The Michigan Recreational Angling Demand Model

Frank Lupi, Department of Agricultural Economics, Michigan State University, East Lansing, MI 48824-1039
John P. Hoehn, Department of Agricultural Economics, Michigan State University, East Lansing, MI
Heng Zhang Chen, American Express Travel Related Services, Phoenix, AZ
Theodore D. Tomasi, Entrix Environmental Consulting, New Castle, DE

Abstract. The paper reports on a large-scale demand model for recreational fishing in Michigan. The model is based on the travel cost method and is specified as a four-level nested-logit. Seasonal participation is modeled by repeating the site choice logit over the course of a season. Data on anglers' trips and site choices were collected using a year-long telephone panel survey of over 1,900 Michigan residents. The model distinguishes among a broad range of fishing trip types including trips of different lengths; trips for different species; and trips to Great Lake, inland lake, and inland stream sites. For Great Lakes fishing, anglers' site choices are related to catch rates which vary spatially and temporally (i.e., time-varying site quality characteristics).

Keywords: Recreation Demand, Nonmarket Values, Nested Logit, Random Utility Model, Repeated Logit

1. INTRODUCTION

Recreational demand models based on the travel cost method are widely used to establish the economic use-values for recreational fisheries. Modern multiple-site travel cost methods allow the spatial patterns of recreation demand to be linked to quality characteristics of recreation sites. This linkage can be used to estimate values of changes in site characteristics. The most common multiple-site travel cost model is the random utility model (RUM) which was popularized in the recreation demand literature by the work of Bockstael et al. (1984, 1987). This paper reports on a large-scale spatial model of the demand for recreational angling in Michigan. The model is a repeated RUM of recreational fishing in Michigan and differs from others in its breadth and scale. The geographic scope of the model is all of the state of Michigan, and the model includes the broad range of fishing activities available in a state with abundant water and fishery resources. Because descriptions of the theory of the RUM and the statistical models used here are readily available in the literature (McFadden, 1981; Train, 1986; Morey, 1999), the exact specification of the site choice probabilities and the likelihood function are relegated to the Appendix. Instead, this paper focuses on describing the model structure and presenting the estimated parameters.

2. MODEL STRUCTURE

The model is a nested repeated random utility travel cost model that permits a wide variety of trip types. Trips are distinguished by trip duration, target species, water body, and destination. Single-day and multiple-day trips are separate types of trips. Fishing trips targeting warm-water species such as walleye are differentiated from trips targeting cold-water species such as salmon. Fishing at Great Lakes, inland lakes, and inland rivers are distinct types of trips. Finally, trips are distinguished by the destination county for the fishing trip. Thus, the model distinguishes between several distinct fishing opportunities available in any given county in Michigan, as well as distinguishing among counties in Michigan.

Two categories of trip duration, single-day trips and multiple-day trips, were chosen to make the best use of the available data. Less than 20% of the trips reported by survey respondents were multiple-day trips. Within each trip length, the different types of fishing trips are called "product lines." Each product line (PL) describes a generalized combination of water body type (Great Lakes, inland lakes, and river/streams) and fish species (cold and warm-water species), plus anadromous runs. This follows Jones and Sung (1993) who use a similar structure in their earlier RUM for Michigan fishing. Based on factor-analytic work by Kikuchi (1986), they categorize fishing activities into seven product lines: (1) Great Lakes warm species, GLwarm; (2) Great Lakes cold species, GLcold; (3) inland lakes warm species, ILwarm; (4) inland lakes cold species, ILcold; (5) river and streams warm species, RSwarm; (6) river and streams non-anadromous cold species, RScold; and (7) river and streams anadromous species, Anad.

Cold-water species include salmon and trout, and warm-water species include bass, yellow perch, panfish, walleye, and pike. In addition, species on anadromous runs are separated from other cold-water river species; anadromous run species include salmon and steelhead. Each of the product lines is available as a single-day trip and as a multiple-day trip. As a result, there are seven fishing product lines within each trip length.

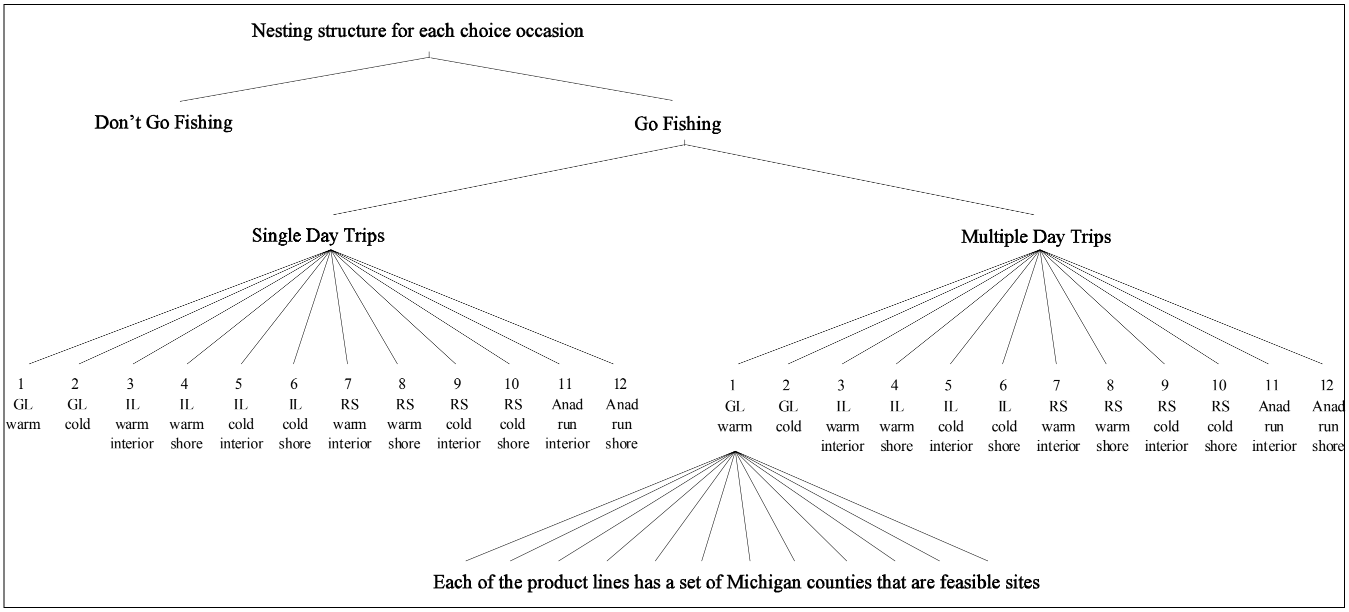

Figure 1: Four-level Nesting Structure for Each Choice Occasion: Participation, Trip, Product Line, and Site Leves.

The sites within each product line are the counties that offer fishing opportunities within that product line. For the Great Lakes product lines, the sites are the stretch of Great Lake shoreline in the county. Each of Michigan's 83 counties can appear several different places as a site in the model. The number of counties that are feasible for each PL are: 41 in GLcold, 40 in GLwarm, 83 in ILwarm, 67 in ILcold, 83 in RSwarm, 69 in RScold, and 44 in Anad. Combining these with the two trip durations yields 854 combinations of trip and site types that are in the model.

In RUMs, the set of sites which can be chosen is called the feasible choice set. The feasible choice set in the model varies across individuals *and* over time. The choice set varies over time because the anadromous run PL is not available during the summer months. The choice set varies for individuals because there is a constraint on how far an individual can travel for a particular trip length. The choice set for day trips is composed of feasible counties within 150 miles of an individual's permanent residence. From the survey data, only about two percent of the observed one-day trips exceed the 150 miles limit. The choice set for multiple-day trips consists of feasible counties within 600 miles of an individual's permanent residence -- the maximum observed one-way driving distance in the sample. Due to the distance constraint, each individual has about 600 feasible fishing activity/site combinations to choose from on each choice occasion with about 60 fewer feasible alternatives in the summer months when anadromous runs are not available.

The model is a *repeated* RUM, where the season is divided into a series of choice occasions (Morey). In each occasion, anglers decide whether or not to fish. The length of the season and the length of a choice occasion jointly determine the number of choice occasions in a season. The model considers only those trips made during the "open water" fishing season, defined as the period from April 1 to October 31. While the determination of the season length is straightforward, many factors affect the selection of the length of the choice occasion. The more choice occasions there are and the larger the number of choices per occasion, the greater the computational burden of the model. On the other hand, if the model contains few choice occasions, there will be some individuals who have taken more trips than there are choice occasions, and some of their trips will need to be "trimmed" from the data. The model permits at most one trip per choice occasion. The length of the choice occasion was set at 3.5 days, or two choice occasions per week. This period was selected because about 96% of all trips were for 4 days or less (3 nights away or less). Of the multiple-day trips, about 80% lasted three nights or less with the average nights away being just less than three. Thus, by setting the choice occasion length at 3.5 days, most of the multiple-day trips "fit" into the choice occasion period.

In the model, catch rates vary by month. Time-varying quality attributes such as these are easily accommodated in the repeated RUM framework. Since catch rates vary monthly, the month of a trip must be known in order to estimate the model. In this sense, the number of trips and choice occasions per month is what drives the data trimming decisions in the model. If there are more trips in any month than there are choice occasions in that month, excess trips must be trimmed. For the model, with the choice occasion length at 3.5 days there are 8 to 9 choice occasions in any month with a total of 61 occasions for the season. For any individual in the sample, if the sum of their reported single-day and multiple-day trips in any month exceeds the number of choice occasions in that month, a random selection of the trips are trimmed so that the month contains no more trips than the number of choice occasions in that month. Less than 2% of the reported trips in any month exceeded the number of choice occasions in that month, and in most months less than 1% of the trips were trimmed.

The model is specified as a four-level nested-logit (Morey 1999; McFadden 1981). This specification permits error terms for alternatives within groups (nests) to be more correlated with one another than with alternatives in other groups which partially relaxes the more restrictive patterns of multinomial logit. The nesting structure closely resembles the types of fishing activities in the model. The first level distinguishes among participation and non-participation; the second level differentiates by trip duration; the third involves the product lines; and the bottom level is for specific sites (Figure 1). The nesting structure is simply a way of grouping elemental alternatives that might be correlated with one another. Whether or not the alternatives are is revealed by parameter estimates. In the nesting structure used in the model, the inland lakes, rivers and streams, and anadromous run categories have been further subdivided to distinguish whether the site is in a county that has Great Lakes shoreline. Here the Jones and Sung nesting structure has been extended for non-Great Lakes product lines to distinguish between counties that do and do not have Great Lakes shoreline. Within each trip length there are a total of twelve nested groups corresponding to shoreline and product line combinations (Figure 1).

3. THE SURVEY DATA

This section provides an overview of the telephone panel survey that was used to collect the data on anglers choices. An important goal of the survey was to obtain accurate data on the number and types of trips individual anglers take in Michigan over the course of a fishing season. Because of the potential difficulties remembering the details of what one does over the course of a season, especially if there are many events to recall, a panel survey was developed which followed a sample of anglers throughout the 1994 fishing season. Recall difficulties have been shown to increase with the length of the recall period and the number of intervening fishing events (WESTAT). In addition, the Michigan Department of Natural Resources (MDNR) has tested alternative mail survey formats, and a comparison of three month and annual recall periods demonstrated that frequent anglers' recall of trips is substantially biased upward with the longer recall period (D. Jester, MDNR, personal correspondence). In the panel survey, the length of time between individuals' panel interviews was varied depending on the anticipated frequency of an individual's fishing. Thus, recall periods were shorter for anglers who fished often.

The telephone survey was implemented using Computer Assisted Telephone Interviewing (CATI). With CATI surveys, the instrument can be programmed to utilize complex skip patterns without having to depend on the interviewer or the respondent to follow the appropriate skip patterns. Also, questions can be programmed to utilize information provided in response to previous questions and/or earlier interviews. Tailoring the survey instrument to each individual can improve the accuracy of respondents' answers, reduce the length of the interview, and reduce the cognitive burden of the interview on respondents.

The survey development included four focus groups and an extensive pilot survey. The pilot survey was a small-scale version of the full survey. The pilot survey was conducted during the fishing season of 1993, and the full survey was conducted in 1994 to1995.

The full survey consisted of two phases: a screening interview to recruit potential anglers into the panel and the subsequent panel interviews. The screening interviews were conducted from late March through early May, 1994. The sample was drawn from the phone numbers for the general population of Michigan residents. To improve the efficiency of the screening interviews, the sample of telephone numbers was stratified so that the proportion of numbers per county matched the proportion of licensed anglers per county. In the initial telephone contact, a random adult (age 18 or older) respondent was selected from the household. Males were over-sampled to improve the efficiency of the screening interviews because male are more likely to fish than females.

The screening interview was very brief with a few questions about fishing and demographics. Anyone indicating they fished in the previous year or they were “likely” to fish in the upcoming year was asked to participate in the panel. The definition of these “potential” anglers was based on an analysis of the pilot survey results. The response rate for completing the screening interviews was between 62% and 75%, depending upon the method of calculating response rates. Of the 6,342 individuals who completed the screening, 3,415 respondents were identified as potential anglers and asked to join the panel. Of these, 2,668, or 78%, agreed to participate. Of those that agreed to participate in the panel, 2,135, or 80%, completed the entire panel. In the RUM model, only data for individuals who completed the panel was used.

Weights were created to appropriately adjust for the stratification based on county population of licensed anglers, the male-female ratio, the number of adults in the household, and the number of telephone lines. After correcting for the sample stratification scheme used in the screening, there was some evidence that persons responding to the screening interview were slightly different than the Michigan population as a whole. To correct for these differences, case weights were created for each sampled person. These case weights were calculated so that the screening sample matched census data on the joint distribution of Michigan adults by regions, age, education, and gender.

The timing of panel interviews was designed so that frequent anglers were called more often than infrequent one.

Using screening questions, panel members were partitioned into three groups based on their anticipated frequency of fishing. The group of frequent anglers was called six times in the period from April through November. The middle range group was called four times while the group of infrequent anglers was called twice. The grouping of respondents and the number of waves for each group was based on an analysis of the pilot data. The goal was to obtain the highest quality data on as many trips as possible, keeping research budget constraints in mind. The scheduling of the panel waves balanced the cost of the panel against the desire to reduce the recall period since the last interview. Another factor that was taken into account was feedback from the pilot survey indicating that infrequent anglers did not want to be called frequently, even if the interview was short. Some detailed information such as the trip length or date was obtained on about 88% of all trips taken by the survey respondents.

Each time panel members were called they were asked how many times, if any, they had fished since their last interview. If they fished they were asked a set of questions about each trip – location, target species, duration, etc. The final panel interview also included questions about their usual travel practices, cabin ownership, and employment characteristics. To avoid double counting of trips and to help respondents answer the question, interviewers would remind respondents of the date of their last interview along with the date and location of the last trip they took – a technique called bounded recall. To minimize any recall bias in the trip counts, respondents who were unable or unwilling to provide details of each of the fishing trips they initially reported were given an opportunity to revise their total number of trips for that interview period. Further, respondents who were in the more avid angling groups were sent fishing logs (diaries) to serve as memory aids when completing the phone interviews.

The analysis sample, upon which the model was estimated, consisted of a subset of the survey respondents and the trips they reported. An eligible case satisfied the following five conditions: (1) the respondent completed the panel, (2) the month of every trip the respondent took was known, (3) the respondent's demographic variables were known, and (4) the case was not flagged for errors (e.g., a respondent who was later identified as under 18 years of age). There were 1,902 respondents (of the 2,135 who completed the panel) that met the above criteria. All of these cases were used in the participation level of the model, and all of their valid trips were used in the trip levels of the model. Additional data was needed for a trip to get used in the trip level of the model. Valid observations for the trip stage estimation were those trips by valid cases that satisfied the following criteria: (1) the destination county was known; (2) the product line was known; (3) the trip duration (single or multiple day) was known; (4) the trip occurred between April 1, 1994 and October 31, 1994; (5) the purpose of the trip was fishing; (6) the product line was feasible for the county visited; (7) the trip had not been flagged for errors; (8) the trip distance was less than 150 miles for day trips; and (9) the trip's random integer was not greater than the number of choice occasions in that month (see above section on choice occasions).

Before these conditions were imposed, there were 5,425 trips taken by the 1,902 cases. Of these, 4,269 trips meet the conditions for use at the trip level of the model, though all 5,425 are valid trips at the participation level. It is possible that implementing these exclusions might result in a sample that is not representative of the overall population of potential anglers. Therefore, a set of weights was created for the analysis sample of 1,902 cases that matched it to the (weighted) sample that was originally recruited to the panel, the "potential" anglers identified in the screening interview. This weighting process ensured that the distribution of angler characteristics in the analysis sample matched the distribution of characteristics in the original sample of recruited anglers. The characteristics matched included the angler's avidity group, the region of the state the angler lived in, the anglers age, and some additional demographic variables.

4. VARIABLES

In this section, the variables that enter the nested logit are briefly described. The discussion begins with the trip price variable (travel cost). Travel costs are the sum of driving, lodging, and time costs. To the extent possible, the travel costs were tailored to the individual sample members.

Driving costs per mile were calculated using a regression-based prediction of per mile fuel costs plus a per mile depreciation charge. Predicted fuel costs vary across individuals in the model. The dependent variable in the fuel cost regression was the reported fuel cost for the fishing trip. The equation was a tobit model with multiplicative heteroskedasticity. The explanatory variables were *distance* (the round-trip distance) and three variables interacted with distance: *share*, *tow*, and *truck* which represented whether or not expenses were shared with other households, whether a boat was towed, and whether the vehicle was a truck or camper. The *distance*, *tow*, and *truck* variables were significant in explaining fuel costs, while the *truck* and *share* variables were significant in explaining heteroskedasticity. Because the repeated RUM requires travel costs to all sites for all choice occasions, regardless of whether a trip was actually taken, the fuel regression was used to predict fuel costs for all choice occasions. Individuals were asked about their usual vehicle type, whether or not they usually towed a boat, and whether they usually shared expenses with other households. Their answers to these questions yield the explanatory variables for fuel cost predictions for all choice occasions.

The complete per mile vehicle cost was calculated by adding to each individual's predicted fuel cost a 14.1 cent

Table 1: Summary of Individual Specific Costs.

Predicted costs†	**Mean**	**Std.Dev.**	**Min.**	**Max.**
Hourly wage	15.2	5.2	5.5	42.9
Driving cost per mile	0.25	0.03	0.22	0.29
Lodging cost per night	19.2	11.1	13.1	47.2

† N = 1,902 cases with complete panel data

per mile depreciation charge, based on the AAA depreciation charge of $141 for each 1,000 miles in excess of 15,000 miles driven annually for a full size passenger vehicle (AAA 1993). This figure does not take into account annual depreciation based on normal driving. Nor does it include any of the fixed costs of vehicle ownership such as insurance. The charge is aimed at capturing the marginal cost of depreciation. Table 1 presents some summary statistics for the predicted per mile driving costs.

For overnight trips, respondents were asked about actual lodging types and lodging costs. The average per-night lodging costs are based on the sample average per night cost of lodging for four types of lodging: camping, hotel, cabin, and other. "Other" includes staying on a boat, fishing all night long, and staying in the car. Individuals were assigned the per night cost that corresponded to the type of lodging that they indicated they usually use when they take overnight trips. The per night costs were multiplied by 3 to reflect the cost of lodging for an average multi-day trip. Lodging costs are zero for single-day trips. Table 1 provides summary statistics for the predicted lodging costs.

Time costs are the predicted value of an individual's time multiplied by the time for the trip length. It was assumed that one-day trips require 8 hours of time and multi-day trips use 3.5 eight hour days or 28 total hours of time. The predicted value of an individual's time came from a regression of wages on demographic variables. A shadow wage equation which accounted for the potential differences between time values and wages for individuals with fixed work hours was tested, but it did not significantly outperform the wage equation. The wage regression was specified as a semi-log model with multiplicative heteroskedasticity. The wage variable was derived from the employment information collected in the survey. For individuals who worked on an hourly basis, wage was specified as their reported wage rate. For those who were on a salary, wage was derived by dividing their annual salary by their annual hours worked. Annual hours worked was calculated using information on the months and hours per week that the respondent worked. The independent variables used in the wage regression were years of education, age, gender, experience, the number of kids in the household, the number of adults in the household, and a dummy variable for those who live in the Detroit metro counties of Monroe, Oakland, and Wayne. Except for adults and kids, all variables were significant in both portions of the wage model. Table 1 presents some summary statistics for the predicted wages.

The rest of this section describes the variables which are used to describe the quality of fishing sites. Catch rates for various species in the GL warm, GL cold, and Anad product lines are included in the model. In the GL warm product line, catch rates were computed for yellow perch, walleye, northern pike, bass (which includes smallmouth bass, largemouth bass, bluegill, and pumpkinseed), and carp (which includes carp, freshwater drum, catfish, and suckers). In GL cold, catch rates are included for chinook salmon, coho salmon, lake trout, and rainbow trout. For the Anad product line, catch rates are included for chinook salmon, coho salmon, and rainbow trout. These catch rates were were estimated using Poisson regressions on MDNR creel survey party interview data (D. Jester, MDNR, personal correspondence). Catch rates are specie and site specific, and they vary monthly throughout the open-water season.

In the IL warm product line, the total surface area (in acres) of warm-water inland lakes within the county is included in the model. For the IL cold product line, the total surface area (in acres) of cold-water lakes within the county is included. The acres for two-story lakes, with warm water on top and cold water below, are included in both warm and cold inland lakes product lines. In the RS cold (RS warm) product lines, the miles of cold (warm) stream in various quality categories in the county are included. The categories are top quality and second quality and are based on the determinations of the MDNR. These stream quality variables reflect the overall ability of the stream to produce a high-quality fishery.

A dummy variable was used to identify whether an individual had a cabin, cottage, or vacation home in a particular county. The effect of this variable was allowed to differ by trip duration but did not differ across product lines within a given trip length. In addition, there are 22 dummy variables distinguishing the combinations of product lines and Great Lake/non-Great Lake counties. There is also a constant at the trip duration level and another at the participation level. The use of constants ensures that the estimated model fits the average of the observed sample data at the level of the constant; the baseline model predictions at the level of a constant will match the sample shares.

At the participation level of the model, demographic variables include sex, age, and years of education. At the trip duration level, the model includes a variable that allows individuals who were employed to have a different value of time than those who did not have a paying job. The variable serves as a shifter for the predicted value of time for individuals without a paying job. The variable was created by interacting an employment dummy with each individual's predicted wage. The dummy variable took the value 1 if the person *did not* have a paying job and this dummy was multiplied by the person's predicted wage; its parameter is only identified at the trip duration level.

At each level of the nesting tree in Figure 1, an inclusive value index is included from the next lower level

Table 2: Parameter Estimates for the Trip Stage of Model.

Variable	Single Day Coef.	Single Day t-stat.	Multiple Day Coef.	Multiple Day t-stat.
Trip Duration Level*				
Trip duration IV	0.043	3.28		
Duration constant	-2.145	-11.8		
Time value shifter	-0.003	-7.50		
Product Line & Site Level				
Product line IV	0.937	28.6	0.617	4.44
Trip cost	-0.143	-58.8	-0.015	-13.3
Cabin at site	1.892	7.86	4.424	19.2
Great Lakes Warm				
Walleye catch rate (CR)	2.531	6.63	1.782	1.52
Bass CR	5.912	1.45	1.652	0.14
Pike CR	1.620	0.36	6.609	0.75
Perch CR	-0.160	-2.75	0.041	0.32
Carp CR	0.972	0.87	0.231	0.09
Great Lakes Cold				
Constant, GL cold	-2.177	-14.8	-0.773	-2.97
Chinook CR	9.170	5.14	15.28	6.05
Coho CR	12.69	5.45	13.62	5.27
Lake trout CR	4.570	3.23	3.036	1.05
Rainbow CR	10.91	2.19	10.25	1.34
Inland Lake Warm				
Constant, ILwarm shore	-1.398	-14.1	0.128	0.63
Constant, ILwarm interior	-0.777	-7.80	0.054	0.24
Warm lake acres/1000	0.067	21.6	0.055	11.7
Inland Lake Cold				
Constant, ILcold shore	-3.185	-11.4	-2.79	-4.92
Constant, ILcold interior	-3.368	-18.5	-2.933	-6.26
Cold lake acres/1000	0.076	3.73	0.065	1.68
Rivers/streams Warm				
Constant, RSwarm shore	-1.448	-10.1	-0.908	-3.41
Constant, RSwarm interior	-1.771	-11.2	-1.519	-4.92
Top quality miles/100	0.761	5.35	1.149	3.15
Second quality miles/100	-0.221	-3.58	-0.413	-2.29
Rivers/streams Cold				
Constant, RScold shore	-2.927	-15.2	-1.763	-5.42
Constant, RScold interior	-3.146	-19.2	-1.562	-5.36
Top quality miles/100	1.552	5.09	1.520	3.40
Second quality miles/100	0.005	0.05	-0.035	-0.24
Anadromous Runs				
Constant, Anad shore	-1.480	-10.57	-0.475	-1.63
Constant, Anad interior	-1.582	-7.78	-0.913	-2.42
Chinook CR	2.795	3.37	4.756	6.37
Coho CR	-0.878	-0.30	6.876	4.28
Rainbow CR	6.997	8.04	6.498	4.58

* Parameters at the trip duration level are listed in the single day column, though they apply to both trip lengths.

of the nesting structure. Since the model has four levels, there are three inclusive value indices -- at the participation level, at the trip duration level, and at the product line level. The inclusive value index is not a separate variable. Rather, it is used to identify parameters of the statistical distribution of nested logit models. If the inclusive value parameter is significantly less than one, the nested logit is preferred to multinomial logit. A necessary condition for the model to be globally consistent with McFadden's (1981) axioms of probabilistic choice requires the estimated inclusive value parameters to lie between zero and one.

5. ESTIMATION RESULTS

In all, the model consists of 78 parameters to be estimated. Each of the 1,902 individuals has about 600 feasible alternatives in their choice sets on each of the 61 choice occasions. Some efficiency in the overall size of the estimation problem was achieved by observing that the variables only vary on a monthly basis, though the estimation problem remains large. The model was estimated in two stages. The lower three levels of nesting (trip length, product line, and site) were estimated using full information maximum likelihood methods. The participation level was estimated using a sequential estimation method for nested logits. In this approach, a trip that did not have complete details of the location of a fishing trip could be used at the participation level of the model even though the information would not be suitable for the site choice level. The parameter estimates will be presented in two tables corresponding to the two stages of model estimation. Details of the likelihood function are provided in the Appendix.

Table 2 presents the estimated parameters associated with variables in the trip duration, product line, and site choice portions of the model. Whenever possible, parameters were allowed to differ for single and multiple-day trips. Thus, in Table 2 there are two sets of columns representing the estimated coefficient and t-statistic for single-day and multiple-day trips. The rows represent the variables which are grouped according to their role in the model. Comparing estimated parameters across trip lengths is complicated by differences in the scale parameters, although parameter ratios are not affected by this.

The first set of rows in Table 2 reports the values of parameters from the trip duration level of the model. The trip duration variables are the inclusive value, a constant, and a time value shifter. While these three parameters are presented in the single day column of Table 2, there is no distinction between parameters for single and multiple-day trips at this model level. The parameter on the trip level inclusive value index is consistent with theory. Since the coefficient is significantly less than one, separating the trip lengths into different nests was an improvement over not doing so. Put differently, the unmeasured characteristics of a trip of a particular duration are more correlated with those of a trip of the same duration than with a trip of a different duration. That the estimated coefficient on the time value shifter is negative suggests that those without jobs have a higher cost of using their time to go fishing than those who have a job. In effect, individuals without jobs are less likely to take multiple-day trips.

After the trip duration choice, the next set of variables in Table 2 represent the product line and the site levels. The first of these are the inclusive value indexes for the product line level nesting. The coefficients on these inclusive values are between zero and one for both trip lengths. Using a one-tailed test, the single-day coefficient is significantly less than one at the 5% level but not at 1%. Each of the combinations of product lines and Great Lake/non-Great Lake counties has a nest-specific constant term. While these are identified at the product line level of the model, in Table 2 they are grouped within each PL.

The next two rows of Table 2 present two site-level variables which share the same parameter across all product lines: trip cost and cabin. Having a cabin at a site makes a person more likely to visit the site, with the influence greater for multiple-day trips. Both the single and multiple-day cabin effect is highly significant. For both day and multiple-day trips, the travel cost variable is negative (as expected) and significantly different than zero. The farther away a site is (all else constant), the less likely it is that it will be visited. This effect is larger in the day trip branch than the multi-day branch, indicating that for single-day trips, the cost of travel is relatively more important than it is for multiple-day trips. These variables were initially constrained to be the same across trip durations, but that specification was clearly rejected using a likelihood ratio test. The implication is that the marginal utility of income differs by trip length.

In the Great Lakes warm product line, the estimated parameters are for catch rates for the individual species. Taken as a group, these catch rates are statistically significant contributors to the model for day trips. Catch rates are not significant at typically-employed significance levels for multiple-day trips for warm-water Great Lakes trips. The coefficients on the catch rates for the single-day part of the model tell us that higher bass and walleye catch rates are highly sought after (all else constant), while northern pike, carp and yellow perch catch rates are less sought after. The negative sign on yellow perch catch rates is unexpected; one might guess *a priori* that yellow perch would have a positive influence. This kind of result might arise because catch rates for perch are correlated with other factors that influence choice. For example, sites with high yellow perch catch rates might be composed primarily of smaller fish, whereas sites with lower catch rates might have a greater share of larger, more desirable fish.

In the Great Lakes cold product line, the estimated parameters are for catch rates for the individual species. The catch rates for each species have a positive influence on both single and multiple-day trips. Within the single-day trips, all of the species significantly contribute to the model's explanatory power. For multiple-day trips, the salmon catch rates are significant, yet the trout catch rates are not. As a group, the multiple-day trip parameters on catch rates are significant. For both single and multiple-day trips, chinook salmon, coho salmon, and rainbow trout are relatively more desirable than lake trout.

Table 3: Participation Choice Level Parameters

Variable	Coefficient	t-statistic
Inclusive value	0.09	1.28
ln (Age)	-0.44	-7.03
ln (Education)	-0.84	-9.29
Sex (male=0)	-0.88	-22.35
April	-7.62	-20.16
May	-7.08	-18.76
June	-7.17	-19.00
July	-7.47	-19.79
August	-7.65	-20.24
September	-8.07	-21.25

For the inland lakes warm-water product line, lake acres is a highly significant variable. The estimates indicate that all else equal, a county with more acres of warm-water inland lakes is more likely to be the destination of single and multiple-day trips. Similarly, in the inland lakes cold product line, acres of cold-water lakes has a positive effect on the chance of a county being selected for either trip length. The parameter on cold lake acres in the IL cold PL is significant for single-day trips, but for multiple-day trips it is only significant at the 10% level.

For the river and streams warm product line, the variables for the miles of top quality and the miles of second quality stream are both significant for single and for multiple-day trips. Top quality stream miles positively influence site choice, while second quality stream miles negatively influence site choice. For the RS cold product line, top quality stream miles again have a significant and positive influence on site choice for both the single and multiple-day trips. Second quality stream miles are not significant for either the single or multiple-day trips portions of the RS cold product line. For the anadromous run product line most of the catch rate variables are significant and positive for single and multiple-day trips. The coho catch rate for single-day anadromous trips is negative, although it is not significant.

The final level of the model is the participation choice, whether to fish or not on a choice occasion. The estimated parameters for the participation model are given in Table 3. This model has several variables, in addition to a set of month-specific constants. First, there is the inclusive value index which summarizes information from the trip choice model. While the participation-level inclusive value is not significantly different from zero, it is significantly less than one indicating the nesting structure is significant. The demographic variables show that, all else equal, older individuals, more educated individuals, and females have a lower probability of taking a trip than do younger, less educated, males. For each individual, the variables at the

Table 4: Statewide Estimates of Resident's Fishing Trips and User Days in Michigan, April to October.

Product Line	Single Day Trips by PL		Multiple Day Trips by PL		Total User Days by PL†	
	(thous.)		(thous.)		(thous.)	
GLwarm	2,082	29%	180	14%	2,776	23%
GLcold	300	4%	162	12%	922	8%
ILwarm	3,092	44%	629	48%	5,513	46%
ILcold	113	2%	22	2%	198	2%
RSwarm	972	14%	125	10%	1,452	12%
RScold	225	3%	94	7%	588	5%
Anad	279	4%	100	8%	663	5%
Totals*	7,062		1,312		12,111	

† User days = multiple-day trips×3.85 + single-day trips.
* Totals may not add up due to rounding.

participation level that vary over time are the inclusive value index and the monthly constants.

6. STATEWIDE TRIP PREDICTIONS

In this section, the estimated model is used to predict total statewide fishing trips. The first step of this process is to combine each individual's data with the site data and the estimated parameters to compute each individual's probability of visiting each site on each choice occasion. Summing these site probabilities up across the choice occasions in the season gives each individual's predicted demand for trips to each site within each fishery type. Next, the weighted average of these seasonal shares is calculated across individuals. These shares are then extrapolated to the state by multiplying by the estimated population of potential anglers in Michigan. The population of potential anglers was estimated from the screening sample, and it too was weighted to be representative of the state population of potential anglers. The results are statewide predictions of trips to each site within each product line. These are added up within a product line and a trip length to produce aggregate estimates of trips within each product line.

The model predicts that during the open-water season (April through October), Michigan residents made about seven million single-day trips and 1.3 million multiple-day trips in Michigan for the primary purpose of fishing. The distribution of these trips across product lines is presented in Table 4. The final columns provide an estimate of total user days by product line. A rough calculation of user days was made by multiplying multiple-day trips by 3.85 (the average nights away plus one), and adding this to single-day trips. This yields an estimate of 12 million user days for fishing in Michigan by state residents. Of the 12 million estimated user days, 58% are due to single-day trips.

How do these estimates compare to other sources of information? The share of user days by fishery types compares favorably with estimates from the 1991 National Survey of Fishing, Hunting, and Wildlife Associated Recreation ("national survey"), with Great Lake trips being slightly lower (36%) for the Michigan panel survey than for the national survey (41%). However, the national survey's 22.8 million user days in Michigan by state residents is higher. The key difference in the sampling frames is that the national estimate refers to all trips that include any fishing, it includes fishing during November through March, and it includes 16 and 17 year-olds. In contrast, Table 4 includes only trips by residents 18 or older during April thru October with a primary purpose of fishing. The panel survey results indicated that fishing was the primary purpose for 97% of the single day trips, but this was the case for only 67% of multiple day trips. Rough adjustments for all of these differences would bring the two estimates closer, but the Michigan survey estimate of total user days would still be lower than the national estiamte. For another view, Bence and Smith (1999) compare estimates of Great Lake fishing effort based on on-site creel surveys performed by Great Lakes states to the 1991 national survey, and they find that the on-site creel surveys tend to predict about half the effort of the national survey.

7. CONCLUSIONS

The purpose of this paper was to summarize the development and estimation of the Michigan recreational angling demand model. While it is not the first repeated nested-logit model of recreation demand, the model is novel in its size and breadth. The model involved 1,900 anglers, each with over 60 choice occasions and with the average choice set on each occasion containing about 600 feasible alternatives. The use of the repeated nested-logit framework allowed us to incorporate site quality characteristics and choice sets that varied over time for each individual. In addition to the usual practice of differentiating choices by location, trips were distinguished by duration, water body fished at, and species sought. The combined geographic scale and broad range of trips types contained in the model makes the research unique. The extent of the model is a reflection of the abundant and complex array of water and fishery resources in Michigan.

Finally, we stress the importance of the trip counts collected through the survey. Since intervening events and longer time periods adversely affect anglers' ability to accurately recall the number of fishing trips they have taken over some time interval, these factors need to accounted for during the design of the surveys used to collect behavioral information. While the accuracy of the recalled number of trips may have little impact on the site and activity portions of such a model, they will directly affect the estimated aggregate number of trips and the aggregate values. There are many examples in the literature of the differences that model specification strategies may have on sample results, yet the process of extrapolating results to more general populations is often ignored. Of course, it is the aggregate results that serve as the basis for water resource and fishery management decisions.

8. ACKNOWLEDGMENTS

We thank Wiktor L. Adamowicz, Richard T. Carson, W. Michael Hanemann, Douglas B. Jester, Carol A. Jones, Edward R. Morey, George Parsons and W. Douglass Shaw for their input during project development. We also thank the Michigan Department of Natural Resources and Department of Environmental Quality for research support. The authors are solely responsible for the contents.

9. REFERENCES

American Automobile Assoc., *Your driving costs*, 1993.

Bence, J, and K. Smith, "An overview of recreational fishing in the Great Lakes," in *Great Lakes fisheries policy and management: a binational perspective*, (W. Taylor and C.P. Ferreri, eds.), East Lansing, Michigan State Univ. Press, 1999.

Bockstael, N., W.M. Hanemann, and I. Strand, "Measuring the benefits of water quality improvements using recreation demand models," EPA Contract No. CR-811043-01-0, 1984.

Bockstael, N., W.M. Hanemann and C. Kling, "Estimating the value of water quality improvements in a recreational demand framework," *Water Resources Research*, 23(5):951-960, May, 1987.

Jones, C., and Y. Sung, *Valuation of environmental quality at Michigan recreational sites: methodological issues and policy applications*, Final Report, EPA Contract No. CR-816247-01-2, September, 1993.

Kikuchi, H., *Segmenting Michigan's sport fishing market: evaluation of two approaches*, PhD Dissertation, Michigan State University, 1986.

McFadden, D., "Econometric models of probabilistic choice models," in: *Structural analysis of discrete data with applications*, (C. Manski and D. McFadden, eds.), Cambridge: MIT Press, 1981.

Morey, E., "Two RUMs uncloaked: nested-logit models of site choice and nested-logit models of participation and site choice," in *Valuing recreation and the environment*, (J. Herriges and C. Kling, eds.), Northhampton: Edward Elgar, 1999.

Train, K., *Qualitative choice analysis*, Cambridge, Mass: The MIT Press, 1986.

U.S. Department of the Interior, Fish and Wildlife Service and Department of Commerce, Bureau of the Census. *1991 National survey of fishing, hunting, and wildlife-associated recreation: Michigan*. 1993.

WESTAT, Inc., *Investigation of possible recall/reference period bias in national surveys of fishing, hunting, and wildlife-associated recreation*, U.S. DOI, Washington, D.C., 1989.

10. APPENDIX

This appendix provides the detailed model specification. On a choice occasion, U_{kstp} is the indirect utility of taking a trip to county k, product line/shoreline s, and trip length t, given a trip is taken (participation) p. Let U_h be the utility of not fishing. Under the hypothesis of utility maximization, the decision on an occasion to take a fishing trip to county k^d, in shoreline/product line s^d, of length t^d implies:

$$U_{k^d s^d t^d p} > U_{kstp} \quad \text{and} \quad U_{k^d s^d t^d p} > U_h \tag{A1}$$

for all feasible $\{k,s,t\} \neq \{k^d,s^d,t^d\}$. Similarly, if the "not fishing" is observed on the occasion, we have the inequality $U_h > U_{kstp}$ for all feasible $\{k,s,t\}$. Since all the factors that influence the individuals' choices cannot be observed, the utility is treated as stochastic and specified as $U_{kstp} = V_{kstp} + \varepsilon_{kstp}$. V_{kstp} is the deterministic indirect utility of taking a trip of type kst, which we specify as a linear function. The "not fishing" utility is $U_h = V_h + \varepsilon_h$ where V_h is a function of demographic variables. The ε's represent the stochastic utility terms and are assumed to follow a generalized extreme value (GEV) distribution. Repeating the utility maximization for each occasion in a season results in a repeated-logit model.

On an occasion in month m, let Pr^{im}_{kstp} be the joint probability of taking a trip to county k in product line/shoreline combination s of trip length t for individual i during month m. Let Pr^{im}_h be the probability of not taking a trip on the occasion in month m. For individual i, the likelihood function during month m is the product of the joint choice probabilities Pr^{im}_{kstp} for the occasions in which the trips are observed and the probabilities Pr^{im}_h for the occasions when there are no trips. That is

$$L^{im} = N_{OC}{}^{im} \times Pr^{im}_h \times \left\{ \prod_{kst \in I^m} Pr^{im}_{kstp} \right\} \tag{A2}$$

where $N_{OC}{}^{im}$ is the number of choice occasions in month m where no trip is observed, and I^m is the set of sites/trip types (kst) for individual i during month m. As a result, the product of the L^{im} over all individuals and months yields the sample likelihood function, $L = \prod_{im} L^{im}$.

To facilitate estimation of the nested-logit model by maximizing L, the choice probabilities of taking a trip on an occasion in month m for individual i (Pr^{im}_{kstp}) were expressed as the product of the conditional probabilities (the superscript i is suppressed in what follows):

$$Pr^m_{kstp} = Pr^m_{k|stp}\, Pr^m_{s|tp}\, Pr^m_{t|p}\, Pr^m_p \qquad \text{(A3)}$$

where $Pr^m_{k|stp}$ is the probability of a trip to county k in month m, conditional on product line/shoreline combination s, trip length t, and participation p; $Pr^m_{s|tp}$ is the probability of a trip to product line/shoreline combination s in month m, conditional on a trip of length t, and participation p; $Pr^m_{t|p}$ is the probability of a trip of length t in month m, conditional on participation p; and Pr^m_p in the probability of taking a trip on a choice occasion in month m.

The random terms were assumed be jointly distributed according to the following generalized extreme value (GEV) distribution

$$F(\varepsilon) = \exp\left(-e^{-\varepsilon_h}\right)\exp\left(-\left[\sum_t\left\{\sum_s\left(\sum_k e^{-\frac{1}{\rho_t}\varepsilon_{kstp}}\right)^{\frac{\rho_t}{\omega}}\right\}^{\frac{\omega}{\eta}}\right]^{\eta}\right) \qquad \text{(A4)}$$

where the subscripts k, s, t, p index the feasible counties, product line/shoreline combinations, trip lengths, and the participation, respectively. The k, s, t, p subscripts index groups of alternatives within the four-level hierarchy of the nesting structure (see McFadden 1981 for more on the GEV distribution). For flexibility, there are separate scale parameters at the level of the product line/shoreline combinations (ρ_t) for the single-day trip branch and the multiple-day trip branch. The other scale parameters ω and η remain the same within their respective levels.

At the bottom level of the model, the conditional probability of taking a trip to county k in product line/shoreline combination s of length t in month m is

$$Pr_{k|stp} = \frac{\exp(\alpha^t_1 p_{kt} + \alpha^t_2 q^m_{ks} + \alpha^t_3 Cbn_{kt})}{\sum_{k|stp}\exp(\alpha^t_1 p_{kt} + \alpha^t_2 q^m_{ks} + \alpha^t_3 Cbn_{kt})} \qquad \text{(A5)}$$

where $\sum_{k|stp}$ means that the summation is over the feasible counties within the product line/shoreline combination s, trip length t, and participation p in month m. α represents parameters to be estimated. There are three types of explanatory variables at this level: (i) p_{kt} are the individual specific travel costs; (ii) q^m_{ks} are the site quality variables, and the superscript m reflects the monthly catch rates; and (iii) Cbn_{kt} is a dummy for individuals with a cabin in county k. The t superscript on α indicates the separate parameter vectors were used to index single and multiple-day trips, allowing the variables to have different effects on the utility of trips of different length.

On the third level, the conditional probability of taking a trip to the product line/shoreline s of length t in month m is:

$$Pr_{s|tp} = \frac{\exp(\beta^t_1 iv^{im}_{stp} + \beta^t_2 D_s)}{\sum_{s|tp}\exp(\beta^t_1 iv^{im}_{stp} + \beta^t_2 D_s)} \qquad \text{(A6)}$$

where $\Sigma_{s|tp}$ indicates summation over all feasible product line/shoreline combinations. There are two types of explanatory variables at this level: (i) Eleven dummy variables (D_s) for the combinations of product lines and counties with/without GL shoreline, s. GLwarm is normalized at 0 for identification during the model estimation; and (ii) The inclusive values iv^m_{stp} of taking a trip to s, calculated as the log of the denominator of (A5). β^t is the vector of parameters to be estimated. The t superscript indicates that the parameter vectors differ by trip lengths.

At the second level, the conditional probability of taking a trip of length t in month m is:

$$Pr_{t|p} = \frac{\exp(\gamma_1 iv^m_{tp} + \gamma_2 D_t + \gamma_3 Tc_t Jb)}{\sum_{t|p}\exp(\gamma_1 iv^m_{tp} + \gamma_2 D_t + \gamma_3 Tc_t Jb)} \qquad \textbf{(7)}$$

where $\Sigma_{t|p}$ means that the summation is over the single-day and multiple-day trip types, given participation. The conditional trip length probability has three variable types: (i) A dummy with D_t=1, if it is a multiple-day trip, 0 otherwise; (ii) The trip's time cost interacted with a dummy for individuals without a job ($Tc_t \times Jb$); and, (iii) The inclusive values iv^m_{tp} calculated as the log of the denominator of (A6).

At the top level, the probability of taking a trip in month m is

$$Pr_p = \frac{\exp(\delta_1 iv^m_p + \delta_3 D_m)}{\exp(\delta_2 sd) + \exp(\delta_1 iv^m_p + \delta_3 D_m)} \qquad \textbf{(8)}$$

where δ are the parameters to be estimated. In any month m, there are two types of variables at the participation level: (*i*) seven dummy variables; D_m=1 if the trip was observed in month m for April through October, and (*ii*) the inclusive values, iv^m_p calculated as the log of the denominator of (A7). Finally, the probability of not fishing on an occasion in month m is

$$Pr_h = \frac{\exp(\delta_2 sd)}{\exp(\delta_2 sd) + \exp(\delta_1 iv^m_p + \delta_3 D_m)} \qquad \textbf{(9)}$$

where δ are parameters. The utility of not fishing is indexed by three demographic variables sd: (1) the natural logarithm of age, (2) the natural logarithm of years of education, and (3) gender with 1 for male, 0 for female.