Journal of Leisure Research
1998, Vol. 30, No. 3, pp. 301-318

Articles

Recreation Demand and the Influence of Site Preference Variables

Christos Siderelis and Roger L. Moore
Department of Parks, Recreation & Tourism Management
North Carolina State University

Leisure researchers' conceptions of how individuals make decisions involve an understanding of the personal tastes, motives, and private decisions that directly affect an individual's utility (satisfaction) of substitute site choices. Following recent theoretical advances in recreation demand, the participant is first viewed as deciding on the number of site trips to take per season and next, as deciding on how to allocate the trips across substitute sites. Decisions are analyzed by linking the results from a nested logit model to separate count-data models using consumer demand theory and trip-price indexes. Our purposes are two-fold. The first is to determine if the inclusion of the respondents' perceived importance of lake attributes in a nested logit model of lake boating improves the allocation of trips to the various lakes in a geographic region. The second is to demonstrate an empirical application of discrete-count modeling and to compare annual trip-counts of lake boating from a discrete-count model with a traditional pooled lakes model.

KEYWORDS: *Recreation choice behavior, outdoor recreation demand, recreation modeling, lakes, discrete choice*

Introduction

Recent advances in recreation modeling are motivated by the need to examine how changes in site quality affect outdoor participation (Bockstael, Hanemann, & Kling, 1987). In contrast, leisure researchers with their conceptions of how participants make choices about activities and site trips require insight into participant choice behaviors when specifying site demand (Ditton, Loomis, & Choi, 1992; Williams, 1984).

Our research is motivated by the fact that published discrete-count models of recreation demand fail to adequately identify and integrate individual preferences, like the importance of site characteristics, that shape trip choices. Using recent theoretical advances in the specification of discrete-count models, we speculate that the inclusion of data regarding the per-

This study was funded in part from a lake research project sponsored by the Carolina Power and Light Company. We gratefully acknowledge the comments from Don English of the Outdoor Recreation and Wilderness Assessment Research Program (NFS), and three anonymous referees.
The authors can be contacted at Box 8004, Raleigh, NC, 27695-8004.

ceived importance of site attributes to participants should improve a choice model's predictive power (Adamowicz, 1994). Clark and Downing (1984), for example, believe that explanatory variables like the importance of site attributes to a participant might influence the marginal choice of a recreation site in a particular geographical area.

We begin with a recent review of recreation demand theory, which leads us to specify a discrete-count model for lake boating trips. We then report on the benefits gained from specifying a nested logit model to explain lake choices. We end with a discussion on the implications of the discrete-count method in estimating outdoor recreation demand.

Related Research

Recent articles advance competing theories of recreation demand that allow analysts to link independent discrete site choices to the aggregate demand for seasonal trip-counts (Hausman, Leonard, & McFadden, 1995; Feather, Hellerstein, & Tomasi, 1995; and Parsons & Kealy, 1995). The main purpose of the advances is to explain users' recreation behaviors when faced with environmental threats to site quality. The demand theories, although different in their hypotheses about individual decision processes, support discrete-count empirical applications. Specifically, each theory differs with respect to a trip-price index that links the allocations of trips among substitute sites (discrete choices) to the seasonal aggregate demand for the seasonal counts of trips. It must be emphasized that the estimation of a discrete-count demand model cannot be completed in a single statistical process. Rather, discrete choice and the trip-count models are two different types of travel cost models described in previous JLR literature reviews (e.g., Fletcher, Adamowicz, & Tomasi, 1990). Trip-counts refer to the quantity of seasonal trips by individuals, with the analysis of trip-counts following a count-data or Poisson distribution.

Feather et al. (1995) follow a household production function where recreation opportunities are produced and consumed by a household, constrained by such scarce resources as the amounts of leisure time, money, and effort. Unknown to the analyst, and therefore to be estimated, are the proportions of scarce resources that are necessary to produce a recreation trip and a participant's expectation of site quality. Feather and colleagues suggest multiplying a participant's probability of visiting regional recreation sites by the travel costs and the measures of site quality to compute an expected cost, expected time, and expected quality per trip. Substituting the computed values into a participant's budget and time constraints results in a single, expected full-income constraint. Maximizing the recreation utility function for site trips, subject to the expected full-income constraint, yields the ordinary recreation demand function for seasonal trips.

Hausman et al. (1995) propose a budgeting model to support their trip demand theory. They view the participant first as budgeting a number of seasonal trips and second as allocating trips across substitute sites. The solution to the household budgeting problem is a Gorman generalized polar form that includes the prices faced by participants in travelling to and from recreation sites.[1] The budgeting process imposes a decision tree on household utility to partition the total production of household services into separate services (e.g., food, vacation, lake trips), so that preferences within each group can be described independently from other services (Hausman et al.). The degree of separability between household services varies because the household devotes different fractions of its resources to the separate services. This concept allows for the analysis of lake fishing trips to be weakly separated from, say, a national park visit, so that we may derive a fishing demand function.

Feather et al. (1995) compute an expected trip-price per user by combining site choice probabilities and travel costs, conditional on users taking at least one trip per season to that site. Hausman et al. (1995) compute compensated surplus values from each respondent's probabilities of choosing the various substitute sites, and combine them with a measure of the marginal utility of income to form a trip-price index.[2] In both travel cost applications, the trip-price index and other socioeconomic variables become a function of the seasonal trip-counts, the dependent variable, in the estimation of the aggregate recreation demand. The resulting estimates of aggregate demand and corresponding consumer surplus values are useful to recreation planners in computing the economic benefits of recreation sites and, on occasion, evaluating site pricing decisions and site development options (Fletcher et al., 1990).

Methodology

Data were obtained from mail questionnaires sent during 1995 to a random sample of registered boat owners who resided in a 17 lake region, approximately 90 miles in diameter, located in the north central portion of North Carolina and extending north into Virginia. Major metropolitan areas in the region included Raleigh-Durham, NC; Greensboro-Winston Salem, NC; Martinsville, VA; and Danville, VA.

Three sets of 700 questionnaires were mailed to registered boat owners during the Spring, Summer, and Fall months from a pool of 63,366 boat owners in NC (73%) and VA (27%). Of the 2,100 questionnaires, 1,079 (51%) were returned, 178 respondents of which indicated that they either did not visit any of the 17 lakes or did not go boating in the past year. A

---

[1]The aggregation of sites into a group (e.g., a region) is represented by the sub-utility function $V_G(q_G)$. Since the site vector q is partitioned into N such groups, the preferences are said to be weakly separable and represented by the utility function, $u = u[v_1(q_1), v_2(q_2),..., v_G(q_G),..., v_N(q_N)]$, for sub-vectors q_1 to q_N. The correspondence between weak separability, two-stage budgeting, and the existence of the sub-utility function allow the authors to adopt an indirect utility function $v(\cdot)$ for grouped site visits.

[2]Compensated surplus values per trip occasion are negative because calculations are based on the hypothetical denial of access to participants of the alternative substitute sites.

participant's choice set of 17 substitute lakes and the option of not choosing a lake were examined for a total of the 18 alternative choices per boating occasion. The no-lake choice was modeled with zero values for the explanatory and lake quality variables (Adamowicz, 1994). While the no-lake choice does little to explain the reasons for not selecting a lake, it allows for the more precise estimates of discrete choice probability outcomes and the subsequent accounting of recreation benefits (Morey, Rowe, & Watson, 1993; Adamowicz, 1994).

Approximately 22% of the 901 respondents who went boating chose to visit multiple lakes. Consequently, each of the respondent's lake choices was treated as a separate observation, which resulted in 1,158 discrete choices. Lake choices were combined with the 178 no-lake choices for a total of 1,336 sample observations. The final data set contained 24,048 separate records (18 alternatives · 1,336 cases).

Rather than representing rows in the data set as 1,336 individual observations, a discrete choice analysis requires panel data with multiple rows of data per observation. For example, a respondent reports taking a boating trip to one lake with that choice occasion being the separate observation. We "explode" this one observation into 18 separate records, one record for each of the 17 lakes and the no-lake choice (Greene, 1993). The dependent variable is specified by giving that alternative chosen a value of one and the remaining 17 choices zeros. The creation of panel data was done with the commercial program LIMDEP (Version 7).

Lake attractiveness. The 17 lakes are managed by various power companies and the Army Corps of Engineers. Their water surface areas range from 2,800 to 50,000 acres. After repeated contacts with managing authorities and visits to lakes, we were able to obtain partial data from direct and secondary sources on only 11 of the 17 lakes regarding the percent of shoreline development, the number of boating ramps, water level fluctuations, and water quality. Preliminary analyses with the partial data found high correlations (> .90) between water surface areas and support facilities like the number of parking spaces and acres of public access. Consequently, we did not compile a composite attractiveness index due to the incomplete data on lake quality variables, the interrelationship among lake surface acres and the quantities of lake support facilities, and descriptive measures that did not vary across lakes (e.g., creel limits had "yes" responses from all lake managers). We instead used the number of water surface acres as a surrogate measure of each lake's attractiveness (e.g., Cesario & Knetsch, 1976).

Travel Costs. The marginal trip-cost includes a person's per hour opportunity cost of travel time plus a vehicle's round-trip operating expense (round-trip miles · $.16 per mile) as set by the NC Dept. of Motor Vehicles. Wage rates were based on a respondent's classification of their employment and the results from Smith's (1983) hedonic wage model, re-scaled to 1993 with the consumer price index. Following the McConnell-Strand procedure for the estimation of the opportunity cost of travel time, we estimated the constant opportunity cost of travel time as 27% of an individual's hourly wage rate (Hof & Rosenthal, 1987)[3]. The constant value is within the 25% to 50% value used in computing the opportunity cost of travel time in travel cost studies (Hof & Rosenthal). The mean travel cost was $21.62 per trip.

Importance of lake attributes. Lin, Peterson, and Rogerson (1988), experimenting with a nested urban recreation site choice model, had respondents rate 31 site characteristics as to their desirability. Lin and associates factor analyzed the agree and disagree responses. The individual factor scores discriminated among urban recreation site choices. In a similar fashion, we recorded respondents' preferences on each of 20 lake attributes with the following question, "When you choose a lake for a boating trip, how important are each of the following features to you?" The attributes examined are listed in Table 1. Respondents rated the importance of each attribute on a seven-point scale from 1, not at all important, to 7, extremely important.

Macro-site decision variables from the marginal choice stage of a nested model are specific to each participant (Hensher & Johnson, 1981). Directly entering individuals' ratings of the 20 lake attributes separately into a discrete choice model is undesirable for two reasons. First, there would likely be a high degree of multicollinearity among the 20 rated values. Second, respondents are essentially informing us about attributes that most importantly influence their preferred choice of a lake. The resulting ratings however do not necessarily relate to the selection of one particular lake, and consequently when analyzed will not vary in importance across the 17 lake choices.

We constructed three factors from respondents' ratings and principal component analysis that summarized the perceived importance of lake attributes to respondents (Table 1). We retained three importance factors with eigenvalues greater than one (Hamilton, 1992). A regression scoring method was followed to estimate one score per respondent from each of three factors, which we labeled as lake use conditions, natural surroundings, and customer support services. Missing values on one or more importance items were uncovered for 48 observations, which the computer dropped from factor analysis. We imputed estimates of index values with a regression of the estimate for 20 preference-importance items where $\hat{y}_j$ was the predicted value of estimate and $\hat{v}_j$ was the square of the standard error of the predication (Hamilton, 1992). We re-scaled the scores to have a mean value of 100 and a standard deviation (σ) of 10 to simplify demand analysis, that is, $\text{index}_i = (\text{score}_i/\sigma_i) \cdot 10 + 100$ for $i = 1, 2, 3$ indexes.

---

[3]Hof and Rosenthal (1987) review and comment on the McConnell and Strand argument, which defines the constant opportunity cost of travel time. Briefly, we regress the cost of travel time per trip (multiplying each participant's round-trip travel time to a lake by the participant's hourly wage rate) and the round-trip vehicle operating expenses against the number of annual boating trips. Next, the opportunity cost results by dividing the coefficients for travel time by the coefficient for travel expense (Hof & Rosenthal).

TABLE 1
Factor Analysis of the Perceived Importance of Lake Attributes

Importance Measures	Lake Conditions	Natural Features	Customer Services
Size	−.2407	.6145	−.0650
Scenery	−.3033	.6156	−.1765
Lake shape	−.1433	.5986	−.0344
Water clarity	−.0688	.6425	.0656
Conveniently located	.1835	.4079	.0452
Natural shoreline	.1523	.4461	.0674
Water temperature	.1267	.3870	−.0139
Absence of aquatic weeds	.1675	.2547	.0244
Smooth vs choppy surface	.0106	.4762	−.0791
Low fees	.4052	−.0977	.1651
Quality of fishing	.4264	.2147	.0598
Boaters behave well	.3317	.2986	.0138
Uncrowded on water	.3617	.3157	.1735
Uncrowded at boat ramps	.8476	−.1101	−.0230
Location of ramps	.8510	−.0916	−.1231
Parking at ramps	.8021	−.1830	−.1933
Places to buy food stuffs	.1507	.0544	−.5721
Bathrooms	.1850	−.0157	−.7741
Helpful staff	.0762	.0640	−.7989
Camping facilities	.1219	.1368	−.2597

Notes. Variables are measured on an importance scale of 1, not at all important, to 7, extremely important. The three factors retained have eigenvalues of 1.162 or more. Correlations among factors are: Lake conditions and natural features (.39), lake conditions and customer services (−.37), and natural features and customer services (−.42).

Macro-site Choice Decision

A key concern in understanding lake choices is the substitutability of 17 lakes in the geographic region (Stynes & Peterson, 1984). The relative probability of choosing one of the 17 lakes must be independent of the remaining lakes, if we are to satisfy the independence of irrelevant alternatives property of a random utility model (Zhang & Hoffman, 1993). If the 17 lakes are perceived by users to be independent from one another, then the probabilities of the observed lake choices will be valid. Where any combinations of the 17 lakes are close substitutes for one another, the independence from irrelevant alternatives property will be implausible and parameter estimates from logit analysis will be inappropriate (Stynes & Peterson).

The close proximity of the 17 lakes to one another appeared to violate the independence of irrelevant alternatives assumption. The proper way to cluster lakes into boating regions is at the analyst's discretion providing that the alternative nesting structures satisfy the independence of irrelevant alternatives test (Hausman et al., 1995). Using data-clustering techniques, an obvious way for us to organize the 17 lakes into 5 boating regions was by their geographical proximity to one another. We adopted the Hausman-McFadden nested logit test (Zhang & Hoffman, 1993). The nesting of substitute lakes into 5 boating regions and a no-lake choice allows us to relax the independence of irrelevant alternatives assumption and to test this property directly (Zhang & Hoffman)[4].

As a result, the participant is viewed on each choice occasion as making a marginal choice among boating regions in a geographical area or the no-lake option (Clark & Downing, 1984). This is followed by the "micro-site" choice of a particular lake, which is conditional on the previous macro-site choice of a boating region (Clark & Downing).

The nesting design is common in the literature. Parsons and Needelman (1992) cluster lakes by regional boundaries along county lines in Wisconsin (e.g., North, South). Kaoru, Smith, and Liu (1995) cluster sampled locations on the Albermarle and Pamlico Sounds in North Carolina into 35, 23, and 11 launching-site groups to test nesting assumptions.

A Discrete Choice Model

The random utility model for the allocation of boating trips to lakes follows from choice theory and the availability of boating participation data. The model is estimated as a conditional logit (conditional on the participant's choice of a site) where the specification of the indirect utility includes a mix of data on both lake choice attributes and individual characteristics (Zhang & Hoffman, 1993)[5].

Indirect utility specification. We begin with the underlying assumption that each participant considers all the specified lake choices, and each lake choice has a non-zero probability of being selected a priori (Kaoru, Smith, & Liu, 1995). This assumption and the theoretical specification of consumer choices are necessary for the analysis of lake choices (Hensher & Johnson, 1981).

To begin, conditional on the choice of lake j from among the set of available lakes, a participant i derives a certain level of recreation utility from visiting the lake, which is specified as $U_{ij} = V_{ij} + \epsilon_{ij}$. The indirect utility, V_{ij}, is a measurable expression that predicts the satisfaction derived from lake j.

---

[4]Zhang and Hoffman (1993) offer a variety of tests for the independence from irrelevant alternatives or IIA property, which are programmed with LIMDEP commands. Hausman and McFadden's specification error test examines the validity of the IIA property by analyzing the change in the model structure and parameters when choice is analyzed on a restricted subset of the full-choice set. The resulting similarity of coefficients from the full set and restricted subset are compared to justify the IIA property.

[5]The main difference between a multinomial logit and a conditional logit model is that the multinomial logit makes the choice probabilities dependent on participant characteristics only, whereas the conditional logit can consider the effects of both participant and choice characteristics. Algebraically, both models are equivalent (Zhang & Hoffman, 1993).

The individual is assumed to choose lake j because the utility from lake j, U_{ij}, is greater than from lake k, U_{ik}. We specify the indirect utility function as follows,

$$V_{ij} = \beta x_i + \gamma z_j + \theta w_{ij} \tag{1}$$

The column vectors are the exploratory variables—x, z, and w; and the row vectors are the greek letters—β, γ, and θ (i.e., the coefficients to be estimated) (Zhang & Hoffman, 1993). Vector x_i contains variables that describe participants, but do not vary over choice alternatives (e.g., a participant's age). The β coefficients must vary over the substitute choices, and one of the coefficients must be set equal to 0 to achieve identification (Bockstael, McConnell, & Strand, 1991)[6]. The variables in z_j vary across choices, but are the same for all participants (e.g., site attractiveness or water surface area). The vector w_{ij} contains variables that describe an interaction or a relation between lake choices and the participant (e.g., round-trip travel costs).

In attempting to fit the indirect utility function to data that best models an individual's lake choices, the statistical deviation from predicted to actual behaviors is attributable to a random error term, ϵ_{ij}. The error reflects omitted variables due to our inability to fully represent individual decisions (Hensher & Johnson, 1981; Bockstael et al., 1991)[7]. Using the two-lake example, if the choice is between lakes j and k, the choice of j over k implies that $(V_{ij} + \epsilon_{ij})$ is greater than $(V_{ik} + \epsilon_{ik})$. By rearranging the observable and random error components so they are together, the difference between observable components, $(V_{ij} - V_{ik})$, is assumed to be greater than the difference in the errors, $(\epsilon_{ik} - \epsilon_{ij})$.

Because it cannot be determined with certainty that the differences in the observable utilities will be greater than the differences in random errors, probabilities (π) are assigned to particular choices that reflect the likelihood of the random errors being less than the observable utilities, so that π_{ij} is equal to $\pi\{(\epsilon_{ik} - \epsilon_{ij}) < (V_{ij} - V_{ik})\}$ for all j not equal to k. Or, the probability that an individual will choose lake j equals the probability that the difference between the random errors of lakes k and j is less than the difference between the utilities of lakes j and k (Hensher & Johnson, 1981).

Nested logit structure. Using the nested logit procedure for analyzing discrete choice models, as described in the LIMDEP econometric software, we evaluate the nested model structure by first estimating the parameters at the conditional stage or where the choice of a substitute lake is conditional on a boating region. Next, the inclusive values per participant are computed

[6]For example, to record a person's age for a three-choice problem, the assigned values would be 26, 0 for the first choice alternative; 0, 26 for the second choice alternative; and 0, 0 for the third alternative.

[7]The ϵ_{ij}'s are assumed to be independent and identically distributed with an extreme-value distribution such that $f(\epsilon) = \exp[-\epsilon - \exp(-\epsilon)]$ or the distribution looks like a normal curve that is skewed to the right, with a thinner tail on the left and a thicker tail on the right. It has mode 0, mean .58, and standard deviation of 1.28.

from a participant's conditional lake choice. The inclusive values summarize the degree of substitutability among lakes in each of the five boating regions. Finally, we use the inclusive values and importance of lake attributes to estimate the inclusive value parameters and to compute the participants' marginal choices of boating regions (Greene, 1993).

Overall, the adequacy of the nested logit model is evaluated by predicting the lake choices of participants and comparing them with observed frequencies. Also, model adequacy is tested by the model's consistency with utility maximization theory in satisfying the independence of irrelevant alternatives condition if, and only if, the five inclusive parameters are between 0 and 1 in the marginal choices of boating regions (Hausman et al., 1995).

The probability of a participant choosing lake j from J possible choices, conditional on the choice of boating region h first, is

$$\pi_{j|h} = \frac{\exp(\beta a_{hj} - \gamma p_{hj})}{\sum_j \exp(\beta a_{hj} - \gamma p_{hj})}, \tag{2}$$

and the participant is not subscripted to simplify the amount of notation. At the conditional choice stage (Equation 2), the independent variables that define a participant's utilities for lakes are the amounts of water surface acres a and the participant's round-trip travel costs p to lakes in the choice set (Greene, 1993).

The inclusive value I_h for a boating region summarizes the overall attractiveness of that region to participants. The inclusive value coefficients measure the similarity of lakes in regions. The inclusive values are computed for each participant from the resulting values at the conditional lake choice stage (Equation 2) as

$$I_h = \log\left[\sum_j \exp(\beta a_{hj} - \gamma p_{hj})\right]. \tag{3}$$

The results from Equation 3 are passed forward to the marginal choice stage as

$$\pi_h = \sum_j \pi_{hj} = \frac{\exp(\alpha w_h + (1 - \sigma)\, I_h)}{\sum_h \exp(\alpha w_h + (1 - \sigma)\, I_h}, \tag{4}$$

where the w_h's are respondents' scores for lake use conditions, natural surroundings, and customer support services. The inclusive value parameters $(1 - \sigma)$, estimated with Equation 4, measure the similarity of lakes for each of the five boating regions. An inclusive parameter (σ) close to one denotes that the lakes are perfect substitutes, whereas a parameter value close to zero denotes that the lakes are not perfect substitutes and no violation of the independence of irrelevant alternatives property (Hausman et al., 1995).

The parameter estimates—α, β, γ, and δ—in Equations 2, 3, and 4 result from the full information maximum likelihood (FIML) estimation technique (Feather, Hellerstein, & Tomasi, 1995; Greene, 1995). The FIML produces more reliable estimates than a sequential estimator, which passes forward the inclusive values from the conditional to the marginal choice stages in a series of separate steps (Feather, Hellerstein, & Tomasi, 1995). The FIML estimates all parameters simultaneously by maximizing the unconditional log-likelihood function of the nested discrete choice model (Greene, 1995).

Trip-price indexes. The computations of price and quality indexes involve the resulting utilities from the analysis of lake choices and the probabilities that lake choices are realized (Freeman, 1993). The expected price $E(P)$ is the weighted average or the conditional probabilities multiplied by the travel costs to all substitute sites in the lake choice set (Feather et al., 1995). It is computed as

$$E(P) = \sum_j \sum_h \pi_{j|h} \, \pi_h \, P_j. \tag{5}$$

Similarly, the expected quality $E(Q_h)$ is the weighted average of lake surface acres, and is computed as

$$E(Q_h) = \sum_j \sum_h \pi_{j|h} \, \pi_h \, a_{hj}. \tag{6}$$

The price index by Hausman and associates (1995) is a per trip consumer surplus cs that equates to the amount of utility a respondent realizes from the substitute lakes as

$$cs = \frac{1}{\gamma} \ln(\Sigma_h \exp(\alpha w_i + (1 - \sigma) \, I_h)). \tag{7}$$

The parameter $1/\gamma$ is the marginal utility of income, where γ is the travel cost coefficient from the conditional lake choice stage.

Count-data Model

The price index is passed forward as an independent variable into the estimation of annual trip-counts across all lakes. The demand analysis includes the 1,336 original individual observations. The statistical specification of a respondent's annual trip decision is a Poisson regression. Trip-counts (λ) are integer values of the trip occurrences during the past year with $\lambda \geq 0$. Respondents recorded trips that ranged in number from 1 to over 300 per annum. A descriptive analysis of trip occurrences exhibited an extra-Poisson variation where the variance exceeded the mean trip-counts (mean = 9.68, s = 22.16). The over-dispersion of variance about the conditional mean is allowed in the negative binomial regression, which is an extension of the Poisson regression model. The negative binomial, $\ln\lambda = \beta_{xi} - \gamma P_i + \epsilon$, includes x_i explanatory variables, P_i the trip-price index computed from discrete choice results, and ϵ an error term (Greene, 1993)[8].

[8]The error follows a gamma distribution with mean 1 and variance alpha. Alpha is the over-dispersion parameter and the larger the alpha, the greater the dispersion.

Since we are particularly interested in participants who took at least one trip during the past year, we censored the zero trip-counts of the 178 respondents who chose the no-lake choice alternative with a LIMDEP statistical procedure that works like a Tobit model for count-data (Greene).

Independent variables. Types of launching facilities are included as dummy (0,1) variables in the boating participation model to capture systematic differences in the quality of public, marina, and private pier facilities at lake entry points. Another independent variable is the payment of launching fees by respondents, which we hypothesize might affect decisions to go boating as well as the weekly patterns of trips. Payment vehicles may have been one-day launching fees or annual marina payments for boat storage with launching ramp rights. In the latter case, the payment of an annual fee for the storage of a boat may be thought of as being independent of the direct use made of a lake, and is therefore an adjustment to income (Kaoru, Smith, & Liu, 1995). This distinction can affect the interpretation of a discrete-count model specification.

The dummy variables of fishing, pleasure cruising, and water-skiing reflect the different experiences of boating activities and the specialized interests of participants. Fishing can be a consumptive activity, pleasure cruising requires a relatively low-level of specialization, and water-skiing involves a degree of athletic skill.

Results

The parameter estimates for the two lake choice models are presented in Table 2. The difference in choice models is that the saturated specification displays three additional parameters where the constrained model does not. The additional parameters relate to the factor scores for natural features (e.g., lake crowding, availability of parking, location of launching facilities), lake use conditions (e.g., scenery, water clarity, lake shape), and customer support services (e.g., availability of restrooms, helpful staffs). Both nested logit models included parameter estimates for round-trip travel costs, water surface acres, and the inclusive values.

The significance ($p < .001$) of chi-squared (χ^2) values for the saturated model and the constrained model were rejected for the equality of lake destination choices. A test of the difference between saturated and constrained log-likelihood values resulted in a χ^2 (3) = 70, ($p < .01$), with the saturated model being more accurate.[9] Overall, the saturated model correctly predicted the assignments of participants to actual lake choices 88% of the time. We retained the saturated model with the importance of lake attributes in the marginal choice stage to derive lake choice probabilities and to compute trip-price indexes.

[9]The chi-square value is equal to the product of a constant of −2 multiplied by the difference between the constrained and saturated log-likelihood values, and the degrees of freedom (*df*) is equal to the *df* of the saturated minus the *df* of the constrained models (Greene, 1993).

TABLE 2
Nested Lake Choice Models

	FIML Parameter Estimates	
	Constrained	Saturated
Travel Costs	−.09999	−.08998
	(−23.832)	(−20.301)
Water Surface Acres	.32818	.55247
(natural log)	(15.840)	(13.774)
Natural features		−.02916
		(−3.134)
Lake conditions		.01495
		(2.129)
Support services		−.01495
		(−2.649)
Inclusive values $(1 - \sigma)$[a]	.73978	.88448
Boating Region 1	(12.028)	(15.756)
Boating Region 2	.53834	.77096
	(17.174)	(16.303)
Boating Region 3	1.1402	.87575
	(8.118)	(11.523)
Boating Region 4	.82657	.86593
	(13.292)	(16.078)
Boating Region 5	.62572	.78921
	(7.970)	(10.204)
Log Likelihood	−2,697.595	−2,662.91

Notes. *z* values (the parameters divided by the standard error) are in parentheses. Standard errors are asymptotic (large samples). FIML is the full-information maximum likelihood estimation technique for analyzing a nested choice structure. Lakes in the boating region are grouped as follows: Region 1—Chesdin and Gaston; region 2—John Kerr, Hyco, Mayo, Falls, Jordan, Harris; region 3—Tillery, Badin, Norman; region 4—Claytor, Philpott, Smith Mountain, Leesville; region 5—W. Kerr Scott, Belews.
[a]Inclusive value parameters $(1 - \sigma)$ summarize the substitutability of lakes within each of the 5 boating zones. An inclusive parameter (σ) close to one denotes perfect substitutes, whereas a parameter close to zero are not perfect substitutes with no violation of the independence of irrelevant alternatives property.

At the conditional lake choice stage, the travel costs and the logarithms of lake size parameters were significant ($p < .01$) and consistent in signs and magnitudes (Table 2). As expected, as travel costs increased, the probability of a trip to a lake decreased, and as the water surface size increased, the probability of a boating trip increased.

From Table 2, the probable choices of a boating trip to substitute lakes in particular regions increased 1.5% as the importance of lake condition (e.g., lack of crowds) increased among participants. Alternatively, the probable choice of boating trips to lakes in particular regions increased as the importance of natural features and customer support services decreased 2.9% and 1.5%, respectively, among participants. Apparently, increases in the occurrences of boating trips to lakes in regions were more likely to occur if lake use conditions were of increased importance to participants as opposed to the natural lake features or the supporting customer services.

Inclusive value coefficients ($1 - \sigma$) were significantly different from zero, and varied in magnitudes (Table 2). Recall that inclusive parameters (σ) satisfy the independence of irrelevant alternatives assumption if the parameter values are between 0 and 1 (Hausman et al., 1995). The saturated model from the inclusion of importance of lake attributes resulted in five inclusive value coefficients ($1 - \sigma$) ranging in value from .77 to .88. Therefore, estimates of σ's were between .12 and .23. The σ estimates were significantly different from 0. This indicated that there were gains from nesting in the model and a lake choice model with the 17 lake alternatives would violate the independence of irrelevant alternatives property (Bockstael & McConnell, 1989). Also, the estimates of σ's were closer to zero, which indicated that the lake alternatives within boating regions were not perfect substitutes (Bockstael & McConnell; Hensher & Johnson, 1981).

Regression Results

Regression results are displayed in Table 3 comparing the parameters and the t-ratios for two discrete-count models and one traditional demand model. The first column presents results of the Hausman, Leonard, and McFadden (1995) price index model, where the participants' consumer surplus values per trip from Equation 7 are the trip-prices. The second column displays results of the expected trip-prices model, computed with Equation 5 (Feather et al., 1995). The third column describes a "pooled" lakes model for comparison with the two models in the first and second columns. The pooled lakes model is simply the business-as-usual approach, where all the observations of boating trips to substitute lakes are treated as belonging to a single recreation demand equation, rather than separate lake demand equations, and uses the conventional measures of travel costs as the trip-prices (Pollak and Wales, 1981).

The significance ($p < .05$) of the alpha (α) values in the three participation models indicated an over-dispersion in the data, which upholds the assumption that it was highly unlikely that we would have observed the data conditional on the process being Poisson. Price parameters were statistically significant and were of the appropriate signs. For each additional dollar increase in trip-price, the counts of annual trips increased from a low of 2.6% in the pooled lakes model to a high of 5.1% in the expected price model, holding all other variables constant.

The quality parameter in the price/quality model has the appropriate positive sign, but has no affect on annual trip-counts and consequently was not reported. Similarly, the boating activity parameters, and the parameter

TABLE 3
Truncated Count-data Models of Regional Boating Participation

Independent Variables	HLM Price[a] Index Model	FHT Expected[a,c] Price Model	Pooled Model
Trip-prices	−.028297	−.051251	−.026069
	(−4.651)	(−5.543)	(−10.583)
Weekend-only[b]	−.39110	−.36298	−.30605
	(−3.803)	(−3.611)	(−3.152)
Annual income	−.0000077	−.0000076	−.0000067
	(−7.436)	(−7.321)	(−8.221)
Fishing[b]	−.18167	−.15882	−.10565
	(−1.704)	(−1.393)	(−.976)
Boating[b]	−.17787	−.22340	−.13718
	(−1.307)	(−1.621)	(−1.058)
Public ramps[b]	.11612	.064356	.14272
	(.931)	(.495)	(1.304)
Private piers[b]	1.0605	.98717	1.1332
	(6.188)	(6.093)	(7.524)
Pay fees[b]	.27901	.28172	.26325
	(2.805)	(2.794)	(2.792)
Constant	2.6354	3.6073	3.0040
	(12.656)	(12.898)	(16.750)
alpha (α)	2.9408	2.7276	2.4486
	(9.449)	(9.722)	(10.546)
Log Likelihood	−3644.69	−3631.29	−3612.68

Notes. Regression models are truncated at trip-counts equal to zero. Numbers in parentheses are the t-ratios ($Pr \leq .05$ at $t \pm 1.96$). Alpha (α) is a measure of the over-dispersion in the trip-counts and all α's are significant. We reject the assumptions that the count-data are Poisson.
[a]The price and attractiveness index variables are computed from the full information estimates with a nested structure of lake choice decisions. HLM is Hausman, Leonard, McFadden and FHT is Feather, Hellerstein, Tomasi. The pooled model treats the 17 substitute lakes as belonging to a single recreation demand equation.
[b]Dummy variables are equal to 1 if the variable is observed and 0 otherwise.
[c]Expected attractiveness (quality) parameter is .03556, and is clearly not significant (t = .454).

on the boat launchings from public ramps has no statistically significant influence on trip-counts.

Not surprisingly, weekend-only participants demanded fewer number of annual trips than did the daily lake users, or 32% (=100[exp(−.391) − 1]) fewer trips with the price index model and 26% fewer trips with the pooled lakes model, holding all other variables constant. Using the price index model results, participants who made payments of either daily or annual access fees demanded roughly 32% more trips. Participants who launched boats from private piers demanded almost twice the number of trips than did participants who used public access facilities. Annual income parameters were significant and negative in signs. Respondents reporting less annual income demanded a greater number of boating trips, which was contrary to the economic assumption that participants with higher annual incomes demanded greater numbers of trips. It does suggest that the travel costs associated with boating day trips are nominal. In fact, time costs are often considered more of a constraint to increasing trip frequency than are annual incomes (Bockstael, McConnell, & Strand, 1991).

In summary, it is difficult to choose one "best" estimator of the two discrete-count models. Both equations are consistent in parameter signs and statistical significance. The results of the truncated count regression reveal the presence of five significant parameters that influence the demand for trips—patterns of use, trip-prices, annual incomes, private pier access, and access payments (Table 3). However, there is a greater similarity between the parameters of the pooled lakes model in magnitude and sign with the price index model in column 1, Table 3, than the expected price model in column 2.

Discussion

The inclusion of the importance of lake attributes in the marginal choice analysis provides for a more complete understanding of the macrosite choice process, which in our opinion has been neglected by researchers in their applications of discrete-count models to environmental settings. Other candidate variables might include (a) the variety of physiographic, topographic, and landscape features; (b) season and elevation; (c) the availability of various resource-dependent opportunities; and (d) recreation opportunity setting conditions (Clark & Downing, 1984). The introduction of additional measures of physical characteristics and environmental quality variables, in lieu of just water surface acres, may have improved the predictive power of the discrete choice model by accounting for more of the within-group variance at the conditional choice stage of the logit analysis.

If a planner is concerned about regional impacts of a decrease in public site access or an increase in site congestion as in the computation of welfare estimates for a benefit-cost analysis, other site-specific measures of site access (e.g., number of entry points, parking spaces) might be good candidates for inclusion at the conditional choice stage. Or, a planner may elect to apply a discrete-count model to simply evaluate the potential loss in participation and compensated welfare benefits from perturbing the nested choice model with the hypothetical closure of substitute sites (Feather, Hellerstein, & Tomasi, 1995). Or, the planner may examine the hypothetical addition of a proposed recreation site or future access points at a site on the distribution of participation to existing sites. Or, the planner may apply a discrete-count model when evaluating the resulting effects on recreation participation rates from the proposed decision to impose access fees at one or many sites.

Even if planners are not interested in evaluating policy or environmental quality changes at recreation sites, but are interested in the allocation of trips among substitute lakes, the discrete-count technique is superior to regional travel cost applications since the results explicitly allocate trips to recreation sites. The predicted probabilities and accompanying standard errors are estimated for alternative site choices with the nested logit, rather than having the planner rely on simple trip frequencies to apportion trips among the substitute sites. We display the alternative lakes, trip probabilities, and the allocation of trips to lakes in Table 4 using the saturated model results.

Planners not interested in the distribution of trips among lake choices might continue to find the pooled lakes model appealing (i.e., specifying participant visits to all regional lakes as one "typical" lake model). Overall, the pooled parameter values and the resulting estimation of annual trip-counts are comparable to discrete-count results. Like the traditional pooled

TABLE 4
Discrete-Count Results by Lakes

	Alternative			
Lake Choices	Probability	[Frequency]	Mean Annual Trip-counts	Trips[b]
Badin	.0149 (.0265)[a]	[.0202]	7.90 (4.7)[a]	7,459
Belews	.0571 (.0863)	[.0352]	7.73 (6.1)	27,969
Claytor	.0064 (.0128)	[.0060]	8.59 (4.7)	3,484
Falls	.1298 (.1061)	[.1138]	7.11 (2.8)	58,479
Harris	.0569 (.0542)	[.0509]	6.67 (2.6)	24,049
Hyco	.0428 (.0341)	[.0487]	10.1 (6.2)	27,392
John H. Kerr	.1440 (.1274)	[.1692]	8.86 (5.4)	80,845
Jordan	.1423 (.1220)	[.1168]	7.27 (2.7)	65,553
Gaston	.0721 (.0124)	[.0763]	7.82 (2.2)	35,727
Tillery	.0137 (.1117)	[.0112]	9.63 (5.4)	8,360
Chesdin	.0031 (.0144)	[.0045]	8.17 (4.7)	1,605
Norman	.0085 (.0235)	[.0112]	9.49 (6.5)	5,111
Leesville	.0291 (.0493)	[.0187]	7.66 (3.4)	14,125
Mayo	.0301 (.0233)	[.0337]	7.51 (2.7)	14,324
Philpott	.0285 (.0546)	[.0337]	6.87 (2.8)	12,407
Smith Mtn.	.0841 (.1364)	[.0868]	10.2 (6.1)	54,356
W. Kerr Scott	.0043 (.0117)	[.0299]	7.55 (3.7)	2,057
No-lake Choice	.1324 (.0535)	[.1332]		

Notes. The mean individual probabilities are derived from nested logit analysis. Annual trip-counts are estimated with a negative binomial regression for truncated count-data. The no-choice alternative means that the persons either did not boat in our boating region or did not go boating during the past year.

[a]Standard errors are in parentheses and are asymptotic (large-sample).

[b]Total allocations of trips are computed by multiplying the lake probabilities by a constant of 63,366 registered boat owners in 1995 in the region and by the mean annual trip-counts per lake.

lakes model, discrete-count models enable the planner to make computations of recreation benefits in the conventional manner. For example, the consumer surplus per trip with the price index model (Hausman et al., 1995) is $35.33, the expected price/quality model (Feather et al., 1995) is $19.51, and pooled lakes model is $38.36.

In conclusion, discrete-count models provide a comprehensive approach for estimating the total demand for the annual counts of recreation trips and for predicting trip allocations among substitute sites. A specification of the importance of lake attributes did enhance the nested logit results. However, a most time-consuming aspect in applying the nested logit process is the need to correctly specify the lake choice model in a manner consistent with utility maximization theory. Theory stipulates that alternative choices be grouped in such a manner as to avoid the independence from irrelevant alternatives property, which requires that the inclusive value parameters be on a one unit interval. If parameter estimates are greater than one, we would reject the model specification and re-estimate the choice model with different explanatory variables or with different combinations of lakes grouped into different geographical areas. In this study, we could have ended our analysis with the constrained choice model, displayed in Table 2, and interpreted the results in a purely statistical sense since one of the inclusive value parameters is greater than one. That is, we could have represented the magnitudes of inclusive parameters as just the degrees of substitutability between alternative lakes with no concern toward being consistent with utility maximization theory (Kling & Herriges, 1995).

As a final point, the linking of site selection decisions on independent choice occasions to aggregate recreation demand is controversial (Smith, 1996). Concerns center on the diminishing marginal value of trips and the variety of trips, which may have an impact on the pattern of seasonal outcomes across sites and ultimately on the aggregate demand (Smith). Technically, the linkage is not theoretically consistent because the quantity index used does not match the price index. At this time however, and to our knowledge, no alternative travel cost model is offered that avoids the criticism about the approximate nature of the proposed solutions of the price indexes. Withstanding this criticism, we demonstrated the discrete-count approach with an empirical application to regional boating participation because of its usefulness to recreation planners in allocating trips among sites, estimating regional recreation use, and computing consumer surplus values.

References

Adamowicz, W. L. (1994). Habit formation and variety seeking in discrete choice model of recreation demand. *Journal of Agricultural and Resource Economics 19*(1), 19-31.

Bockstael, N. E., Hanemann, W. M., & Kling, C. L. (1987). Estimating the value of water quality improvements in a recreation demand framework. *Water Resources Research 23*(5), 951-960.

Bockstael, N. E., McConnell, K. E., & Strand, I. E. (1991). Recreation. In J. B. Braden & C. D. Kolstad (Eds.), *Measuring the demand for environmental quality.* B. V. North-Holland: Elsevier Science Publishers.

Bockstael, N. E., K. E. McConnell, & Strand, I. E. (1989). A random utility model for sport fishing: Some preliminary results for Florida. *Marine Resource Economics 6,* 245-260.

Cesario, F. J., & Knetsch, J. L. (1976). A recreation site demand and benefit estimation model. *Regional Studies 10,* 97-104.

Clark, R. N., & Downing, K. B. (1984). Why here and not there: The conditional nature of recreation choice. In G. H. Stankey & S. F. McCool (Eds.), *Proceedings—Symposium on Recreation Choice Behavior* (Gen. Tech. Report INT-184, pp. 31-37). Inter-mountain Research Station, Ogden, UT: Forest Service.

Ditton, R. B., Loomis, D. K., & Choi S. (1992). Recreation specialization: Re-conceptualization from a social worlds perspective. Journal of Leisure Research 21(3), 33-51.

Feather, P., Hellerstein, D., & Tomasi, T. (1995). A discrete-count model of recreation demand. *Journal of Environmental Economics and Management 29,* 214-227.

Freeman, A. M. (1993). *The measurement of environmental and resource values.* Washington DC: Resources for the Future.

Greene, W. H. (1993). *Econometric analysis* (2nd ed.). NY: MacMillian Publishing Co.

Greene, W. H. (1995). *LIMDEP user's manual* (Ver. 7). NY: Econometric Software, Inc.

Hamilton, L. C. (1992). *Regression with graphics.* Belmont, CA: Wadsworth, Inc.

Hausman, J. A., Leonard, G. K., & McFadden, D. (1995). A utility-consistent, combined discrete choice and count-data model. *Journal of Public Economics 56,* 1-30.

Hensher, D. A., & Johnson, L. W. (1981). *Applied Discrete-Choice Modelling.* London: Halsted Press.

Hof, J. G., & Rosenthal, D. H. (1987). Valuing the opportunity cost of travel time in recreation demand models: An application to aggregate data. *Journal of Leisure Research 19*(3), 174-188.

Kaoru, Y., Smith, V. K., & Liu, J. L. (1995). Using random utility models to estimate the recreational value of estuarine resources. *American Journal of Agricultural Economics 77,* 141-151.

Lin, Y., Peterson, G. L., & Rogerson, P. A. (1988). A nested urban recreation site choice model. *Leisure Sciences 10,* 1-15.

Morey, E. R., Shaw, W. D., & Rowe, R. D. (1991). A discrete-choice model of recreational participation, site choice, and activity valuation when complete trip data are not available. *Journal of Environmental Economics and Management 20,* 181-201.

Parsons, G. R., & Kealy, M. J. (1995). A demand theory for number of trips in a random utility model of recreation. *Journal of Environmental Economics and Management 29,* 357-367.

Parsons, G. R., & Needelman, M. S. (1992). Site aggregation in a random utility model of recreation. *Land Economics 68*(4), 418-433.

Pollak and Wales (1981). Demographic variables in demand analysis. *Econometric 49,* 1533-1559.

Smith, V. K. (1983). The role of site and job characteristics in hedonic wage models. *Journal of Urban Economics 13,* 296-321.

Stynes, D. J., & Peterson, G. L. (1984). A review of logit models with implications for modeling recreation choices. *Journal of Leisure Research 16,* 295-310.

Williams, D. R. (1984). A developmental model of recreation choice behavior. In G. H. Stankey & S. F. McCool (Eds.), *Proceedings—Symposium on Recreation Choice Behavior* (Gen. Tech. Report INT-184, pp. 31-37). Inter-mountain Research Station, Ogden, UT: Forest Service.

Zhang, J., & Hoffman, S. D. (1993). Discrete-choice logit models. *Sociological Methods & Research 22,* 193-213.