Recreational Swimming Benefits of New Hampshire Lake Water Quality Policies: An Application Of A Repeated Discrete Choice Model

Michael S. Needelman and Mary Jo Kealy

Water pollution control policies generally direct sources (i.e., industry, agriculture) to reduce loadings of certain pollutants. Thus, evaluating the relative net recreation benefits of policies to improve water quality requires establishing a linkage between the sources, the resultant water quality degradation at the affected water bodies, and, ultimately, the effect on recreation behavior. This linkage is rarely present in the empirical literature which is, thus, deficient for water pollution control policy assessment purposes. In this paper, we estimate the relative recreational swimming benefits that may result from controlling point and nonpoint sources of pollution, respectively, in New Hampshire's lakes. We use a repeated discrete choice framework to model swimming behavior as a function of each lake's level of eutrophication, bacteria, and oil and grease. For each pollutant, at each affected lake, we identify which source is responsible for the pollution, and we conduct scenarios controlling each pollution source independently, and then, taken together. Seasonal benefit estimates are presented for each scenario. Coupled with information on the most cost effective means of generating the scenarios, these estimates provide a useful starting point for a quantitative assessment of the net recreation benefits of policies to improve the quality of New Hampshire lakes.

For research on valuing environmental amenities, la raison d'etre is to provide information on changes in non-market values related to changes in policy. Yet, in the case of valuing enhanced water based recreation opportunities, the specified recreation demand models often cannot be related to water pollution control policy. Smith and Kaoru (pp. 419) make this point in their survey piece when they say ". . . the available benefit estimates fall short of what is needed for an increasing array of policy related activities." Matulich et al. also suggest this when they say, ". . . recreation economics is a policy-oriented sub-discipline and we ask specialists on either the benefit side or the cost side to serve policy analysis more fruitfully by structuring analyses to fit or to be consistent with some form of managerial model."

Water pollution control policies generally direct sources (i.e., industry, agriculture) to reduce loadings of certain pollutants. Thus, evaluating the relative net recreation benefits of policies to improve water quality requires establishing a linkage between the sources, the resultant water quality degradation at the affected water bodies, and, ultimately, the effect on recreation behavior.

We know of only one study, a recreational fishing application by Karou, Smith, and Liu, that attempts to identify all these linkages. Instead, the bulk of the recreational fishing literature has emphasized the values of access (Freemen). Comparatively few fishing demand studies (e.g., Karou, Smith, and Liu; Karou and Smith; Karou; Cameron) added the important step of demonstrating the sensitivity of fishing behavior to changes in water quality. Furthermore, the literature on relating changes in water quality to the other forms of outdoor recreation (i.e. swimming, boating, and other near shore activities) is even more sparse (e.g., Bockstael et al. 1987a; Bockstael et al. 1988; Parsons and Kealy).

Michael S. Needelman is with the University of Delaware and Mary Jo Kealy is with the Economic Analysis and Research Branch, EPA. The authors wish to thank George Parsons, Frank Lupi, Douglass Shaw, Dan Hellerstein, and two anonymous referees for their insightful comments. We would also like to thank Jim Lockhart for providing valuable research assistance and the people at the New Hampshire Department of Environmental Services for providing us with the water quality data, and for helping us to understand the information contained in these data.

The present paper focuses on estimating the relative recreational swimming benefits that may result from controlling point and nonpoint sources of pollution, respectively. Our application involves lake swimming in New Hampshire. To determine whether or not a lake is suitable for swimming, New Hampshire has developed standards based upon eutrophication and the level of bacteria found in the lakes. We model swimming behavior as a function of these pollutants as well as a third pollutant, oil and grease, which has been found to affect swimming behavior (e.g., Bockstael, Hanemann, and Strand). For each pollutant, at each lake, we identify which source is responsible for the pollution thus completing the source—water quality—recreation behavior linkage discussed earlier.

We conduct the point and nonpoint source control scenarios over the state's set of high priority lakes because these lakes are most likely to be targets for improvement.[1] Mean and aggregate seasonal benefits are presented for each policy scenario. Also, for some scenarios, we present mean seasonal benefits by socioeconomic status. Last, for illustrative purposes, we estimate the recreational swimming benefits of achieving the Clean Water Act goal of improving all New Hampshire lakes to swimmable quality.

The Model

We estimate a repeated discrete choice model (Morey, Rowe, and Watson). On any given day during the recreation season, an individual is faced with choosing to swim at one of J sites, or choosing not to swim. This process is then "repeated" over the T days of the season.

In modeling this decision process, we assume that on a given day, an individual has a utility associated with each of these alternatives. In each case this utility is divided into a systematic and a random component. The utility associated with individual n choosing to visit site j on day i is:

$$(1) \qquad U_{ijn} = V_{ijn} + \epsilon_{ijn}$$

Where:

V_{ijn} = The systematic component of utility measurable by the researcher.

ϵ_{ijn} = The random component of utility unknown to the researcher, but known to the individual on day i.

Commonly, it is assumed that the systematic component of utility is a linear function of its parameters so that:

$$(2) \qquad V_{ijn} = \beta_M(M_{in} - P_{ijn}) + \beta X_{ijn}$$

Where:

M_{in} = The income individual n has to spend on day i.

P_{ijn} = The price individual n has to pay to visit site j on day i.

X_{ijn} = The characteristics of site j as perceived by individual n on day i.

β_M = The marginal utility of income.

β = Parameters to be estimated along with β_M.

Included within X_{ijn}, are measures of a site's water quality, different amenities which may or may not be available at a site, and characteristics of the individual interacted with alternative specific variables. The characteristics of the individual only become relevant when interacted with variables which vary by alternative. This is because an individual's characteristics do not vary across the different site utilities, and, thus, do not influence site choice. An example may be helpful. In our survey, we gathered information on an individual's likelihood of boating while on a swimming trip. We call this variable *BOATLIKE*. Entered by itself into equation 2, this information cannot exert any influence over the site choice of an individual because it does not vary across the different sites. The information is clearly important in determining site choice because boating is an option at some sites, but not at others. Thus, we create a variable indicating whether or not boating is available at each site. Then, by taking the product of this variable and *BOATLIKE*, we create an "interaction variable," *BOATINT*, which captures the increase in utility that a person interested in boating experiences when boating is available at a swimming site.

The utility associated with individual n not swimming on day i is:

$$(3) \qquad U_{in} = V_{in} + \epsilon_{in}$$

where V_{in} takes the form of:

$$(4) \qquad V_{in} = B_M M_{in} - \alpha Z_{in}$$

[1] The State's criteria for inclusion in this set is based upon the accessibility of the lakes, the designated uses for the lakes, the overall commitment and interest that New Hampshire residents have in protecting the lakes, the source of the pollution problem, and the feasibility of restoring the lakes.

Where:

Z_{in} = Characteristics of the individual that are likely to influence whether or not she takes a swimming trip. These characteristics are interacted with a dummy variable set to one only for this alternative.

α = Parameters to be estimated.

Since the term $\beta_M M_{in}$ is present in the utilities of each of the J + 1 alternatives, it will not exert any influence on which alternative is chosen. As a result, there is no need to include the term in the actual estimation.

On each day, it is assumed that the individual will choose the alternative which yields the maximum utility. We further assume that the error terms in equations 1 and 3 are identically distributed Type 1 Extreme Value. Given this assumption, it can be shown (Morey) that the probability of individual n choosing site j on day i is:

$$(5)\quad Pr_{ijn} = \frac{e^{\frac{1}{\mu}I_{in}}}{e^{\frac{1}{\mu}I_{in}} + e^{-\alpha z_{in}}} \cdot \frac{e^{\mu\beta X_{ijn} - \mu\beta_M P_{ijn}}}{\sum_k e^{\mu\beta X_{ikn} - \mu\beta_M P_{ikn}}}$$

Where: $I_{in} = \ln \sum_k e^{\mu\beta X_{ikn} - \mu\beta_M P_{ikn}}$

This is the product of the probability of taking a trip on day i and the conditional probability of visiting site j given that a trip is taken. I_{in} is called the "inclusive value," and it represents the individual's expected maximum utility of taking a trip on day i.[2] In deriving equation 5, it is assumed that the error terms in equations 1 and 3 are independent across all individuals, days, and alternatives with the exception of a component that is shared among swimming alternatives, but not with the no-swim alternative. That is, the error terms in the swimming alternatives are assumed to be correlated with each other, but not with the no-swim alternative. The term, $1/\mu$, in equation 5 measures the degree of this correlation. It is bounded by zero and one, with a value close to zero indicating a high degree of correlation.

The probability of individual n choosing not to take a swimming trip on day i is:

$$(6)\quad Pr_{in} = \frac{e^{-\alpha Z_{in}}}{e^{\frac{1}{\mu}I_{in}} + e^{-\alpha Z_{in}}}$$

Then, the likelihood function is expressed as:

$$(7)\quad \prod_n^N \prod_i^T \prod_j^J (Pr_{ijn})^{Y_{ijn}} (Pr_{in})^{Y_{in}}$$

Where: Y_{ijn} = 1 if individual n chooses to swim at site j on day i, 0 otherwise.

Y_{in} = 1 if individual n chooses not to swim on day i, 0 otherwise.

Ideally, one would obtain estimates for α, β_M, β, and $1/\mu$ in (7) using maximum likelihood. Unfortunately, the standard discrete choice software packages do not maximize this function. Instead, the model is estimated in stages. In the first stage, the conditional probability of visiting site j is estimated. This is referred to as the site choice stage. The likelihood function is specified as:

$$(8)\quad \prod_n^{N'} \prod_i^T \prod_j^J \left(\frac{e^{\mu\beta X_{ijn} - \mu\beta_M P_{ijn}}}{\sum_k e^{\mu\beta X_{ikn} - \mu\beta_M P_{ikn}}} \right)^{Y_{ijn}}$$

Where N' = The number of people in our sample who took trips.

The combined parameters $\mu\beta$ and $\mu\beta_M$ are chosen so that this function is maximized.[3] Then, the estimates of $\mu\beta$ and $\mu\beta_M$ are used to construct I_{in} which is carried forward to the trip frequency stage. That is, the probability of taking a trip (and of not taking a trip) is used in a second likelihood function:

$$(9)\quad \prod_n^N \prod_i^T \left(\frac{e^{\frac{1}{\mu}I_{in}}}{e^{\frac{1}{\mu}I_{in}} + e^{-\alpha Z_{in}}} \right)^{(1 - Y_{in})} \cdot \left(\frac{e^{-\alpha Z_{in}}}{e^{\frac{1}{\mu}I_{in}} + e^{-\alpha Z_{in}}} \right)^{Y_{in}}$$

By maximizing this function, estimates of $1/\mu$ and α are obtained. The likelihood functions shown in equations 8 and 9 are maximized using standard programs. The parameter estimates are consistent, but are less efficient than if obtained from maxi-

[2] Actually, the expected maximum utility for individual n taking a trip on day i is I_{in} + Euler's constant (.557). However, this constant term cancels out in estimation.

[3] At this stage, it is impossible to separate the μ from the β's.

mizing equation 7 directly (Brownstone and Small).[4]

To estimate the benefits of water quality improvements, it can be shown (Hanemann) that a compensating variation measure for individual n on day i is:

$$CV_{in} = \frac{\ln\left(e^{\frac{1}{\mu}I_{in}^{2}} + e^{-\alpha Z_{in}}\right) - \ln\left(e^{\frac{1}{\mu}I_{in}^{1}} + e^{-\alpha Z_{in}}\right)}{\beta_M}$$

Where: I_{in}^2 = Inclusive value with the improvement (degradation) in site quality.

I_{in}^1 = Inclusive value without the improvement (degradation) in site quality.

This measure is then summed over all days in the season to get a per season measure of compensating variation for each individual, including both participants and non-participants.

The Data

We developed a database from several sources. First are the data on New Hampshire residents and the trips that they look, if any, during the swimming season. Second are the characteristics (including the water quality) of all public New Hampshire lakes. Last are the distance and time costs (i.e. the "price") for each individual to visit each lake. We are not considering river and stream swimming in this analysis because it is not a close substitute for lake swimming in New Hampshire. The state discourages river and stream swimming for safety reasons, and we observe that only 5.6% of our trips are taken to rivers and streams.

Our data on New Hampshire residents and their swimming trips are from a survey conducted in 1989 in support of the National Acid Precipitation Assessment Program (NAPAP). Using a population based sample, this survey provided data on the demographic and freshwater recreation characteristics of 5,724 individuals in Maine, New Hampshire, New York, and Vermont.[5] Of these, 519 individuals completed the survey in New Hampshire. A screening survey collected demographic information on the individuals and divided them into separate activity panels for fishing, boating, and swimming. Individuals who participated in more than one activity were placed into one of the panels according to the following criteria: those who fished were placed in the fishing panel, boaters who did not fish were placed in the boating panel, and swimmers who neither boated nor fished were placed in the swimming panel. This procedure continued until each panel in each state met a statistically determined target.

Each panelist was administered a questionnaire that collected data on: the individual's characteristics related to the particular panel (e.g. his/her swimming, fishing, or boating ability, whether or not he/she owned a boat, etc . . .), the individual's perceptions of the characteristics on each site that he/she visited, and the details of each trip that was taken. For each individual, a first questionnaire was administered in July and August of 1989, and covered trips taken from April 1st to the time of the interview. This was followed by a second questionnaire which was administered in September and October to cover trips taken during the rest of the season (Shankle et al.).

Fifty-three people taking 1,021 trips are included in the swimming panel for New Hampshire. However, 288 people (approximately 56% of interviewed New Hampshire residents) indicated that they took at least one swimming trip prior to the interview date. For our fifty-three panel members, we had information on the total number of trips taken during the season, and on where they were taken. Thus, these people were used in the site choice as well as the trip frequency models. For the remaining 466 individuals, we had information only on the total number of trips taken (if any) up to the time of the screener interview. These people were used in the trip frequency decision only.[6]

Data on the lake characteristics were obtained from three different sources. The primary water quality data were provided by the Biology Bureau of New Hampshire's Department of Environmental Services (NHDES). The Biology Bureau has collected data on various aspects of a lake's morphology, chemistry, biology, and on its trophic status.

A database of lake amenities was constructed

[4] There are software packages such as Gauss & LIMDEP which maximize any specified function. However, the likelihood function in equation 7 becomes difficult to specify as the number of alternatives increases. Although not much has been published on this issue, there has been discussion as to whether or not the gain in efficiency is worth the added complexity. In this paper, we estimate the model in stages.

[5] The 5,724 responses were out of 11,979 attempts, resulting in a response rate of 48 percent. We used census data to weight our sample of New Hampshire residents so that it would be representative of the state's population in 1989.

[6] A technical appendix is available from the authors which describes this process in more detail.

from the *Inventory of Outdoor Recreation Facilities* published by the New Hampshire Office of State Planning. The inventory lists all private and public facilities located throughout the state along with the amenities (i.e. a swimming beach, boating, picnicking) available at each facility. These amenities are defined to be associated with a particular lake if the facility is within ¼ mile of the lake.

A third source of data was the *Nonpoint Source Pollution Assessment Report* which was released in 1989 by the NHDES. This assessment lists all water bodies believed to be affected by nonpoint source pollution as well as the causes and sources of the problem. A majority of these data are "evaluated" rather than monitored. That is, for many of the lakes, this information is deduced by means other than by taking measurements at the site. Despite this limitation, these data are the best available for nonpoint source pollution in New Hampshire. Also, these data are used by the state, in conjunction with the primary water quality data base described above, to aid in environmental policy decisions.

To compute the price term we used the HYWAYS/BYWAYS software package (New Directions Software, Inc.) to construct a matrix of road distances (and travel times) between each individual's hometown and the set of New Hampshire lakes. This software computes road distances and travel times between towns. The "location" of each lake was determined to be the nearest town recognized by the software. In computing travel time, the traffic patterns of the different roads are taken into account so that travel time is not a linear function of travel distance. To measure the price in dollars, we multiply distance by $.25/mile and travel time by ¼ of the household's hourly income.[7]

Modeling Methodology

In this application, we consider trips taken for one day or less. Overnight trips are omitted because the overnight trip experience is different. That is, a day trip to go swimming is generally taken primarily to go swimming, whereas an overnight trip is likely to be taken for many different purposes.

[7] We tried following both the McConnell and Strand (1981) and the Bockstael et al. (1987b) approaches to treating the opportunity cost of travel time, but obtained imprecise parameter estimates due to multicollinearity. We adopted the fraction (¼) of the wage rate as a conservative lower bound. The tendency would be to underestimate the benefits of quality improvements.

There are 1,071 lakes in New Hampshire, but they are not all available for swimming. First, several of these lakes are surrounded by privately owned land or are designated for the water supply. Also, many lakes are designated as wetlands. Once these lakes are deleted from the choice set, eight-hundred-thirty-seven remain (no trips were observed at any of the deleted lakes). Also, since we are modeling day trips only, we assume that an individual will not travel more than two hours (one way) to reach a swimming destination (nobody in our sample travelled more than 2 hours).

After these adjustments, many swimmers could still choose from among five hundred lakes. Modeling all five-hundred lakes as swimming alternatives is cumbersome. One modeling approach involves constructing aggregate sites based upon some regional denomination (see, for example, Bockstael, McConnell, & Strand, and Wegge, Carson, & Hanemann). While this eases estimation, it could seriously inhibit the model's ability to determine the importance of water quality in swimming site choices (See Parsons and Needelman for details on aggregation bias).

As an alternative to aggregating, we randomly assign nineteen lakes (Parsons and Kealy) to each swimmer's choice set. This random assignment is "representative" of the complete set of choices. Then, we include the lake actually visited on a given trip to give a total of twenty lakes.[8] The random draws are done without replacement, and the actually visited lake is not included in the choice set from which the randomly drawn lakes are picked. McFadden has shown that this procedure provides unbiased estimates of the parameters. The efficiency of the parameters will increase as the number of randomly drawn alternatives increases.

Model Specifications

The lake characteristics that influence the site choice decision are shown in Table 1. We include a price term, several characteristics of the lakes, and the three pollution variables described below. In addition, we include the *BOATINT* discussed earlier.

MESEUT, *BACTPROB*, and *OILGREAS* are each measures of water quality at a site. *MESEUT*

[8] The random draw procedure is used in the site choice stage only. When computing the inclusive value used in the trip frequency stage, we include all the lakes within the choice set of the individual.

Table 1. Variables Used in the Site Choice Stage[1]

PRICE (6.43, 7.76)	Individual's travel costs * \$.25/mile + opportunity costs of time * ¼ daily wage.
BEACH (0.226, 0.418)	= 1 if a beach is present at the site, 0 otherwise.
LNAREA (3.07, 1.53)	Natural log of the acreage of the lake.
LNDEPTH (1.23, 0.582)	Natural log of the depth (in meters) of the lake.
BOATINT (0.070, 0.170)	Product of BOATLIKE and BOATAVAL where BOATLIKE is the observed percentage of all swimming trips to each visited site during which the individual boated as well. BOATAVAL equals one when boating is available at a site, zero otherwise.
ELEVTN (817.88, 475.03)	Elevation, in meters, of the site above sea level.
MESEUT (0.754, 0.431)	= 1 if not oligotrophic, 0 otherwise.
BACTPROB (0.096, 0.294)	= 1 if a bacteria problem exists at the site, 0 otherwise.
OILGREAS (0.045, 0.208)	= 1 if an oil and grease problem exists at the site, 0 otherwise.

[1]Numbers in parentheses below each variable are the mean and std. deviation. These statistics are computed over all trips for: PRICE and BOATINT, and are weighted (using census data) to represent the entire population. All other statistics are computed over all lakes.

Table 2. Variables Used in the Trip Frequency Stage[1]

AGE (43.31, 17.28)	The age of the individual.
AGE2 (2173.77, 1669.82)	The square of the age of the individual.
KIDS5–16 (0.30, 0.46)	= 1 if the individual has children between the ages of 5 and 16, 0 otherwise.
KIDSL5 (0.17, 0.38)	= 1 if the individual has children aged less than 5 years, 0 otherwise.
NOHS (0.18, 0.39)	= 1 if the respondent has not completed high school, 0 otherwise.
HS (0.32, 0.47)	= 1 if the respondent has completed high school, but did not go any further, 0 otherwise.
FTPT (0.66, 0.47)	= 1 if the individual is employed outside the home (either full or part time), 0 otherwise.
ONE	A constant term added to each individual's utility associated with not taking a swimming trip.
IV1 (4.37, 1.85)	The inclusive value term.

[1]Numbers in parentheses below each variable are the mean and std. deviation of the variable. In each case, these statistics were computed over all New Hampshire residents in our sample (both participants and non-participants) and were weighted (using census data) to represent the entire population.

indicates trophic status. It takes a value of 1 if a lake is either mesotrophic or eutrophic, and 0 if it is oligotrophic. *BACTPROB* takes a value of one if a lake has a bacteria problem, and *OILGREAS* indicates an oil and grease problem at a lake.

The majority of the data used to construct *BACTPROB* and *OILGREAS* is evaluated.[9] On the other hand, *MESEUT* is defined using monitored data. It is based upon the level of dissolved oxygen, the secchi disk transparency, the plant abundance, and the level of chlorophyll.[10] These components are correlated with more subjective measures like taste, aesthetics, and odor and, at the same time, are affected directly by environmental policy. By using the objective policy variables, we avoid having to establish links between the subjective measures and changes in the policy variables.

Due to budgetary constraints, the state is able to obtain measures of these components for only about 40 lakes each year.[11] As a result, for many lakes, we used measures taken in years other than 1989, the year of our survey data.[12] Our measurements ranged from 1976 to 1991. A little more than half of the lakes had measurements taken in 1985 or later, and only 64 lakes had measurements taken prior to 1980. This is, nevertheless, a wide range and will undoubtedly lead to measurement error in these components as they are likely to exhibit year to year variability. However, the use of the more aggregate variable, *MESEUT*, limits this variability while still picking up the water quality differences among sites that affect swimming behavior.

The variables included in the trip frequency stage are listed in Table 2. We include information on characteristics of the individual that are likely to influence the decision to take a trip. In addition, we include the "inclusive value" variable, which

[9] An evaluated assessment is based upon non-measurement criteria such as: land use, location of sources, and citizen complaints. To the extent that people are influenced *directly* by these factors rather than through their affect on water quality, we will be overstating the benefits of the water quality improvement.

[10] Aquatic biologists at the NHDES weighted and summed these measures to construct a composite variable indicating trophic status. We tried other less aggregate specifications, but MESEUT provided the best statistical fit.

[11] This is a non-statistical sample so that inferences cannot be drawn about the lakes not sampled in a given year.

[12] Approximately 35 percent of the lakes did not have information for one or more of these components for any year. For these lakes, mean values were used for each component missing.

Table 3. Parameter Estimates in the Site Choice Stage[1]

Variable	Parameter Estimate	Standard Error	T-Ratio
PRICE	−0.250554	0.0113110	−22.151
BEACH	1.31201	0.123321	10.639
LNAREA	0.389023	0.0442483	8.792
LNDEPTH	0.663858	0.110394	6.014
BOATINT	4.21630	0.508546	8.291
ELEVTN	−0.00158300	0.000248736	−6.364
MESEUT	−0.496543	0.126965	−3.911
BACTPROB	−2.38083	0.411922	−5.780
OILGREAS	−0.599820	0.272314	−2.203

[1]The dependent variable is equal to 1 if the individual visited a particular site, 0 otherwise. The above model was estimated using a weighted sample of 53 people taking 1,021 trips. The sample was weighted, using census data, so that it would be representative of the entire state. The goodness of fit measure, ρ^2, is equal to 0.749.

Table 4. Parameter Estimates in the Trip Frequency Stage[1]

Variable	Parameter Estimate	Standard Error	T-Ratio
AGE	0.0126168	0.00691249	1.825
AGE2	−0.000401646	0.00007829	−5.130
KIDS5–16	0.627082	0.0388342	16.148
KIDSL5	−0.437503	0.0504843	−8.666
NOHS	−0.260292	0.0574256	−4.533
HS	−0.433422	0.0415808	−10.424
FTPT	−0.269681	0.0426612	−6.321
ONE	−3.36440	0.150014	−22.427
IV1	0.179268	0.0108579	18.510

[1]The dependent variable is equal to 1 if the individual took a swimming trip on a particular day, 0 otherwise. This model was estimated using the sample of 519 New Hampshire residents. This sample was weighted, using census data, so that it would be representative of the entire state. The goodness of fit measure, ρ^2, is equal to 0.746.

represents the expected maximum utility of taking a trip, thus linking the trip frequency model with the site choice model. Recognizing that New Hampshire has a small coastline (about 18 miles), we attempted to include a variable measuring the minimum distance of each individual to the coast. The inclusion of this variable had an insignificant effect, thus providing evidence that the beach sites are not substitutes for lake swimming in New Hampshire.

Estimation Results

The parameters estimated in the site choice stage are shown in Table 3. A positive parameter associated with a characteristic indicates that the presence of that characteristic increases the individual's chances of visiting a site, while a negative coefficient indicates the opposite. All of the coefficients are significant and have the expected sign. That is, people tend to swim at the closer, larger (and deeper) lakes with a swimming beach and with better water quality. Also, people who enjoy boating while swimming, tend to visit sites with boating available. Last, people tend to avoid the highly elevated lakes indicating their preference for lakes that are more easily accessible, and that have slightly higher water and surrounding air temperatures.

The parameters estimated in the trip frequency stage are shown in Table 4. Here, a positive parameter indicates that the variable increases an individual's chances of taking a swimming trip. We find that an individual's probability of taking a trip increases with age, up to the age of 28, after which it decreases. Also, people with older children in the household and people who do not work outside the home are more likely to take a trip, while people with young children and people who were not educated beyond high school are less likely to take a trip. Last, the coefficient on the inclusive value, $1/\mu$, falls between zero and one, and is significantly different from one. This is a strong indica-

Table 5. Seasonal Benefits Estimates for the Elimination of Pollution Problems in High Priority Lakes[1]

	Mean Seasonal Benefits			Aggregate Seasonal Benefits		
Scenario	All Sources	Nonpoint Sources	Point Sources	All Sources	Nonpoint Sources	Point Sources
Eliminate eutrophication	$1.40	$1.33	$0.09	$1,163,000	$1,105,000	$75,000
Eliminate bacteria	$1.82	$1.82	$0.00	$1,512,000	$1,512,000	$ 0.00
Eliminate oil & grease	$0.09	$0.09	$0.00	$ 75,000	$ 75,000	$ 0.00
Eliminate eutrophication and bacteria problems	$4.09	$3.71	$0.08	$3,397,000	$3,081,000	$75,000
Eliminate all pollution problems	$4.30	$3.93	$0.08	$3,571,000	$3,264,000	$75,000

[1]The numbers in this table were estimated using a weighted sample of the 519 New Hampshire residents (including both participants and nonparticipants). The sample was weighted, using census data, so that it would be representative of the entire state.

Table 6. Initial Pollution Problems in New Hampshire's High Priority Lakes[1]

Pollution Problem	Source		Total
	Nonpoint	Point	
Eutrophication	40	4	44
Bacteria	8	0	8
Oil and grease	3	0	3

[1]Each cell gives the number of lakes with that pollution problem and source. There are 51 lakes on New Hampshire's high priority list.

tion that the error terms among the swimming alternatives are more correlated with each other than with the no-swim alternative, confirming the maintained assumption in our model.

Benefit Estimates

To estimate the benefits of water quality improvements, it is important that the pollution variables not only be significant in explaining behavior, but also be policy relevant. The variables used to construct *MESEUT* and *BACTPROB* are closely related to the variables used by the state to determine if a lake is impaired for swimming. Therefore, except for some extreme cases, if a lake is impaired according to the state's criteria, this will be reflected by the *MESEUT* and/or *BACTPROB* variables.[13] Although *OILGREAS* is not used by the state for this purpose, it is nonetheless significant in explaining swimming behavior, and thus, can be an impediment to achieving swimmable quality water.

We use our model to estimate the swimming benefits that result from restoring the state's highest priority lakes to varying levels of quality.[14] The benefit estimates are shown in Table 5. For each scenario, we distinguish between eliminating point sources and nonpoint sources of pollution.

In scenarios 1 through 3, we estimate the benefits of eliminating eutrophication, bacteria, and oil & grease, respectively. In the fourth scenario, we bring the lakes up to swimmable quality as defined by the State. That is, we simultaneously eliminate all eutrophication problems and bacteria problems from the lakes. Finally, in the last scenario, we simultaneously remove all three pollution problems.

For all 519 New Hampshire residents in our sample, we use equation 10, to compute per-day benefit estimates for each scenario.[15] By summing over the total number of days in the swimming season, we get a per-season estimate of benefits for each individual. We calculate the sample mean per-season estimates of benefits for each scenario and present them in Table 5. We aggregate these benefits up to the state level by multiplying by the number of people in New Hampshire who are 18 years or old, 830,497 (U.S. Department of Commerce, 1990).

In Table 5, we find that most of the benefits come from the elimination of nonpoint sources of pollution. This is not surprising since, in Table 6, we see that almost all of the pollution problems in New Hampshire's high priority lakes are from nonpoint sources. What is surprising is the relatively large benefits accrued when bacteria problems are eliminated. These benefits are over thirty percent larger than the benefits from eliminating eutrophication, and there are over five times as many high priority lakes affected by eutrophication. Another interesting point shown in Table 5 is that the sum of the benefits from each of the independent scenarios is less than the benefits we get when the scenarios are considered collectively. This is an indication that the scenarios are complementary (Hoehn).

In Table 7, we present the benefits estimates by demographic group.[16] We find that people 29 and over tend to receive smaller benefits from water quality improvements. This follows from our previous result in Table 4 where we found that people in this age group, all else equal, were less likely to take a trip. We also see in Table 7 that, with the exception of the oil and grease scenario, people in the lowest income group in New Hampshire receive smaller benefits than their counterparts. Since income effects are not included in our model, this indicates that lower income communities are located relatively far from the high priority polluted sites.[17] It is also reflecting the fact that higher income individuals pay a higher price (in

[13] Our criteria are somewhat stricter than the state's criteria so that an impaired lake in our model may not be designated as being impaired by the state. We have found that swimming behavior is responsive to this stricter criteria. Therefore, to use the state's criteria would result in an understatement of the true benefits.

[14] The variables used in the prioritization of the lakes are given above, in footnote 2.

[15] The benefits to non-swimmers come about as a result of their having an increased probability of using the lakes for swimming once water quality has improved. The model does not capture the non-use benefits of either the participants or the non-participants.

[16] We present the mean per-season estimates from eliminating both point and nonpoint sources of pollution.

[17] Remember that income cancels out of the site choice decision since it is the same for all alternatives. We attempted to include income in the trip frequency decision, but it was not significant in affecting the probability of taking a trip.

Table 7. Mean Seasonal Benefits for the Elimination of Pollution Problems in High Priority Lakes Broken Down by Demographic Characteristics

	Age[1]			Income[1]			Race[2]	
Scenario	18–28 (23, 3)	29–57 (41, 8)	58–86 (68, 8)	$20,000 or Below (12, 5)	$20,001 to $50,000 (33, 7)	Above $50,000 (69, 35)	White (98.7)	Non-white (1.3)
Eliminate eutrophication	$1.91	$1.64	$0.38	$0.51	$1.48	$2.52	$1.41	$0.96
Eliminate bacteria	$2.31	$2.04	$0.85	$0.87	$2.13	$1.90	$1.83	$1.03
Eliminate oil & grease	$0.15	$0.08	$0.05	$0.17	$0.07	$0.04	$0.09	$0.15
Eliminate etrophication and bacteria	$5.39	$4.61	$1.62	$1.84	$4.62	$5.25	$4.10	$2.53
Eliminate all pollution problems	$5.71	$4.86	$1.68	$2.00	$4.86	$5.45	$4.32	$2.81

[1]Numbers in parentheses below each age and income classification are the mean and standard deviation in each group, rounded to the nearest year for age and rounded to the nearest thousand dollars for income. All numbers in this table are weighted, using census data, to be representative of the entire state.
[2]The number below each race classification is the percentage of our weighted sample which falls into that classification.

terms of time costs) to get to substitute sites. Last, there is a tendency for Caucasians to receive higher benefits than their counterparts. This is surprising since race was not significant in predicting participation.[18]

Conclusions

Since the introduction of the discrete choice model, the empirical literature on valuing water quality improvements has grown substantially. Despite this growth, much of the literature is still not well suited to answer specific policy questions. This is due, in part, to the considerable effort required for developing the databases and establishing the linkages among water quality control programs, ambient measures of water quality, and recreation behavior.

In this paper, we use the discrete choice framework to estimate a model of swimming behavior, a popular outdoor recreation activity that has received comparatively little attention in the valuation literature. The recreational swimming benefits from policies to improve water quality in New Hampshire's lakes are estimated. We link the changes in water quality to the elimination of particular sources (point and nonpoint) of pollution in order to assess the relative benefits of source specific policies to eliminate or reduce each type of pollution.

New Hampshire residents may choose a swimming destination from as many as 500 lakes within a two hour drive of their home. Nonetheless, we estimate substantial economic benefits (i.e., over $3.5 million) from eliminating pollution problems (i.e., eutrophication, bacteria, and oil and grease) that impede swimming in New Hampshire's 51 highest priority lakes while holding water quality fixed in all remaining lakes. The available data indicate that most of the high priority lakes are polluted by nonpoint sources. Thus, policies that target nonpoint sources will be necessary to achieve substantial water quality improvements. In addition, it appears that relatively large benefits may come from policies to eliminate bacterial problems. They are found in only eight of the fifty-one high priority lakes, and yet account for the largest part of the economic benefits. Finally achieving swimmable quality water in all New Hampshire lakes, a Clean Water Act goal, generates $18 million in swimming benefits.[19]

These economic benefit estimates relate only to swimming activities and do not include boating, fishing, near shore activities, or non-use benefits. Nonetheless, they provide a useful starting point for a quantitative economic benefit assessment of policies to improve the quality of New Hampshire lakes. Coupled with information on the most cost effective means of generating the water quality improvements, the information on the potential economic benefits resulting from specific policies will facilitate State watershed management and planning.

[18] We caution against using these results to draw general conclusions about race and its affects on benefits in New Hampshire since even with the weighting scheme, there were very few Non-Caucasians in our sample.

[19] In this scenario, we define a "swimmable" water body to be a water body without any of the pollutants we have found to impede swimming. It is equivalent to the last scenario in table 5 done over all lakes.

References

Bockstael, N.E., K.E. McConnell, and I.E. Strand. "Benefits From Improvements In Chesapeake Bay Water Quality." *Benefit Analysis Using Indirect Or Imputed Market Methods,* Vol. 2, Washington, DC, EPA, Agreement CR-811043-01-0, University of Maryland, 1988.

Bockstael, N.E., K.E. McConnell, and I.E. Strand. "A Random Utility Model for Sport Fishing: Some Preliminary Results for Florida." *Marine Resource Economics* 6(1989):245–60.

Bockstael, N.E., W.M. Hanemann, and C.L. Kling. "Modeling Recreation Demand in a Multiple Site Framework." *Water Resources Research* 23 No. 5(1987a):951–60.

Bockstael, N.E., W.M. Hanemann, and I.E. Strand. "Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models." *Benefit Analysis Using Indirect or Imputed Methods,* Vol. 2, Washington, DC, EPA Agreement CR-811043-01-0, University of Maryland, 1984.

Bockstael, N.E., I.E. Strand, and W.M. Hanemann. "Time and the Recreation Demand Model." *American Journal of Agricultural Economics* 69(1987b):213–32.

Brownstone, D., and K.A. Small. "Efficient Estimation of Nested Logit Models." *Journal of Business and Economic Statistics* 7 No. 1(1989):67–74.

Cameron, T.A. "The Determinants of Value for a Marine Estuarine Sportfishery: The Effects of Water Quality in Texas Bays," presented at the AERE Workshop on Marine and Sports Fisheries: Economic Valuation and Management, June 1988.

Freeman, A.M. III. "The Economics of Valuing Marine Recreation: A Review of the Empirical Evidence." Economics Working Paper 93-102, Dept. of Economics, Bowdoin College, Brunswick, ME 04011, 1993.

Hanemann, W. Michael. "Applied Welfare Analysis With Qualitative Response Models." California Agricultural Experiment Stations working paper No. 241, 1982.

Hoehn, J.P. 1991. "Valuing the Multidimensional Impacts of Environmental Policy: Theory and Methods." *American Journal of Agricultural Economics* 73 No. 2(1991):239–99.

Kaoru, Y. "Valuing Marine Recreation by the Nested Random Utility Model: Functional Structure, Party Composition and Heterogeneity," unpublished, Woods Hole Oceanographic Institution, 1991.

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

Kaoru, Y., V.K. Smith, J.L. Liu. "Using Random Utility Models to Estimate the Recreational Value of Estuarine Resources." RFF discussion paper 94-04 (Washington, DC, Resources for Future), 1993.

Matulich, S.C., W.G. Workman, and A. Jubenville. "Recreation Economics: Taking Stock." *Land Economics* 63 No. 3(1987):310–6.

McConnell, K.E., and I.E. Strand. "Measuring the Cost of Time in Recreation Demand Analysis." *American Journal of Agricultural Economics* 63(1981):153–6.

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

Morey, E.R. "Two RUMS Uncloaked: Nested-Logit Models of Site Choice and Nested-Logit Models of Participation and Site Choice," presented at the W-133 Meetings, Tucson Arizona, February, 1994.

Morey, E.R., R.D. Rowe, and M. Watson. "A Repeated Nested-Logit Model of Atlantic Salmon Fishing." *American Journal of Agricultural Economics* 75 No. 3(1993): 578–92.

New Hampshire Department of Environmental Services. *Nonpoint Source Pollution Assessment Report,* 1989.

New Hampshire Office of State Planning. *Inventory of Outdoor Recreation Facilities New Hampshire 1981.*

Parsons, G.R., and M.J. Kealy. "Randomly Drawn Opportunity Sets in a Random Utility Model of Lake Recreation." *Land Economics* 68, No. 2(1992):93–106.

Parsons, G.R., and M.S. Needelman. "Site Aggregation in a Random Utility Model of Recreation." *Land Economics* 68 No. 4(1992):418–33.

Shankle, S.A., J.E. Englin, W.H. Monroe, D.M. Beck, M.F. Casinelli, H. Saunders, and S.E. Grover. "Freshwater-Based Recreation in the Upper Northeast: Data Collection Documentation." Pacific Northwest Laboratory, Battelle Memorial Institute, 1990.

Smith, V.K., and Y. Kaoru. "Signals or Noise? Explaining the Variation in Recreation Benefit Estimates." *American Journal of Agricultural Economics* 72 No. 2(1990):419–33.

U.S. Department of Commerce, Economics and Statistics Administration. *1990 Census of Population and Housing.*

Wegge, T.C., R.T. Carson, and W.M. Hanemann. "Site Quality and the Demand for Sport Fishing in Southcentral Alaska," presented at the Sportfishing Institute Symposium, Charleston, SC, March 14–15, 1988.