Recreation Demand Using Physical Measures of Water Quality

Kevin J. Egan, Joseph A. Herriges, Catherine L. Kling, and John A. Downing

Working Paper 04-WP 372
October 2004

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu

Kevin Egan is a postdoctoral research associate at the Center for Agricultural and Rural Development (CARD); Joseph Herriges is a professor in the Department of Economics; Catherine Kling is a professor of economics and head of the Resource and Environmental Policy Division at CARD; and John Downing is a professor in the Department of Ecology, Evolution, and Organismal Biology; all at Iowa State University.

Funding for this project was provided by the Iowa Department of Natural Resources and the U.S. Environmental Protection Agency's Science to Achieve Results (STAR) program. Although the research described in the article has been funded in part by the U.S. Environmental Protection Agency's STAR program through Grant R830818, it has not been subjected to any EPA review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

For questions or comments about the contents of this paper, please contact Kevin Egan, 569 Heady Hall, Iowa State University, Ames, IA 50011-1070; Ph: 515-294-0670; Fax: 515-294-6336; E-mail: kegan@iastate.edu.

Abstract

This paper incorporates a rich set of physical water quality attributes, as well as site and household characteristics, into a model of recreational lake usage in Iowa. Our analysis shows individuals are responsive to physical water quality measures. Willingness-to-pay estimates are reported based on improvements in these measures.

1 Introduction

More than three decades have elapsed since the passage of the 1972 Clean Water Act (CWA), yet progress toward meeting the standards set forth in the CWA has been slow in the area of nonpoint source pollution. The most recent National Water Quality Inventory (USEPA,[16]) categorizes forty-five percent of assessed lake acres in the U.S. as impaired, with the leading causes of these impairments being nutrients and siltation. Moreover, few states have developed the priority ranking of their impaired waters or determined the Total Maximum Daily Loads (TMDLs) as required under Section 303(d) of the CWA.[1] Legal actions by citizen groups have prompted renewed efforts toward developing both the priority listing and associated TMDL standards.[2] However, the task facing both the EPA and state regulatory agencies remains a daunting one. The prioritization process alone, which is all the more important given current tight budgets, requires information on the cost of remediation and the potential benefits that will flow from water quality improvements. Both types of information are in short supply. The purpose of this paper is to help fill this gap by providing information on the recreational value of water quality improvements as a function of detailed physical attributes of the water bodies involved. The water quality values are obtained from a recreation demand model of lake usage in the state of Iowa, combining trip and socio-demographic data from the Iowa Lakes Valuation Project and an extensive list of physical water quality measures collected by Iowa State University's Limnology Laboratory.

Recreation demand models have long been used to value water quality improvements, but studies typically rely on limited measures of water quality. The most commonly used indicators are fish catch rates (e.g., [3], [10]). However, catch rates are themselves endogenous,

[1] TMDLs specify the amount of a pollutant that a water body can receive and still meet existing water quality standards.

[2] As of March 2003, there have been approximately 40 legal actions taken against the USEPA in 38 states concerning the implementation of Section 303(d) of the CWA.

depending on both fishing pressure and the abilities of the anglers, and provide only indirect measures of the underlying water quality. Physical water quality measures, such as secchi depth and bacteria counts, are used only sparingly, in large part because of limitations in available data. Phaneuf, Kling, and Herriges [13] use fish toxin levels in their model of Great Lakes fishing, but the toxin levels were available only for a limited number of aggregate sites in the region. Parsons and Kealy [12] use dummy variables based on dissolved oxygen levels and average secchi depth readings to capture the impact of water quality on Wisconsin lake recreation. Similarly, Parsons, Helm, and Bondelid [11] construct dummy variables indicating *high* and *medium* water quality levels for use in their analysis of recreational demand in six northeastern states. These dummy variables are based on pollution loading data and water quality models, rather than on direct measurements of the local water quality. In all of these studies, the physical water quality indicators are found to significantly impact recreation demand, but, because of the limited nature of the measures themselves, provide only a partial picture of value associated with possible water quality improvements.

Bockstael, Hanemann, and Strand's [2] analysis of beach usage in the Boston-Cape Cod area has perhaps one of the most extensive lists of objective physical water quality attributes included in a model of recreation: oil, fecal coliform, temperature, chemical oxygen demand (COD), and turbidity. However, the study also points out one of the frequently encountered problems in isolating the impact of individual water quality attributes - multicollinearity. Seven additional water quality measures were available to the analysts: color, pH, alkalinity, phosphorous, nitrogen, ammonia, and fecal coliform. These latter variables were excluded from the analysis becuase of correlations among the various groups of water quality measures. The five water quality variables used were chosen because they were either directly observable by recreationists or highly publicized. While these choices are certainly reasonable given limitations in the available data, the lack of direct information on how nutrient levels (phosphorous and nitrogen) impact recreational usage is unfortunate in the context of setting TMDL standards in midwestern states, where nutrient loadings are of particular concern.

The contribution of the current paper lies in our ability to incorporate a rich set of physical water quality attributes, as well as site and household characteristics, into a model of recreational lake usage in Iowa. Trip data for the study are drawn from the 2002 Iowa Lakes Survey, the first of a four-year project aimed at valuing recreational lake usage in Iowa. The survey was sent to a random sample of 8,000 Iowa households, eliciting information on their recreational visits to Iowa's 129 principal lakes, along with socio-demographic data and attitudes toward water quality issues. The unique feature of the project, however, is that a parallel inventory of the physical attributes of these lakes is being conducted by Iowa State University's Limnology Laboratory.[3] Three times a year, over the course of a five-year project, eleven distinct water quality measurements are being taken at each of the lakes, providing a clear physical characterization of the conditions in each lake. Moreover, because of the wide range of lake conditions in the state, Iowa is particularly well suited to identifying the impact of these physical characteristics on recreation demand. Iowa's lakes vary from a few clean lakes with up to fifteen feet of visibility to other lakes having some of the highest concentrations of nutrients in the world, and roughly half of the 129 lakes included in the study are on the EPA's list of impaired lakes.

The remainder of the paper is divided into five sections. Section 2 provides an overview of the two data sources. A repeated mixed logit model of recreational lake usage in Iowa is then specified in Section 3. The mixed logit model allows for a wide variety of substitution patterns among the recreational sites and for heterogeneity among households in terms of their reaction to individual site characteristics. (See, e.g., [6],[9], and [15].) Parameter

[3] The limnological study is funded by the Iowa Department of Natural Resources.

estimates are reported in Section 4. In Section 5, we illustrate not only the implications of the model in terms of recreational value of meeting the objectives of the CWA (i.e., removing all of the lakes in the state from the impaired water quality list) but also how the model can be used to prioritize the remediation task. Conclusions of the paper are provided in Section 6.

2 Data

Two principal data sources are used in developing our model of recreational lake usage in Iowa: the 2002 Iowa Lakes Survey and the physical water quality measures collected by Iowa State University's Limnology Laboratory. As noted earlier, the 2002 Iowa Lakes Survey is the first survey in a four-year study of lake usage in the state. The focus of the survey was on gathering baseline information on the visitation patterns to Iowa's 129 principal lakes, as well as socio-demographic data and attitudes towards water quality issues. After initial focus groups and pre-testing of the survey instrument, the final survey was administered by mail in November 2002 to 8,000 randomly selected households in the state. Standard Dillman procedures ([5]) were used to ensure a high response rate.[4] Of the 8,000 surveys mailed, 4,423 were returned. Allowing for the 882 undeliverable surveys, this corresponds to an overall response rate of sixty-two percent.

The survey sample was initially paired down to 3,859 households as follows. Those individuals who returned the survey from out of state were excluded (thirty-eight observations). It is not feasible to ascertain whether these respondents have permanently left the state or simply reside elsewhere for part of the year. Respondents who did not complete the trip questions or did specify their numbers of trips (i.e., they simply checked that they had visited a given lake) were excluded (224 observations). Lastly, anyone reporting more than fifty-two

[4] Complete details of the survey design and implementation can be found in [1].

total single-day trips to the 129 lakes were excluded (133 observations). In the analysis that follows, only single-day trips are included to avoid the complexity of modeling multiple-day visits. Defining the number of choice occasions as fifty-two allows for one trip per week to one of the 129 Iowa lakes. While the choice of fifty-two is arbitrary, it seems a reasonable cut-off for the total number of allowable single-day trips for the season.[5] This last step eliminated approximately three percent of the returned surveys. Finally, because of the large number of respondents, the overall sample was randomly divided into three segments; specification, estimation, and prediction portions. The analysis reported here comes from the specification stage using 1,286 observations. Once the estimation stage is reached, the results will be free from any form of pretest bias and the standard errors will be not be biased by the extensive specification search.[6]

Table 1 provides summary statistics for trip and the socio-demographic data obtained from the survey. The average number of total single-day trips for all 129 lakes is 6.68, varying from some respondents taking zero trips and others taking fifty-two trips. In general, the survey respondents are more likely to be older, male, have a higher income, and to be more educated than the general population. Schooling is entered as a dummy variable equaling one if the individual has attended or completed some level of post-high school education.

The physical water quality measures used in modeling recreational lake usage in Iowa were gathered by Iowa State University's Limnology Laboratory. Table 2 provides a listing of the water quality attributes and 2002 summary statistics for the 129 lakes used in our analysis. All of the physical water quality measures are the average values for the 2002 season. Samples were taken from each lake three times throughout the year, in spring/early summer, mid-summer, and late summer/fall to include seasonal variation.

[5] Sensitivity analysis, raising the allowable number of trips per year above fifty-two, indicated that the results were not sensitive to the choice of this cut-off.

[6] Creel and Loomis [4] use a similar procedure in investigating alternative truncated count data estimators.

Each of the water quality measures help to characterize a distinct aspect of the lake ecosystem. Secchi depth indicates the lake depth at which the bottom of the lake can still be seen, providing an overall water clarity measure. Chlorophyll is an indicator of plant biomass or algae, which in turn leads to greenness in the water. Three nitrogen levels are gathered. In addition to total nitrogen, NH3+NH4 measures particular types of nitrogen, such as ammonia, that can be toxic, whereas NO3+NO2 measures the nitrate level in the water. Total phosphorous is an important indicator of water quality in Iowa, as it is usually the principal limiting nutrient which determines algae growth. Silicon is important to diatoms, a key food source for marine organisms. The acidity of the water is measured by "pH" with levels below 6 or above 8 indicating unhealthy lakes. As Table 2 notes, all of the pH levels in this sample are tightly clustered between 7.3 and 10. Alkalinity is the concentration of calcium or calcium carbonate in the water. Plants need carbon to grow and all carbon comes from alkalinity; therefore, alkalinity is an indication of the abundance of plant life. Inorganic suspended solids (ISS) consist basically of soil and silt in the water due to erosion, where as volatile suspended solids (VSS) consists of organic matter. Increases in either ISS or VSS levels will decrease water clarity. With the exception of pH levels, Table 2 demonstrates that there is considerable variation in water quality conditions throughout the state. For example, secchi depth varies from a low of 0.09 meters (or 3.5 inches) to a high of 5.67 meters (over 18 feet). Total phosphorus varies from 17 to 453 ug/L, some of the highest concentrations in the world.

In addition to trip and water quality data, two other data sources were used. First, the travel costs, from each survey respondent's residence to each of the 129 lakes, were needed. The out-of-pocket component of travel cost was computed as the round-trip travel distance multiplied by \$0.25 per mile.[7] The opportunity cost of time was calculated as one-third the

[7] *PCMiler (Streets Version 17)* was used to compute both round-trip travel distance and time.

estimated round-trip travel time multiplied by the respondent's wage rate. Table 3 provides summary statistics for the resulting travel cost variable. The average price of a recreational trip to a lake is \$136, although perhaps a more meaningful statistic is the average price of a lake visit, \$85.

Second, lake site characteristics were obtained from the Iowa Department of Natural Resources [8]. Table 3 provides a summary of these site characteristics. As Table 3 indicates, the size of the lakes varies considerably, from 10 acres to 19,000 acres. Four dummy variables are included to capture different amenities at each lake. The first is a "ramp" dummy variable which equals one if the lake has a cement boat ramp, as opposed to a gravel ramp or no boat ramp at all. The second is a "wake" dummy variable which equals one if wakes are allowed and zero otherwise. About sixty-six percent of the lakes allow wakes, whereas thirty-four percent of lakes are "no wake" lakes. The "state park" dummy variable equals one if the lake is located in a state park, which is the case for 38.8 percent of the lakes in our study. The last dummy variable is the "facilities" dummy variable. Facilities include things like restrooms, picnic tables, or vending machines. A concern may be that facilities would be strongly correlated with the state park dummy variable. However, while fifty of the lakes in the study are located in state parks and fifty have accessible facilities, only twenty six of these overlap.

3 The Model

The mixed logit model was chosen because it exhibits many desirable properties, including that "...it allows for corner solutions, integrates the site selection and participation decisions in a utility consistent framework, and controls for the count nature of recreation demand" (Herriges and Phaneuf, [6]).

Assume the utility of individual i choosing site j on choice occasion t is of the form

$$U_{ijt} = V\left(X_{ij}; \beta_i\right) + \varepsilon_{ijt},\ \ i = 1, ..., N;\ \ j = 0, .., J;\ \ t = 1, ..., T \tag{1}$$

where V represents the observable portion of utility, and from the perspective of the researcher, ε_{ijt}, represents the unobservable portion of utility. A mixed logit model is defined as the integration of the logit formula over the distribution of unobserved random parameters (Revelt and Train, [14]). If the random parameters, β_i, were known then the probability of observing individual i choosing alternative j on choice occasion t would follow the standard logit form

$$L_{ijt}\left(\beta_i\right) = \frac{\exp\left(V_{ijt}\left(\beta_i\right)\right)}{\sum\limits_{k=0}^{J} \exp\left[V_{ikt}\left(\beta_i\right)\right]}. \tag{2}$$

Since the β_i's are unknown, the corresponding unconditional probability, $P_{ijt}\left(\theta\right)$, is obtained by integrating over an assumed probability density function for the β_i's. The unconditional probability is now a function of θ, where θ represents the estimated moments of the random parameters. This repeated Mixed Logit model assumes the random parameters are $i.i.d.$ distributed over the individuals so that

$$P_{ijt} = \int L_{ijt}\left(\beta\right) f\left(\beta|\theta\right) d\beta. \tag{3}$$

No closed-form solution exists for this unconditional probability and therefore simulation is required for the maximum likelihood estimates of θ.[8]

Following Herriges and Phaneuf [6], a dummy variable, D_j, is included which equals one for all of the one through J recreation alternatives and equals zero for the stay-at-home option ($j = 0$). Including the stay-at-home option allows a complete set of choices, including in the population those individuals who always "stay at home" on every choice occasion and

[8] Randomly shifted and shuffled uniform draws are used in the simulation process (Hess, Train, and Polak, [7]). The number of draws used in the simulation is 750.

do not visit any of the sites. It is convenient to partition the individual's utility into the stay-at-home option or choosing one of the J sites, with

$$U_{ijt} = \begin{array}{l} \beta^{z'} z_i + \varepsilon_{i0t} \\ \beta_i' x_{ij} + \alpha_i + \varepsilon_{ijt}, \;\; j = 1, ..., J \end{array}, \tag{4}$$

where α_i is the random parameter on the dummy variable, D_j, which does not appear since it equals one for $j = 1, ..., J$ and zero for $j = 0$. The vector z_i contains socio-demographic data such as income and age, and x_{ij} represents the site characteristics that vary across the lakes, including attributes such as facilities at the lake as well as water quality measures. Notice that the parameters associated with the socio-demographic data are not random as this information does not vary across the sites.[9]

The random coefficient vectors for each individual, β_i and α_i, can be expressed as the sum of population means, b and a, and individual deviation from the means, δ_i and γ_i, which represents the individual's tastes relative to the average tastes in the population (Train, [15]). Therefore, redefine

$$\beta_i' x_{ij} = b' x_{ij} + \delta_i' x_{ij} \tag{5}$$

$$a_i = a + \gamma_i \tag{6}$$

and then the partitioned utility is

$$U_{ijt} = \begin{array}{l} \beta^{z'} z_i + \eta_{i0t} \\ \beta_i' x_{ij} + a + \eta_{ijt}, \;\; j = 1, ..., J \end{array}, \tag{7}$$

where

$$\eta_{ijt} = \begin{array}{ll} \varepsilon_{i0t} & i = 1, ..., N; \;\; t = 1, ..., T \\ \delta_i' x_{ij} + \gamma_i + \varepsilon_{ijt}, \quad j = 1, ..., J; & i = 1, ..., N; \;\; t = 1, ..., T \end{array} \tag{8}$$

is the unobserved portion of utility. This unobserved portion is correlated over sites and trips because of the common influence of the terms δ_i' and γ_i, which vary over individuals. For example, an individual who chooses the stay-at-home option for all choice occasions

[9] It is possible to interact the socio-demographic data with the sites, if one believed, for example, that income would affect which lake was chosen.

would have a negative deviation from a, the mean of α_i, while someone who takes many trips would have a positive deviation from a, allowing the marginal effect to vary across individuals. However, the parameters do not vary over sites or choice occasions; thus, the same preferences are used by the individual to evaluate each site at each time period. Since the unobserved portion of utility is correlated over sites and trips, the familiar IIA assumption does not apply for mixed logit models.

In particular, we model the utility individual i receives from choosing lake j on choice occasion t as

$$U_{ijt} = \begin{array}{l} \beta^{z'} z_i + \varepsilon_{i0t} \\ -\beta^P P_{ij} + \beta^{q'} Q_j + \beta_i^{a'} A_j + \alpha_i + \varepsilon_{ijt},\ j = 1, ..., J \end{array}, \tag{9}$$

where z_i is the socio-demographic data summarized in Table 1, P_{ij} is the travel cost from each Iowan's residency to each of the 129 lakes, as calculated with PCMiler (Table 3). The vector Q_j denotes the physical water quality measures (Table 2) and A_j represents the attributes of the lake (Table 3). As shown in equation (9), notice that the parameters on the lake attributes and the dummy variable, D_j, are random. These six variables are assumed to be independently normally distributed with the mean and dispersion of each variable estimated.

Finally, we estimate two models. The first specification, model A, includes six physical water quality measures. Included are the four paramount variables for nutrient criteria (USEPA [16]): total phosphorus, total nitrogen, chlorophyll, and secchi depth, as well as inorganic suspended solids and organic suspended solids, which we consider to be crucial indicators as well. A second model, model B, includes the complete list of eleven water quality measures. Estimating two models allows us to observe the stability of the parameters across different specifications.

4 Results

The results for model A and B are divided into two Tables, 4a and 4b. For both models, the coefficients for the socio-demographic data, price, and the random coefficients on the amenities are given in Table 4a. Table 4b lists for both models the coefficients for the physical water quality measures. All of the coefficients are significant at the one percent level except for a few of the socio-demographic data. For model B, with eleven physical water quality measures, only the "male" dummy variable is not significant. In model A, income, household size, and the quadratic term on age are insignificant. Note that the socio-demographic data are included in the conditional indirect utility for the stay-at-home option. Therefore, the negative income coefficient indicates that as income rises the respondents are less likely to stay at home and more likely to visit a lake (i.e., lake visits are a normal good). Males, higher-educated individuals, and larger households are all more likely to take a trip to a lake. Age has a convex relationship with the stay-at-home option and therefore has a concave relationship with trips. For model B, the peak occurs at about age 37, which is consistent with the estimate of larger households taking more trips, as at this age the household is more likely to include children.

The price coefficient is negative as expected and identical in both models. Now turning to the amenities parameters, again all of the parameters are of the expected sign. As the size of a lake increases, has a cement boat ramp, gains accessible facilities, or is in a state park, on average leads to increased trips. Notice, however, the large dispersion estimates. For example, in model A the dispersion on the size of the lake indicates 11.1 percent of the population prefers a smaller lake, possibly someone who enjoys a more private experience. The large dispersion on the "wake" dummy variable seems particularly appropriate given the potentially conflicting interests of anglers and recreational boaters. Anglers would possibly prefer "no wake" lakes and recreational boaters would obviously prefer lakes that allow wakes. It seems the population is almost evenly split, with 56.9 percent preferring a lake that allows wakes and 43.1 percent preferring a "no wake" lake. Lastly, the mean of α_i, the trip dummy variable, is negative, indicating that on average the respondents receive higher utility from the stay-at-home option, which is expected considering the average number of trips is 6.7 out of a possible 52 choice occasions.

The physical water quality coefficients are reported in Table 4b and are relatively stable across the two models. For both models A and B, secchi depth is positive and the suspended solids, both organic and inorganic (volatile), are negative, indicating the respondents strongly value water clarity. However, the coefficient on chlorophyll is positive, suggesting that on average respondents do not mind some variation of green water. The negative coefficient on total phosphorus, the most likely principal limiting nutrient, indicates higher algae growth leads to fewer recreational trips.

The only physical water quality coefficient to change qualitatively across the two specifications is total nitrogen, which is positive in model A. Total nitrogen having a positive coefficient is consistent with expectations given the negative sign on total phosphorus. With such large amounts of phosphorus in the water, more nitrogen can actually be beneficial by allowing a more normal phosphorus-to-nitrogen ratio. If the ratio becomes too imbalanced, more problematic blue-green algae blooms become dominant. Total nitrogen is negative in model B, but two other forms of nitrogen are included, with the nitrates form (NO3+NO2) being positive, possibly for the same reason as just discussed.

Continuing with the additional measures in model B, alkalinity has a positive coefficient, consistent with alkalinity's ability to act as a buffering capacity on how much acidity the water can withstand before deteriorating. Since all of the lakes in the sample are acidic (i.e., pH greater than seven), a positive coefficient for alkalinity is expected. The positive coefficient on silicon is also consistent since silicon is important for diatoms, which in turn are an important food source for marine organisms. Lastly, pH is entered quadratically, reflecting the fact that low or high pH levels are signs of poor water quality. However, as mentioned, in our sample of lakes all of the pH values are normal or high. The coefficients for pH show a convex relationship (the minimum is reached at a pH of 8.2) to trips, indicating that as the pH level rises above 8.2, trips are predicted to increase. This is the opposite of what we expected and further specifications will consider this fact.

5 Welfare Calculations

Given the random parameters β_i, the conditional compensating variation associated with a change in water quality from Q to Q' for individual i on choice occasion t is

$$CV_{it}(\beta_i) = \frac{-1}{\beta^p}\left\{\ln\left[\sum_{j=0}^{J}\exp\left(V_{ijt}\left[Q';\beta_i\right]\right)\right] - \ln\left[\sum_{j=0}^{J}\exp\left(V_{ijt}\left[Q;\beta_i\right]\right)\right]\right\}$$

which is the compensating variation for the standard logit model. The unconditional compensating variation does not have a closed form, but it can be simulated by

$$CV_{it} = \frac{1}{R}\sum_{r=1}^{R}\frac{-1}{\beta^p}\left\{\ln\left[\sum_{j=0}^{J}\exp\left(V_{ijt}\left[Q';\beta_i^r\right]\right)\right] - \ln\left[\sum_{j=0}^{J}\exp\left(V_{ijt}\left[Q;\beta_i^r\right]\right)\right]\right\}$$

where R is the number of draws and r represents a particular draw from its distribution. The simulation process involves drawing values of β_i and then calculating the resulting compensating variation for each vector of draws, and finally averaging over the results for many draws. Following Von Haefen [17], 2,500 draws were used in the simulation.

Three water quality improvement scenarios are considered with the results from model A used for all the scenarios. The first scenario improves all 129 lakes to the physical water quality of West Okoboji Lake, the cleanest lake in the state. Table 5 compares the physical water quality of West Okoboji Lake with the average of the other 128 lakes. All of West Okoboji Lake's measures are considerably improved over the other 128. For example, West Okoboji Lake has slightly over five times the water clarity, measured by secchi depth, of the other lakes. Given such a large change, the annual compensating variation estimate of $208.68 for every Iowa household seems reasonable (Table 7). Aggregating to the annual value for all Iowans simply involves multiplying by the number of households in Iowa, which is 1,153,205.[10] Table 7 also reports the average predicted trips before and after the water quality improvement. Improving all 128 lakes to the physical water quality of West Okoboji Lake leads to a reasonable 14.1 percent increase in average trips. As expected, the predicted trips to West Okoboji Lake fall by 19.8 percent, from 0.39 average trips per Iowa household to 0.31. Iowans can now choose the nearest lake with the attributes they prefer, instead of traveling further to West Okoboji Lake.

The next scenario is a less ambitious, more realistic plan of improving nine lakes to the water quality of West Okoboji Lake (see Table 5 for comparison). The state is divided into nine zones with one lake in each zone, allowing every Iowan to be within a couple of hours of a lake with superior water quality. The nine lakes were chosen based on recommendations by the Iowa Department of Natural Resources for possible candidates of a clean-up project. The annual compensating variation estimate is $39.71 for each Iowa household. As expected, this estimate is 19.0 percent of the value if all lakes were improved, even though the scenario involves improving only 7.0 percent of the lakes. This suggests location of the improved lakes is important and to maximize Iowan's benefit from improving a few lakes, policymakers should consider dispersing them throughout the state.

The last scenario is also a policy-oriented improvement. Currently of the 129 lakes, 65 are officially listed on the EPA's impaired waters list. TMDLs are being developed for these lakes and by 2009 the plans must be in place to improve the water quality at these lakes enough to remove them from the list. Therefore, in this scenario the 65 impaired lakes are

[10] Number of Iowa households as reported by Survey Sampling, Inc., 2003.

improved to the median physical water quality levels of the 64 non-impaired lakes. Table 6 compares the median values for the non-impaired lakes to the averages of the impaired lakes. The table indicates that the median values of the non-impaired lakes seem an appropriate choice, with physical water quality measures higher than the averages of the 65 impaired lakes but much below those of West Okoboji Lake. This scenario is valued considerably lower than the first two water quality improvement scenarios. The estimated compensating variation per Iowa household is \$4.87. Consistent with this, the predicted trips only increase 0.3 percent over the predicted trips with no improvement in water quality. A reasonable conclusion is that Iowans have an abundance of lakes at this threshold level, and bringing the low-quality lakes up to this level is not much of a benefit.

6 Conclusions

The first-year survey of the Iowa Lakes Project gathered information about the recreational behavior of Iowans at 129 of Iowa's principal lakes. This data was combined with extensive physical water quality measures from the same set of lakes gathered by the Iowa State University Limnology Lab. Our analysis, which employs the repeated mixed logit framework, shows that individuals are responsive to physical water quality measures, and it is possible to base willingness-to-pay calculations on improvements in these physical measures. In particular we considered three improvement scenarios, with the results suggesting that Iowans value more higly a few lakes with superior water quality rather than all recreational lakes that have only adequate levels (i.e., sufficient to not be listed as impaired by the Environmental Protection Agency).

A number of important practical findings come directly from this work. Limnologists and other water quality researchers should be interested in the results of this paper, since the general belief is that visitors care about water clarity as measured by secchi depth (how many meters beneath the surface of the water a secchi dish is visible) or water quality in general. By estimating the partial effects of a list of physical measures, we have determined which measures significantly affect recreationists' behavior. Limnologists and water resource managers can use this information about what physical lake attributes visitors' trip behavior responds to in designing projects for water quality improvements. Our results indicate water clarity is very important as evidenced by the secchi dish and suspended solids parameters. Also, high nutrients measures in general are found to decrease recreational trips.

The findings of this study also have direct relevance for environmental protection managers and citizens concerned with water quality in that they can be used to prioritize clean-up activities to generate the greatest recreational benefits for a given expenditure. Not only can the findings be used to determine which lakes to target and in what order to clean them but also the most efficient levels of improvement can be identified.

Table 1. 2002 Iowa Lakes Survey Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.
Total Day Trips	6.68	10.46	0	52
Income	$56,140	$37,436	$7,500	$200,000
Male	0.67	0.46	0	1
Age	53.36	16.47	15	82
School	0.66	0.47	0	1
Household Size	2.61	1.32	1	12

Table 2. Water Quality Variables and 2002 Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.
Secchi Depth (m)	1.17	0.92	0.09	5.67
Chlorophyll (ug/l)	41	38	2	183
NH3+NH4 (ug/l)	292	159	72	955
NO3+NO2 (mg/l)	1.20	2.54	0.07	14.13
Total Nitrogen (mg/l)	2.20	2.52	0.55	13.37
Total Phosphorous (ug/l)	106	81	17	453
Silicon (mg/l)	4.56	3.24	0.95	16.31
pH	8.50	0.33	7.76	10.03
Alkalinity (mg/l)	142	41	74	286
Inorganic SS (mg/l)	9.4	17.9	0.6	177.6
Volatile SS (mg/l)	9.4	7.9	1.6	49.9

Table 3. Summary Statistics for Lake Site Characteristics

Variable	Mean	Std. Dev.	Min.	Max.
Travel Cost	135.79	29.47	94.12	239.30
Acres	672	2,120	10	19,000
Ramp	0.86	0.35	0	1
Wake	0.66	0.47	0	1
State Park	0.39	0.49	0	1
Facilities	0.39	0.49	0	1

Table 4a. Repeated Mixed Logit Model Parameter Estimates (Std. Errs in Parentheses)[a]

Variable	Model A: 6 Water Quality Measures		Model B: 11 Water Quality Measures	
	Mean	Dispersion	Mean	Dispersion
Income	-0.008^*		-0.12^*	
	(0.007)		(0.007)	
Male	-4.98^*		-0.31	
	(0.42)		(0.42)	
Age	-0.24^*		-0.58^*	
	(0.07)		(0.08)	
Age^2	0.0001		0.0078^*	
	(0.00006)		(0.0007)	
School	-4.45^*		-3.44^*	
	(0.40)		(0.40)	
Household	-0.41		-1.24^*	
	(0.17)		(0.17)	
Price	-0.17^*		-0.17^*	
	(0.0006)		(0.0007)	
Log(Acres)	4.60^*	3.81^*	5.13^*	4.05^*
	(0.064)	(0.057)	(0.067)	(0.06)
Ramp	11.60^*	17.85^*	14.87^*	18.79^*
	(0.78)	(0.51)	(0.89)	(0.59)
Facilities	1.18^*	18.09^*	3.54^*	16.78^*
	(0.26)	(0.28)	(0.24)	(0.25)
State Park	8.00^*	15.15^*	6.67^*	13.99^*
	(0.26)	(0.27)	(0.24)	(0.27)
Wake	2.76^*	15.81^*	-1.64^*	15.57^*
	(0.30)	(0.33)	(0.30)	(0.29)
α	-8.97^*	3.01^*	-9.19^*	3.12^*
	(0.05)	(0.04)	(0.05)	(0.04)

* Significant at 1% level.

[a] All of the parameters are scaled by 10, except α (which is unscaled) and the income coefficient (which is scaled by 10,000).

Table 4b. Repeated Mixed Logit Model Parameter Estimates (Std. Errs. in Parentheses)[a]

Variable	Model A: 6 Water Quality Measures	Model B: 11 Water Quality Measures
Secchi Depth (m)	0.78*	0.84*
	(0.05)	(0.07)
Chlorophyll (ug/l)	0.054*	0.06*
	(0.03)	(0.003)
NH3+NH4 (ug/l)		−0.002*
		(0.0006)
NO3+NO2 (mg/l)		3.16*
		(0.19)
Total Nitrogen (mg/l)	0.31*	−3.21*
	(0.01)	(0.19)
Total Phosphorous (ug/l)	−0.0033*	−0.016*
	(0.001)	(0.001)
Silicon (mg/l)		0.81*
		(0.02)
pH		−136.72*
		(5.83)
pH^2		8.35*
		(0.34)
Alkalinity (mg/l)		0.038*
		(0.002)
Inorganic SS (mg/l)	−0.010*	−0.089*
	(0.008)	(0.009)
Volatile SS (mg/l)	−0.18*	−0.28*
	(0.01)	(0.02)
LogLik	-47,740.38	-47,494.17

*Significant at the 1% level.

[a] All of the parameters are scaled by 10.

Table 5. West Okoboji Lake vs. the other 128 Lakes

	West Okoboji Lake	Averages of the other 128 Lakes	Averages of the 9 Zone Lakes
Secchi Depth (m)	5.67	1.13	1.23
Chlorophyll (ug/l)	2.63	41.29	40.13
Total Nitrogen (mg/l)	0.86	2.22	3.64
Total Phosphorous (ug/l)	21.28	106.03	91.11
Inorganic SS (mg/l)	1.00	9.49	9.52
Volatile SS (mg/l)	1.79	9.43	8.42

Table 6. 64 Non-impaired Lakes vs. the 65 Impaired Lakes

	Median of the 64 Non-impaired Lakes	Averages of the 65 Impaired Lakes
Secchi Depth (m)	1.27	0.70
Chlorophyll (ug/l)	23.25	56.76
Total Nitrogen (mg/l)	1.11	2.77
Total Phosphorous (ug/l)	58.79	153.70
Inorganic SS (mg/l)	3.51	20.42
Volatile SS (mg/l)	6.02	15.49

Table 7. Annual Compensating Variation Estimates using Model A

Average CV	All 128 Lakes Improved to W. Okb.	9 Zone Lakes Improved to W. Okb.	65 Impaired Lakes Improved to Median
Per choice occasion	\$4.01	\$0.76	\$0.09
Per Iowa household	\$208.68	\$39.71	\$4.87
For all Iowa households	\$240,649,000	\$45,788,092	\$5,612,219
Predicted Trips (9.80 with current water quality)	11.18	10.06	9.83

References

[1] Azevedo, C.D., K.J. Egan, J.A. Herriges, and C.L. Kling (2003). *Iowa Lakes Valuation Project: Summary and Findings from Year One.* Final Report to the Iowa Department of Natural Resources, August.

[2] Bockstael, N.E., W.M. Hanemann, and I.E. Strand (1986). *Measuring the Benefits of Water Quality Improvements Using Recreationd Demand Models*, report presented to the Environmental Protection Agency under cooperative agreement CR-811043-01-0, Washington, D.C.

[3] Chen, H. Z., F. Lupi, and J. P. Hoehn (1999). "An Empirical Assessment of Multinomial Probit and Logit Models for Recreation Demand," in Herriges, J.A., and C. L. Kling, *Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice*, Cheltenham, UK: Edward Elgar, pp. 141-161.

[4] Creel, M.D., and J.B. Loomis (1990). "Theoretical and Empirical Advantages of Truncated Count Data Estimators for Analysis of Deer Hunting in California." *American Journal of Agricultural Economics* **72**(2): 434-41.

[5] Dillman, D. A. (1978) *Mail and Telephone Surveys – The Total Design Method*, New York: Wiley.

[6] Herriges, J., and D. Phaneuf (2002). "Introducing Patterns of Correlation and Substitution in Repeated Logit Models of Recreation Demand." *American Journal of Agricultural Economics* **84**: 1076-1090.

[7] Hess, S., K. Train, and J. Polak (2003). "On the use of randomly shifted and shuffled uniform vectors in the estimation of a Mixed Logit model for vehicle choice." Working paper.

[8] Iowa Department of Natural Resources (2004), *Fishing Guide for Iowa Lakes*.

[9] McFadden, D., and K. E. Train (2000). "Mixed MNL Models for Discrete Response." *Journal of Applied Econometrics* **15**(September/October): 447-70.

[10] Morey, E.R., R. D. Rowe, and M. Watson (1993). "A Repeated Nested-Logit Model of Atlantic Salmon Fishing." *American Journal of Agricultural Economics* **75**(August): 578-592.

[11] Parsons, G.R., E. C. Helm, and T. Bondelid (2003). "Measuring the Economic Benefits of Water Quality Improvements to Recreational Users in Six Northeastern States: An Application of the Random Utility Maximization Model." Working paper, University of Delaware, July.

[12] Parsons, G.R., and M.J. Kealy (1992). "Randomly Drawn Opportunity Sets in a Random Utility Model of Lake Recreation." *Land Economics* **68**(1): 93-106.

[13] Phaneuf, D.J., C.L. Kling, and J.A. Herriges (2000). "Estimation and Welfare Calculations in a Generalized Corner Solution Model with an Application to Recreation Demand." *The Review of Economics and Statistics* **82**(1) (February): 83-92.

[14] Revelt, D., and K. Train (1998). "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." The Review of Economics and Statistics 80: 647-57.

[15] Train, K. (1998). "Recreation Demand Models with Taste Differences Over People." *Land Economics* **74** (2) (May): 230-239.

[16] U.S. Environmental Protection Agency (USEPA) (2000). "Nutrient Criteria Technical Guidance Manual: Lakes and Reservoirs." Office of Water, Office of Science and Technology, Report EPA-822-B00-001, Washington, D.C.

[17] Von Haefen, R. (2003). "Incorporating observed choice into the construction of welfare measures from random utility models." Journal of Environmental Economics and Management 45: 145-65.