Randomly Drawn Opportunity Sets in a Random Utility Model of Lake Recreation

George R. Parsons and Mary Jo Kealy

ABSTRACT. *Random Utility Models are widely applied in studies of recreation demand. The model is particularly useful when the number of recreation sites from which individuals may choose is large. Yet, when the number gets too large, say in the hundreds, estimation becomes burdensome. We present an analysis suggested by McFadden (1978) for dealing with large numbers of sites. We estimate a model using randomly drawn opportunity sets. We use each person's chosen site plus a random draw of as few as eleven other sites (when hundreds are available) to estimate a plausible behavioral model.*

I. INTRODUCTION

The application of Random Utility Models to recreation decisions has increased considerably in the past decade. Most of these applications have been for the purpose of measuring the benefits of improvement in water quality or fish catch.[1] The model appears to have emerged as the preferred model of recreation decisions when the number of alternatives (recreation sites) facing an individual is large and site substitution is important. Yet, when the number of alternatives gets too large (thousands of possible lakes for boating or hundreds of possible access points on the ocean for fishing) estimation becomes burdensome, both in computation and data set-up. In this paper, we present an approach suggested by McFadden (1978) for estimation when there are many sites. An individual's opportunity set of sites is represented by a random draw from his or her full set of sites, say 12 sites drawn from a set of 300. Estimation is then done with the set of randomly drawn sites. Under plausible assumptions the estimated model is unbiased.

We offer the random draw approach as an alternative to site aggregation for dealing with large opportunity sets. Site aggregation involves defining sites as a region or a county where an averaging of characteristics of sites within the region or county serve as the site characteristics. Aggregation reduces the number of alternatives in the choice set and simplifies estimation. Unfortunately, ease of estimation comes at the expense of aggregation bias.

Our application is to lake recreation in the state of Wisconsin. We estimate benefits for water quality improvements for four independent user groups: boaters, anglers, swimmers, and viewers. Our water quality measures are level of dissolved oxygen and clarity of water. These measures are good indicators of quality as perceived by individuals and are measures that policymakers often control in regulations. The data we analyze were previously studied by Caulkins (1982) and Caulkins, Bishop, and

---

College of Marine Studies and Department of Economics, University of Delaware; and U.S. Environmental Protection Agency, Washington, DC and Department of Economics, Colgate University, respectively.

This research was funded through EPA Cooperative Agreement CR-815427-01-0. Mike Needelman and Steve Hess provided valuable research assistance. Pete Caulkins provided the data and insight throughout the project. An earlier version of this paper (see Parsons and Kealy 1990) was rewritten based on the comments of Yoshi Kaoru and two anonymous referees. We also benefitted from conversations with Nancy Bockstael, Jeff Englin, Reed Johnson, Ted McConnell, Greg Michaels, Bill O'Neil, V. Kerry Smith, and Ivar Strand.

[1]Among the studies are non-nested models of beach use in Boston by Hanemann (1978), Binkley and Hanemann (1978), and Feenberg and Mills (1980); lake visits in Wisconsin by Caulkins (1982) and Caulkins, Bishop, and Bouwes (1986); recreational fishing on the Albemarle-Pamlico Estuary in North Carolina by Kaoru and Smith (1990); and ocean fishing off the Oregon coast by Morey, Shaw, and Rowe (1991). Studies using nested models include beach use in Boston by Bockstael, Hanemann, and Kling (1987); beach use on the Chesapeake Bay by Bockstael, McConnell, and Strand (1988); ocean fishing by the same authors (1989); sport fishing in Alaska by Hanemann, Carson, Gum, and Mitchell (1987); and ocean fishing off the Florida coast by Milon (1988).

Bouwes (1986). Their analysis was, however, limited to a small subset of the data—45 of 1,200 observations. We use all observations in our analysis.

We begin with a description of the Random Utility Model (RUM) and follow with a section on the use of random draws to represent the opportunity set. Last, we present an example of the welfare effects of water quality changes using our estimated model. Our analysis pertains only to day trips.

II. THE RANDOM UTILITY MODEL

An individual takes total number of day trips to lakes during the year as predetermined and decides which lake to visit on each trip. He or she has utility for a trip to lake (ai) of $U_{ai} = V_a + V_{ai} + \varepsilon_{ai} + \varepsilon_a$. V_a is a systematic component of utility common to all lakes in area a of Wisconsin ($a = 1$ if the lake is located in the north and $a = 0$ if located in the south). V_{ai} is a systematic component for lake i in area a ($i = 1, \ldots, N$ if in the north and $i = 1, \ldots, S$ if in the south). The term $\varepsilon_{ai} + \varepsilon_a$ is a random element capturing excluded characteristics of lakes. The part ε_a specifically incorporates excluded characteristics common to all lakes in area a. We define $V_{ai} = V(x_{ai}, p_{ai})$ where x_{ai} is a vector of lake characteristics such as size, presence of commercial facilities nearby, and water quality of the lake, and p_{ai} is the price of visiting the lake which includes opportunity cost of time and travel cost. We assume a linear utility function with $V_{ai} = \beta z_{ai}$ where $z_{ai} = (x_{ai}, p_{ai})$.

Northern Wisconsin lakes share a natural setting and nearness to outdoor amenities that distinguish them from southern Wisconsin lakes. These distinguishing characteristics are not measured in our study. For this reason, we divide the state by north and south and define $V_a = \alpha' d_a$, where $d_a = 1$ for a lake in the north and $d_a = 0$ for a lake in the south. V_a captures an "average" contribution to utility for any trip taken to a northern relative to southern lake. The remaining characteristics common to each area which we do not control for are captured in ε_a. Random utility for a visit to lake (ai) on a given choice occasion then is

$$U_{ai} = \alpha' d_a + \beta z_{ai} + \varepsilon_{ai} + \varepsilon_a. \qquad [1]$$

We estimate the parameters in equation [1] using a nested model. First, an individual decides whether to visit a lake in the north or south. Second, he or she decides which lake in that region to visit. We assume the ε_{ai} are independent and identically distributed Gumbel random variables with scale parameter $\mu' = 1$. It follows that an individual's probability of visiting lake i' given that he or she makes a trip to the north or south is the simple logit

$$Pr(i' | a = 1) = \exp(\beta z_{1i'}) \Big/ \left\{ \sum_{\{1i \in N\}} \exp(\beta z_{1i}) \right\} \qquad [2]$$

$$Pr(i' | a = 0) = \exp(\beta z_{0i'}) \Big/ \left\{ \sum_{\{0i \in S\}} \exp(\beta z_{0i}) \right\}. \qquad [2']$$

N is the set of lakes in his or her opportunity set in the north, S is the set in the south. It follows that $\mathrm{MAX}(\beta z_{11} + \varepsilon_{11}, \ldots, \beta z_{1N} + \varepsilon_{1N})$ and $\mathrm{MAX}(\beta z_{01} + \varepsilon_{01}, \ldots, \beta z_{0S} + \varepsilon_{0S})$ are Gumbel random variables with

$$E[\mathrm{MAX}(\beta z_{11} + \varepsilon_{11}, \ldots, \beta z_{1N} + \varepsilon_{1N})] = I_1 = \ln\left[\sum_{\{1i \in N\}} \exp(\beta z_{1i}) \right] + 0.577$$

$$E[\mathrm{MAX}(\beta z_{01} + \varepsilon_{01}, \ldots, \beta z_{0S} + \varepsilon_{0S})] = I_0 = \ln\left[\sum_{\{0i \in S\}} \exp(\beta z_{0i}) \right] + 0.577.$$

We define $\tilde{\varepsilon}_1 = \mathrm{MAX}(\beta z_{11} + \varepsilon_{11}, \ldots, \beta z_{1N} + \varepsilon_{1N}) - I_1$ and $\tilde{\varepsilon}_0 = \mathrm{MAX}(\beta z_{01} + \varepsilon_{01}, \ldots, \beta z_{0S} + \varepsilon_{0S}) - I_0$. Then, assuming $\tilde{\varepsilon}_1 + \varepsilon_1$ and $\tilde{\varepsilon}_0 + \varepsilon_0$ are iid Gumbel random variables with scale parameter μ, we have the probability of choosing a northern or southern lake of

$$Pr(a = 1) = \exp(\alpha + \mu I_1)/\{\exp(\alpha + \mu I_1) + \exp(\mu I_0)\} \quad [3]$$
$$Pr(a = 0) = \exp(\mu I_0)/\{\exp(\alpha + \mu I_1) + \exp(\mu I_0)\}. \quad [3']$$

Since $\mu' = 1$ in the site choice stage, we have $\mu/\mu' = \mu$ in the area choice stage. Also, note that $\alpha = \alpha'\mu$ in these models.

We assume that the error terms in the area and lake choice stages are independent across an individual's trips as well as across different individuals' trips during the year. If an individual visits site (ai) a total of T_{ai} times during the year, the probability of that occurrence is $(Pr(ai))^{T_{ai}}$. If the individual visits more than one site, the probability of that pattern of visits is $\Pi_a \Pi_i (Pr(ai))^{T_{ai}}$. The likelihood function for the nested logit then is

$$L = \Pi_n \Pi_a \Pi_i ((Pr_n(i|a) \cdot Pr_n(a))^{T_{ain}} \quad [4]$$

where $Pr(i|a)$ is from equation [2] or [2′] and $Pr(a)$ is from equation [3] or [3′]. The n denotes the nth individual. The parameters α and β of the RUM are estimated by selecting those values that maximize this likelihood. This can be done by maximizing [4] simultaneously or, with less efficiency, sequentially by first maximizing the product of the conditional probabilities ($Pr(i|a)$ in the likelihood function) and then the product of the marginal probabilities ($Pr(a)$ in the likelihood function). For ease of estimation we use the sequential approach.

The *ex ante* equivalent variation (and compensating variation) per choice occasion for an improvement in water quality in our linear nested model is

$$\Delta w = [1/(\mu\beta_y)] \cdot [\ln\{\exp(\alpha + \mu\tilde{I}_1) + \exp(\mu\tilde{I}_0)\} - \ln\{\exp(\alpha + \mu I_1) + \exp(\mu I_0)\}] \quad [5]$$

where $\tilde{I}_0$ and $\tilde{I}_1$ refer to values of I_0 and I_1 with the improvements in water quality, and β_y is the marginal utility of income, the coefficient on p_{ai} in the RUM. (See Small and Rosen 1981 and Hanemann 1982 for derivations of compensating and equivalent variation in the RUM.)

III. RANDOM DRAWS TO REPRESENT OPPORTUNITY SETS

Because Wisconsin has thousands of lakes, each individual's opportunity set is potentially large. This makes estimation of the nested logit model complicated. Maximizing a likelihood function with thousands of alternatives per person is burdensome, both for computation and data set-up. Our approach is to estimate the model allowing all sites to enter individually into each person's opportunity set but to estimate the model using a subset randomly drawn from the full set.

Our draw works as follows. If an individual visited a northern lake, 23 lakes are randomly drawn from the set of all lakes in the north that are within a day's drive (180 miles) of that person's home. We add to that set the lake actually visited. (The lake actually visited is not included in the set from which we make our draw.) These 24 lakes form the individual's opportunity set of northern lakes for estimation. If the individual visited a southern lake, we follow the same procedure except that we draw 23 lakes from the south within 180 miles of the person's home and add the actually visited lake. A separate draw is done for each person. For purposes of comparison, we estimate the model using randomly drawn opportunity sets of 3, 6, 12, and 24 lakes.

McFadden has shown that estimating a model using random draws can give unbiased estimates of the model with the full set of alternatives. His approach exploits the maintained hypothesis of independence of irrelevant alternatives within any given nest in a model. Take an individual's probability of visiting lake i given that he or she makes a trip to the north. In the context of that model, McFadden has shown that his or her probability of visiting site (i') given a subset of alternatives D, where D is drawn from his or her full opportunity set of northern lakes within 180 miles, is

$$Pr(i'|D) = [Pr(D|i') \cdot \exp(\beta z_{1i'})] / \left[\sum_{\{1i \in D\}} Pr(D|i) \cdot \exp(\beta z_{1i}) \right]. \quad [6]$$

Each of the lakes in the subset of lakes drawn for the individual is weighted by $Pr(D|i)$ in the simple logit. $Pr(D|i)$ is the probability of drawing the subset D had lake i actually been chosen.

In our simple random draw of 23 lakes without replacement, we have $Pr(D|i) = \binom{N-1}{23}^{-1}$ for all i, where $\binom{N-1}{23}$ is the number of combinations of $N - 1$ lakes (all northern lakes less the chosen lake) taken 23 at a time. In the case of a simple random draw $Pr(D|i') = Pr(D|i)$ for all i. It follows that the $Pr(D|i)$ cancels out of equation [6] and that we can estimate the original model using the drawn subset of lakes for each individual.

McFadden defined the uniform conditioning property as sampling whereby $Pr(D|i') = Pr(D|i)$ for all i. He also defined the positive conditioning property as sampling with $Pr(D|i) > 0$ for all i. Equation [6] could still be estimated with sampling strategies that satisfy the positive conditioning property, but it requires that the probabilities be known and incorporated in the likelihood function. Since we use random sampling, our strategy satisfies the uniform conditioning property and weighting is unnecessary.

After estimating the site choice stage of our nested model using a randomly drawn subset of alternatives, we estimate the area choice stage of the model. In that stage we estimate the inclusive values (I's in equations [3] and [3']) using the parameter estimates from the site choice stage and using all the lakes in each person's opportunity set. The inclusive value is not constructed from the subset of alternatives. For a given individual, his or her inclusive value for the north is $\ln \Sigma_{\{1i \in N\}} \exp(\beta z_{1i})$. β is estimated from the site choice model of random draws in equation [6] and N is the set of all northern lakes within 180 miles of the person's home. The same is done for the inclusive value for the southern lakes.

Implicit in our random draws is an assumption that every individual is aware of all lakes within a day's drive from his or her home. For most Wisconsin residents this is hundreds of lakes. Undoubtedly, most residents will know of only a fraction of this full set. This raises the possibility of bias due to erroneously including alternatives that individuals may not consider in their site choice. In recreation studies we are usually concerned about erroneously excluding alternatives that are believed to be important substitutes. Here our problem is not limited observation but excessive observation—too many substitutes.

Suppose each person in our sample lives next to the largest and cleanest lake in the state but that all are unaware it exists (certainly an unlikely occurrence). If we include this "prize" lake in each person's opportunity set, we observe erroneously each person choosing a smaller, dirty, and more distant lake over the prize lake. In fact, each person knows nothing about the prize lake and, hence, has not revealed such preference. In estimation this has the effect of erroneously reducing the measured importance of travel cost, size, and cleanliness in each person's choice. The model of behavior is biased.[2]

If an erroneously included alternative has a high probability of selection when known to the individual, the bias of inclusion may be large. (Our prize lake example fits this scenario.) In such cases we dramatically misrepresent behavior—believing the prize lake is undesirable when it is not. As the probability of selecting the erroneously included alternative falls, the bias of its inclusion falls. We do not misrepresent be-

[2] It is tempting to appeal to the property of independence from irrelevant alternatives (iia) here and argue that the inclusion or exclusion of the prize lake is irrelevant for the parameter estimates. But, this is a misinterpretation of the iia property which permits consistent estimates with a subset of all known alternatives. It does not permit consistent estimation with an expanded set of alternatives which includes alternatives unknown to an individual.

havior dramatically by including alternatives that would never be chosen.

The degree of bias could be measured if we knew the actual lakes known to each individual in our sample. Comparing parameter estimates using the known alternatives with parameter estimates using the full set of alternatives is an appropriate method of investigation. Unfortunately, the set of known alternatives would be unknown to us in the Wisconsin sample. Short of having this information, we believe the random draw procedure is best. Unknown lakes with a high probability of selection (if known about) are likely to be uncommon. We expect most of the bias to come from unknown lakes with low probabilities of selection, and this type of inclusion presents less bias.

We believe that the random draw procedure represents behavior quite well in our context. Most people know about lakes nearby, lakes of considerable size, and perhaps lakes with dramatic qualitative features. People also know that there are many other smaller lakes scattered throughout the state and in considerable density in the northern counties. These latter sets of lakes are not known by name but rather in a collective sense. On a given choice occasion a person knows there are many possible lakes that could be chosen but may only know a few by name. Representing a set of lakes that have a collective identity by random draw is appealing, it captures the breadth of availability understood by the individual without having to identify specific lakes which the individual could not identify anyway.

IV. THE DATA SET

Our analysis is based on a 1978 random phone survey of Wisconsin residents, the "Statewide Water Quality Survey," and two supplementary data sources, one on lakes and their characteristics and the other on travel distances and times. The survey was conducted by the Wisconsin Survey Research Laboratory and funded by the U.S. Environmental Protection Agency. The survey was done in the fall and questioned people about their use of 1,133 Wisconsin lakes during the preceding 12 months. Lakes less than 100 acres, Lake Michigan, and Lake Superior were excluded.

Nearly 1,200 individuals 18 years old and older were interviewed. All were asked a list of questions pertaining to their socioeconomic status—age, income, years living in Wisconsin, ownership of property on a lake, hometown, and so on. In addition, each person was asked to identify his or her primary use of Wisconsin lakes. The primary use categories are: boating, fishing, swimming, viewing (including picnicking and hiking), or no use.

Approximately 60 percent of the people surveyed made at least one visit to a Wisconsin lake during the year. These persons were asked to identify and estimate the number of trips taken to each of their visited lakes. No person was questioned about more than six lakes. For each individual one of the visited lakes was randomly drawn for detailed questioning. For that lake the interviewer asked about a typical trip—number of people on the trip, expenses, length of trip, distance traveled to the site, and if they owned property at the site.

The lake characteristic data set is from the Water Resources Center at the University of Wisconsin (see Uttormark and Wall 1975) and the Wisconsin Department of Natural Resources. It includes information on acreage, depth, water quality, and measures of access such as presence of boat ramps. From the set of water quality variables we use the measures of dissolved oxygen and turbidity for reasons explained in the next section.

Finally, we constructed a matrix of road distances and travel times between each interviewed person's hometown and the set of 1,133 lakes using the software, HYWAYS/BYWAYS. This software computes road distances and travel times between more than 500 towns in Wisconsin. The travel-time measure in HYWAYS/BYWAYS accounts for different average

TABLE 1
VARIABLE DEFINITIONS FOR THE NESTED LOGIT MODEL

Variable Name	Description
*PRICE**	Opportunity Cost of Time Plus Transit Costs: [(1/3) * (Annual income/2080) * (Travel Time)] + [.10 * 2 * (Distance to Lake)]
LNACRES	Log of acreage of the lake
CF	= 1 if commercial facilities are present; = 0 if not (Commercial facilities include restaurants, bait shops, hotel, boat services, and so on)
REMOTE	= 1 if lake is in a remote location; = 0 if not (Remote means that the lake can only be reached by navigable water or is located in a public wilderness area without a road or defined trail within 200 feet)
NORTH	= 1 if lake is located in a northern county; = 0 if not (Northern counties include: Douglas, Bayfield, Ashland, Vilas, Forest, Florence, Burnett, Washington, Sawyer, Price, Oneida, Marinetta, Polk, Barren, Rusk, Lincoln, Langlade, Iron, and Oconto)
LNMXD	Log of maximum depth of the lake
BR	= 1 if a boat ramp is present at lake; = 0 if not
INLET	= 1 if lake has an inlet; = 0 if not
DONO	= 1 if the entire hypolimnion is void of oxygen at times; = 0 if not
DOYES	= 1 if dissolved oxygen in hypolimnion is greater than 5 ppm virtually all the time; = 0 otherwise
CLEAR	= 1 if average secchi depth reading is at least 3 meters; = 0 if not

*Annual income was missing for approximately 15 percent of the sample. For these individuals we predicted income using a linear regression. The following regression was estimated using the portion of the sample that report their income:

$$y = \underset{(7)}{-138} + \underset{(9)}{7.6} * AGE - \underset{(9)}{.08} * AGE^2 + \underset{(8)}{8.3} * EDUCATION$$

$$+ \underset{(3)}{110} * d1 + \underset{(2)}{52} * d2 + \underset{(3)}{126} * d4 + \underset{(10)}{54} * MARRIED$$

where y = (annual income) * 100; *AGE* is age of interviewee; *EDUCATION* is years of schooling, $d1$ = 1 if individual is a lawyer and 0 if not, $d2$ = 1 if individual is an engineer and 0 if not, and $d3$ = 1 if individual is a physician and 0 if not; *MARRIED* is 1 if individual is married (t-statistics are in parentheses below coefficients). Annual income was predicted using this equation for those not reporting.

speeds over different routes—travel on interstates is faster than travel on county roads, open road travel faster than city traffic, and so on. Travel time is not converted from a distance measure as is common in recreation demand studies. The location of each lake was assumed to be the nearest town recognized by the software. (This undoubtedly has introduced some measurement error.) In this way the price of a trip could be constructed for all lakes for each person. Using the travel price information and lake characteristic data, we were able to conduct our random draw experiment.

V. ESTIMATION

The definition of the variables used as arguments in the utility function in our model are presented in Table 1. *PRICE* is the sum of transit costs and opportunity cost of time. We assume the transit cost per mile is \$.10 (1978 dollars). We assume an individual's trip time is just the travel time to the site. (Since on-site time is assumed to be constant across sites it can be ignored in estimation.) Each individual is assumed to value an hour at (1/3) · (annual income)/(2080).

We have two measures of size of the site: acres (*LNACRES*) and depth (*LNMXD*). We take the logarithm of both because we expect a diminishing effect—the difference between lakes that are 101,000 and 104,000 acres is not the same as the difference between lakes that are 1,000 and 4,000 acres. The depth variable is expected to matter only to those persons using lakes primarily for fishing or boating.

We have three measures of access in the models: *BR*, *INLET*, and *REMOTE*. *BR* is a measure of the presence of boat ramps at the lake, *INLET* is a measure of the presence of an inlet on the lake, and *REMOTE* is a dummy variable identifying those lakes that may be reached only by a stream flowing into or out of the lake or that are in a public wilderness area without an adjacent road or defined trail. This last variable cannot be signed a priori. Less accessible sites are more costly to visit and yet remote sites may offer other highly desirable features such as solitude or an undisturbed natural setting. We expect *BR* and *INLET* to matter only to the boating and fishing population.

CF is a measure of the commercial facilities available near the lake—stores, bait shops, restaurants, hotels, and so forth. *NORTH* is a dummy variable denoting whether the lake is in a northern county. The lakes in the northern counties are generally different than those in the south. Most are in a wilderness setting and many are in a state or national forest with the protection of natural amenities that such a designation provides. Our dummy variable is intended to capture that distinction.

We use two indicators of water quality: dissolved oxygen and water clarity. We chose dissolved oxygen because it is a good overall indicator of water quality (Welch 1989) and is related to pollutants that are regulated by the Environmental Protection Agency (EPA), e.g., biological oxygen demand (BOD). We consider clarity only for swimming and viewing uses of lakes. It is not apparent that clarity affects boating or fishing decisions so we did not consider it in these cases. Clarity is related to total suspended solids which are also regulated by EPA.

We have two measures of dissolved oxygen (DO): *DOYES* and *DONO*. Both are dummy variables. For any given lake *DOYES* = 1 if the dissolved oxygen in the lake's hypolimnion is greater than 5ppm at virtually all times. This DO level is adequate to support the most desirable game fish in the state and to assure the water is not anaerobic. *DONO* = 1 if the entire hypolimnion is devoid of oxygen at critical times during the year. This implies water that cannot support aquatic life and is anaerobic. These waters are inclined to excessive weed and plant growth and unpleasant odor. In terms of DO then, every lake falls into one of three groups: high dissolved oxygen (*DOYES* = 1 and *DONO* = 0), moderate dissolved oxygen (*DOYES* = 0 and *DONO* = 0), and low dissolved oxygen (*DOYES* = 0 and *DONO* = 1). Approximately 5 percent of Wisconsin lakes fall in the high DO grouping and 25 percent in the low DO grouping.

For clarity, we use one measure, also a dummy variable, *CLEAR*. For a given lake *CLEAR* = 1 if the average secchi depth reading during 1978 was at least 3 meters. About 5 percent of Wisconsin's lakes have *CLEAR* = 1.

We estimated four separate models by individuals' primary use of the lakes: boating, fishing, swimming, and viewing.[3] All models are for day trips. (See the Appendix for our definition of a day trip. Because some trips are reported without trip length, we predicted length for a portion of our sample.) The boating regression was estimated using all trips by individuals who identified boating as their primary use of Wisconsin lakes, the swimming regression using all trips by individuals who identified swimming as their primary use, and so on. For this reason the regressions are not specific to type of recreation trip. A boater, after all, may take some trips for fishing, swimming, or viewing. For each recreation group we estimated the model using randomly drawn opportunity sets of 3, 6, 12, and 24. The results are in Tables 2 through 5.

We find that people are more likely to visit cheaper, larger, and cleaner lakes that

[3]We considered a nested structure in which individuals first choose to boat, swim, fish, view, or not use lakes, then which site to visit given their chosen use. The results in the use stage of the model were poor. We used individual characteristics (age, sex, income, and so on) to explain type of recreation use but gained little understanding of behavior. This may be because individuals did not declare their intended use on each trip, rather they declared their primary use of lakes during the entire year. Because the model provided little behavioral insight, we have not used it in our analysis.

TABLE 2
MAXIMUM LIKELIHOOD ESTIMATES OF THE NESTED LOGIT MODEL FOR PERSONS USING LAKES PRIMARILY FOR *BOATING*

	Number of Lakes in Randomly Drawn Opportunity Set				
Variable	3	6	12	24	24(Visited)
PRICE	−.47 (11.9)	−.19 (15.6)	−.27 (19.0)	−.30 (24.4)	−.28 (24.7)
LNACRES	2.5 (11.5)	.85 (4.3)	.08 (2.1)	.05 (1.7)	.21 (9.4)
CF	4.6 (4.2)	3.5 (9.2)	3.3 (10.7)	3.7 (10.8)	2.6 (9.0)
REMOTE	10.1 (9.3)	−2.5 (3.6)	−1.5 (2.9)	−2.5 (4.9)	−1.0 (1.8)
LNMXD	−.83 (6.0)	1.5 (11.2)	.25 (4.0)	.33 (5.7)	−.05 (0.9)
BR	4.7 (9.8)	−1.0 (8.4)	.52 (5.1)	−.18 (1.9)	.05 (0.5)
INLET	7.1 (10.8)	3.5 (9.0)	3.4 (10.6)	3.6 (10.0)	3.3 (9.7)
DONO	−3.4 (9.3)	−.75 (2.2)	−.34 (2.2)	−.54 (3.8)	−.20 (1.9)
DOYES	6.2 (10.1)	3.6 (8.5)	.41 (2.6)	1.2 (8.8)	.28 (2.1)
CLEAR	—	—	—	—	—
INC VALUE (μ)	.10 (5.6)	.19 (5.5)	.17 (5.6)	.16 (5.6)	.18 (5.4)
NORTH	.94 (2.3)	.28 (1.0)	.34 (1.1)	.28 (1.0)	.51 (1.6)
Number of Visits	1,334	1,334	1,334	1,334	1,334
Number of Individuals	71	71	71	71	71
First Stage:					
Log-Likelihood	−218	−574	−1,135	−1,752	−2,154
Slopes = 0 Log-L	−1,466	−2,392	−3,317	−4,242	−4,242
Second Stage:					
Log-Likelihood	−74	−76	−75	−76	−71
Slopes = 0 Log-L	−118	−118	−118	−118	−107

Notes: t-statistic for test that coefficient equals zero is given in parentheses beside coefficient estimate. Standard errors on *INCLUSIVE VALUE* and *NORTH* are uncorrected but given the size of our sample are likely to differ little from the corrected errors.

TABLE 3
MAXIMUM LIKELIHOOD ESTIMATES OF THE NESTED LOGIT MODEL FOR PERSONS USING LAKES PRIMARILY FOR *FISHING*

	Number of Lakes in Randomly Drawn Opportunity Set				
Variable	3	6	12	24	24(Visited)
PRICE	−.20 (24.8)	−.27 (33.8)	−.25 (45.5)	−.23 (52.4)	−.24 (55.1)
LNACRES	.69 (15.3)	.61 (21.1)	.55 (24.0)	.56 (30.0)	.38 (22.6)
CF	−.22 (1.7)	.19 (1.9)	.18 (2.4)	.31 (4.5)	.20 (2.9)
REMOTE	.34 (1.6)	−.47 (3.4)	−.33 (2.7)	−.48 (4.3)	−.11 (1.1)
LNMXD	.45 (5.6)	.54 (8.8)	.44 (9.9)	.38 (10.5)	.26 (7.0)
BR	−.43 (3.0)	−.61 (6.9)	−.46 (6.1)	−.28 (4.3)	−.22 (3.5)
INLET	.86 (5.2)	.90 (6.3)	.93 (8.7)	.63 (7.0)	.64 (7.0)
DONO	−.85 (4.7)	−.88 (6.7)	−.84 (8.2)	−.79 (8.8)	−.82 (10.0)
DOYES	−.15 (0.8)	−1.02 (6.5)	.30 (2.5)	.11 (1.0)	.33 (4.0)
CLEAR	—	—	—	—	—
INC VALUE (μ)	.20 (9.9)	.16 (9.8)	.17 (9.9)	.18 (9.9)	.18 (9.6)
NORTH	.54 (3.5)	.55 (3.6)	.52 (3.5)	.53 (3.5)	.59 (3.8)
Number of Visits	3,598	3,598	3,598	3,598	3,598
Number of Individuals	239	239	239	239	239
First Stage:					
Log-Likelihood	−790	−1,599	−2,670	−4,154	−5,162
Slopes = 0 Log-L	−3,954	−6,448	−8,943	−11,437	−11,437
Second Stage:					
Log-Likelihood	−338	−338	−338	−338	−313
Slopes = 0 Log-L	−438	−438	−438	−438	−401

Notes: t-statistic for test that coefficient equals zero is given in parentheses beside coefficient estimate. Standard errors on *INCLUSIVE VALUE* and *NORTH* are uncorrected but given the size of our sample are likely to differ little from the corrected errors.

TABLE 4

MAXIMUM LIKELIHOOD ESTIMATES OF THE NESTED LOGIT MODEL FOR PERSONS USING LAKES PRIMARILY FOR *SWIMMING*

	Number of Lakes in Randomly Drawn Opportunity Set				
Variable	3	6	12	24	24(Visited)
PRICE	−.30 (19.9)	−.26 (27.5)	−.25 (36.7)	−.25 (42.8)	−.23 (46.9)
LNACRES	.85 (13.9)	.72 (17.8)	.62 (20.3)	.62 (25.5)	.31 (14.7)
CF	2.1 (7.7)	1.1 (7.4)	.87 (8.2)	1.0 (10.6)	.45 (4.6)
REMOTE	−.25 (0.4)	.11 (0.5)	.10 (0.6)	−.34 (2.2)	−.54 (3.7)
LNMXD	—	—	—	—	—
BR	—	—	—	—	—
INLET	—	—	—	—	—
DONO	−1.6 (6.3)	−2.1 (11.8)	−2.7 (16.7)	−1.7 (12.2)	−2.0 (15.7)
DOYES	−1.2 (4.9)	.60 (2.1)	−.12 (0.6)	1.0 (5.0)	.63 (4.0)
CLEAR	2.7 (9.0)	2.1 (11.6)	1.1 (7.0)	.31 (3.0)	.37 (4.2)
INC VALUE (μ)	.16 (6.8)	.18 (6.9)	.19 (6.9)	.19 (7.0)	.21 (6.9)
NORTH	.63 (2.5)	.61 (2.5)	.65 (2.6)	.69 (2.7)	.68 (2.7)
Number of Visits	2,296	2,296	2,296	2,296	2,296
Number of Individuals	126	126	126	126	126
First Stage:					
Log-Likelihood	−342	−1,030	−1,772	−2,664	−3,664
Slopes = 0 Log-L	−2,524	−4,116	−5,696	−7,300	−7,300
Second Stage:					
Log-Likelihood	−128	−128	−127	−127	−120
Slopes = 0 Log-L	−176	−176	−176	−176	−168

Notes: t-statistic for test that coefficient equals zero is given in parentheses beside coefficient estimate. Standard errors on *INCLUSIVE VALUE* and *NORTH* are uncorrected but given the size of our sample are likely to differ little from the corrected errors.

TABLE 5

MAXIMUM LIKELIHOOD ESTIMATES OF THE NESTED LOGIT MODEL FOR PERSONS USING LAKES PRIMARILY FOR *VIEWING*

	Number of Lakes in Randomly Drawn Opportunity Set				
Variable	3	6	12	24	24(Visited)
PRICE	−.38 (21.0)	−.31 (30.0)	−.31 (36.7)	−.34 (45.3)	−.35 (47.7)
LNACRES	.74 (15.6)	.80 (20.1)	.67 (23.7)	.71 (30.5)	.49 (23.6)
CF	.20 (1.2)	−.47 (4.0)	.21 (2.4)	.28 (3.6)	.06 (0.9)
REMOTE	−4.4 (11.4)	−1.3 (7.3)	−.73 (4.7)	−.82 (5.4)	−.84 (6.1)
LNMXD	—	—	—	—	—
BR	—	—	—	—	—
INLET	—	—	—	—	—
DONO	1.4 (8.5)	.13 (1.3)	−.36 (3.8)	−.27 (3.6)	−.52 (7.4)
DOYES	−.39 (.82)	.67 (1.7)	.51 (2.7)	.24 (1.3)	−.27 (1.6)
CLEAR	2.1 (10.9)	2.6 (14.8)	1.1 (11.3)	1.8 (19.5)	1.2 (15.4)
INC VALUE (μ)	.16 (8.0)	.19 (8.1)	.20 (8.1)	.18 (8.1)	.17 (8.0)
NORTH	.22 (0.9)	.17 (0.7)	.21 (0.8)	.20 (0.8)	.23 (0.9)
Number of Visits	2,985	2,985	2,985	2,985	2,985
Number of Individuals	187	187	187	187	187
First Stage:					
Log-Likelihood	−563	−1,137	−1,977	−2,928	−3,619
Slopes = 0 Log-L	−3,280	−5,350	−7,420	−9,490	−9,490
Second Stage:					
Log-Likelihood	−154	−154	−154	−154	−150
Slopes = 0 Log-L	−293	−293	−293	−293	−279

Notes: t-statistic for test that coefficient equals zero is given in parentheses beside coefficient estimate. Standard errors on *INCLUSIVE VALUE* and *NORTH* are uncorrected but given the size of our sample are likely to differ little from the corrected errors.

have commercial facilities nearby, have relatively easy access, and are in the north. This basic story holds for most of the regressions and, with some exceptions, is supported by coefficient estimates that are significantly different than zero. Moreover, all models have large likelihood ratios—a test of our specification against the restricted specification that assumes equal probability of visiting any site indicated our specification was preferred. In each regression we also find that the coefficient on the inclusive value from the site choice stage is less than and significantly different from one at a 95 percent level. This is grounds for rejecting the simple logit in favor of our nested version.

We also find that as the number of randomly drawn alternatives in each person's opportunity set increases from 3 to 24, our confidence in the parameter estimates increases—standard errors on the coefficients tend to drop and expected signs emerge. Yet, for many of the coefficients we have a considerable degree of confidence with only 3 lakes in the opportunity set. The *PRICE* and *LNACRES* coefficients are particularly stable and significant across all regressions. So are many other coefficients. For example, *CF* and *INLET* in the boating regression or *LNMXD* and *INLET* in the fishing regression change little as the number of alternatives increase.

With respect to the water quality variables, the *DONO* and *CLEAR* coefficients are reasonably stable across all regressions—we expect a negative sign on *DONO* and a positive sign on *CLEAR*. The *DOYES* coefficient, on the other hand, is less stable and at times has the wrong sign—we expect a positive sign on *DOYES*. The size of the parameters tends to increase modestly in some of the regressions as the number of draws drops. This may be evidence of correlation among the random utilities and hence of specification error (see Ben-Akiva and Lerman 1985, 189). However, we must keep in mind that these are estimates of parameters on a utility function which is of significance only in an ordinal sense. If the predicted preference ordering of sites is unchanged as the number of alternatives used in estimation declines, the model is stable even though variation in the parameter estimates is observed.

Finally, we estimated an alternative regression. We formed an opportunity set of 24 lakes drawn from a full set that included only lakes visited by at least one person in the sample. The set of visited lakes include 367 of the 1,133 lakes. We then estimated our four basic models using this new draw. The experiment is relevant because researchers at times are limited to using opportunity sets of alternatives reported by respondents in a survey. These results are given in Tables 2 through 5 for each recreation group under the column 24(visited). We find the parameter estimates are quite robust with respect to this exercise.[4,5]

VI. WELFARE ANALYSIS

We consider welfare changes for two scenarios:

1. Improving water quality at all lakes to a *low standard:* no lakes have periods where its hypolimnion is void of oxygen (set $DONO = 0$ for all lakes).
2. Improving water quality at all lakes to a *high standard:* dissolved oxygen in the hypolimnion is maintained at 5

---

[4]For sake of comparison we also estimated a model similar to that of Caulkins et al. (1986) in the original analysis of these data. For each person the relevant opportunity set was defined as the set of lakes that he or she actually visited. Only persons visiting two or more lakes were included. Since each lake in a person's opportunity set appears as a success (the chosen lake) at least once, the force of this model comes through the frequency with which a person visits the different lakes in his or her set. More frequent visits to large, near, and clean lakes, for example, would give the expected signs on the *PRICE*, *LNACRES*, and water quality variables. We find that this model works well for the price and size variables but beyond that is quite erratic and even poor for the water quality variables. The results are not presented here but are available upon request.

[5]While we have not formally tested for independence from irrelevant alternatives here, the rather stable parameter estimates across the model drawing from 1,133 possible lakes and the model drawing from 367 lakes lends some support to that maintained hypothesis within the nests.

ppm at all times for all lakes (set *DOYES* = 1 and *DONO* = 0 for all lakes) and clarity is maintained at 3 meters or greater for all lakes (set *CLEAR* = 1 for all lakes).

The scenarios have the same counterfactual: water quality of lakes in 1978. The estimates are per person and for an improvement across all lakes in the state that fall below the standard. We estimate equivalent variation using the results of our five random draw experiments: 3, 6, 12, 24, and 24 (visited) lakes. In the first four experiments all lakes within 180 miles of an individual's home are used to estimate the equivalent variation—the opportunity sets are the same across these four experiments. In the 24(visited) experiment we used the set of lakes visited by at least one person in the sample and within 180 miles of the person's home to calculate the equivalent variation—the opportunity set is less than half as large as in the first four cases. The average per choice occasion values for persons in our sample (Δw in equation [5]) are in Table 6 for all models for both scenarios. The average, minimum, and maximum values for Δw for persons in our sample for the case with 24 alternatives are given in Table 7.

For the low standard, persons using the lakes primarily for swimming have the highest equivalent variation per choice occasion followed by fishing, boating, and viewing. For the cases of 24 draws these values are \$0.83 for swimming, \$0.50 for fishing, \$0.19 for boating, and \$0.15 for viewing. This pattern holds roughly across all experiments with 6 or more alternatives in the drawn opportunity set. Indeed, there appears to be little lost in estimating average equivalent variation with 6 instead of 24 alternatives or by using only visited lakes instead of all lakes to form the opportunity sets.

The latter result is perhaps the most surprising. We are reducing each person's opportunity set by more than half and yet we see little change in their value of water quality benefits per choice occasion. This result is not so surprising when we consider that equivalent variation is just the natural

TABLE 6
Sample Average Equivalent Variation Per Choice Occasion Per Individual
(1978 Dollars)

Number of Lakes in Randomly Drawn Opportunity Set	Low Standard[a]				High Standard[b]			
	Primarily Boating	Primarily Fishing	Primarily Swimming	Primarily Viewing	Primarily Boating	Primarily Fishing	Primarily Swimming	Primarily Viewing
3	\$.97	\$.65	\$ .70	\$.00	\$13.64	\$ .65	\$7.71	\$4.70
6	.15	.48	.89	.00	17.72	.48	9.61	9.10
12	.14	.49	1.01	.22	1.62	1.64	4.87	5.06
24	.19	.50	.83	.15	4.22	.94	5.82	5.53
24(Visited)	.10	.47	.89	.26	1.08	1.76	4.87	3.35

[a]This standard improves lakes with the lowest dissolved oxygen levels to moderate oxygen levels. Approximately 25 percent of the lakes in the state are affected. (All lakes with *DONO* = 1 are set to *DONO* = 0.)

[b]This standard improves all lakes with low or moderate dissolved oxygen levels to high dissolved oxygen levels and improves all lakes with poor clarity to good clarity. Approximately 95 percent of the lakes are affected by the dissolved oxygen improvement and 90 percent by the clarity improvement. (All lakes are set to *DONO* = 0, *DOYES* = 1, and *CLEAR* = 1.)

TABLE 7
SAMPLE AVERAGE, MINIMUM, AND MAXIMUM CHOICE OCCASION VALUES PER INDIVIDUAL FOR CASE OF 24 ALTERNATIVES RANDOMLY DRAWN
(1978 Dollars)

	Low Standard			High Standard		
	Minimum	Average	Maximum	Minimum	Average	Maximum
Primarily Boating	\$.00	\$.19	\$ .56	\$2.35	\$4.22	\$4.73
Primarily Fishing	.05	.50	1.96	.38	.94	2.44
Primarily Swimming	.04	.83	3.22	3.19	5.82	8.15
Primarily Viewing	.01	.15	.60	2.49	5.53	6.64

logarithm of a ratio

$$(1/\beta_y) \cdot \ln\left\{ \sum_{\{i \in T\}} \exp(\tilde{x}_i \beta) \Big/ \sum_{\{i \in T\}} \exp(x_i \beta) \right\},$$

where T is the full set of sites (within 180 miles) over which the sum is calculated, $\tilde{x}$ is the measure of lake characteristics when the water is clean, and x is the measure of characteristics when water is dirty. We ignore nesting for the moment. The 24(visited) experiment gives a Δw with a similar ratio

$$(1/\beta_y) \cdot \ln\left\{ \sum_{\{i \in T'\}} \exp(\tilde{x}_i \beta) \Big/ \sum_{\{i \in T'\}} \exp(x_i \beta) \right\}$$

where T' is a subset of T. T' is the set of lakes visited by at least one person in the sample and within 180 miles of an individual's home. Define the ratios $R_T = \{\Sigma_{\{i \in T\}} \exp(\tilde{x}_i\beta)/\Sigma_{\{i \in T\}}\exp(x_i\beta)\}$ and $R_{T'} = \{\Sigma_{\{i \in T'\}} \exp(\tilde{x}_i\beta)/\Sigma_{\{i \in T'\}}\exp(x_i\beta)\}$. Assume we know β. When $R_T = R_{T'}$, it must be the case that

$$R_T = R_{T'} = \left\{ \sum_{\{i \in T - T'\}} \exp(\tilde{x}_i \beta) \Big/ \sum_{\{i \in T - T'\}} \exp(x_i \beta) \right\}.$$

$\{T - T'\}$ is the set of lakes visited by no one in the sample. If $R_T = R_{T'}$ the per choice occasion values using the full set of lakes and the set of visited lakes are equal. Thus, roughly speaking, if the relative effect on the utilities at sites excluded from the opportunity set is the same as at the sites included in the opportunity set, little is lost by excluding the former from the individual's opportunity set in estimating equivalent variation. Such an assumption appears to be reasonable in our sample.

In the case of the high standard, the stability of the per choice occasion values found above is not as strong. The loss here reflects the variability in the coefficient estimates on *DOYES*. (Note that whenever a coefficient on a water quality variable has the wrong sign, we set its value to zero in computing the equivalent variation.) That variability is magnified in the high standard scenario because over 90 percent of the lakes are affected by the change in water quality. In the low standard only 25 percent of the lakes are affected. Nevertheless, the parameter estimates for the experiments with 12, 24, and 24(visited) draws are reasonably stable. Persons using the lakes primarily for swimming and viewing have the largest values—nearly \$5 in the cases with 12 draws or greater. These are nearly 4 to 5 times larger than the values estimated for fishing and boating.

VII. CONCLUSION

We have demonstrated an approach to estimating a Random Utility Model of recreation when the number of potential sites in the opportunity set is large, say greater than 100. It is a random draw approach based on a suggestion by McFadden (1978). The approach is widely applicable because opportunity sets are large in many recreation demand analyses. We showed a case where parameter estimates and benefit measures for a water quality improvement may be estimated with reasonable confidence with as few as six alternatives included in a person's opportunity set when over 1,000 alternatives are possible. Finally, we presented some welfare estimates that may be of practical use to policymakers. Notable here is our finding that those who visited lakes primarily for boating and viewing had considerable use value for water quality. These groups are often thought to benefit little from water quality improvements.

We suggest three additional experiments with randomly drawn opportunity sets. First, the estimates from a random draw model should be compared with the estimates from an aggregate site model. Aggregation could be done along county lines or some other regional lines. In theory, the aggregate site model introduces an aggregation bias that is alleviated in the random draw approach. The greater the heterogeneity of site characteristics the greater the bias. A comparison would give an empirical test and suggest the potential bias in parameter and benefit estimates. Kaoru and Smith (1990) have initiated such tests in a study with 35 possible sites and with various levels of aggregation.

Second, experiments with weighted random draws should be investigated. If several common, perhaps frequently visited, sites are known to enter most persons' opportunity sets, there may be gains in estimation accuracy to include these in every draw. Opportunity sets for estimation might include a core set of sites plus a random draw from outside that set. Appropriate adjustment of $Pr(D|i)$ in equation [6] would allow estimation without bias.

Finally, using information on individuals' "perceived" opportunity set may be promising. If an individual reveals by questionnaire that he or she is unaware of certain sites, should these sites be included in the opportunity set for estimation? Perhaps we can improve estimates by advance screening. Comparison of estimates with "perceived" versus full opportunity sets would tell.

APPENDIX

Each interviewee reported trips to as many as six lakes during the year. However, the interviewee reports length of trip (number of days) for only one of those lakes (randomly chosen by the interviewer). For that lake, which we will refer to as the interviewee's "key" lake, the trip length is reported as "the usual trip length for a visit to that lake."

For interviewees that visited more than one lake, we do not know the usual trip length to those lakes which are not key lakes. For these lakes we predicted the probability of the usual trip being for a day. We used the following procedure to predict these probabilities.

We formed a sub-sample of trips including each interviewee's "key lake" trip. We created a dummy variable for each observation.

$$DAY = \begin{cases} 0 \text{ if usual trip length was a day} \\ 1 \text{ if usual trip length was for more than} \\ \text{a day} \end{cases}$$

Then, we estimated a logit regression to predict the probability that a trip to his or her key lake was for a day. The dependent variable was *DAY*, the parameter estimates were

	Coefficient	Std. Error
INTERCEPT	−3.86	.39
DISTANCE	.046	.003
INCOME	.002	.001
BOATER	.865	.397
ANGLER	−.223	.320
SWIMMER	−1.09	.458
OWNCABIN	2.46	.808

DISTANCE is miles from interviewee's hometown to key lake; *INCOME* is annual income/100; *BOATER*, *ANGLER*, and *SWIMMER* are dummy variables for the interviewee's primary use of lakes; and *OWNCABIN* = 1 if interviewee owns property on a lake and = 0 if not. For each trip taken to a lake that was not a key lake, the probability of that trip being for a day was estimated using this regression equation.

Then, we estimated the day-trip likelihood function using the following rules:

1. For a key lake where the typical trip length is reported as being for a day, all trips to that lake are assumed to be for a day.
2. For a key lake where the typical trip length is greater than a day, all trips are

assumed to be for greater than a day and are not included in estimation.

3. For a lake that is not key and is located more than 180 miles from the interviewee's hometown, all trips are assumed to be for greater than a day and not included in estimation.
4. For a lake that is not key and is located less than 180 miles from the interviewee's site, the total number of day trips is assumed to be $T \cdot Pr(DAY)$. $Pr(DAY)$ is the predicted probability of a usual trip to that lake being for a day using the equation reported above. T is the actual number of trips reported by the interviewee to the lake.

References

Ben-Akiva, Moshe, and Steven R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge: The MIT Press.

Binkley, Clark S., and W. Michael Hanemann. 1978. *The Recreation Benefits of Water Quality Improvements: An Analysis of Day Trips in an Urban Setting*. Washington, DC: E.P.A.-600/5-79-010.

Bockstael, N. E., W. M. Hanemann, and C. L. Kling. 1987. "Modeling Recreational Demand in a Multiple Site Framework." *Water Resources Research* 23:951–60.

Bockstael, Nancy E., K. E. McConnell, and Ivar E. Strand. 1988. "Benefits from Improvements in Chesapeake Bay Water Quality." In *Measuring the Benefits of Water Quality Improvements Using Recreational Demand Models*, Vol. 3, Washington, DC, E.P.A., Agreement CR-81143-01-1, University of Maryland.

———. 1989. "A Random Utility Model for Sportfishing: Some Preliminary Results for Florida." *Marine Resource Economics* 6:245–60.

Caulkins, Peter P. 1982. "An Empirical Study of the Recreational Benefits Generated by Water Quality Improvement." Ph.D. diss., Agricultural Economics, University of Wisconsin, Madison.

Caulkins, Peter P., Richard C. Bishop, and Nicolaas W. Bouwes, Sr. 1986. "The Travel Cost Model for Lake Recreation: A Comparison of Two Methods for Incorporating Site Quality and Substitution Effects." *American Journal of Agricultural Economics* 68 (2):291–97.

Feenberg, Daniel, and Edwin Mills. 1980. *Measuring the Benefits of Water Pollution Abatement*. New York: Academic Press.

Hanemann, W. Michael. 1978. "A Methodological and Empirical Study of Recreation Benefits from Water Quality Improvement." Ph.D. diss., Department of Economics, Harvard University.

———. 1982. "Applied Welfare Analysis with Qualitative Response Models." California Agricultural Experiment Station Working Paper No. 241. Sacramento, CA.

Hanemann, W. Michael, Richard T. Carson, Russell Gum, and Robert Mitchell. 1987. "Southcentral Alaska Sport Fishing Economic Study." Prepared for Alaska Department of Fish and Game, Anchorage, AK.

Kaoru, Y., and V. K. Smith. 1990. "Black Mayonnaise and Marine Recreation: Methodological Issues in Valuing a Cleanup." Resources for the Future Discussion Paper (QE91-02), Washington, DC.

McFadden, Daniel. 1978. "Modelling the Choice of Residential Location." In *Spatial Interaction Theory and Planning Models*, eds. A. Karlqvist et al., 75–96. Amsterdam: North-Holland Publishing.

Milon, Walter J. 1988. "A Nested Demand Shares Model of Artificial Marine Habitat Choice by Sport Anglers." *Marine Resource Economics* 5 (3):191–213.

Morey, E. R., W. Douglass Shaw, and R. D. Rowe. 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 (Mar.):181–201.

Parsons, George R., and Mary Jo Kealy. 1990. "Measuring Water Quality Benefits Using a Random Utility Model of Lake Recreation in Wisconsin." University of Delaware, Department of Economics, Working Paper Series No. 90-14.

Small, Kenneth A., and Harvey S. Rosen. 1981. "Applied Welfare Economics with Discrete Choice Models." *Econometrica* 49 (1): 105–30.

Uttormark, Paul D., and J. Peter Wall. 1975. "Lake Classification for Water Quality Management." Second printing, 1976. Water Resources Center, University of Wisconsin, Madison.

Welch, E. B. 1989. "Alternative Criteria for Defining Lake Quality for Recreation." *Enhancing States' Lake Management Programs:* 7–15.