ESTIMATION AND WELFARE CALCULATIONS IN A GENERALIZED CORNER SOLUTION MODEL WITH AN APPLICATION TO RECREATION DEMAND*

by

Daniel J. Phaneuf
Department of Agricultural and Resource Economics, North Carolina State University

Catherine L. Kling, and Joseph A. Herriges
Department of Economics, Iowa State University

Abstract

The Kuhn-Tucker model of Wales and Woodland (1983) provides a utility theoretic framework for estimating preferences over commodities for which individuals choose not to consume one or more of the goods. Due to the complexity of the model, however, there have been few applications in the literature and little attention has been paid to the problems of welfare analysis within the Kuhn-Tucker framework. This paper provides an application of the model to the problem of recreation demand. In addition, we develop and apply a methodology for estimating compensating variation, relying on Monte Carlo integration to derive expected welfare changes.

JEL Classification: C25, Q26

July 1, 1998

Center for Agricultural and Rural Development
Iowa State University
Working Paper 99-WP 207

---

*The authors would like to thank Arne Hallam, Yanis Bilias, Douglas Miller, Noel Cressie, and three anonymous reviewers for their helpful comments on earlier drafts of this paper. We are also indebted to Richard Bishop, Audrey Lyke, and Mike Welsh, who gathered the Wisconsin Great Lakes angling data used in our analysis and generously shared it with us. As usual, any remaining errors are our own. This research was supported in part by the U.S. Environmental Protection Agency and by the Western Regional Research Project W-133. Although the research described in this article has been funded in part by the United States Environmental Protection Agency through R82-3362-010 to the authors, it has not been subject to the Agency's required peer review policy and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred.

I. INTRODUCTION

Traditional econometric methods for modeling consumer demand rely upon the specification of an indirect utility function, Roy's Identity, and the assumption of an interior solution to the consumer's utility maximization problem in order to derive an estimable system of demand equations. There are many applications, however, in which the assumption of an interior solution is unrealistic and, instead, corner solutions prevail. For example, in modeling recreation demand, it is typical to find that most households visit only a small subset of the available sites, setting their demand for the remaining sites to zero.[1] Similar corner solutions emerge in studies of both labor supply (e.g., Ransom (1987a, b), Lacroix and Fortin (1992), and Fortin and Lacroix (1994)) and food demand (e.g., Wales and Woodland (1983) and Yen and Roe (1989)).[2] In these situations, it is well known that failure to allow for the possibility of zero expenditure on one or more goods can lead to inconsistent estimates of consumer preferences.

Two broad strategies have emerged in the literature to deal with corner solutions. The first strategy, labeled the Amemiya-Tobin model by Wales and Woodland (1983), proceeds by initially deriving systems of demand equations without regard to non-negativity restrictions. The model then enforces these restrictions by employing an extension of Tobin's (1958) limited dependent variable model for single equations, later generalized by Amemiya (1974) for systems of equations. In particular, a truncated distribution for the random disturbances is used to ensure non-negative expenditure shares, while allowing for a non-trivial proportion of the sample to have zero expenditure on one or more goods. Applications of the Amemiya-Tobin model have been implemented for a variety of goods. A sampling includes Wales and Woodland's (1983)

analysis of meat demand and Heien and Wessells' (1990) study of general food consumption.

This approach has dominated the recreation demand literature. Single demand models or systems of demands for recreation have been estimated using a variety of estimators, including the tobit, Heckman, and Cragg models (Bockstael, Strand, McConnell, and Arsanjani (1990), Ozuna and Gomez (1994), Smith (1988), and Shaw (1988)), and a variety of count data models (Smith (1988), and Englin and Shonkwiler (1995)). Morey (1984) estimates a system of share equations that adopts a density function assuring strictly positive shares. The strand of this literature that has focused on multiple recreation sites has taken the Amemiya-Tobin model one step further. A two-stage budgeting argument has been used to separately analyze the total number of trips and the allocation of those trips among the available recreation sites.[3] The first stage site selection models use a discrete choice random utility framework. Corner solutions are then explicitly controlled for in the second stage model of the total number of trips using estimators that correct for censoring alone (Bockstael, Hanemann, and Kling (1989); Morey, Waldman, Assane, and Shaw (1990); and Morey, Shaw, and Rowe (1991)) or in combination with count models (Creel and Loomis (1990); Feather, Hellerstein, and Tomasi (1995); Hausman, Leonard, and McFadden (1995); and Yen and Adamowicz (1994)). Although representing a range of estimation approaches, these models all share the Amemiya-Tobin reliance on statistical adjustments to represent corner solutions.

The second strategy for dealing with corner solutions takes a more direct approach to the problem beginning with the consumer's maximization problem. Dubbed the Kuhn-Tucker model by Wales and Woodland (1983), it assumes that individual preferences are

randomly distributed over the population. The standard Kuhn-Tucker conditions characterizing individual utility maximization are then also randomly distributed, providing the basis for probabilistic statements regarding when corner conditions will occur and for constructing the likelihood function used in estimation. Initially developed by Wales and Woodland (1983) and Hanemann (1978) starting with the direct utility function, the approach has subsequently been extended to a dual form starting with the specification of the indirect utility function (Lee and Pitt (1986a) and Bockstael, Hanemann, and Strand (1986)).

The appeal of the Kuhn-Tucker strategy lies in the unified and internally consistent framework it provides for characterizing the occurrence of corner solutions. Since it begins explicitly with a utility function, all of the restrictions of utility theory are automatically satisfied. In addition, the behavioral implications of corner solutions are automatically incorporated.[4] However, due to the complexity of the model, there have been few applications (e.g., Wales and Woodland (1983), Lee and Pitt (1986b), Srinivasan and Winer (1994), and Ransom (1987a)) and none in the area of recreation demand.[5] Furthermore, little attention has been paid to the problem of welfare analysis within the Kuhn-Tucker framework. Due to the non-linearity of the model, closed form solutions for compensating or equivalent variation will typically not be available, requiring instead the use of Monte Carlo integration techniques.

The purpose of this paper is two-fold. First, we provide an empirical application of the Kuhn-Tucker model to the problem of recreation demand and site selection, modeling the demand for fishing in the Wisconsin Great Lakes region. Federal and state agencies are actively involved in management of the local fish populations and

environmental conditions in this region. Understanding the demand for the resulting recreation opportunities will allow regulators to better evaluate existing programs and the impact of potential policy changes. Second, we develop and apply a methodology for estimating compensating variation in the context of the Kuhn-Tucker model, relying on Monte Carlo integration to derive the expected welfare effects of several hypothetical policy changes in the Great Lakes region.

II. MODEL SPECIFICATION

A. Behavioral Model

The Kuhn-Tucker model begins with the assumption that consumers preferences over a set of M+1 commodities can be represented by a random utility function, which they maximize subject to a budget constraint and a set of non-negativity constraints. In particular, each consumer solves:

$$\underset{\mathbf{x},z}{Max}\ U(\mathbf{x},z,\mathbf{q},\gamma,\varepsilon) \tag{1}$$

s.t.

$$\mathbf{p}'\mathbf{x}+z\le y \tag{2a}$$

and

$$z\ge 0, x_j\ge 0,\ j=1,\ldots,M \tag{2b}$$

where $U(\cdot)$ is assumed to be a quasi-concave, increasing, and continuously differentiable function of $(\mathbf{x},z)$, $\mathbf{x}=(x_1,\ldots,x_M)'$ is a vector of goods to be analyzed, z is the numeraire good, $\mathbf{p}=(p_1,\ldots,p_M)'$ is a vector of commodity prices, y denotes income, and $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_M)'$ is a vector of random disturbances capturing the variation in preferences in the population. The disturbance vector is assumed to be known to the individual, but

unobservable by the analyst. The vector $\mathbf{q} = (q_1, \ldots, q_M)'$ represents attributes of the M commodities.[6] The inclusion of commodity attributes is particularly important in recreation demand studies since policy analysis is often interested in the welfare implications of changing the environmental quality of a site.

The first-order necessary and sufficient Kuhn-Tucker conditions for the utility maximization problem are then given by:

$$U_j(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon) \equiv \frac{\partial U(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon)}{\partial x_j} \leq \lambda p_j, \; x_j \geq 0, \; x_j\left[U_j(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon) - \lambda p_j\right] = 0 \quad j = 1, \ldots, M, \tag{3a}$$

$$U_z(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon) \equiv \frac{\partial U(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon)}{\partial z} \leq \lambda, \;\; z \geq 0, \;\; z\left[U_z(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon) - \lambda\right] = 0, \tag{3b}$$

and

$$\mathbf{p}'\mathbf{x} + z \leq y, \lambda \geq 0, \; (y - \mathbf{p}'\mathbf{x} - z)\lambda = 0, \tag{3c}$$

where λ denotes the marginal utility of income. For simplicity, we assume that the numeraire good is a necessary good, so that equation (3b) can be replaced by

$$\lambda = U_z(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon). \tag{3b′}$$

In addition, since $U(\cdot)$ is increasing in $\mathbf{x}$ and z, the budget constraint will be binding, with

$$z = y - \mathbf{p}'\mathbf{x}. \tag{3c′}$$

Substituting equations (3b′) and (3c′) into (3a) yields the M first-order conditions associated with the commodities of interest:

$$U_j(\mathbf{x}, y - \mathbf{p}'\mathbf{x}; \mathbf{q}, \gamma, \varepsilon) \leq p_j U_z(\mathbf{x}, y - \mathbf{p}'\mathbf{x}; \mathbf{q}, \gamma, \varepsilon), \; x_j \geq 0, x_j\left[U_j - U_z p_j\right] = 0 \;\; j = 1, \ldots, M. \tag{3a′}$$

Finally, we assume that $U_{z\varepsilon} = 0$, $\partial U_j / \partial \varepsilon_k = 0 \ \forall k \neq j$, and $\partial U_j / \partial \varepsilon_j > 0 \ \forall j = 1,\ldots,M$, so that[7]

$$U_j(\mathbf{x}, y - \mathbf{p}'\mathbf{x}; \mathbf{q}, \gamma, \varepsilon) = \tilde{U}_j(\mathbf{x}, y - \mathbf{p}'\mathbf{x}; \mathbf{q}, \gamma, \varepsilon_j), \ j = 1,\ldots,M, \tag{4a}$$

with $\partial \tilde{U}_j / \partial \varepsilon_j > 0 \ \forall j = 1,\ldots,M$ and

$$U_z(\mathbf{x}, y - \mathbf{p}'\mathbf{x}; \mathbf{q}, \gamma, \varepsilon) = \tilde{U}_z(\mathbf{x}, y - \mathbf{p}'\mathbf{x}; \mathbf{q}, \gamma). \tag{4b}$$

Let $g_j = g_j(\mathbf{x}, y, \mathbf{p}; \mathbf{q}, \gamma)$ (j=1,…,M) be implicitly defined as the solution to

$$\tilde{U}_j(\mathbf{x}, y - \mathbf{p}'\mathbf{x}; \mathbf{q}, \gamma, g_j) - \tilde{U}_z(\mathbf{x}, y - \mathbf{p}'\mathbf{x}; \mathbf{q}, \gamma) p_j = 0. \tag{5}$$

The first-order conditions in equation (3a′) can be then be rewritten as:

$$\varepsilon_j \leq g_j(\mathbf{x}, y, \mathbf{p}; \mathbf{q}, \gamma), \ x_j \geq 0, \ x_j\left[\varepsilon_j - g_j(\mathbf{x}, y, \mathbf{p}; \mathbf{q}, \gamma)\right] = 0 \quad j = 1,\ldots,M. \tag{6}$$

Equation (6), along with the specification of the joint density function $f_\varepsilon(\varepsilon)$ for ε, provides the necessary information to construct the likelihood function for estimation. Consider an individual who chooses to consume positive quantities for only the first k commodities (i.e., $x_j > 0, \ j = 1,\ldots,k$ and $x_j = 0, \ j = k+1,\ldots,M$). The complementary slackness condition in equation (6) implies that $\varepsilon_j = g_j(\mathbf{x}, y, \mathbf{p}; \mathbf{q}, \gamma)$ for the consumed commodities (i.e., $j = 1,\ldots,k$), while for the remaining commodities (i.e., j=k+1,…,M) we know only that $\varepsilon_j \leq g_j(\mathbf{x}, y, \mathbf{p}; \mathbf{q}, \gamma)$. Thus, this individual's contribution to the likelihood function is given by the probability

$$\int_{-\infty}^{g_{k+1}} \cdots \int_{-\infty}^{g_M} f_\varepsilon(g_1,\ldots,g_k,\varepsilon_{k+1},\ldots,\varepsilon_M)\, abs|J_k| \, d\varepsilon_{k+1} \cdots d\varepsilon_M \tag{7}$$

where J_k denotes the Jacobian for the transformation from ε to $(x_1,\ldots,x_k,\varepsilon_{k+1},\ldots,\varepsilon_M)'$. There are 2^M possible patterns of binding non-negativity constraints for which a probability

statement such as (7) can be constructed. The likelihood function can then be formed as the product of the appropriate probabilities and maximum likelihood can be used to recover estimates of the utility function's parameters.

B. Conditional Utility Functions and the Computation of Welfare Effects

A common reason for estimating the structure of consumer preferences over a set of commodities is to provide a basis for welfare analysis. In particular, policymakers may be interested in the welfare implication of changing the price or quality characteristics of the existing set of alternatives, or of reducing the number of alternatives available. Formally, let $V(\mathbf{p}, y; \mathbf{q}, \gamma, \varepsilon)$ denote the solution to the utility maximization defined in equations (1) and (2) above. The compensating variation (C) associated with a change in the price and attribute vectors from $(\mathbf{p}^0, \mathbf{q}^0)$ to $(\mathbf{p}^1, \mathbf{q}^1)$ is implicitly defined by

$$V(\mathbf{p}^0, y; \mathbf{q}^0, \gamma, \varepsilon) = V(\mathbf{p}^1, y + C(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma, \varepsilon); \mathbf{q}^1, \gamma, \varepsilon)\,. \tag{8}$$

There are several important attributes of the compensating variation measure that are worthy of note. First, from the analyst's perspective, $C(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma, \varepsilon)$ is a random variable. Policy makers will typically be interested in the average value of this measure in the population, $\overline{C}(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma)$. Second, the non-linearity of the utility maximization problem will typically preclude a closed form solution for C or its average. As a result, numerical techniques will be required.[8]

The process of computing C can be clarified by considering the utility maximization as a two-stage process, in which the individual maximizes his or her utility conditional on a set of binding non-negativity constraints and then chooses among the resulting conditional indirect utility functions.[9] Formally, let

$$A = \{\varnothing, \{1\}, \ldots, \{M\}, \{1,2\}, \{1,3\}, \ldots, \{1,2,\ldots, M\}\} \tag{9}$$

denote the collection of all possible subsets of the index set $I = \{1,\ldots, M\}$. A conditional indirect utility function $V_\omega(\mathbf{p}_\omega, y; \mathbf{q}, \gamma, \varepsilon)$ can then be defined for each $\omega \in A$ as the maximum utility level the consumer can achieve when they are restricted to the commodities indexed by ω. Formally:

$$V_\omega(\mathbf{p}_\omega, y; \mathbf{q}, \gamma, \varepsilon) = \underset{\mathbf{x},z}{Max} U(\mathbf{x}, z, \mathbf{q}, \gamma, \varepsilon) \tag{10}$$

s.t.

$$\sum_{j\in\omega} p_j x_j + z \leq y \tag{11a}$$

and

$$z \geq 0, x_j = 0,\ j \notin \omega,\ x_j \geq 0\ \ j \in \omega, \tag{11b}$$

where $\mathbf{p}_\omega = \{p_j : j \in \omega\}$ is the vector of commodity prices that have not been constrained to zero. Let $\mathbf{x}_\omega(\mathbf{p}_\omega, y; \mathbf{q}, \gamma, \varepsilon)$ denote the conditional demand levels solving this utility maximization problem. Notice that, since the prices associated with those commodities that have been forced to zero do not enter the budget constraint in (11a), V_ω and $\mathbf{x}_\omega$ are both functions of $\mathbf{p}_\omega$ and not $\mathbf{p}$. However, both the conditional indirect utility function and conditional demand equations will depend on the entire vector of quality attributes, $\mathbf{q}$, and not simply $\mathbf{q}_\omega = \{q_j : j \in \omega\}$, unless the property of weak complementarity is imposed (Maler, 1974).[10]

Constraining a subset of the commodities to have zero consumption provides, of course, no assurance that the optimal consumption levels for the remaining commodities will be positive. Let

$$\tilde{A} \equiv \tilde{A}(\mathbf{p}, y; \mathbf{q}, \gamma, \varepsilon) = \left\{\omega \in A: x_{\omega j}(\mathbf{p}_\omega, y; \mathbf{q}, \gamma, \varepsilon) > 0, \forall j \in \omega\right\} \tag{12}$$

denote the collection of ω's for which the corresponding conditional utility maximization problem yields an interior solution. The original consumer utility maximization problem can then be viewed as a two-stage problem in which conditional indirect utility functions are computed for each $\omega \in A$ and then the consumer chooses the V_ω that maximizes his or her utility. That is[11]

$$V(\mathbf{p}, y; \mathbf{q}, \gamma, \varepsilon) = \underset{\omega \in A}{Max}\left\{V_\omega(\mathbf{p}_\omega, y; \mathbf{q}, \gamma, \varepsilon)\right\} = \underset{\omega \in \tilde{A}}{Max}\left\{V_\omega(\mathbf{p}_\omega, y; \mathbf{q}, \gamma, \varepsilon)\right\}. \tag{13}$$

The computation of the compensating variation in equation (8) then corresponds to implicitly solving for $C(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma, \varepsilon)$ in

$$\underset{\omega \in \tilde{A}^0}{Max}\left\{V_\omega(\mathbf{p}^0_\omega, y; \mathbf{q}^0, \gamma, \varepsilon)\right\} = \underset{\omega \in \tilde{A}^1}{Max}\left\{V_\omega(\mathbf{p}^1_\omega, y + C(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma, \varepsilon); \mathbf{q}^1, \gamma, \varepsilon)\right\}. \tag{14}$$

Notice that the index collection $\tilde{A}$ may change as a result of the changing price and/or quality attribute levels.[12]

There are three difficulties associated with computing $\overline{C}(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma)$ in practice. First, for any given ε and γ, $C(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma, \varepsilon)$ is an implicit function for which no closed form solution typically exists. However, numerical procedures, such as numerical bisection, can be readily applied to solve this problem.

Second, given $C(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma, \varepsilon)$ and γ, $\overline{C}(\mathbf{p}^0, \mathbf{q}^0, \mathbf{p}^1, \mathbf{q}^1, y; \gamma)$ does not have a closed form solution. However, Monte Carlo integration can be used, resampling from the

underlying distribution of ε, $f_\varepsilon(\varepsilon)$, and averaging $C(\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,y;\gamma,\varepsilon)$ over the draws of ε. For many error distribution assumptions it is possible to resample directly from the underlying probability density function to perform the Monte Carlo integration.[13] As is discussed below, we assume that the ε's are drawn from a Generalized Extreme Value (GEV) distribution, for which it is not possible to resample directly from the density function. However, McFadden (1995) has recently developed a Monte Carlo Markov Chain approach that we will employ in this paper. The Markov Chain Simulator does not draw directly from a GEV distribution, but rather constructs a sequence of variates that asymptotically exhibit the properties of a GEV distribution.[14]

Third, given an algorithm for computing $\overline{C}(\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,y;\gamma)$, the analyst does not typically have available γ, but instead must rely upon an estimator $\hat{\gamma} \sim g_{\hat{\gamma}}$ (e.g., the maximum likelihood estimator of γ). Thus, any computation of $\overline{C}$ will itself be a random variable, dependent upon the distribution of $\hat{\gamma}$. We bootstrap the data to approximate the statistical properties of $\hat{C}$, our estimate of $\overline{C}$. Formally, the above elements are combined into the following numerical algorithm:

- Resample with replacement (n=1,…,N) observations from the original data set. Re-estimate the model using this pseudo data and repeat this procedure N_γ times to yield a total of N_γ parameter vectors (i.e., $\gamma^{(i)}$, $i = 1,\ldots,N_\gamma$).

- For each $\gamma^{(i)}$ and each observation in the sample ($n = 1,\ldots,N$), McFadden's Monte Carlo Markov Chain Simulator is implemented to generate N_ε vectors of random disturbance terms (i.e., $\varepsilon^{(ink)}$, $k = 1,\ldots,N_\varepsilon$).

- Substituting $\gamma^{(i)}$ and $\varepsilon^{(ink)}$ for γ and ε in equation (14), numerical bisection can then be used to solve for C, with the result labeled $C^{(ink)}$.

- Averaging $C^{(ink)}$ over the N_ε draws from the disturbance distribution and the N observations in the sample yields $\hat{C}^{(i)}$, a Monte Carlo integration evaluation of $E_\varepsilon\left[C(\mathbf{p}^0,\mathbf{q}^0,\mathbf{p}^1,\mathbf{q}^1,y;\gamma^{(i)},\varepsilon)\right]$.

- The distribution of $\hat{C}^{(i)}$'s provides the basis for characterizing the distribution of the mean compensating variation of interest ($\overline{C}$) in light of our uncertainty regarding γ. The mean value of $\hat{C}^{(i)}$ over the N_γ parameter draws provides a consistent estimate of $\overline{C}$. The distribution of the $\hat{C}^{(i)}$'s can be used to construct standard errors for our estimate of $\overline{C}$.

C. Empirical Specification

In our application below, we employ the empirical specification suggested by Bockstael, Hanemann, and Strand (1986). In particular, we assume that the consumer's direct utility function is a variant of the linear expenditure system, with

$$U(\mathbf{x},z;\mathbf{q},\gamma,\varepsilon)=\sum_{j=1}^{M}\Psi_j(q_j,\varepsilon_j)\ln(x_j+\Omega)+\ln(z) \tag{15}$$

and

$$\Psi_j(q_j,\varepsilon_j)=\exp\left(\sum_{k=1}^{K}\delta_k q_{jk}+\varepsilon_j\right)\quad j=1,\ldots,M \tag{16}$$

where $\gamma=(\delta,\Omega)$ and q_{jk} denotes the k^{th} quality attribute associated with commodity *j*. The Ψ_j's can be thought of as quality indices associated with each good.[15]

One advantage of the above utility function is that the implicit equations for the ε_j's in equation (4) that result from the Kuhn-Tucker conditions can be explicitly solved, yielding the following equivalent first-order conditions:

$$\varepsilon_j\le g_j(\mathbf{x},y,\mathbf{p};\mathbf{q},\gamma),\ \ x_j\ge 0,\ \ x_j\left[\varepsilon_j-g_j(\mathbf{x},y,\mathbf{p};\mathbf{q},\gamma)\right]=0\ \ j=1,\ldots,M, \tag{17}$$

where

$$g_j(\mathbf{x}, y, \mathbf{p}; \mathbf{q}, \gamma) = \ln\left[\frac{p_j(x_j + \Omega)}{y - \sum_{j=1}^{M} p_j x_j}\right] - \sum_{k=1}^{K} \delta_k q_{jk} \quad j = 1, \ldots, M. \tag{18}$$

Specifying a joint distribution for the random disturbances (i.e., $f_\varepsilon(\varepsilon)$) completes the empirical model. As mentioned above, we assume that the ε's are drawn from a generalized extreme value (or GEV) distribution. An important feature of this specification is that it allows for correlation among the alternative sites, while still yielding closed form equations for the probabilities in the likelihood function. The sites are grouped into "nests" of alternatives that are assumed to have correlated error terms, with $s(j)$ denoting the nest to which alternative *j* is assigned and *S* denoting the total number of nests.[16] Given the specification of a nesting structure, the cdf associated with ε is given by:

$$F(\varepsilon_1, \ldots, \varepsilon_J) = \exp\left[-\sum_{r=1}^{S}\left(\sum_{\substack{j \in \omega \\ s(j)=r}} \exp\left(\frac{-\varepsilon_j}{\mu\theta_r}\right)\right)^{\theta_r}\right] \tag{19}$$

where μ is a scale parameter determining a common variance for the error terms and the θ_k's are the dissimilarity coefficients measuring the degree of correlation between the errors in a given nest.[17] The log-likelihood function is then obtained by forming choice probabilities for each consumption pattern (i.e., each $\omega \in \tilde{A}$), integrating the corresponding pdf for ε as indicated by equation (7) above.[18]

III. DATA

Our empirical application of the Kuhn-Tucker model focuses on angling in the Wisconsin Great Lakes region. The data are drawn primarily from two mail surveys of angling behavior conducted in 1990 by Richard Bishop and Audrey Lyke at the University of Wisconsin-Madison.[19] The surveys provide detailed information on the 1989 angling behavior of Wisconsin fishing license holders, including the number and destination of fishing trips to the Wisconsin Great Lakes region, the distances to each destination, the type of angling preferred, and the socio-demographic characteristics of the survey respondents. A total of 509 completed surveys were available for analysis, including 266 individuals who had fished the Wisconsin Great Lakes region for lake trout or salmon and 247 who fished only inland waters of Wisconsin (i.e., non-users from the perspective of the Great Lakes region). While the surveys provide data on 22 distinct Great Lake fishing destinations, we have combined these destinations into four aggregate "sites":

- Site 1: Lake Superior
- Site 2: South Lake Michigan
- Site 3: North Lake Michigan, and
- Site 4: Green Bay.

This aggregation divides the Wisconsin portion of the Great Lakes into distinct geographical zones consistent with the Wisconsin Department of Natural Resources' classification of the lake region.[20]

The price of a single trip to each of the four fishing sites consists of two components: the cost of getting to the site (i.e., direct travel cost) and the opportunity cost

of the travel time. Round trip direct travel costs were computed for each destination and each individual by multiplying the number of round trip miles for a given individual-destination combination by the cost per mile for the vehicle class driven, as provided by the American Automobile Association. The cost of the travel time was constructed using one-third of the individual's wage rate as a measure of the hourly opportunity cost of recreation time and assuming an average travel speed of forty-five miles per hour to compute travel time.[21] The price of visiting a destination p_j is then the sum of the direct travel cost and the cost of the travel time. Total annual income was collected and used for the income variable in estimation.

Two types of quality attributes (i.e., q_{js}'s) are used to characterize the recreation sites: fishing catch rates and toxin levels. Catch rates are clearly important site characteristics since the anticipated success of fishing is likely to be a major determinant in the recreation decision. Furthermore, state and federal agencies currently spend large amounts of time and money to influence catch rates in the region through stocking programs and regulations. The inclusion of catch rates as a quality attribute in the model will allow it to be used to conduct welfare analyses of existing and/or alternative fishery management programs.

In constructing the catch rate variables, we focus our attention on the catch rates for the four aggressively managed salmonoid species: lake trout, rainbow (or steelhead) trout, Coho salmon, and Chinook salmon. Creel surveys by the Wisconsin Department of Natural Resources provide 1989 catch rates for each of these species at each of the 22 disaggregate destinations used in the angling surveys. Furthermore, these catch rates are broken down by angling method, including private boat, charter fishing, and pier/shore

angling. Data from the Wisconsin angling survey were used to match the mode-specific catch rates to each individual anglers based upon their most frequent mode of fishing.

We include toxin levels as an additional quality attribute of each site since the presence of environmental contaminants is likely to influence the recreation decision and they provide a proxy for the overall level of water quality at the site (De Vault *et al.* (1996)). Toxins are found in varying levels in fish, water, and sediments throughout the Great Lakes and are routinely responsible for health warnings in the regions. De Vault *et al.*(1989) provide a study of toxin levels in lake trout during the relevant time period, with samples taken from locations throughout the Great Lakes. We use the average toxin levels (ng/kg-fish) from this study, matched on the basis of proximity to our four aggregate sites, to form a basic toxin measure T_j $(j = 1,\ldots,4)$ for each site.[22] However, toxin levels are likely to influence visitation decisions only if the consumer perceives that the toxins create a safety issue. The Wisconsin angling survey asked respondents if the toxin levels in fish were of concern to them. We use this information to form an "effective toxin level" variable $E_j = T_j D$ $(j = 1,\ldots,4)$, where $D = 1$ indicates that the respondent was concerned about the toxin levels in fish and $D = 0$ otherwise.

With both catch rates and toxins included as quality variables, the quality index terms from equation (16) become

$$\Psi_j(\mathbf{q}_j, \varepsilon_j) = \exp\left[\delta_0 + \delta_{lk} R_{lk,j} + \delta_{ch} R_{ch,j} + \delta_{co} R_{co,j} + \delta_{rb} R_{rb,j} + \delta_E E_j + \varepsilon_j\right],\ j = 1,\ldots,4, \quad (20)$$

where $R_{k,j}$ denotes the catch rate for species k and site j, with k=*lk* for lake trout, *ch* for Chinook salmon, *co* for Coho salmon, and *rb* for rainbow trout.

Tables 1 and 2 provide summary statistics for the data. Table 1 focuses on the mean and standard deviation of the usage, price, and quality characteristics for the four sites used in our analysis. Table 2 characterizes the trip usage patterns (i.e., ω) found in the Wisconsin angling survey data. Note that, while many (72%) of the visitors to the Great Lakes sites visit only one of the sites, a substantial percentage (28%) visit more than one site. Thus, neither an extreme corner solution (Hanemann (1984)) nor an interior solution model could accurately depict this group of consumers' choices.

Finally, to implement the model using the GEV distribution (19), we must choose the nesting structure (implicitly specifying which set of sites exhibit correlated behavior). There are several plausible possibilities for nesting structures. For example, we may expect that Green Bay and North Lake Michigan would be correlated, since they are geographically close. Similarly, it may be reasonable to include North Lake Michigan and South Lake Michigan in a nest, since they exhibit similar physical characteristics. One could also include Lake Superior and Green Bay in a nest, since they are the more remote sites with respect to the population centers in southern Wisconsin. Combinations of these can be specified using multiple nests. In fact, we estimated the model with many combinations of nesting patterns and report results for the model that yielded the best fit of the data (based on the likelihood dominance criteria).[23] The final nesting structure used places North Lake Michigan and Green Bay in one nest and South Lake Michigan and Green Bay in a second nest. We estimate separate dissimilarity coefficients for each nest to allow the degree of correlation between the alternatives within a nest to differ.

IV. RESULTS

A. Model Estimation

Two Kuhn-Tucker models of Wisconsin Great Lakes angling were estimated using maximum likelihood, yielding the parameter estimates provided in Table 3.[24] In the first model, the dissimilarity coefficients in the GEV distribution are constrained to equal one, yielding a simple extreme value distribution and implying independence among the alternative site choices. All of the parameters have the expected signs and, with the exception of the coefficient on lake trout catch rates, are statistically different from zero at a 5% critical level or less. For example, one would expect, and we find, that higher toxins reduce the perceived quality of a site (i.e., $\delta_E < 0$). On the other hand, higher catch rates should enhance site quality (i.e., $\delta_k > 0$). This is the case for each of the fish species considered. Furthermore, the small and statistically insignificant coefficient on lake trout is not unexpected, since among anglers lake trout are typically considered a less desirable species. The other salmon species have a "trophy" status not shared by lake trout. In addition, the eating quality of lake trout is generally considered inferior to that of other species. Finally, the parameter Ω is estimated to be 1.76 and is significantly different from 1.00, indicating that weak complementarity (Maler) can be rejected.

The second model reported in Table 3 allows for correlation among the alternative sites by employing a GEV specification for the distribution of the error terms. Separate dissimilarity coefficients are estimated for the {North Lake Michigan, Green Bay} nest (θ^N) and the {South Lake Michigan, Lake Superior} nest (θ^S).[25] Indeed, a simple likelihood ratio test indicates that the GEV model provides a significantly better fit to the data. However, the resulting coefficient estimates are generally of the same magnitude as

the estimates from the extreme value model. Again, the parameters are estimated to be of the expected sign and generally significant except for the parameter of Lake Trout which is negative, but insignificant.

B. Welfare Analysis

One of the motivations for estimating models of recreation demand is to provide policy makers with estimates of the welfare implications of changing environmental quality or site availability. A primary advantage of the Kuhn-Tucker model is that it permits the construction of these welfare estimates in an internally consistent and utility-theoretic framework. The model simultaneously predicts changes to the sites visited and the total number of trips taken, which in turn determines changes in consumer utility. In this subsection, we use the estimated Kuhn-Tucker models in Table 3, along with the numerical procedures developed above, to evaluate a series of policy scenarios for the Wisconsin Great Lakes region.

The Great Lakes region provides many opportunities for policy-relevant welfare experiments as the lakes are heavily managed. The fishery itself is, in many ways, artificially created and maintained. Of the major species included in the model, only lake trout are native to both Lake Superior and Lake Michigan. Rainbow trout were introduced around the turn of the century, while the salmon species were not present until the 1950's. These species now reproduce naturally in the lakes, but are heavily augmented with stocking programs. The lakes have also been invaded by exotic species, including the sea lamprey. A parasite accidentally introduced in the 1930's, the sea lamprey decimated lake trout populations in the lakes. Efforts to reintroduce naturally reproducing lake trout to Lake Superior have been successful, while in Lake Michigan the population is completely

maintained through stocking. Expensive sea lamprey control efforts continue to this day. Finally, there are ongoing efforts throughout the Great Lakes region to improve the fisheries by reducing the level of toxins entering the food chain from commercial and industrial sources. For each of these forms of intervention, the natural policy question arises as to whether the benefits of these programs are sufficient to offset the corresponding costs. Our Kuhn-Tucker models can be used to assess program benefits. As an illustration of this capability, we estimate welfare loss under three policy scenarios:

- Scenario A: Loss of Lake Michigan Lake Trout. Under this first policy scenario, state and local efforts to artificially stock lake trout in Lake Michigan and Green Bay would be eliminated. It is assumed that this would drive lake trout catch rates ($R_{lk,j}$) to zero for sites 2, 3, and 4, since the species is only naturally reproducing in Lake Superior (site 1).[26]

- Scenario B: Loss of Lake Michigan Coho Salmon. Under this policy scenario, state and local efforts to artificially stock Coho salmon in Lake Michigan and Green Bay would be suspended. Again, it is assumed that the corresponding Coho catch rates ($R_{co,j}$) would be driven to zero for sites 2, 3, and 4.

- Scenario C: Reduced Toxin Levels. Under the final policy scenario, we consider the welfare implications of a twenty percent reduction in toxin levels (i.e., $E_j, \ \ j = 1,2,3,4$).

Of particular interest from a policy perspective is Scenario A, given the intense ongoing efforts to rehabilitate the lake trout population in Lake Michigan. Without these efforts and the continuing sea lamprey control programs, catch rates for lake trout at the three sites would be zero, as analyzed in the scenario. Although the elimination of Coho salmon stocking programs would not reduce the catch rates to zero due to natural reproduction in the fishery, we nonetheless use this assumption in Scenario B for comparison purposes. Finally, we include Scenario C to demonstrate the ability of the

model to measure the benefits of improvements in general environmental quality, given that toxins in fish flesh can act as a proxy for many other types of pollution.

For each of these scenarios, mean compensating variation ($\overline{C}$) was estimated using GAUSS and the procedures outlined in Section IIB above. In particular,

- A total of $N_\gamma = 200$ parameter vectors (i.e., $\gamma^{(i)}$, $i = 1,\ldots,N_\gamma$) were generated from bootstrapping the original data as described above.
- For each $\gamma^{(i)}$ and each observation in the sample ($n = 1,\ldots,509$), a total of $N_\varepsilon = 2000$ vectors of random disturbance terms (i.e., $\varepsilon^{(ink)}$, $k = 1,\ldots,N_\varepsilon$) were formed via the Monte Carlo Markov Chain Simulator.[27] Specifically, at iteration *ink*, a pseudo-random number generator is used to draw 5 independent (0,1) uniform random variables, $\zeta_j^{(ink)}$ $(j = 1,\ldots,4)$ and $\eta^{(ink)}$. Four extreme value random variates (one for each site) are then formed using the transformation $\tilde{\varepsilon}_j^{(ink)} = -\log\left(-\log\left(\zeta_j^{(ink)}\right)\right)$. Finally, the following Markov chain is used to construct:

$$\varepsilon^{(ink)} = \begin{cases} \tilde{\varepsilon}^{(ink)} & \text{if } \eta^{(ink)} \le \dfrac{f\left(\tilde{\varepsilon}^{(ink)}\right)/g\left(\tilde{\varepsilon}^{(ink)}\right)}{f\left(\varepsilon^{(in,k-1)}\right)/g\left(\varepsilon^{(in,k-1)}\right)} \\ \varepsilon^{(in,k-1)} & \text{otherwise} \end{cases} \tag{21}$$

where $f(\cdot)$ and $g(\cdot)$ denote the GEV and EV probability density functions, respectively.

- Substituting $\gamma^{(i)}$ and $\varepsilon^{(ink)}$ for γ and ε in equation (14), numerical bisection was then used to solve for *C*, with the result labeled $C^{(ink)}$.
- Averaging $C^{(ink)}$ over the N_ε draws from the disturbance distribution and the *N* observations in the sample yields an estimate ($\hat{C}^{(i)}$) of the mean compensating variation for the j[th] draw from the estimated parameter distribution.

The distribution of the $\hat{C}^{(i)}$'s provides the basis for characterizing the distribution of the mean compensating variation of interest ($\overline{C}$) in light of our uncertainty regarding the parameter estimates in Table 3. The mean value of the $\hat{C}^{(i)}$ over the 200 parameter draws provides a consistent estimate of $\overline{C}$ and is reported in column two of Table 4 for each scenario and model, with the corresponding standard deviations reported in parentheses.[28]

The compensating variations in Table 4 have the expected signs and relative magnitudes, given the parameter estimates in Table 3.[29] As expected, the loss of Coho salmon (Scenario B) has a greater impact on consumer welfare than the loss of lake trout (Scenario A). Focusing on the GEV results, an average of $310 per angler per season would be required to compensate for the loss of Coho salmon in the Lake Michigan and Green Bay sites, whereas the loss of lake trout would actually yield benefits of about $58 per season. It is important to note however, that these lake trout values are not statistically different from zero using any reasonable confidence level. The lake trout results are particularly interesting from a policy perspective, since so much effort has gone into rehabilitating the lake trout fishery during the past three decades.

Turning to Scenario C, we find that a twenty-percent reduction in toxin levels would have a substantial and statistically significant impact on angler welfare. Based on the GEV model, anglers would be willing to pay, on average, $111 per season for such a reduction.

V. SUMMARY AND CONCLUSIONS

In this study, we have provided an empirical application of the Kuhn-Tucker model to the problem of recreation demand, estimating the demand for fishing in the Wisconsin Great Lakes region. We have developed a methodology for estimating

compensating variation within the context of the model and have applied this methodology to measure the welfare effects associated with changes in site catch rates and toxin levels. Although the Kuhn-Tucker model is not new, there have been few applications due to the computational complexity. We have demonstrated that with the availability of faster and cheaper computing power, the model can now be applied to questions of policy interest. This is of particular importance in recreation demand modeling, since the Kuhn-Tucker model is appealing in that it deals with the abundance of general corner solutions in recreation data in an internally consistent and utility theoretic framework. The same model drives both the site selection choice and the total number of trips taken by recreationists. This feature is particularly important to the task of assessing welfare changes.

In our application to the Great Lakes region, we estimate the lost value to anglers of eliminating lake trout from Lake Michigan and Green Bay, the loss of Coho Salmon from Lake Michigan and Green Bay, and the welfare improvements associated with reduced toxin levels in the lakes. In addition to providing point estimates of these welfare measures, we provide information on the reliability of the estimates in the form of standard errors.

There are two areas where improvements to the model estimated here could be made. First, it would be desirable to explore alternative functional forms in the specification of individual utility. The trade-off here, of course, is in identifying forms that are both flexible and yet yield Kuhn-Tucker conditions that generate closed-form probabilities for the likelihood function. Second, it would be desirable to experiment with

error distributions other than the Generalized Extreme Value to investigate the robustness of the results to the assumed error structure.

VI. REFERENCES

Amemiya, T., "Multivariate Regression and Simultaneous Equation Models When the Dependent Variables are Truncated Normal," *Econometrica* 42 (September 1974), 999-1012.

Bockstael, N.E., W.M. Hanemann, and C.L. Kling, "Estimating the Value of Water Quality Improvements in a Recreation Demand Framework," *Water Resources Research* 23 (May 1987), 951-60.

Bockstael, N.E., W.M. Hanemann, and I.E. Strand, *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 (1986).

Bockstael, N.E., I.E. Strand, K.E. McConnell, and F. Arsanjani, "Sample Selection Bias in the Estimation of Recreation Demand Functions: An Application to Sportfishing," *Land Economics* 66 (1990), 40-49.

Creel, M., and J.B. Loomis, "Theoretical and Empirical Advantages of Truncated Count Data Estimators for Analysis of Deer Hunting in California, *American Journal of Agricultural Economics* 72 (May 1990), 434-45.

De Vault, D.S., R. Hesselberg, P.W. Rodgers and T.J. Feist, "Contaminant Trends in Lake Trout and Walleye from the Laurentian Great Lakes," *Journal of Great Lakes Research* 22 (1996), 884-895.

De Vault, D.S., D. Dunn, P. Bergqvist, K. Wiberg and C. Rappe, "Polychlorinated Dibenzofurans and Ploychlorinated Dibenzo-p-Dioxins in Great Lakes Fish: A Baseline and Interlake Comparison," *Environmental Toxicology and Chemistry* 8 (1989), 1013-1022.

Efron, B., and R.J. Tibshirani, *An Introduction to the Bootstrap*, New York: Chapman and Hall, 1993.

Englin, J., and S. Shonkwiler, "Estimating Social Welfare Using Count Data Models: An Application to Long-Run Recreation Demand Under Conditions of Endogenous Stratification and Truncation," *Review of Economics and Statistics* 77 (1995), 104-112.

Feather, P., "Sampling and Aggregation Issues in Random Utility Model Estimation," *American Journal of Agricultural Economics* **76** (1994), 722-779.

Feather, P., D. Hellerstein, and T. Tomasi, "A Discrete-Count Model of Recreation Demand," *Journal of Environmental Economics and Management* 30 (1996), 316-322.

Fortin, B. and G. Lacroix, "Labor Supply, Tax Evasion and the Marginal Cost of Public Funds - An Empirical Investigation," *Journal of Public Economics* 55 (1994), 407-431.

Freeman, A. M., *The Measurement of Environmental and Resource Values: Theory and Methods*, Washington, D.C.: Resources for the Future, 1993.

Fulginiti, L., and R. Perrin, "The Theory and Measurement of Producer Response Under Quotas," *Review of Economics and Statistics* 75 (1993), 97-106.

Geweke, J., "Monte Carlo Simulation and Numerical Integration," in H. Amman, D. Kendrick, and J. Rust, eds., *Handbook of Computational Economics* (Amsterdam: Elsevier Science B.V., 1996).

Hanemann, W.M., *A Methodological and Empirical Study of the Recreation Benefits from Water Quality Improvement*, Ph.D. dissertation, Department of Economics, Harvard University (1978).

Hanemann, W. M., "Discrete-Continuous Models of Consumer Demand," *Econometrica* 52 (1984), 541-561.

Hausman, J. A., G.K. Leonard, and D. McFadden, "A Utility-Consistent, Combined Discrete Choice and Count Data Model: Assessing Recreational Use Losses Due to Natural Resource Damage," *Journal of Public Economics* 56 (Winter 1995), 1-30.

Heien, D., and C. R. Wessells, "Demand Systems Estimation with Microdata: A Censored Regression Approach," *Journal of Business & Economic Statistics* 8 (1990), 365-371.

Herriges, J. A., and C. L. Kling, "Nonlinear Income Effects in Random Utility Models," *Review of Economics and Statistics*, forthcoming, (1997).

Kaoru, Y., V. K. Smith, and J. Liu, "Using Random Utility Models to Estimate the Recreation Value of Estuarine Resources," *American Journal of Agricultural Economics* **77** (1995), 141-151.

Kling, C. L. *Measuring the Recreational Benefits of Environmental Amenities Using Multiple Site Models: An Evaluation of Techniques*, Ph.D. dissertation, University of Maryland, 1986.

Lacroix, G., and B. Fortin, "Utility-Based Estimation of Labor Supply Functions in the Regular and Irregular Sectors," *Economic Journal* 102 (1992), 1407-1422.

Lee, L.F., and M.M. Pitt, “Microeconometric Demand Systems with Binding Nonnegativity Constraints: The Dual Approach,” *Econometrica* 54 (November 1986a), 1237-42.

Lee, L. F., and M.M. Pitt, “Specification and Estimation of Consumer Demand with Many Binding Non-Negativity Constraints,” Center for Economic Research Discussion Paper No. 236, Department of Economics, University of Minnesota, October, 1986.

Lee, L.F., and M.M. Pitt, “Microeconometric Models of Rationing, Imperfect Markets, and Non-Negativity Constraints,” *Journal of Econometrics* 36 (1987), 89-110.

Lyke, A.J., *Discrete Choice Models to Value Changes in Environmental Quality: A Great Lakes Case Study*, Ph.D. dissertation, Department of Agricultural Economics, University of Wisconsin-Madison, 1993.

Maler, K., *Environmental Economics*, Baltimore: Johns Hopkins, 1974.

McConnell, K.E., and I.E. Strand, “Measuring the Cost of Time in Recreation Demand Analysis: An Application to Sport Fishing,” *American Journal of Agricultural Economics* 63 (1981), 153-156.

McFadden, D., "Econometric Models of Probabilistic Choice," in Charles F. Manski and Daniel L. McFadden (eds.), *Structural Analysis of Discrete Choice Data* (Cambridge, MA: MIT Press, 1981).

McFadden, D., “Computing Willingness-to-Pay in Random Utility Models,” working paper, Department of Economics, University of California, Berkeley, December 1995.

Morey, E.R., “The Choice of Ski Areas: Estimation of a Generalized CES Preference Ordering with Characteristics, Quadratic Expenditure Functions, and Non-Additivity,” *Review of Economics and Statistics* 66 (1984), 221-36.

Morey, E.R., W.D. Shaw, and R.D. Rowe, “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 (1991), 181-201.

Morey, E.R., D. Waldman, D. Assane, and D. Shaw, “Specification and Estimation of a Generalized Corner Solution Model of Consumer Demand: An Amemiya-Tobin Approach,” working paper, May 1990.

Morey, E.R., D. Waldman, D. Assane, and D. Shaw, “Searching for a Model of Multiple-Site Recreation Demand that Admits Interior and Boundary Solutions,” *American Journal of Agricultural Economics* 77 (1995), 129-140.

Ozuna, T., and I. Gomez, "Estimating a System of Recreation Demand Functions Using a Seemingly Unrelated Poisson Regression Approach," *Review of Economics and Statistics* 76 (1994), 356-60.

Parsons, G.R., and M.J. Kealy, "Site Aggregation in a Random Utility Model of Recreation Demand," *Journal of Environmental Economics and Management* **68** (1992), 418-433.

Parsons, G.R., and M.J. Kealy, "A Demand Theory for Number of Trips in a Random Utility Model of Recreation," *Journal of Environmental Economics and Management* 29 (1996), 418-433.

Ransom, M.R., "The Labor Supply of Married Men - A Switching Regressions Model," *Journal of Labor Economics* 5 (1987a), 63-75.

Ransom, M.R., "An Empirical Model of Discrete and Continuous Choice in Family Labor Supply," *Review of Economics and Statistics* 69 (1987b), 465-472.

Shaw, D. "On-Site Sample' Regression: Problems of Non-negative Integers, Truncation, and Endogenous Stratification," *Journal of Econometrics* 37 (1988), 211-23.

Smith, V.K., "Selection and Recreation Demand," *American Journal of Agricultural Economics* 70 (1988), 29-36.

Srinivasan, T. C., and R. S. Winer, "Using Neoclassical Consumer-Choice Theory to Produce a Market Map From Purchase Data," *Journal of Business and Economic Statistics* **12** (1994), 1-9.

Tobin, J., "Estimation of Relationships for Limited Dependent Variables, *Econometrica* 26 (1958), 24-36.

Wales, T.J., and A.D. Woodland, "Estimation of Consumer Demand Systems with Binding Nonnegativity Constraints," *Journal of Econometrics* 21 (1983), 263-285.

Yen, S. T., and W. L. Adamowicz, "Participation, Trip Frequency, and Site Choice: A Multinomial Poisson Hurdle Model of Recreation Demand," *Canadian Journal of Agricultural Economics* 42 (1994), 65-76.

Yen, S.T., and T.L. Roe, "Estimation of a 2-Level Demand System with Limited Dependent-Variables," *American Journal of Agricultural Economics* 71 (1989), 85-98

Table 1. – Average Site Characteristics (Standard Deviations in Parentheses)

	Lake Superior	North Lake Michigan	South Lake Michigan	Green Bay
1989 Fishing Trips (x_j)	2.75 (13.33)	1.56 (6.32)	2.35 (8.92)	0.65 (3.07)
Price (p_j)	177.84 (172.59)	123.70 (172.92)	85.88 (139.62)	129.11 (173.54)
Lake Trout Catch Rate ($R_{lk,j}$)	.046 (.059)	.022 (.030)	.029 (.045)	.001 (.002)
Chinook Salmon Catch Rate ($R_{ch,j}$)	.010 (.014)	.048 (.030)	.027 (.024)	.036 (.032)
Coho Salmon Catch Rate ($R_{co,j}$)	.028 (.021)	.005 (.005)	.040 (.053)	.005 (.008)
Rainbow Trout Catch Rate ($R_{rb,j}$)	.001 (.001)	.018 (.026)	.012 (.013)	.001 (.002)
Effective Toxin Level (E_j)	.597 (.491)	2.270 (1.866)	3.464 (2.847)	2.270 (1.866)

Notes: Catch rates are measured in terms of fish per person-hour of effort.

Table 2. – Distribution of Trips

Sites Visited	Number of Observations
All four sites, ω={1,2,3,4}	3
Lake Superior, North and South Lake Michigan, ω={1,2,3}	1
Lake Superior, North Lake Michigan, and Green Bay, ω={1,2,4}	7
Lake Superior, South Lake Michigan, and Green Bay, ω={1,3,4}	0
North and South Lake Michigan and Green Bay, ω={2,3,4}	13
Lake Superior and North Lake Michigan, ω={1,2}	10
Lake Superior and South Lake Michigan, ω={1,3}	8
Lake Superior and Green Bay, ω={1,4}	2
North and South Lake Michigan, ω={2,3}	13
North Lake Michigan and Green Bay, ω={2,4}	19
South Lake Michigan and Green Bay, ω={3,4}	4
Lake Superior, ω={1}	49
North Lake Michigan, ω={2}	46
South Lake Michigan, ω={3}	85
Green Bay, ω={4}	11
No sites visited, ω=∅	243

Table 3. – Parameter Estimates

	Extreme Value		Generalized Extreme Value	
Parameter	Estimate	P-Value	Estimate	P-Value[a]
δ_0 (Intercept)	-8.53	<.001	-8.43	<.001
δ_{lk} (Lake Trout)	0.10	.953	-0.70	.667
δ_{ch} (Chinook Salmon)	13.39	<.001	11.11	<.001
δ_{co} (Coho Salmon)	3.12	.023	3.71	.007
δ_{rb} (Rainbow Trout)	8.61	.035	13.96	<.001
δ_E (Effective Toxin Level)	-0.06	.018	-0.07	.007
Ω	1.76	<.001	1.82	<.001
θ^N (North Lake Michigan and Green Bay)	1.00	Not estimated	0.57	<.001
θ^S (South Lake Michigan and Lake Superior)	1.00	Not estimated	0.92	<.001
μ	1.29	<.001	1.31	<.001
Log Likelihood	-1935.8		-1890.2	

[a] The P-values associated with θ^N and θ^S correspond, respectively, to tests of the hypotheses $H_0^N: \theta^N = 1$ and $H_0^S: \theta^S = 1$. If both of these hypotheses are imposed, then the extreme value model results.

Table 4. – Welfare Estimates (Standard Errors in Parentheses)

Policy Scenario	Mean Compensating Variation ($\overline{C}$)	
	EV Model	GEV Model
Scenario A: Loss of Lake Trout Species at Sites 2, 3 and 4	15.97 (269.19)	-37.10 (272.78)
Scenario B: Loss of Coho Salmon at Sites 2, 3 and 4	274.18 (123.18)	304.82 (192.20)
Scenario C: A 20% Reduction in Toxins at all Sites	-89.35 (54.37)	-108.13 (51.73)

VII. ENDNOTES

[1] See Bockstael, Hanemann, and Strand (1986) and Morey, *et al*. (1995) for general discussions of non-participation and corner solution problems in the context of recreation demand.

[2]Corner solutions can also emerge for producers, both due to non-negativity constraints (e.g., Lee and Pitt (1987) and to upper bounds externally imposed by quotas (e.g., Fulginiti and Perrin (1993)).

[3] See, for example, Bockstael, Hanemann, and Kling (1987), Hausman, Leonard, and McFadden (1995), Parsons and Kealy (1995), and Feather, Hellerstein, and Tomasi (1995).

[4] However, as a reviewer noted, there may be instances when a sequential or dynamic random utility model may be more suitable such as when exogenous variables change within a season.

[5] Morey, Waldman, Assane, and Shaw (1995) describe the Kuhn-Tucker model in the context of recreation demand, suggesting that it is the preferred approach, Bockstael, Hanemann, and Strand (1986) provide specifications appropriate for recreation demand, and Kling (1986) employs a form of the model to generate simulated data. However, none of these authors estimate the model or suggest how such a model could be used to compute welfare estimates.

[6] In general, a vector of attributes may characterize each commodity. However, we have used a scalar attribute here to simplify notation.

[7] Wales and Woodland (1983) accomplish this by assuming that the errors enter the utility function such that $U_j(\mathbf{x}, z; \mathbf{q}, \gamma, \varepsilon) = \overline{U}_j(\mathbf{x}, z; \mathbf{q}, \gamma) + \varepsilon_j \quad j = 1, \ldots, M.$ See Bockstael, Hanemann, and Strand (1986) and Morey *et al.* (1995) for more general treatments of the error term.

[8] This problem is similar to the one encountered in nonlinear site selection models and recently addressed by McFadden (1995) and Herriges and Kling (1997).

[9] Hanemann (1984) originally detailed this argument in the case of extreme corner solutions (i.e., when only one of the commodities is consumed). Bockstael, Hanemann, and Strand (1986) extend the argument for the general case.

[10] Imposing weak complementarity implies that there is only "use value" associated with the commodities. In the absence of weak complementarity, individuals may also assign "non-use" value to a

commodity (i.e., the individual perceives utility from the availability of a good without actually consuming it). Here, we adopt Freeman's (1993) definitions of use, non-use, and existence values and note, as an aside, that models based on observed behavior cannot elicit information on existence value.

[11] The second equality follows from the fact that for all $\omega \notin \tilde{A}$, the associated conditional utility maximization problem yields a binding non-negativity constraint for some $j \in \omega$. The solution is, therefore, redundant, being equivalent to another utility maximization problem (defined by $\tilde{\omega} \subset \omega$ with $\tilde{\omega} \in \tilde{A}$) where that good has been constrained to zero *a priori*.

[12] Policy changes may also involve the elimination of initially available sites. Such changes can be reflected in the make-up of the index collection $\tilde{A}$.

[13] See Geweke (1996) for a useful review of Monte Carlo integration.

[14] Furthermore, McFadden's (1995) Theorem 3 states that the GEV simulator can be used to construct a consistent estimator of any real-valued function that is integrable with respect to the distribution of the ε's. See Herriges and Kling (1997) for a recent application.

[15] As noted by one reviewer, a limitation of this LES system is that the resulting function is convex in the quality attributes of the sites. Further research is needed to develop more flexible functional forms that relax such restrictions while remaining manageable from an estimation perspective.

[16] One limitation of the GEV specification is that it requires the analyst to specify the nesting structure a priori. In our empirical application below, while a variety of nesting structures were investigated, we only report the results from that nesting structure that best fit the data on basis of a likelihood dominance criterion.

[17] The θ_r's are required to lie within the unit interval in order to satisfy consistency with McFadden's (1981) random utility maximization hypotheses. The degree of correlation among alternatives within nest *r* increases as θ_r diminishes towards zero, whereas the alternatives become independent if $\theta_r = 1$.

[18] Details of the log-likelihood function for the four site model presented in the empirical section below, along with the associated Jacobian transformations J_ω, while not difficult, are algebraically tedious and relegated to an Appendix available from the authors upon request.

[19] Details of the sampling procedures and survey design are provided in Lyke (1993).

[20] For papers that consider issues related to aggregation in recreation demand, see Feather (1994), Parsons and Needelman (1992), and Kaoru, Smith, and Liu (1995).

[21] There is an extensive debate on appropriate measure of the opportunity cost of travel time. Since it is not a purpose of this study to enter into this debate, we have chosen this relatively simple means of accounting for the travel time cost, drawing on research results of McConnell and Strand (1981).

[22] While there are a variety of toxins reported in the De Vault *et al.*(1989) study, we use the levels of toxins 2,3,7, 8-TCDD, which are generally responsible for the fish consumption advisories issued by states in the region.

[23] The results for the alternative nesting structures are available from the authors upon request.

[24] Since it was not known whether or not the likelihood function is globally concave, numerous starting values were tried and the maximum likelihood routine always resulted in the reported parameter estimates.

[25] The dissimilarity coefficients (i.e., the θ^r's) were constrained to lie in the unit interval in order to insure consistency with McFadden's (1981) random utility maximization hypothesis.

[26] Under this scenario, it is assumed that the catch rate for lake trout in Lake Superior is unchanged, either because of ongoing stocking programs or the natural replenishment capabilities of the fishery.

[27] The choice of $N_\varepsilon = 2000$ was selected on the basis of a Monte Carlo experiment in which the process of estimating $\overline{C}$ for scenario B using N_ε iterations and the maximum likelihood parameters was repeated 100 times. This exercise was conducted using various choices of N_ε. The simulation results indicated that the standard deviation of $\overline{C}$ was reduced to \$10 once $N_\varepsilon = 2000$. Thus, the standard

deviation associated with the GEV simulator is roughly five percent of the standard deviation reported in Table 4 (192.20).

[28] Some caution should be exercised in using the standard deviations to construct confidence intervals. The $\hat{C}^{(i)}$'s are unlikely to be symmetrically distributed and, hence, two-standard deviation confidence intervals will be inappropriate. While the construction of asymmetric confidence intervals is conceptually straightforward, a substantially larger N_γ would be needed to precisely construct the necessary tail statistics (See, e.g., Efron and Tibshirani (1993)).

[29] Since our estimates of Ω are significantly different from 1.00, consumers value quality changes even when they do not engage in fishing. Thus, these welfare estimates can be said to encompass both "use" and "non-use" values. However, these values are quite distinct from "existence" value that cannot be estimated from revealed preference data.