—provide some interesting implications. About 83% of day visitors are estimated to dislike sites with high "Degree of difficulty" of hiking activities. Only about 17% show a positive *WTP* value for this attribute. Similarly, a large number of "Ferrata" at the site is attractive to only about 16% of the population of day-out visitors. The presence of "Alpine shelters" is preferred by the vast majority of visitors: only 5% of visitors prefer sites without the shelters. For most members of the Italian Alpine Club, the site becomes more attractive as the percent of trails that are classified as easily walkable and hard walkable (as opposed to those with mixed classification) rises. Finally, visitors are found to be willing to pay more to visit the Dolomites than the "Prealps," which—given the popularity of these sites—is a perhaps foregone conclusion, but it is nevertheless confirmed by the negative sign of the alternative specific constant for "Prealps."

Further research and applications of models specified in the *WTP* space should address their performance in models that include nested recreation choices (e.g., participation decisions) and in the derivation of consumer surplus measures for site attribute changes that vary across sites.[10] This will lead to a more complete understanding of this category of models

[10] We are grateful to an anonymous referee for this suggestion.

which we believe enrich further the tool kit of analysts interested in nonmarket valuation.

[Received August 2007; accepted February 2008.]

Appendix

Utility in WTP Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps

The following tables collect auxiliary estimates for the above mentioned study.

Table A.1 reports the summary statistics for the ML estimates of the basic MNL model. As can be seen by comparing the log-likelihood value at the maximum with those reported in the paper and obtained by MSL (or simulated at the posterior in the case of HB), the RPL models produce a large improvement in fit.

Table A.2 reports the estimated ML parameters of the MNL model. The WTP estimates show similar magnitudes to the means of their RPl counterparts.

Table A.3 reports the estimated Cholesky matrix for the MSL estimate in WTP space. From this one can derive the variance-covariance matrix of the multivariate distribution of WTPs, and the associated correlation matrix.

Table A.4 reports the estimated Cholesky matrix for the MSL estimate in preference space. From one can derive the variance-covariance matrix of the multivariate distribution of taste intensities for site attributes. These, along with the mean estimates can this be used to simulate draws which in turn can be sued to compute WTP distributions.

Table A.5 reports the estimates of the WTP space model with bounded distributions for ln(λ) and number of Alpine shelters.

Table A.6 reports the estimates of the preference space model with bounded distributions for ln(λ) and number of Alpine shelters.

Table A.7 reports the correlations of the latent variables for both the bounded specifications.

Table A.1. Summary of MNL Model

Model	:	Multinomial Logit
Number of estimated parameters	:	7
Number of observations	:	9,221
Number of individuals	:	9,221
Null log-likelihood	:	−26,652.12
Init log-likelihood	:	−50,407.59
Final log-likelihood	:	−21,754.39
Likelihood ratio test	:	9,795.45
Rho-square	:	0.1838
Adjusted rho-square	:	0.1835
Final gradient norm	:	+1.739e−003
Variance-covariance	:	From analytical hessian

Table A.2. Estimates of MNL Model

Variable Number	Description	Coeff. Estimate	Robust Asympt. Std. Error	*t*-Stat	*p*-Value	*WTP*
1	Travel cost	−0.2835	0.0057	−49.3	0.00	
2	Degree of difficulty	−0.5600	0.0208	−26.8	0.00	−1.975
3	Ferrata	−0.0793	0.0046	−17.3	0.00	−0.280
4	% of easy trails	0.0157	0.0013	12.3	0.00	0.055
5	Alpine shelters	0.0885	0.0032	27.2	0.00	0.312
6	% of hard trails	0.0797	0.0033	24.1	0.00	0.281
7	Prealps ASC	−0.8917	0.0619	−14.4	0.00	−3.145

Table A.3. Cholesky Matrix from MSL Estimates in *WTP* Space

Parameters	$\ln \lambda$	Degree of Diff.	Ferrata	% Easy Trails	Alpine Shelters	% Hard Trails	Prealps ASC
$\ln \lambda$	−0.043						
	(21.5)						
Degree of difficulty	0.193	−2.977					
	(1.7)	(19.4)					
Ferrata	0.067	−0.291	0.220				
	(2.9)	(11.1)	(9.3)				
% of easy trail	−0.007	0.060	0.015	−0.043			
	(1.1)	(7.5)	(1.3)	(12.5)			
Alpine shelters	−0.037	0.148	−0.149	−0.003	−0.081		
	(2.2)	(8.8)	(8.5)	(0.3)	(7.0)		
% of hard trail	0.011	0.279	0.070	−0.038	0.024	−0.244	
	(0.5)	(11.2)	(2.2)	(4.3)	(2.1)	(10.5)	
Prealps ASC	2.520	−4.517	2.449	1.605	−0.014	−1.312	2.490
	(7.9)	(11.4)	(7.2)	(5.3)	(1.6)	(4.2)	(14.2)

Note: |z-values| in brackets.

Table A.4. Cholesky Matrix from MSL Estimates in Preference Space

Parameters	$\ln \lambda$	Degree of Diff.	Ferrata	% Easy Trails	Alpine Shelters	% Hard Trails	Prealps ASC
$\ln \lambda$	0.92						
	(20.4)						
Degree of difficulty	−0.19	0.70					
	(3.9)	(19.7)					
Ferrata	−0.06	0.05	−0.08				
	(5.5)	(6.5)	(7.3)				
% of easy trail	0.00	0.00	0.00	0.01			
	(0.4)	(1.3)	(1.0)	(0.7)			
Alpine shelters	0.06	−0.02	0.06	−0.00	−0.00		
	(8.1)	(3.5)	(9.6)	(0.1)	(0.7)		
% of hard trail	0.01	−0.01	−0.02	0.00	−0.03	0.06	
	(2.8)	(2.2)	(2.0)	(0.9)	(6.1)	(10.8)	
Prealps ASC	−1.29	0.92	−0.34	−0.07	−0.02	1.08	−0.01
	(7.3)	(8.2)	(2.4)	(0.5)	(4.0)	(14.8)	(0.04)

Note: |z-values| in brackets.

Table A.5. Estimates of *WTP* Space Model with S_b. $\ln \mathcal{L} - 20,177.50$

Site Attributes	HB Estimates				
DISTR.	PARAM.	Mean	St.dev.	Var.	St.dev. Var.
$S_b[0, 2]$	$\lambda \times c$	0.292	0.188	0.604	0.076
Normal	Degree of difficulty	−3.341	3.359	10.957	1.624
Normal	Ferrata	−0.450	0.419	0.176	0.024
Normal	% of easy trails	0.119	0.156	0.023	0.003
$S_b[0, 1.5]$	Alpine shelters	0.417	0.240	0.632	0.116
Normal	% of hard trails	0.429	0.402	0.162	0.022
Normal	Prealps ASC	−5.952	8.132	62.996	8.949

Table A.6. Estimates of Preference Space Model with S_b. ln $\mathcal{L}$ − 20, 706.25

Site Attributes	HB Estimates				
DISTR.	PARAM.	Mean	St.dev.	Var.	St.dev. Var.
$S_b[0, 2]$	λ	0.383	0.291	1.076	0.128
Normal	Degree of difficulty	−0.920	0.932	0.877	0.105
Normal	Ferrata	−0.151	0.153	0.023	0.002
Normal	% of easy trails	0.031	0.076	0.006	0.000
$S_b[0, 2]$	Alpine shelters	0.133	0.097	0.572	0.082
Normal	% of hard trails	0.119	0.142	0.021	0.002
Normal	Prealps ASC	−2.196	2.284	4.937	0.613

Table A.7. Correlations from HB Estimates of Models with Bounded Distributions

Site Attributes	Correlation Matrix for Random *WTP* for *WTP* Space Model with S_b.						
PARAM.	$\ln\hat{\lambda}$	Deg. of Diff.	Ferrata	% Easy Trails	Alp. Shelters	% Hard Trails	Prealps
$\ln\hat{\lambda}$	1	−0.2737	−0.1885	0.042	0.079	0.046	−0.4014
Degree of diff.	−0.2737	1	0.6441	−0.3248	−0.4825	−0.5496	0.7706
Ferrata	−0.1885	0.6441	1	−0.2434	−0.7887	−0.3309	0.6533
% of easy trails	0.042	−0.3248	−0.2434	1	0.1551	0.6308	−0.4181
Alpine shelters	0.079	−0.4825	−0.7887	0.1551	1	0.1682	−0.5409
% of hard trails	0.046	−0.5496	−0.3309	0.6308	0.1682	1	−0.3489
Prealps ASC	−0.4014	0.7706	0.6533	−0.4181	−0.5409	−0.3489	1
Site Attributes	Correlation Matrix for Utility Coefficients of Preference Space Model with S_b						
PARAM.	$\ln\hat{\lambda}$	Deg. of Diff.	Ferrata	% Easy Trails	Alp. Shelters	% Hard Trails	Prealps
$\ln\hat{\lambda}$	1	−0.1956	−0.3122	0.0237	0.6219	0.1775	−0.4204
Degree of diff.	−0.1956	1	0.4955	−0.0801	−0.4014	−0.3038	0.6971
Ferrata	−0.3122	0.4955	1	−0.1066	−0.5936	−0.2441	0.6077
% of easy trails	0.0237	−0.0801	−0.1066	1	0.1282	0.3609	−0.2097
Alpine shelters	0.6219	−0.4014	−0.5936	0.1282	1	0.2207	−0.6711
% of hard trails	0.1775	−0.3038	−0.2441	0.3609	0.2207	1	−0.2447
Prealps ASC	−0.4204	0.6971	0.6077	−0.2097	−0.6711	−0.2447	1

References

Albert, J., and S. Chib. 1993. "Bayesian Analysis of Binary and Polychotomous Response Data." *Journal of the American Statistical Association* 88:669–79.

Allenby, G. 1997. "An Introduction to Hierarchical Bayesian Modeling." Tutorial notes, Advanced Research Techniques Forum, American Marketing Association.

Allenby, G., and P. Rossi. 1999. "Marketing Models of Consumer Heterogeneity." *Journal of Econometrics* 89:57–78.

Bierlaire, M. 2002. "An Introductory Tutorial to BIOGEME. Federal Polytechnic School of Lausanne, Switzerland." http://rosowww.epfl.ch/mbi/biogeme.

——. 2003. "BIOGEME: A Free Package for the Estimation of Discrete Choice Models." Paper presented at the 3rd Swiss Transport Research Conference Monte Verita, Ascona, March 2003.

Bockstael, N.E., W.M. Haneman, and C.L. Kling. 1987. "Estimating the Value of Water Quality Improvements in a Recreational Demand Framework." *Water Resources Research* 23:951–60.

Bowman, A.W., and A. Azzalini. 1997. *Applied Smoothing Techniques for Data Analysis: The Kernel Approach with S-Plus Illustrations*. Oxford, UK: Oxford University Press.

Breffle, W.S., and E.R. Morey. 2000. "Investigating Preference Heterogeneity in a Repeated Discrete-choice Recreation Demand model of Atlantic Salmon Fishing." *Marine Resource Economics* 15:1–20.

Cameron, T. 1988. "A New Paradigm for Valuing Non-market Goods using Referendum Data:

Maximum Likelihood Estimation by Censored Logistic Regression." *Journal of Environmental Economics and Management* 15:355–79.

Cameron, T., and M. James. 1987. "Efficient Estimation Methods for Closed-ended Contingent Valuation Survey Data." *Review of Economics and Statistics* 69:269–76.

Casella, G., and E. George. 1992. "Explaining the Gibbs Sampler." *The American Statistician* 46:167–74.

Geweke, J. 1992. "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments." In J. Bernardo, J. Berger, A. Dawid, and F. Smith, eds. *Bayesian Statistics*. New York: Oxford University Press, pp. 169–93.

Goett, A., K. Train, and K. Hudson. 2000. "Customer Choice Among Retail Energy Suppliers: The Willingness-to-Pay for Service Attributes." *Energy Journal* 21:1–28.

Hanemann, W.M. 1984. "Welfare Evaluations in Contingent Valuation Experiments with Discrete Response Data." *American Journal of Agricultural Economics* 66:332–41.

——. 1989. "Welfare Evaluations in Contingent Valuations Experiments with Discrete Response Data: A Reply." *American Journal of Agricultural Economics* 71:1057–61.

Hess, S., K.E. Train, and J.W. Polak. 2006. "On the Use of a Modified Latin Hypercube Sampling (MLHS) Method in the Estimation of a Mixed Logit Model for Vehicle Choice." *Transportation Research B* 40:147–67.

Koop, G. 2003. *Bayesian Econometrics*. New York: Wiley.

Lawrence, C., J. Zhou, and A. Tits. 1997. "User's Guide for CFSQP Version 2.5: A C Code for Solving (large scale) Constrained Nonlinear (minimax) Optimization Problems, Generating Iterates Satisfying all Inequality Constraints." Technical Report TR-94-16r1, Institute for Systems Research, University of Maryland, College Park, MD 20742, 1997.

Layton, D.F., and G. Brown. 2000. "Heterogeneous Preferences Regarding Global Climate Change." *The Review of Economics and Statistics* 82(4):616–24.

Layton, D., and R. Levine. 2005. "Bayesian Approaches to Modeling Stated Preference Data." In R. Scarpa and A. Alberini, eds. *Applications of Simulation Methods in Environmental and Resource Economics*. Dordrecht, The Netherlands: Springer Publisher, chapter 10, pp. 187–208.

McConnell, K. 1995. "Consumer Surplus from Discrete Choice Models." *Journal of Environmental Economics and Management* 29:263–70.

McCulloch, R., and P. Rossi. 1994. "An Exact Likelihood Analysis of the Multinomial Probit Model." *Journal of Econometrics* 64:207–40.

Morey, E., and K. Rossmann. 2003. "Using Stated-Preference Questions to Investigate Variations in Willingness to Pay for Preserving Marble Monuments: Classic Heterogeneity, Random Parameters, and Mixture Models." *Journal of Cultural Economics* 27(3/4):215–29.

Morey, E., R. Rowe, and M. Watson. 1993. "A Repeated Nested-Logit Model of Atlantic Salmon Fishing." *American Journal of Agricultural Economics* 75:578–92.

Provencher, B., K. Baerenklau, and R.C. Bishop. 2002. "A Finite Mixture Model of Recreational Angling with Serially Correlated Random Utility." *American Journal of Agricultural Economics* 84(4):1066–75.

Revelt, D., and K. Train. 1998. "Mixed Logit with Repeated Choices." *Review of Economics and Statistics* 80(4):647–57.

Rigby, D., and M. Burton. 2006. "Modeling Disinterest and Dislike: A Bounded Bayesian Mixed Logit Model of the UK Market for GM Food." *Environment and Resource Economics* 33:485–509.

Rossi, P.E., G. Allenby, and R. McCulloch. 2005. *Bayesian Statistics and Marketing*. New York: Wiley.

Ruud, P. 1996. "Simulation of the Multinomial Probit Model: An Analysis of Covariance Matrix Estimation." Working paper, Department of Economics, University of California.

Sawtooth Software. 1999. "The CBC/HB Module for Hierarchical Bayes." Technical paper available on the web site of Sawtooth Software.

Scarpa, R., and M. Thiene. 2005. "Destination Choice Models for Rock Climbing in the Northeastern Alps: A Latent-Class Approach Based on Intensity of Preference." *Land Economics* 85(3):426–44.

Scarpa, R., M. Thiene, and K. Train. 2008. AJAE Appendix: "Utility in *WTP* Space: A Tool to Address Confounding Random Scale Effects in Destination Choice to the Alps." Available at: http://agecon.lib.umn.edu/.

Sonnier, G., A. Ainslie, and T. Otter. 2007. "Heterogeneity Distributions of Willingness-to-Pay in Choice Models." *Quantitative Marketing Economics* 5(3):313–31.

Train, K. 1998. "Recreation Demand Models with Taste Differences over People." *Land Economics* 74:185–94.

——. 1999. "Mixed Logit Models for Recreation Demand." In J. Herriges and C. Kling, eds.

Valuing Recreation and the Environment, pp. 121–40. Northampton, MA: Edward Elgar.

——. 2003. *Discrete Choice Methods with Simulation*. New York: Cambridge University Press.

Train, K., and G. Sonnier. 2005. "Mixed Logit with Bounded Distributions of Correlated Partworths." In R. Scarpa and A. Alberini, eds. *Applications of Simulation Methods in Environmental and Resource Economics*. Dordrecht, The Netherlands: Springer Publisher, chapter 7, pp. 117–34.

Train, K., and M. Weeks. 2005. "Discrete Choice Models in Preference Space and Willing-to-Pay space." In R. Scarpa and A. Alberini, eds. *Applications of Simulation Methods in Environmental and Resource Economics*. Dordrecht, The Netherlands: Springer Publisher, chapter 1, pp. 1–16.