Marine Resource Economics, Volume 18, pp. 1–14
Printed in the U.S.A. All rights reserved

0738-1360/00 $3.00 + .00

An Application of the Kuhn-Tucker Model to the Demand for Water Trail Trips in North Carolina

DANIEL J. PHANEUF
CHRIS SIDERELIS
North Carolina State University

Abstract *The Kuhn-Tucker demand model is an attractive, recent addition to the methods available for analyzing seasonal, multiple-site recreation demand data. We provide a new application of the approach to the demand for sea paddling trips in eastern North Carolina and calculate welfare measures for changes in site characteristics. In addition, we present a non-technical, intuitive overview of the model and a stepwise derivation of the estimation and welfare calculation algorithms.*

Key words Kuhn-Tucker model, seasonal recreation demand, paddling.

JEL Classification Codes Q26, C34, C35.

Introduction

In this paper we present a non-technical, stepwise overview of the Kuhn-Tucker (KT) demand system model as it is used for estimating seasonal recreation demand and calculating welfare measures. We demonstrate the workings of the model with a new application to water trail paddling trips in eastern North Carolina. In addition to presenting the application, our primary goal is to increase the accessibility of the KT approach to researchers analyzing seasonal, multiple-site recreation data. As such, the paper adopts a tutorial approach. The derivation of the estimator is presented in stepwise detail, and the numerical algorithm necessary for computing welfare effects is derived and explained.

The recreation demand literature has provided many useful tools for estimating preferences for recreation resources. The most widely applied is the random utility maximization (RUM) model, which has proven to be an effective tool for understanding the substitutability between recreation sites and their attributes on a given choice occasion. Because the unit of observation in the model is a single choice occasion, however, simple RUM models are not well suited for modeling behavior over a longer time horizon, such as a recreation season. In response to this, researchers have pursued methods capable of describing seasonal demands and the associated corner solutions, as characterized by zero levels of consumption of some of the available goods.

Daniel J. Phaneuf is an assistant professor in the Department of Agricultural and Resource Economics, North Carolina State University, Box 8109, Raleigh, NC 27695, email: dan_phaneuf@ncsu.edu. Chris Siderelis is a professor in the Department of Parks, Recreation, and Tourism Management, North Carolina State University, Box 8004, Raleigh, NC 27695, email: chris_siderelis@ncsu.edu.

This research was supported in part by grants from North Carolina Sea Grant and Watersports Company. We also acknowledge the contributions of Joe Herriges and Roger von Haefen in designing the programs used for estimation and welfare calculations in Kuhn-Tucker recreation demand models.

The Kuhn-Tucker model is an example of this trend.[1] The KT approach models the demand for trips to a set of recreation sites over the time horizon of a season. It combines aspects of a traditional system of demand equations approach with the RUM model's emphasis on capturing the substitutability between sites and their attributes while allowing for zero consumption. Unlike the RUM model, the KT model allows for generalized corner solutions. Recreation visitors may visit only a subset of the available sites, yet visit these sites multiple times over the season. In this sense, the model integrates the two aspects of choice within a single framework: the choice of which sites to visit during a season, and how many trips to make to each site. Importantly, the characterization of this decision process is made in a manner consistent with utility theory that allows a smooth integration of the behavioral and econometric models.

This utility-consistent integration of the behavioral and econometric models has been acknowledged as an attractive feature of the KT framework. Furthermore, recent modeling advances by von Haefen, Phaneuf, and Parsons (forthcoming) have demonstrated that the approach can be applied with an arbitrarily large number of sites. To date, however, there have been relatively few applications in the recreation literature,[2] and existing papers tend to examine technical aspects of the model rather than the intuition and basic steps necessary for application in other settings. This paper augments the existing literature with a new application and a comparably non-technical presentation. In the following sections, we present an intuitive overview of the model followed by a description of the eastern North Carolina paddling application. A step-wise derivation of the estimation and welfare calculation algorithms and presentation of results follows. We conclude with discussion and suggestions for further research.

Overview of KT Model

In this section, we adopt a tutorial approach to explaining the workings of the KT model. As with the RUM model, we begin by specifying a random utility function. The consumer's direct utility function is $u(x, z; q, \varepsilon, \gamma)$, where u has the normal curvature properties, x is trips to a recreation site, q is a vector representing characteristics of that site, and z is spending on all other goods, with price normalized to one. For pedagogical purposes, we consider only one recreation good, although all derivations generalize to multiple sites. The term ε is known to the individual, but random to the researcher; thus the problem is stochastic from the analyst's point of view, and our intent will be to characterize the probability of the consumer's outcome. Finally, the term γ represents parameters of the utility function that are to be estimated.

The model assumes consumers maximize utility over a season subject to their

[1] Other models capable of modeling seasonal recreation demand include repeated random utility models (*e.g.*, Morey 1999), linked models (*e.g.*, Parsons, Jakus, and Tomasi 1999), systems of count data demand models (*e.g.*, Englin, Boxall, and Watson 1998), and share models (*e.g.*, Morey, Breffle, and Green 2001). Parsons, Jakus, and Tomasi (1999), von Haefen and Phaneuf (forthcoming), and Herriges, Kling, and Phaneuf (1999) provide discussion and comparisons of subsets of these models.

[2] Wales and Woodland (1983) and Lee and Pitt (1986) were the first to suggest the estimation technique, and Bockstael, Hanemann, and Strand (1986) were the first to discuss the model in the context of recreation. Phaneuf, Kling, and Herriges (2000) provide an application of the model to Great Lakes fishing in Wisconsin, Phaneuf and Herriges (1999) and von Haefen and Phaneuf (forthcoming) estimate the model for visits to wetlands in Iowa, and von Haefen, Phaneuf, and Parsons (forthcoming) estimate the model for beach visitation in the mid-Atlantic region.

budget constraint and a non-negativity constraint (*i.e.*, the number of trips can be zero but not negative). Stating the formal maximization problem and first-order conditions for this problem is useful because it motivates the link between the behavioral and empirical models. The choice problem is given by:

$$\max_{x,z} u(x, z; q, \varepsilon, \gamma) \qquad s.t. \qquad y = z + xp, \qquad x \geq 0, \tag{1}$$

where y is annual income and p is the price (travel cost, access fees, *etc*.) of visiting the recreation site. Because of the non-negativity constraint, the first-order conditions for maximization take the form of Kuhn-Tucker conditions. Assuming $z > 0$, the KT first-order and complementary slackness conditions are:

$$\frac{u_x(x, y - px; q, \varepsilon, \gamma)}{u_z(x, y - px; q, \varepsilon, \gamma)} \leq p \tag{2}$$

$$x \geq 0$$

$$x\left[u_x(x, y - px; q, \varepsilon, \gamma) - pu_z(x, y - px; q, \varepsilon, \gamma)\right] = 0,$$

where subscripts on the utility function denote a first partial derivative. These equations show that if positive trips are chosen, it will be the case that the marginal rate of substitution between trips and other spending is equal to the price ratio, while this equality breaks down if x is equal to zero. If positive trips are made, the number of trips is determined by the demand equation $x(p, y, q, \gamma, \varepsilon)$, the form of which derives from solving the first equation in (2). Our task now is to convert these equations, derived from utility theory, into a form that is useful for stating the probability of observing individual behavior.

This is accomplished by assuming the utility function is of a convenient and particular form so that the first-order conditions in equation (2) can be readily rearranged to:

$$\varepsilon \leq g(x, p, y, q, \gamma) \tag{3}$$

$$x \geq 0$$

$$x\left[\varepsilon - g(x, p, y, q, \gamma)\right] = 0,$$

where $g(x, p, y, q, \gamma)$ denotes the solution to the equation $u_x(x, y - px; q, g, \gamma) - pu_z(x, y - px; q, g, \gamma) = 0$. This step makes transparent the link to the estimation of the model. Given an assumption on the form of $f(\varepsilon)$, the distribution of ε, we are able to characterize the probability of observing an individual's outcome by noting that the probability of a corner solution (*i.e.*, $x = 0$) is $pr(X = 0) = pr[\varepsilon < g(0, z, q, \gamma)]$, while the probability of observing some level of trips conditional on positive trip taking is given by $pr(X = x) = pr[\varepsilon = g(x, z, q, \gamma)]$. Dependent upon the choice of error distribution, the form of these probabilities can be relatively simple or quite complex. In either case, a probability can be calculated for each individual in the sample and maximum likelihood used to recover estimates of the utility function parameters, γ.

Recovery of estimates of γ provides a characterization of the recreation user's preference function up to the unobserved error term. As noted above, this implies that the form of the demand equation is $x(p, y, q, \gamma, \varepsilon)$ if positive consumption occurs, and zero otherwise. In other words, the presence of the binding non-negativity constraint conditions the form of the demand relation based on the *demand regime* — the pattern of positively consumed goods. The binding non-negativity constraint also implies that the form of the indirect utility function, which we are accustomed to working with in applied welfare analysis, will also depend on the demand regime. In our simple example, we recover an indirect utility function conditional on x being positive, and correspondingly an indirect utility function conditional on $x = 0$, simply by plugging the functions $x = x(p, y, q, \gamma, \varepsilon)$ and $x = 0$ into the direct utility function given above. The overall, unconditional indirect utility is then the maximum over these two conditional functions. Formally we can state this as:

$$V(p, y; q, \varepsilon, \gamma) = \max\left\{v_1(p, y, q, \varepsilon, \gamma), v_0(y, \varepsilon, \gamma)\right\}, \tag{4}$$

where v_1 and v_0 are the conditional indirect utility functions for x positive and zero, respectively. Aside from being important for many policy-relevant calculations, stating the indirect utility function in this fashion is useful in that it allows us to make a final link and comparison to the more familiar random utility model. The random utility model assumes recreation visitors make a discrete choice over a portfolio of site options with the goal of maximizing utility. The conditional indirect utility for each discrete option is specified as a function of the price and quality characteristics. The model's description ends here, because this choice involves a single visit at a point in time. The KT analog to this is equation (4): the consumer makes a discrete choice over a set of available demand patterns (*i.e.*, two in this case, but 2^M in the general case of M available sites) to maximize utility. Each conditional indirect utility is a function of the prices of the visited sites and site characteristics. The KT model goes a step further in that conditional on the choice of demand pattern, the associated demand equation informs us of the number of trips over the course of the season. In this sense, the KT model can be thought of as a generalized random utility model that preserves the link to utility theory while allowing us to characterize the season-long, multiple site behavior (and the associated corner solutions) of recreation visitors.

Equation (4) provides a characterization of preferences for the recreation sites and attributes up to the unobserved error term. We turn our attention now to how it can be used, particularly for applied welfare analysis. Typically, we are interested in a measure of the compensating variation (CV) for a change in price and/or quality. A well-known general definition for compensating variation implicitly defines CV in $V(p^0, q^0, y) = V(p^1, q^1, y - CV)$. This form can be used to define CV from our preference characterization by:

$$\begin{aligned} &\max\left\{v_1(p^0, y, q^0, \varepsilon, \gamma), v_0(y, \varepsilon, \gamma)\right\} \\ &= \max\left\{v_1(p^1, y - CV, q^1, \varepsilon, \gamma), v_0(y - CV, \varepsilon, \gamma)\right\}. \end{aligned} \tag{5}$$

Two aspects of equation (5) are apparent upon inspection. First, the presence of the random term implies CV will be a random variable. We are ultimately interested in estimating its expectation for each individual, as in the case of a random utility model. Second, no closed-form solution will exist for the expectation of CV. In this aspect, the model departs from the standard linear-in-income RUM, which provides a convenient, closed-form solution for the expectation of CV.[3] Lacking a closed-

[3] It is similar, however, to the nonlinear-in-income random utility model, which also lacks a closed form solution for willingness to pay. See Herriges and Kling (1999) for a description of this generalization of the standard random utility model and techniques for calculating willingness to pay.

form solution, numerical methods are necessary to calculate an expectation for *CV*. While this can be computationally intense, the intuition of how this is done is fairly straightforward. To begin, multiple realizations for ε are sampled from $\hat{f}(\varepsilon)$, the estimated distribution for the error. Two possibilities exist for obtaining values of ε. The first samples directly from $\hat{f}(\varepsilon)$ unconditionally. The second, recently suggested by von Haefen (forthcoming), samples values for ε conditional on the individual's revealed choice under baseline conditions. In other words, the error is drawn such that the model predicts the person's revealed behavior perfectly under baseline conditions. This approach has the advantage of limiting the support of the distribution of unobserved heterogeneity to the range that is observed in the sample. In practice, conditional sampling means that welfare estimates are less sensitive to extreme draws of the error that may predict implausible economic behavior.

Given a sampled value of the error and a starting guess for *CV*, the left and right sides of equation (5) are calculated. Note this involves solving for the value of each conditional indirect utility function, given the error draw, and selecting the maximum as the value of utility. The guess for *CV* is then updated, and the process is repeated until a value for income adjustment is located that equates the utility levels under the initial and new values for price and quality. This updating occurs in practice via the use of numerical bisection to update the income adjustment. Importantly, the calculation allows for *regime switching*. That is, the person can adjust the pattern of visitation as well as the number of visits in response to the new price and quality. Having completed this for a single draw of the error, the process repeats for additional error realizations. The mean of the income adjustments generated in this fashion provides an estimate of $E(CV)$ for a single individual in the sample. The mean or median of the sample provides an estimate of the welfare effect for the population.

Application — Eastern North Carolina Paddle Data

Our application focuses on using the Kuhn-Tucker model to estimate recreation use and recreation benefits of water trails in eastern North Carolina by recreation paddlers. We model the demand for trips to eight designated paddle zones as a function of travel costs and individual and site-specific attributes, including public amenities, such as miles of marked trails and number of camp sites, and a measure of water quality.

In North Carolina, 13% of households in the state participate in some type of padding activity (NC DEHN 1995), and recreation paddling is of increasing importance in the eastern part of the state. The study area comprises the tidal flat area. Tourism is a major industry in this region, with estimated tourism expenditures on the order of $2,297 million in 1998 (NC Department of Commerce 1998). This region contains many bays, estuaries, islands, and inlets, all of which comprise ideal conditions for sea kayaking and canoeing. The state has designated 1,189 miles of water trails in this area. Recreation managers have created marked trails and routes in the area and have begun the process of creating and improving infrastructure (*i.e.*, launch points, amenities) to support expanded paddling opportunities. The area is under pressure, however, from competing uses, such as development and agricultural operations, that threaten environmental conditions. Information on trail usage and values is useful for informing decisions on continuing efforts to build and maintain the water trail system. As a preliminary step in this process, a survey of canoeing and kayaking visitors to the area was conducted for the 2000 paddling season.

Three methods were used to solicit survey respondents. First, a letter requesting participation in the study was sent to 622 individuals whose names appeared on a mailing list requesting information about coastal paddling from resource development and state park sources. Second, commercial paddling businesses geographically dispersed throughout the coastal area were contacted and agreed to cooperate by contacting customers on their mailing lists. Finally, the sampling process was supplemented by posting the project description and request for survey users on paddle club, association, outfitter, and other email list servers in North and South Carolina and Virginia. Six hundred-one individuals who went canoeing or kayaking during 2000 agreed to participate in the study.

Each of these individuals received a mail survey. The paddling survey elicited information about the number of paddling trips, trip expenses, group characteristics, and trip purposes. The final convenience sample consisted of 491 returned questionnaires. Overall, 426 questionnaires were sufficiently complete to be usable for economic impact and demand analyses. Clearly, this sample is not based on a random population survey, and as such, care should be taken when extending results beyond the sample. Nonetheless, the data set is representative of what is often available to recreation analysts and, therefore, provides a suitable application for the KT model.

The survey solicited information on visits to nine aggregate areas, corresponding to county groupings, in eastern North Carolina during 2000 for paddling activities. For modeling purposes, we have defined eight paddling areas which serve as our site definition for this application. These areas correspond to the division of the eastern part of the state into paddle zones for which marked trails and maps are available online and via requests to the state tourism office.[4] On average, respondents visited two of the available sites during the period of interest and took an average of 9.84 total trips to the region. The median number of total trips to the region is 4. Corner solutions are the typical outcome, and a model allowing for non-consumption is needed to analyze this type of data. Table 1 provides a summary of the sample specifics and percentage of corner solutions.

Out-of-pocket travel costs for accessing each site by respondents were computed from round-trip mileage estimates to paddling areas from the primary addresses of respondents with the software product ZIPFIP (Hellerstein *et al*. 1996). The direct travel cost was the round-trip mileage from a resident's home to each of the paddling areas multiplied by $0.14 per mile. Rather than including the opportunity cost of travel time as part of the travel cost estimate, we follow an alternative approach suggested by Shaw and Feather (1999) based on conditional demands. The fundamental insight behind this suggestion is that short-term recreation demand decisions are made conditional on the longer-term labor supply decisions of each respondent. As such, the demand equations are specified to contain prices based only on the out-of-pocket travel costs and total hours worked in a week. Table 1 provides travel cost summary statistics of access to each of the eight sites and weekly hours worked.

Individual and site characteristics are expected to influence the demand for trips. We include explanatory variables for age and kayak ownership to capture individual effects. The variable *age* is simply the age of the respondent. The variable *kayak* is a dummy variable equal to one if the person owns a sea kayak. Summaries of these variables are also included in table 1. Site characteristics provide a link be-

[4] These eight paddling zones are described in detail on the web at www.ncsu.edu/paddletrails. Maps of trails, descriptions of amenities, and other useful information are available at this site, along with a map of the state that illustrates our choice set definition.

Table 1
Summary Statistics for North Carolina Paddle Data

Site-Specific Summary Statistics

	Site 1	Site 2	Site 3	Site 4	Site 5	Site 6	Site 7	Site 8
Trips	0.74	1.56	0.93	2.20	2.60	1.00	0.30	0.38
	(2.55)	(5.00)	(5.40)	(6.64)	(10.72)	(4.20)	(1.82)	(2.11)
Price	39.34	50.62	40.56	41.70	49.87	45.50	30.64	36.19
	(22.99)	(25.05)	(23.61)	(23.39)	(25.02)	(23.93)	(21.43)	(22.63)
Trail miles	129	0	141	304	351	127	0	158
Water quality	1.75	2	1.6	1.5	2.87	3.5	2.16	3.35
Public camp sites	455	1615	793	1221	1537	379	235	299

Total Trip Summary Statistics

Mean	Standard Deviation	Median	Range
9.84	15.48	4	0–160

Individual Specific Summary Statistics		Sample Summary Statistics	
Income	$76,590	Respondents	426
	(40,094)		
Work hours	37.85	# respondents not visiting study area	47
	(18.42)		
Kayak ownership	0.78	# respondents visiting >1 sites	225
Age	46.61)	# observations visiting all sites	0
	(11.76)		

tween environmental and amenity aspects of the sites and people's visitation decisions. We include one measure of environmental quality and two measures of site amenities in our model.

To characterize water quality in each paddle zone, we calculated a water quality index for each zone based on the Environmental Protection Agency's Index of Watershed Indicators (US EPA 1999). The Index of Watershed Indicators (IWI) provides an index system rating water quality at the level of eight-digit hydrological units (watersheds) throughout the country. This system indexes watersheds with a value between one and six, with one corresponding to the highest water quality. The average IWI values associated with the watersheds contained in each of the paddle zones were used to construct a water quality variable for each area, labeled *water* in our application. The values of *water* are given for each site in table 1. Inspection of these values displays variation in water quality across the study region.

Our two site amenities include miles of marked water trails in the area (*miles*) and the number of public campsites in the area (*camp*). These variables were measured and provided by contract as part of the survey process and are summarized in table 1.[5] *A priori* we expect *trails* and *camp* to positively influence trips. Again, variation in the values of these amenity variables exists across areas, suggesting it may be possible to identify behavioral impacts associated with these characteristics.

[5] The variable *trails* measures marked water trails that are described on the website in the preceding footnote and in eastern North Carolina tourism paddling literature. Note that currently paddle zones 2 and 7 are lacking marked trails, although there are many other unmarked waterways available in these and the other zones for recreational paddling.

Empirical Specification and Results

Derivation of Estimator and Results

Application of the KT model requires specification of the direct utility function and distribution for the error vector. In specifying the individual's direct utility function we follow existing literature (*e.g.*, Phaneuf, Kling, and Herriges 2000) and assume an additively separable functional form. Specifically, utility is:

$$u = \sum_{i=1}^{8} \Psi_i(\mathbf{s}, \varepsilon_i) \ln\left[x_i \quad \varphi(\mathbf{q}_i) + \theta\right] + \ln(z), \tag{6}$$

where x_i is the number of trips taken to the ith site, z is spending on all other goods, $\mathbf{s}$ denotes the individual specific characteristic variables, $\mathbf{q}_i$ denotes the site-specific quality variables, ε_i is a site-specific random error, and θ is a parameter to be estimated. The term Ψ_i is constructed to be a positive aggregator function of the individual-specific variables and the random error term:

$$\Psi_i(\mathbf{s}, \varepsilon_i) = \exp(\delta_0 + \delta_{age} age + \delta_{kayak} kayak + \delta_{hours} hours + \varepsilon_i), \qquad i = 1, \ldots, 8 \tag{7}$$

where the δ's are parameters to be estimated. The term $\varphi(\mathbf{q}_i)$ is a strictly positive aggregator function of the site i specific quality variables, given by:

$$\varphi(\mathbf{q}_i) = \exp(\gamma_{water} water_i + \gamma_{trails} trail_i + \gamma_{camp} camp_i), \tag{8}$$

where the γ's are parameters to be estimated.

The utility function is specified such that weak complementarity between the recreation sites and their associated quality measures holds. Note that if no visits are made to site i, $\varphi(\mathbf{q}_i)$ drops out of the direct utility function, and changes in $\mathbf{q}_i$ have no impact on the level of utility. This implies that the welfare effects estimated from this model contain only use value, and the demand equations will be functions of the visited-site quality variables only.

Maximization of equation (6) subject to income and non-negativity constraints with $z > 0$ implies first-order (KT) conditions:

$$\frac{\Psi_i(\mathbf{s}, \varepsilon_i)}{x_i + \alpha_i} \leq \frac{p_i}{y - \sum_{i=1}^{8} p_i x_i}, \qquad i = 1, \ldots 8, \tag{9}$$

where the substitution $\alpha_i = \theta/\varphi(\mathbf{q}_i)$ has been made. Recall from equation (3) that estimation in the KT model follows directly from the utility maximization conditions. Equation (9), therefore, forms the basis for deriving the estimating equations. Substituting equation (7) into equation (9) and solving for the error term in each equation allows the first-order utility maximization conditions to be rewritten as:

$$\varepsilon_i \leq g_i, \qquad i = 1, \ldots 8, \tag{10}$$

where

$$g_i = \ln p_i + \ln(x_i + \alpha_i) - \ln\left(y - \sum_{k=1}^{8} p_k x_k\right) - \delta_0 - \delta_{age} age - \delta_{kayak} kayak - \delta_{hours} hours.$$

This version of the first-order conditions is the key link between the behavioral and econometric aspects of the model. If x_i is observed to be positive, the condition in equation (10) is $\varepsilon_i = g_i$, while if x_i is zero, the condition is $\varepsilon_i \le g_i$. By assuming a distribution for the error terms, we can state the probability of observing this outcome and derive estimating equations.

To make this operational, we assume that ε_i is an identically and independently distributed type I extreme value for each of the eight sites.[6] Although restrictive, this assumption is useful in that it allows us to state a closed form for the likelihood function that greatly simplifies estimation. The assumption of independent type I extreme value errors implies that the log of the probability (equivalently the individual contribution to the log-likelihood function) of observing an individual's outcome is:

$$\ln pr(\mathbf{x}) = -\sum_{k=1}^{8} I_{x_k>0} \times \frac{g_i}{v} - \sum_{k=1}^{8} \exp\left[-\frac{g_i}{v}\right] - \sum_{k=1}^{8} I_{x_k>0} \times \ln(v) + \ln(|J|), \qquad (11)$$

where $I_{x>0}$ is an indicator function equal to one if $x_i > 0$, v is the extreme value scale parameter, and $|J|$ is the Jacobian transformation from ε to $\mathbf{x}$.[7]

A Jacobian transformation is needed whenever the statement of probability involves a change of variables. In this case, we have specified $f(\varepsilon)$ — the distribution function for the errors — when we are ultimately interested in the probability of observing $\mathbf{x}$. Bain and Engelhardt (1992, p. 206) describe the necessary change of variable technique. Since $\mathbf{x}$ is a function of the underlying error terms, the distribution function for $\mathbf{x}$ is given by $f(\mathbf{x}) = f[\varepsilon(\mathbf{x})]|J|$, where

$$|J| = \begin{vmatrix} \frac{\partial \varepsilon_1}{\partial x_1} & \cdots & \frac{\partial \varepsilon_1}{\partial x_8} \\ \vdots & \ddots & \vdots \\ \frac{\partial \varepsilon_8}{\partial x_1} & \cdots & \frac{\partial \varepsilon_8}{\partial x_8} \end{vmatrix},$$

is the absolute value of the determinant of the matrix of first partial derivatives between ε and $\mathbf{x}$. The actual dimension and form of the Jacobian matrix will depend on the number and pattern of positively consumed trips that is observed. The elements of the Jacobian can be found by differentiating each of the eight equations implied by the right-hand-side of equation (10) with respect to each element of $\mathbf{x}$. The linear form in equation (10) assures relatively simple terms for each element of the matrix. For example, the diagonal elements are:

---

[6] The probability density and cumulative distribution functions for the type I extreme value distribution are given by:

$$f(\varepsilon_i) = \exp\left[-\exp(-\varepsilon_i/v)\right]\exp(-\varepsilon_i/v)/v$$

$$F(\varepsilon_i) = \exp\left[-\exp(-\varepsilon_i/v)\right],$$

where v is the extreme value scale term.

[7] It is well known that the type I extreme value distribution provides the multivariate logit random utility model with its convenient closed form, as it does here for the KT model. In the logit model, the scale parameter, v, is normalized to one, since the qualitative choice nature of the model does not provide scale in the dependent variable from which this term can be identified. In the KT model, the dependent variable (seasonal trips) is quantitative, allowing v to be identified.

$$\frac{\partial \varepsilon_i}{\partial x_i} = \frac{1}{x_i + \alpha_i} + \frac{p_i}{y - \sum_{k=1}^{8} p_k x_k}\,, \tag{12}$$

and the off-diagonal elements are simply:

$$\frac{\partial \varepsilon_i}{\partial x_j} = \frac{p_j}{y - \sum_{k=1}^{8} p_k x_k}\,. \tag{13}$$

These terms form the building blocks for calculation of $|J|$ in equation (11). A closed, but ugly, form solution exists for $|J|$, but actual implementation of the model requires only providing the computer with the correct components of the matrix and allowing the calculation to be made numerically.

Given construction of the probabilities in equation (11) for each respondent in the sample, standard software package search algorithms for maximum likelihood can recover estimates of the utility function parameters. Example code for estimating the model presented here using the computer program GAUSS (with references to the equation numbers given in the derivation) is available upon request.

Estimation of the utility function parameters provides a characterization of the indirect utility function up to the unobserved error term, derived as follows. Conditional on values for the error terms, solving the maximization conditions in equation (9) results in *regime-specific* demand equations of the form:

$$\begin{aligned} x_i^{\omega} &= -\alpha_i + \frac{\Psi_i(\mathbf{s}, \varepsilon_i)}{1 + \sum_{k \in \omega} \Psi_k(\mathbf{s}, \varepsilon_k)} \frac{1}{p_i} \left[y + \sum_{k \in \omega} \alpha_k p_k \right] \quad i \in \omega \\ x_i^{\omega} &= 0 \quad i \notin \omega, \end{aligned} \tag{14}$$

where ω indexes the specific combination of visited sites. Note that the number of trips for all positively consumed sites is given by the top equation in (14), and that the form of this equation is dependent on the demand regime. In our eight-good application, there are 2^8 or 256 unique combinations of demand patterns, including the option of not visiting a site. Substituting the regime-specific demand equations in equation (14) into the direct utility function in equation (6) provides the set of regime-specific indirect utility functions that constitute the elements of the maximum function for the overall indirect utility function described by equation (4). This function is of interest for welfare analysis.

The results of estimation are presented in table 2. The parameter estimates characterize the direct utility function and the associated demand relationships. The signs of the individual characteristic variables for age and kayak ownership suggest older individuals make fewer trips, and people who own their own kayak make more paddling trips to all sites. The parameter estimates for each of these variables are significant at a 5% or better level. In contrast, the hours-worked variable provides no statistical contribution to the characterization of preferences.[8]

[8] Estimating the model using the standard one-third of the wage rate for the opportunity cost of time provided qualitatively similar estimation results in terms of the significant variables.

Table 2
Estimation Results

Parameter	Estimate	Asymptotic t-statistic
θ	3.49	6.78
$\delta_{intercept}$	–7.359	29.59
δ_{age}	–0.0191	5.57
δ_{kayak}	0.155	1.71
δ_{hours}	–0.0002	0.092
$\gamma_{water\ quality}$	–0.175	3.73
$\gamma_{trail\ miles}$	0.001	3.188
$\gamma_{camp\ sites}$	0.0008	11.36
v	0.987	27.88

The signs of the coefficients for the site quality and amenity variables are significant and of the expected sign. We find that water quality is a significant determinant of choice as measured by the water quality index, suggesting that visitors take more trips to sites with better water quality. Likewise, the presence of more water trail miles and public campsites cause paddlers to frequent these sites more often, all else equal.

Welfare Analysis

As noted above, the preference characterization recovered from the parameter estimates can be used for utility-consistent applied welfare analysis. We described the general intuition of welfare measurement in the KT model, suggesting how compensating variation is defined in the model. In this section, we list and explain the steps involved in calculating an individual's expected compensating variation for a change in price or site attributes, and demonstrate the technique by considering three policy scenarios relevant to the application. The measures derived are conceptually and theoretically equivalent to the welfare measures derived in other areas of applied welfare analysis, with the main difference being that compensating variation is computed numerically rather than via a closed form. In deriving the welfare calculation algorithm we follow von Haefen (forthcoming) and focus on drawing values of the error conditional on observed choice. The specific steps for a single individual are as follows:

Step 1: Sample values of the error vector ($\varepsilon_1, \ldots, \varepsilon_8$) conditional on the person's observed choice.

The maximization conditions in equation (10) suggest that $\varepsilon_j = g_j$ for the sites observed to be visited by the individual. Therefore, ε_j is given by equation (10) for the visited sites. For the sites that are not visited, equation (10) implies $\varepsilon_j \leq g_j$, suggesting a draw from the truncated extreme value distribution $f(\varepsilon_j \,|\, \varepsilon_j \leq g_j)$ is needed for each good, j, that is not consumed. A draw from this distribution can be recovered by the function:

$$\varepsilon_j = -\ln\left(-\ln\left\{\exp\left[-\exp(-g_j / v)\right] \times U\right\}\right) \times v, \tag{15}$$

where U is a draw from the (0,1) uniform distribution.[9]

Step 2: Calculate ordinal utility under baseline conditions.

The observed levels of demand and sampled values for ε from Step 1 are plugged into the utility function in equation (6) providing the utility level V^b obtained at baseline conditions.

Step 3: Calculate ordinal utility under changed conditions.

Using the sampled values for ε from Step 1 and equation (14), calculate the demand and utility levels for *every* possible demand regime as a function of the new prices/site attributes. If demand for a site is predicted to be negative, set the value to zero. The maximum of the regime-specific utility values is the utility under changed conditions, denoted V^c. This step is used only as an input to Step 4.

Step 4: Use numerical bisection to calculate compensating variation.

Begin with a guess on the upper and lower bounds for compensating variation. Denote these bounds CV_L^0 and CV_H^0 and set an initial guess for compensating variation, CV^0, as the average of the bounds. Using Step 3, calculate utility under changed conditions as a function of income defined by $y - CV^0$. If $V^c > V^b$, update the lower bound by $CV_L^1 = CV^0$ and set $CV_H^1 = CV_H^0$. If $V^c < V^b$, update the upper bound by $CV_U^1 = CV^0$ and set $CV_L^1 = CV_L^0$. An updated guess for compensating variation is the average of the updated bounds. Repeat this process k times until $CV^{k-1} \approx CV^k$. At this point, CV^k provides an estimate of the individual's compensating variation for the current draw of the error.[10]

Step 5: Repeat the process.

Repeat the process for multiple draws of the error. The average of the compensating variation estimates over the draws of the error provides an estimate of $E(CV)$ for the individual. Example GAUSS code for this process, with references to the steps listed above, is available upon request.

We provide a demonstration of these techniques by considering three welfare scenarios for our application to paddling in eastern North Carolina:

> Scenario A: The addition of 100 miles of marked trails in areas two and seven, which currently lack marked trails.
>
> Scenario B: Improvements in water quality such that all areas have an IWI index value of at least two.
>
> Scenario C: A $30 access fee for paddling in area five, the most heavily visited paddling area.

While these scenarios are hypothetical and intended as a demonstration, they do reflect realistic policy concerns. As noted, there have been increased efforts on the

---

[9] This follows from the fact that the cumulative distribution function $F(\varepsilon_j \mid \varepsilon_j \le g_j)$ is given by:

$$F(\varepsilon_j \mid \varepsilon_j \le g_j) = \exp[-\exp(-\varepsilon_j/\nu)]/\exp[-\exp(-g_j/\nu)].$$

This implies that a draw from the distribution $f(\varepsilon_j \mid \varepsilon_j \le g_j)$ can be obtained by solving $U = \exp[-\exp(-\varepsilon_j/\nu)]/\exp[-\exp(-g_j/\nu)]$ for ε_j, where U is a draw from the (0,1) uniform distribution. Judd (1998, p. 289) describes this process in greater detail.

[10] The numerical bisection routine is explained in more detail in Judd (1998, p. 148). Conceptually, the process is numerically searching for the value of CV that solves an equation of the form $v(p^0, q^0, y) - v(p^1, q^1, y - CV) = 0$. The bisection routine is needed, since no closed form exists for CV in this problem.

part of tourism management agencies in the state to improve infrastructure for paddling, including increasing the number of marked trails. Use of access fees to raise revenue for this purpose is one possibility. Finally, water quality issues and their effects on coastal recreation are of perennial concern in North Carolina.

Table 3 provides estimates of the mean and standard error individual yearly welfare effects of these policies. On average, survey respondents are willing to pay $4.64 for the increase in trail miles, and $24.44 for an increase in water quality in the paddling areas. Likewise the welfare impact of the access fee is –$41.35. In all cases, the Krinsky and Robb (1986) standard errors indicate the statistical significance of the welfare measures.

Table 3
Willingness-to-Pay Results

Policy Scenario	Welfare Estimate[a]
Scenario A: Increase Trail Miles at Sites 2 and 7	$4.64/year (1.50)
Scenario B: Improved Water Quality	$24.44/year (6.01)
Scenario C: $30 Access Fee at Site 5	–$41.35/year (0.34)

[a] Standard errors calculated using 100 Krinsky and Robb [1986] repetitions.

Final Comments

Our intent in writing this paper has been to present a new application of the Kuhn-Tucker model and to present an overview of the method in an intuitive manner that will help promote further application of the approach. Our data on recreational coastal paddling trips in eastern North Carolina is typical of many data sets available to recreation analysts and, therefore, provides a good demonstration of the preference estimation strategy. In an eight-site model, we find environmental and site attribute variables to significantly impact site choice and trip frequency. Furthermore, we find positive and significant willingness-to-pay estimates for increased trail miles and improved water quality.

The version of the Kuhn-Tucker model presented is relatively tractable due to convenient and somewhat restrictive assumptions on the form of the utility function and error structure. Future methodological work on the Kuhn-Tucker model should examine the possibility of employing more flexible functional forms and error distributions to better characterize recreation preferences.

References

Bain, L.J., and M. Engelhardt. 1992. *Introduction to Probability and Mathematical Statistics*. Boston: PWS-Kent Publishing.

Bockstael, N.E., W.M. Hanemann, and I.E. Strand. 1986. *Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models*, Vol. 2. Washington, D.C.: U.S. Environmental Protection Agency, Office of Policy Analysis.

Englin, J., P. Boxall, and D. Watson. 1995. Modeling Recreation Demand in a Poisson System of Equations: An Analysis of the Impact of International Exchange Rates. *American Journal of Agricultural Economics* 80:255–63.

Hellerstein, D., D. Woo, D. McCollum, and D. Donnelly. 1996. *ZIPFIP: A ZIP and FIPS Database*. Washington, DC: Economic Research Service.

Herriges, J.A., and C.L. Kling. 1999. Non-linear Income Effects in Random Utility Models. *Review of Economics and Statistics* 81:62–72.

Herriges, J.A., C.L. Kling, and D.J. Phaneuf. 1999. Corner Solution Models in Recreation Demand: A Comparison of Competing Frameworks. *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice*, J.A. Herriges and C.L. Kling, eds., Aldershot: Edward Elgar Publishing.

Judd, K.L. 1998. *Numerical Methods in Economics*. Cambridge: MIT Press.

Krinsky, I., and A. Robb. 1986. On Approximating the Statistical Properties of Elasticities. *Review of Economics and Statistics* 68:715–19.

Lee, L., and M. Pitt. 1986. Microeconomic Demand Systems with Binding Nonnegativity Constraints: The Dual Approach. *Econometrica* 54:1237–42.

Morey, E. 1999. Two RUMs UnCloaked. *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice,* J.A. Herriges and C.L. Kling, eds., Aldershot: Edward Elgar Publishing.

Morey, E.R., W.S. Breffle, and P.A. Green. 2001. Two Nested CES Models of Recreation Participation and Site Choice. *American Journal of Agricultural Economics* 83:414–27.

North Carolina Department of Commerce. 1998. *Economic Impacts*. Raleigh: NCDC. Available at www.nccommerce.com/tourism/econ.

North Carolina Department of Environment, Health, and Natural Resource, Division of Parks and Recreation. 1995. *North Carolina Outdoor Recreation Plan 1995-2000*. Raleigh, NC: NC DEHNR.

Parsons, G.R., P.M. Jakus, and Ted Tomasi. 1999. A Comparison of Welfare Estimates from Four Models for Linking Seasonal Recreation Trips to Multinomial Logit Models of Site Choice. *Journal of Environmental Economics and Management* 38:143–57.

Phaneuf, D.J., and J.A. Herriges. 1999. Choice Set Definition in a Kuhn-Tucker Model of Recreation Demand. *Marine Resource Economics* 14(4):343–55.

Phaneuf, D.J., C.L. Kling, and J.A. Herriges. 2000. Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand. *Review of Economics and Statistics* 82:83–92.

Shaw, W.D., and P. Feather. 1999. Possibilities for Including the Opportunity Cost of Time in Recreation Demand Systems. *Land Economics* 75:592–602.

U.S. Environmental Protection Agency. 1999. *Index of Watershed Indicators*. Available at www.epa.gov/iwi/.

von Haefen, R.H. Forthcoming. Incorporating Observed Choice into the Construction of Welfare Measures from Random Utility Models. *Journal of Environmental Economics and Management*.

von Haefen, R.H., and D.J. Phaneuf. Forthcoming. Estimating Preferences for Outdoor Recreation: A Comparison of Continuous and Count Data Demand Systems. *Journal of Environmental Economics and Management*.

von Haefen, R.H., D. Phaneuf, and G. Parsons. Forthcoming. Estimation and Welfare Analysis with Large Demand Systems. *Journal of Business and Economic Statistics*.

Wales, T., and A. Woodland. 1983. Estimation of Consumer Demand Systems with Binding Non-negativity Constraints. *Journal of Econometrics* 21:263–85.