VALUATION OF ENVIRONMENTAL QUALITY AT MICHIGAN RECREATIONAL FISHING SITES: METHODOLOGICAL ISSUES AND POLICY APPLICATIONS

Prepared by:

Carol Adaire Jones
Yusen D. Sung

EPA Contract No. CR-816247-01-2
FINAL REPORT
September 1993

DISCLAIMER

ACKNOWLEDGEMENTS

We are grateful to the many people who have contributed to the successful completion of this project. In particular we are grateful to Doug Jester of the Fisheries Division, Michigan Department of Natural Resources (MNDR), and to Ted Graham-Tomasi, of the University of Minnesota. Doug initiated the project and, from the beginning, has been a source of creative modeling ideas for the economic analysis and for the Linkages between biology and economics. Doug has provided much of the data used in the analysis, both from environmental resource surveys and from the Michigan angler survey. Almost as important as the data has been his keen appreciation of the sources and limits of the data series. Ted Graham-Tomasi was the original Principal Investigator, starting the project while he was on leave at the School of Natural Resources (SNR), University of Michigan. Due to funding delays, the money came through as Ted was returning to the University of Minnesota, and responsibility for directing the project was passed on to Carol Jones. Ted continued to contribute to the project, as time permitted.

Sharon Nowlen provided valuable assistance as the MDNR Project Manager during the latter portion of the project, facilitating the acquisition of data and smoothing the grant administration process. Anne Wittenberg provided very able research assistance in the early stages for the project. Wendy Silverman provided effective research assistance in sifting information about potential environmental resource measures during the middle of the project.

In its current form, the report is a modified version of Yusen Sung's dissertation, submitted to the Economics Department of the University of Michigan in 1991. Sung has worked on the project from the beginning, originally providing -research and computing assistance, and over time coming to serve as partner in the research. Sung wrote all the computer programs used in the modeling, with the exception of the multinomial logit algorithm, written by C. Manski.

Preliminary results from the research project. have been reported in a series of earlier working papers, cited in the Bibliography. For the most part, revised versions of the earlier work are incorporated in this document.

The Michigan Department of Natural Resources provided generous financial support, in addition to the extensive research support and data noted above. The US Environmental Protection Agency provided support in the form of a Cooperative Agreement: Contract No. CR-816-217-01-2. Resources for the Future (RFF) provided Carol Jones with a Gilbert F. White Fellowship to carry out the research in residence at RFF for a year. and graciously accommodated her during an additional year of leave from the University of Michigan. Other financial support has been provided by the Rackham Graduate School and the Institute for Science and Technology at the University of Michigan.

In addition to the individuals mentioned above! useful comments and collaborative ideas have been offered by Mary Jo Kealy, USEPA and George Parsons, University of Delaware; and by Gary Solon, Joe Swierzbinski, and Al Jensen of the University of Michigan, who served on Yusen Sung's dissertation committee along with Carol Jones. Nonetheless, we alone are responsible for any errors.

In addition, we would like to thank Dean Jim Crow-foot, and Acting Dean Harry Morton, as well as Wayne Say, the Research Director at SNR, for the support that they provided over the years of the project. And finally, we particularly want to thank the team of people in the School of Natural Resources Business Office at the

University of Michigan, including Barb Branscum, Joan Kipfmiller, Barbara Murphy, Tracy Willoughby, Diana Woodworth, and Carole Shadley. Each in her own way has assisted with grace and immeasurable good will with various aspects of the financial and grants administration; all have eased the process of running the grant while on leave elsewhere and have allowed us to concentrate on the research itself.

Addresses for contacting the authors are:

Carol Adaire Jones
National Oceanic and Atmospheric Administration
1305 East West Highway
Silver Spring MD 20910-3281

Yusen D. Sung
Department of Economics
National Taiwan University
Taipei TAIWAN

TABLE OF CONTENTS

ACKNOWLEDGEMENTS ii

LIST OF FIGURES . . . . , ... vii

LIST OF MAPS viii

LIST OF TABLES. ix

CHAPTER

I. INTRODUCTION. 1

Overview of the Model
Performing Policy Analysis
Methodological Issues
Outline of the Report

II. DISCRETE CHOICE MODELS OF RECREATION DEMAND: A BRIEF REVIEW 11

III. INDIVIDUAL MICRO-LEVEL CHOICE MODELING . 19

Consumer Preferences and Behavior
The Micro-level Product Line/Site Decision
Estimation of the Sequential Multinomial Logit
The Nested Multinomial Logit Specification
The Valuation of Time

IV. THE MACRO-LEVEL PARTICIPATION MODELING . . . 47

Participation as the Sum of Independent Trip Decisions
The Critique and an Alternative Proposal
The Stochastic Renewal Approach

V. **CONSUMER WELFARE MEASURE** . . . . , ... **80**

Welfare Measure for Individual Choice Occasions
Welfare Measure for Multiple Choice Occasions

VI. DATA SOURCES AND DESCRIPTIVE STATISTICS 87

Angler Survey Data
Fish Catch Rate Data from the MDNR Creel Survey
Data on Other Characteristics of Site Quality

VII. EMPIRICAL RESULTS: MULTINOMIAL LOGIT **MODEL** **112**

Choice Set Computation: Implementation Details
The Site Choice MNL Estimation
The Product Line Choice MNL Estimation

VIII. EMPIRICAL RESULTS: THE PARTICIPATION **MODEL.** **127**

Variable Definitions and Analysis Sample
Participation Model Estimates
External Validation of the Participation Model

IX. **POLICY APPLICATION: LUDINGTON PUMPED-STORAGE PLANT** 141

Biological Scenarios
Consumer Surplus Calculation
Comparison with Other Estimates

X. POLICY APPLICATION: KALAMAZOO RIVER CONTAMINATION 150

Biological Scenarios
Consumer Surplus Calculation

APPENDIX 158

BIBLIOGRAPHY . . . . , ... 163

LIST OF FIGURES

Figure		
III.1	The flat micro PL-site decision structure.	28
III.2	The two-stage micro PL-site decision structure.	30
IV.1	The choice occasion participation decision	50
IV.2	The truncated between-trip duration	59
IV.3	Derivation of the age distribution	62
IV.4	Hazard rate with inter-type dependence	71
IV.5	Hazard rate without inter-type dependence	72

LIST OF MAPS

Map		
VI.1	State of Michigan counties	101
VI.2	Great Lakes product line counties	102
VI.3	Anadromous run product line counties	103
VI.4	Inland coldwater product line counties	104
VI.5	Areas of Concern in Michigan	105
IX.1	Michigan counties affected by the Ludington scenario	148
X.1	Michigan counties affected by the Kalamazoo scenario	156

LIST OF TABLES

Table

VI.1	Classification of sample observations	106
VI.2	Angler characteristics of the Day group.	107
VI.3	Angler characteristics of the Wkn group.	107
VI.4	Angler characteristics of the Vac group.	107
VI.5	Means and standard deviations of the GLcd catch rates . , ...	108
VI.6	Means and standard deviations of the Anad catch rates . . .	1.08
VI.7	Means and standard deviations of the GLww catch rates . . .	109
VI.8	Descriptive statistics: Great Lakes site attributes	110
VI.9	Descriptive statistics: Anad site attributes	110
VI.10	Descriptive statistics: LScd site attributes	110
VI.11	Descriptive statistics. ILww site attributes	111
VI. 12	Descriptive statistics: ISww site attributes	111
VII.1	MNL estimates for the GLcd product line	120
VII.2	MNL estimates for the GLww product line	121
VII.3	MNL estimates for the Anad product line	122
VII.4	MNL estimates for the LScd product line	123
VII.5	MNL estimates for the ILww product line	124
VII.6	MNL estimates for the ISww product line	125
VII.7	MNL estimates for the product line choice	126
VIII.1	Attributes of the participation analysis sample	136
VIII.2	Distribution of the age (or censored age) duration length	137
VIII.3	Competing risks exponential model estimates	138
VIII.4	Competing risks Weibull model estimates	139

VIII.5	Predicted number of trips per angler in an open-water season	140
VIII.6	Predicted number of total angler-days in an open-water season	140
IX.1	Ludington: Mean compensating variation per trip in 1984 dollars	149
IX.2	Ludington: Total trips per person with plant operation	149
IX.3	Ludington: Mean change in season trips . . .	149
IX.4	Ludington: Mean season compensating variation in 1984 dollars	149
X.1	Kalamazoo: Mean compensating variation per trip in 1984 dollars	157
X.2	Kalamazoo: Total trips per person before PCB cleanup . .	157
X.3	Kalamazoo: Mean change in season trips	157
X.4	Kalamazoo: Mean season compensating variation in 1983 dollars	157
X.5	MNL estimates for the GLcd-Day sample	160
X.6	MNL estimates for the GLcd-Wkn sample . . .	161
X.7	MNL estimates for the GLcd-Vac sample	162

CHAPTER I

INTRODUCTION

In the research described in this report, we have developed a random utility model of demand for recreational fishing in Michigan, covering all water bodies and all species types throughout all counties in the state. The major study sponsor, the Michigan Department of Natural Resources (MDNR), funded the research to produce a model that could be used to improve fisheries resource management and to perform natural resource damage assessments. One out of every two households in Michigan has a fishing license, suggesting that fishing-related benefits will represent a substantial portion of the total benefits of improvements in water and sediment quality.

The travel cost model was designed to value recreational experiences. In a recent state-of-the-art review of recreation models, Bockstael, McConnell and Strand (1991) conclude that the random utility version of the travel cost model is particularly well-suited to valuing changes in quality at one or more recreation sites. The random utility model allows the researcher to model a wide range of substitution possibilities and, consequently, provides a procedure for estimating the value of changes in environmental quality. Nonetheless, many issues remain regarding the correct specification of these models and the sensitivity of welfare estimation to specification errors.

We identified two major research objectives for this project. The first was to address several key methodological issues associated with implementation of the random utility models. The second was to incorporate in the model sufficient data about the environmental attributes of sites in the State to perform the policy analyses of interest.

Below, we outline the model and the policy analysis we perform with the model. With that background, we will then briefly highlight the methodological issues addressed in the report.

Overview of the Model

To implement the random utility framework for modeling recreational trip demand, economists have identified two levels of consumer decisions: (1) How many recreational trips does each individual take during a year or a season? and (2) What attributes do people seek for each recreational trip? The first question pertains to total demand for recreation, the macro decision. The second question pertains to the *micro* decisions associated with an individual trip.

On any given choice occasion in a sport-fishing season, anglers must decide whether or not to take a fishing trip. For participants, we model three levels of choices they make for an individual trip: fishing site, by county; fishing product line, which captures distinctions by macro-species and water-body type; and trip duration. The anglers' decision structure is shown on page 3, along with the options available and the factors hypothesized to influence each decision.

In our context of recreational fishing, the *macro* decision is the total number of fishing trips anglers take during a fishing season. Since anglers may take trips of different lengths: we model separately total demand for different trip- lengths. Consequently, we handle the third-level choice for individual trips, trip duration, within the *macro-level* participation model.

Though it is theoretically possible to model the discrete product-line/site choices

CHOICE STRUCTURE OF SPORTFISHING ANGLERS

Trip Length	Fish Product Line	Destination Site
Alternatives:		
Day	Great Lakes Coldwater	83 Michigan counties
Weekend: 2-4 days	Great Lakes Warmwater	
Vacation: 5+ days	Anadromous Runs	
	Inland (Lk+Strm) Coldwater	
	Inland Lakes Warmwater	
	Inland Streams Warmwater	
Factors influencing choice;		
Inclusive value of PL	inclusive value of sites	Travel costs
Lodging/food cost	Product line costs	Fish catch rates
Workstatus	Fishing skill/preference	Quantity of resources
Avidness of angler	Demographic attributes	Natural beauty
Household income		Accessibility
Marital status		Contamination
School vacation		

and the total participation decision jointly, the data and computational requirements for the correct treatment of the corner-solutions implied by zero trips of certain categoriest makes an integrated utility-theoretic model practically infeasible. Essentially, researchers face a trade-off: they either implement a utility-theoretic framework that does not properly model the statistics of the corner solutions; or they model the *micro* and *macro* decisions in separate models that may address the corner solution problem but do not form an integrated utility-theoretic framework.

In our analysis we estimate separate models at the *micro* and *macro* levels. We use the nested multinomial logit model (NMNL) to estimate the determinants of site and product line choices on the *micro*- level. Due to severe data limitations at the total participation level, our participation model is somewhat different from the standard treatment in the literature. We do not know the total number of season trips: our *macro* level information is limited to the duration between trips. and this variable is censored because we only observe the duration from last trip to the survey return date, not to the subsequent trip. By incorporating a key result from stochastic renewal theory in our modeling. we are able to estimate the determinants of the between-trip durations with a stochastic renewal model and then to derive the total number of trips in a season from the duration model.

Though necessitated by the data limitations we face. this approach in fact may provide several advantages. The most prominent advantage of the competing risks approach is the capacity for modeling the dependency of choices among trips of different types, which is lacking in most other empirical work with random utility models. Most researchers have limited their analysis to day trips. Another advantage is that we are able to incorporate time-varying covariates to account for changing fishing conditions over the season at individual sites.

Performing Policy Analysis

In order to perform policy analysis with the model, it is important to incorporate appropriate measures of site quality to capture the quality changes associated with the policies. Michigan identified several policies of particular interest. In the resource management area, the key concern was evaluating alternative fish stocking regimes. In the area of natural resource damage assessments, the State wanted the capability to estimate damages from power-plant related fishkills, toxic contamination at state and federal Superfund sites, fishkills from acute toxic episodes, and acid rain contamination.

In order to value these injury scenarios, the determinants of site choice in the model had to include the key measures of environmental quality that change in the scenarios, as they are experienced by anglers. The two key categories of quality change are fish catch rates (to capture the stock effects) and toxic contamination levels. We incorporated detailed information on fish catch rates from the MDNR creel survey for the Great Lakes and anadromous fisheries, and generally found the predicted positive relationships between expected catch rates and anglers' valuation of a site. Due to problems with endogeneity between participation and catch rates for the inland product lines, we were only able to use measures of lake area or stream length, broken down by quality level, for those product lines.

Unfortunately, we were not. able to use a fish consumption advisory measure to capture toxic contamination in the Great Lakes product lines Because fish consumption advisories apply to virtually all of the Great Lakes warmwater and coldwater fisheries (except a few counties with no fish, and a few counties in Lake Superior), the variable lacks the variability required for inclusion in the modeling. We used fish advisory measures for inland product lines, but there were few inland resources with advisories at the time of the angler survey, so there is limited variation in the advisory variable for those product lines also.

Toxic contamination in the Great Lakes product lines is measured by a variable indicating that (selected) water bodies in the county have been designated as part of an Area of Concern by the International Joint Commission. A noteworthy finding in the empirical analysis is that designation of a county as an Area of Concern has a substantial dampening effect on participation, an effect that spills over into water bodies and species (fishing product lines) that are not directly located in the (localized) Area of Concern within the county.

In constructing the model. we estimated how individuals value for fishing at a site varied with the fish catch rates and contamination variables. To carry out a policy analysis with the model, a resource expert must provide the "policy scenario", which specifies how the values of the environmental quality variables will change as a result of the policy.

To illustrate the capabilities of the model for performing policy analysis. we apply the model to two current contexts in which environmental injury is occurring in Michigan. First, we calculate the damages to Michigan-licensed recreational anglers from fish kills due to operation of the largest pumped- storage plant in the US. Second, we calculate the benefits of cleaning up PCB contamination in a river in Michigan, which would allow the State to remove dams currently containing contaminated sediments and to open a substantial reach of the river for anadromous runs. The contamination at this site is sufficient to merit designation of the site as an Area of Concern

Methodological Issues

We identified three key methodological issues raised in implementing the random utility model:

1. modeling total trip participation across the season, given that we have detailed information on a single trip and very limited information about total trip demand;

2. developing a consumer surplus measure that takes into account the changes in predicted number of trips due to policy changes (as well as the change in value per trip); and

3. performing sensitivity analysis of the model to alternative specifications, including alternative treatments of the opportunity costs of time.

Participation modeling

The major methodological challenge is to link a *macro*- level model of total recreational trip demand to the micro-level model of demand for fishing site and fishing product line. Our participation model represents an innovative solution to the estreme limited-data problem we faced. The analytical framework, which develops estimation procedures for a competing risk model with censored duration data and time-varying covariates, has wide applicability beyond the recreational demand contest.

By modeling demand for trips of different durations, we are able to show that two-thirds of the damages in our policy scenarios accrue to anglers taking trips of longer than one day. If we had followed the standard procedure in the literature of analyzing day trips only, we would have seriously underestimated damages.

In order to validate the participation estimates from the model, we compared the estimated trip-days derived from our model against estimated trip-days based on analysis of the MDNR creel survey. Because the procedures and criteria for counting trips and trip-days are different in the two datasets, the comparison is not suited to statistical testing. Though the differences between the surveys limit our ability to compare the estimates, we conclude that the similarity of predicted participation between the model and the annual diary data provides some evidence corroborating the participation model.

Several possible avenues exist for improving model specification. We have not explicitly addressed the “corner solution” problem, as Bockstael, Hanemann, and Strand have labelled it. We need to test to see whether non- participants should be treated differently from participants. Resolution of this issue is more complicated in our dataset than in a more typical survey, where total trips are measured for a fixed time period across all individuals. In our dataset, we observe “no trip” outcomes over very different time periods, ranging from one to fourteen months. To model "no-participation'', we must confront the question, over what length of time must a licensed angler not participate to be considered a different type of person?

Consumer Surplus Measure

Linked with the *macro* modeling issue is the correct specification of the consumer surplus measure. The standard measure employed for discrete choice models is based on the assumption that total trips do not change with policy changes. This measure will result in an under- or over-estimate of “true” consumer surplus, depending upon whether total trips increase or decrease. We develop a consumer surplus measure that incorporates the change in trips predicted by the participation model. Additional complexity is added to the measure with a nested multinomial logit model (NMNL), when the choice occasion income is not observed and the marginal utility of income is not constrained to be constant across alternatives due to the computational complexity of such a procedure. We propose a simplifying procedure that makes the calculation tractable under these circumstances.

Model Specification Issues

Finally, we analyze the sensitivity of model estimates to alternative treatments of the time constraints faced by anglers in making their trip choices. Extensive exploration in conventional (continuous demand) travel cost models has shown that consumer welfare measures are extremely sensitive to the treatment of time, though no consensus has emerged on the appropriate method for valuing time. Discrete choice models have not been subjected to comparable exploration. In this study, we develop a careful accounting of household allocation of time; the accounting highlights the fact that different treatments of the time constraints imply different choice sets of feasible sites: as well as different treatments of the opportunity costs of time in the modeling.

Outline of the Report

The report is organized as follows. Chapter II reviews the literature on random utility models of recreation demand. The emphasis is on highlighting the methodological issues associated with implementing the random utility model. Chapters III through V specify the theoretical framework for modeling the PL-site choice, for modeling total trips in a season, and for calculating the exact seasonal consumer surplus. Chapter VI is a description of the data sources. Chapters VII and VIII present estimation results of the multinomial logit and the participation models, respectively. Chapters IX and X apply the model to two natural resource damage scenarios in Michigan fisheries, one relating to fishkills and the other to toxic contamination, and calculate the loss in consumer value as a result of the injuries. In the Appendix, we report the sensitivity analysis of site choice model estimates with alternative treatments of the value of time.

CHAPTER II

DISCRETE CHOICE MODELS OF RECREATION DEMAND: A BRIEF REVIEW

The purpose of this chapter is to provide a brief overview of random utility models (RUM) of recreation demand, highlighting some key methodological issues that remain in model design and implementation. First used by Luce (1959) to model psychological choice behavior. RUM was shown by McFadden (1974, 1978) to be consistent with underlying consumer utility maximization behavior.[1]

An individual, upon deciding to take a trip on a choice occasion, is assumed to choose the site among the available alternatives that offers him/her the highest utility. The utilities that can be derived from visiting different sites are usually considered deterministic to the individuals, but stochastic to the outside investigators due to unobserved personal/site characteristics, data measurement errors, or simply random elements in human decision-making process.

By assuming *weak complementarity* which posits that a consumer will not care about marginal improvements of a commodity if he/she consumes none of it,[2] i.e.,

$$\frac{\partial u(x, \ldots)}{\partial x} = 0, \text{ if } x = 0,$$

[1] See McFadden (1976, 1961, 1982, 1984), Amemiya (1981), Hensher and Johnson (1981), or Maddala (1983) for surveys and discussions of qualitative response models.

[2] This in effect rules out the non-use value of the commodity. See Maler (1974. p. 134) or Feenberg and Mills (1980. p. 64).

the utility function; conditional on site j being chosen: of individual i can be specified as

$$u_{ij} = v_{ij}(q_j, \bar{y}_i - p_{ij})$$

where q_j is the characteristics vector of site j, p_{ij} is the cost of i travelling to site j, and $\bar{y}_i$ is the budget allocated to the trip duration in question. All the characteristics vectors pertaining to unchosen sites are excluded as a result of the weak complementarity assumption. Note that individual-specific variables can also be omitted if v_{ij} is linear in its parameters since they have the same values across all alternatives and thus will not affect the utility ranking of the feasible sites.[3]

Since the conditional utility appears stochastic to researchers: a disturbance term must be added to form the random utility

$$u_{ij} = v_{ij}(q_j, \bar{y}_i - p_{ij}, \varepsilon_{ij})$$

An individual i will then choose k among a set of feasible sites C_i if

$$u_{ik} > u_{ij}, \ \forall j \neq k, \ j \in C_i. \tag{II.1}$$

By strategically choosing a utility function u and defining the joint probability distribution for ε to make the mathematics tractable, we can calculate the probability of an individual i going to site k. given i's decision of participation:

$$\pi_{ik} = \text{Prob}\{u_{ik} > u_{ij}, \ \forall j \neq k, \ j \in C_i\}.$$

The most widely adopted multinomial response model in the literature is the multinomial logit (MNL) model.[4] because it yields a simple form of π_{ik} as well as other computational advantages. In the MNL model, the random terms ε are assumed to

[3] This is in fact the result of adopting an additively separable utility form usually assumed for estimation convenience.

[4] See Train (1986), Ben-Akiva and Lerman (1985), McFadden (1974, 1976, 1984) and Maddala (1983) for model specification.

be i.i.d. type I extreme value distributed [5] The probability of an individual i choosing site k among a collection Ω_i of sites can then be shown to be

$$\pi_{ik} = \frac{e^{u_{ik}}}{\sum_{j\in\Omega_i} e^{u_{ij}}}.$$

A restrictive feature of the MNL model is the *Independence from Irrelevant Alternatives* (IIA) property. which states that the probability ratio of two sites being chosen will stay the same regardless of the addition or deletion of other sites (or their properties).[6] This can be easily verified since the probability ratio

$$\frac{\pi_{ij}}{\pi_{ik}} = \frac{e^{u_{ij}}}{e^{u_{ik}}},$$

depends only on variables in u_{ij} and u_{ik}. Given the weak complementarity assumption. u_{ij} and u_{ik} consist solely of the quality variables of sites j and k, respectively.

While the multinomial logit models have the IIA property which is not very desirable in many situations, researchers car, circumvent this problem by using the more flexible *generalized extreme valve* (GEV) model,[7] which embodies the correlation among sites within its joint distribution structure of the error terms. The most commonly employed GEV model is the nested multinomial logit (NMNL);[8] which captures the inter-site correlation in the coefficient of the inclusive value index. Derivation of both MNL and NMNL from GEV can be found in Ben-Akiva and Lerman (1985).[9] The NMNL model is particularly useful when the number of

---

[5] Ben-Akiva and Lerman use the *Gumbel* distribution, which is a slightly more general structure than the type I extreme value distribution.

[6] See Maddala (1963, pp. 61- 62) or Amemiya (1985, p. 298). Ben-Bkiva and Lerman (1985. p. 109) point out that any model assuming the independence of all the disturbances would necessarily yield the IIA property.

[7] Introduced by McFadden (1978, 1961).

[8] Some examples of empirical NMNL studies are Carson and Hanemann (1967) and Bockscael et al. (1988)

[9] Pages 127 and 304, respectively.

alternatives is very large but the decision process itself can be properly described by a tree structure to reduce computational complexity. [10]

Like other discrete choice models, the RUM is used to explain the choice of site to visit and possibly other characteristics for a specific trip, which is referred to as the *micro* decision. As discussed below, the total number of trips taken during a season, the *macro* decision, is generally estimated by other means.

Many researchers have estimated models based on the random utility discrete choice approach to explain trip allocation decisions and to measure the welfare effects from environmental quality changes, including Hanemann (1978, 1982, 1984. 1985), Binkley and Hanemann (1978), Feenberg and Mills (1980), Caulkins (1982), Caulkins, Bishop and Bouwes (1986), Rowe, Morey, Ross and Shaw (1985), Bockstael, McConnell and Strand (1988), Morey et al. (1991, 1989). Jones et al. (1988, 1989, 1990), Parsons and Kealy (1990). Smith and Kaoru (1990), and Carson, Hanemann, Gum, and Mitchell (1987).

The multinomial logit model is attractive not only because it can avoid some of the problems of conventional travel cost methods, but also due to its computational tractability and feasibility when the number of alternatives gets large. In a recent state-of-the-art review of recreation models, Bockstael, McConnell, and Strand (1991) conclude that the random utility version of the travel cost model is particularly well-suited to valuing changes in quality at one or more recreation sites. The random utility model allows the researcher to model a wide range of substitution possibilities and, consequently, provides a procedure for estimating the value of changes in environmental quality.

Simulations have been run to show the advantages the random utility method has over other approaches. Kling (1986, 1988) uses Monte Carlo methods to generate var-

---

[10] Conditions to be met for the employment of a nested analysis are explained in Ben-Akiva and Lerman (1985, pp. 291-93) for a three- dimensional case.

ious data sets for a Stone-Geary utility function and compares the welfare estimates of different models with actually known measures. In their review, Bockstael, McConnell, and Strand conclude that Kling's "stylized simulation experiments . . . give preliminary support to the notion that discrete choice models produce better benefit estimates in problems characterized by much substitution among sites, especially when a large portion of the sample is observed to choose more than one site to visit in a season." (p. 256)

Nonetheless several fundamental methodological issues remain. Perhaps the most thorny is to integrate the *micro* and *macro* levels of the modeling, with the correct statistical treatment of the corner-solutions implied by zero trips of certain categories, (otherwise known as the 'corner-solution' problem.) We consider this issue in some detail in Chapter IV.

In this section, we discuss specification issues associated with specifying time constraints and choice sets in the random utility models. One important issue that has not been explored in the random utility context is the valuation of the opportunity costs of time. As pointed out by Bockstael et al. (1987), recreationists often cite time as more constraining than money in their recreation consumption. So the time spent on recreation consumption is, in many cases, an important determinant of the demand.

It has been recognized, since the early period of recreation demand modeling, that the omission of time costs (i.e., the opportunity costs of on-site and travel) in conventional travel cost models biases the parameter estimates and understates the final welfare measures.[11] The time-valuation literature since has focused on the context of conventional travel cost demand models.[12] In the multinomial logit models of recre-

[11] See Clawson and Knetsch (1966) or Cesario and Knetsch (1970).

[12] E.g.. Cesario (1976), Smith et al. (1983), Kealy and Bishop (1986). Bockstael et al. (1987), and McConnell (1990).

ational demand reported in the literature, the treatment of travel time apparently has varied substantially. However. our literature review revealed that authors frequently did not explain how they defined the travel cost/time variables, rarely explained how they defined a choice occasion, and never explained how an individual's choice set of feasible sites related to the time constraint for the choice occasion.

1. In the studies conducted by Bockstael et al. (1986, p. 213; 1987), all we know is that they have a "trip cost" variable. No details are provided.

2. In their MNL model of southcentral Alaska sport fishing, Carson, Hanemann, Gum, and Mitchell (1987) include only a round-trip distance cost[13] variable, computed as round-trip distance multiplied by the individua1 respondent's reported motor vehicle cost per mile. No time cost is included.

3. Morey et al. (1991. p. 4) state only that they have the "cost of a trip to site *j* mode *m* " in their model. No explanation is given as to how this variable is calculated.

4. Bockstael et al. (1988) calculate their travel cost variable as $.10 per mile plus 80 percent of the wage rate for individuals who worked for a wage and could vary their time. A separate travel time variable is used for anglers who cannot vary their work time.

5. McConnell et al. (1990) assume that anglers spend a fixed amount of time fishing at the site, whatever site is chosen. Both the distance cost and the cost of travel time thus enter the angler's site decision. For anglers who work flexible hours, the cost of travel time, valued at the wage rate, is included as part of the total travel cost. For anglers without such discretion. travel time enters the utility function directly.

6. Parsons (1990) allows the recreation period to be longer than the trip duration, and includes the individual's opportunity cost of time, distance cost, as well as other expenses in the price of taking a trip.

7. In their Wisconsin lake recreation study, Parsons and Kealy (1990) measure the travel cost as the sum of transit costs and opportunity cost of time. The transit cost is assumed to be $.10 (1978 dollars) per mile. For the time cost, they assume that all individuals stay on site for a fixed four-hour period. Each

---

[13] We use the term distance cost to refer to the cost of motor vehicle operation for the trip. The term *travel cost* is meant to be the all inclusive measure, which consists of the distance cost and the opportunity time cost of travelling

individual is then assumed to value an hour at one third of his/her wage rate for the travel time and on-site time [14]

8. The travel cost measure in Smith and Kaoru (1990) is the sum of a distance cost plus the opportunity cost of travel time. The former consists of the vehicle operating costs measured as round-trip mileage times $.20 per mile; the opportunity costs of travel is measured as the predicted wage per hour for employed respondents and the minimum wage for non-working individuals times travel time. The travel time is estimated from the round-trip mileage by assuming an average speed of 40 miles per hour.

The multinomial logit literature on recreational demand has not focused on the question of how an individual's choice set of feasible sites is defined. When is a site too far for an individual to reach on a choice occasion? What are the time constraints used for defining the choice sets? None of the papers mentioned above provide enough information to answer these questions.[15] However, as we will show, these specification choices indeed have a large impact on the MNL estimates. [16]

We know of only one study that has analyzed explicitly the sensitivity of model estimates to alternative definitions of choice sets. Smith and Kaoru (1990) consider the geographical resolution of site definitions, evaluating the specification error from increasing levels of aggregation across heterogeneous sites. They observe that site definition does have important implications for specification of the nesting structure of the model and for the benefits measurements associated with quality changes. However: they conclude that their findings provide "rather strong support for using

[14] The wage rate is calculated as annual income divided by 2080, the average number of hours worked in a year of their sample.

[15] Parsons and Kealy (1990) only mention that they include "lakes within a day's drive from an individual's home" in the individual opportunity set. In our case, though all the 83 Michigan counties form the units of our site MNL analysis. not every county is in the choice set of an angler on a certain choice occasion. This is especially true for the day anglers. Some counties are simply beyond reach for a day trip. Some counties may be within reach, but heavy driving may make trips to them infeasible. For example, it is unlikely that people are willing to drive ten hours each way to a distant site in a single day.

[16] As Smith and Kopp (1980) point out , there are spatial limits to the legitimate use of travel cost methods.

random utility models to [estimate] the effects of quality attributes on people's decision to use different recreation sites." They further note that their study "strongly reinforces the Bockstael, McConnell and Strand (1991) conclusions supporting the RUM framework even in cases where the site definition and specification of the set of alternatives is unclear." (p. 27)

CHAPTER III

INDIVIDUAL MICRO-LEVEL CHOICE MODELING

In this chapter we specify a utility-theoretic model to analyze the PL-site choices of recreational anglers. In the following chapter we present the model of the macro-level demand for total recreational trips per season.

Consumer Preferences and Behavior

Consider a consumer *i* who derives utility from two kinds of activity: consuming market goods and taking fishing trips Let Z = $(\mathbf{Z}^1, \mathbf{Z}^2, \ldots, \mathbf{Z}^T)$ denote the numeraire composite market good consumed by *i* in the *T* periods of a fishing season. The periods are determined in such a way that the individual *i* can take no more than one fishing trip in each period $t = 1, 2, \ldots, T$. In each period, individual *i* will decide whether to take a fishing trip for one of the total *M* product lines. The number of feasible sites for product line *m* is J_m which varies with product line choice and individuals. The attributes of all PL-site combinations in all periods are denoted by $\mathbf{Q} = \{q^t_{(l,j)}, \forall\, t,\ l,\ j\}$, where *t, l* and *j* are the indices for periods, PLs and sites, respectively. Also, denote the costs[1] individual *i* has to incur to fish for all possible PL-site choices *(l,j)* in all periods *t* as P $= \{p^t_{(l,j)}, \forall\, t,\ l,\ j\}$. The participation and

[1] The costs of recreational activities generally include license fees, site entrance fees, travel cost. gear purchases. etc.

PL-site decisions made by i are $\delta = \{\delta^t_{(l,j)}, \forall t, l, j\}$, where $\delta^t_{(l,j)}$ has a value of 1 if i decides to visit site j for PL l during period t, and a value of 0 otherwise. The indicator variable $\delta^t_{(l,j)}$ will be zero for all (l, j) alternatives if the individual i does not take a trip in period t. Individual i is assumed to maximize his or her utility, given annual income y. The maximization problem facing i is thus

$$\begin{aligned} \text{maximize} \qquad & U_i(Z, \delta Q) = U(Z, \delta Q, S_i) \qquad \text{(III.2)} \\ \text{subject to} \qquad & \delta P + Z = y \\ & \delta^t_{(l,j)} \cdot \delta^t_{(m,k)} = 0, \ \forall (l,j) \neq (m,k), \ \forall t \\ & \delta^t_{(l,j)} = 0 \text{ or } 1, \ \forall (l,j), \ \forall t \\ & Z \geq 0 \end{aligned}$$

where S_i is the vector of socioeconomic attributes of individual i. The first constraint is the budget constraint, while the other constraints force the corner solution in which the consumer i can only buy at most one of the quality-differentiated fishing trips. Note also that the parameter (δQ) in the utility function U_i embodies the weak complementarity assumption, which asserts that i will only obtain utility from quality attributes Q through realized trips. The indirect utility function can then be derived as $V_i = V_i(P, Q, y)$ or $V_i = V(P, Q, y, S_i)$.

Since the decision indicators $\delta^t_{(l,j)}$ can only take on integer values. a solution to the above maximization problem can only be found by first comparing the utility levels yielded by all possible trip choices over all T periods and then selecting the one that generates the highest utility. The procedure to solve this problem is described in both Kling (1986) and Bockstael et al. (1986). Since solving the problem (III.2) is computationally infeasible, simplification of the model is necessary.

A common practice is to impose further structure on the utility function. A useful and reasonable strategy is to assume that individuals adopt a *two-stage budgeting* process. Individuals, seeking to maximize their utility, are hypothesized first to optimally allocate their season budget y among all T time periods, and then to determine the actual consumption pattern in each period with the period budget y^t.[2]

An implication of the two-stage budgeting process is that the utility function is characterized by weak separability across budget categories, (such as recreation, housing, food etc), where weak separability is defined as follows:[3]

Definition III.1 For a utility function $u = u(q_1, q_2, \cdots, q_K)$ where q_k is the vector of commodities in kth category, the *weak separability* assumption requires that the utility function u be expressible as $u = f(v_1(q_1), v_2(q_2), \ldots, v_K(q_K))$, while strong (or additive) separability further implies the simpler form of $u = f(v_1(q_1) + v_2(q_2) + \cdots + v_K(q_K))$. ∎

Because our dataset (as with most datasets in recreation demand studies) has data only on the most recent fishing trip, we must further assume weak separability across choice occasions within the recreational fishing budget branch. With this restriction, an angler's ranking of possible fishing trips on a particular choice occasion does not depend on how many fishing trips of different types he/she has already taken or will take later in the season.

We further assume weak separability across site choices: the quality of a site only affects an individual's utility if the site is chosen. The recreational fishing sub-utility function is defined over the vector of market goods associated with recreation Z^t and the vector of site characteristics for the chosen site, for each choice occasion t. Therefore, i's season utility function can be written in the form

$$U_i = U\left(u(Z^1, \delta^1 Q^1, S_i), u(Z^2, \delta^2 Q^2, S_i), \ldots, u(Z^T, \delta^T Q^T, S_i) \right).$$

When the allocation of y to each period t is done, the utility maximization problem

---

[2] This assumption can be justified by observing that people frequently form a general price aggregate about market prices in the near future and allocate their long-term income to different time periods accordingly.

[3] See Deaton and Muellbauer (1983, Part 2, Chapter 5) or Morey (1984) for a thorough treatment of this topic. Discussion on two-stage budgeting can also be found in Varian (1984, pp. 146-49).

can be attacked by solving the following maximization problem for each period t

$$\begin{aligned} \text{maximize} \quad & u^t = u(Z^t,\ \delta^t Q^t,\ S_i) && \text{(III.3)} \\ \text{subject to} \quad & \delta^t P^t + Z^t = y^t \\ & \delta^t_{(l,j)} \cdot \delta^t_{(m,k)} = 0,\ \forall\,(l,j) \neq (m,k) \\ & \delta^t_{(l,j)} = 0 \text{ or } 1,\ \forall\,(l,j). \\ & Z^t \geq 0 \end{aligned}$$

By substituting the budget constraint $Z^t = y^t - \delta^t P^t$ into the utility function u^t, we can reformulate the problem as

$$\begin{aligned} \text{maximize} \quad & u = u(y^t - \delta^t P^t,\ \delta^t Q^t,\ S_i) \\ \text{subject to} \quad & \delta^t_{(l,j)} \cdot \delta^t_{(m,k)} = 0,\ \forall\,(l,j) \neq (m,k) \\ & \delta^t_{(l,j)} = 0 \text{ or } 1,\ \forall\,(l,j) \end{aligned}$$

Consider period t where a micro PL-site choice decision has to be made by individual i. Let Ω_i be i's choice set of available PL-site (l,j) alternatives. Upon choosing not to take a trip in period t, individual i will obtain the no-trip utility

$$u^t_0 = u_0(y^t,\ S_i).$$

Otherwise, if the PL-site (l,j) combination is chosen, he or she will, by weak complementarity, receive the following conditional utility

$$\begin{aligned} u^t_{(l,j)} &= u_{(l,j)}\left(Z^t,\ q^t_{(l,j)},\ S_i\right) \\ &= u_{(l,j)}\left(y^t - p^t_{(l,j)},\ q^t_{(l,j)},\ S_i\right) \end{aligned}$$

We can also define the unconditional utility function as

$$V_i(P^t,\ Q^t,\ y^t) \equiv \max\left\{u^t_{(l,j)},\ \forall\,(l,j) \in \Omega_i\right\}$$

One way to incorporate the participation decision in this framework is to compare the no-trip utility u_0^t with the unconditional utility V_i. A no-participation decision will consequently be made if

$$u_0^t > V_i(P^t, Q^t, y^t),$$

and hence $\delta^t_{(l,j)} = 0, \forall (l,j) \in \Omega_i$. On the other hand, the PL-site (m,k) will be chosen if

$$u^t_{(m,k)} = \max\left\{u^t_{(l,j)},\ \forall (l,j) \in \Omega_i\right\} > u_0^t,$$

giving us $\delta^t_{(m,k)} = 1$. and $\delta^t_{(l,j)} = 0$ for all $(l,j) \neq (m,k)$.

Since there exist some unobserved factors affecting PL-site and participation decisions, the utilities u_0^t and $u^t_{(l,j)}$ are random from the analyst's perspective. A PL-site specific disturbance is, hence, introduced into the various utility functions to form the random utility functions

$$\begin{aligned}
\tilde{u}_0^t &= u_0(y^t, S_i) + \epsilon_0 \\
\tilde{u}^t_{(l,j)} &= u_{(l,j)}\left(y^t - p^t_{(l,j)},\ q^t_{(l,j)},\ S_i\right) + \epsilon_{(l,j)} \\
\tilde{V}_i(P^t, Q^t, y^t) &= \max\{\tilde{u}^t_{(l,j)},\ \forall (l,j) \in \Omega_i\}.
\end{aligned} \tag{III.4}$$

The Micro-level Product Line/Site Decision

This section presents a model of the micro PL-site choice given that an individual i has decided to take a trip. A rational individual i will prefer PL-site combination (m,k) to (l,j) if

$$\begin{aligned}
& \tilde{u}_{(m,k)} > \tilde{u}_{(l,j)} \\
\iff\ & u_{(m,k)} + \epsilon_{(m,k)} > u_{(l,j)} + \epsilon_{(l,j)} \\
\iff\ & \epsilon_{(l,j)} < \epsilon_{(m,k)} + u_{(m,k)} - u_{(l,j)}.
\end{aligned}$$

Conditional on participation, consumer i will choose PL-site (m, k) from his or her feasible set of alternatives Ω_i if and only if

$$\tilde{u}_{(m,k)} > \tilde{u}_{(l,j)}, \ \forall (l, j) \neq (m, k), \ (l, j) \in \Omega_i,$$

or equivalently

$$\epsilon_{(l,j)} < \epsilon_{(m,k)} + u_{(m,k)} - u_{(l,j)}, \ \forall (l, j) \neq (m, k), \ (l, j) \in \Omega_i.$$

The probability $\pi_{(m,k)}$ of individual i choosing PL-site (m, k) is then

$$\begin{aligned} \pi_{(m,k)} &= \text{Prob}\left\{\epsilon_{(l,j)} < \epsilon_{(m,k)} + u_{(m,k)} - u_{(l,j)}, \ \forall (l, j) \neq (m, k), \ (l, j) \in \Omega_i\right\} \\ &= \int_{-\infty}^{\infty} \left\{ \left[\prod_{(l,j) \neq (m,k)} F(\epsilon_{(m,k)} + u_{(m,k)} - u_{(l,j)}) \right] f(\epsilon_{(m,k)}) \right\} d\epsilon_{(m,k)} \end{aligned}$$

where $F(\cdot)$ and $f(\cdot)$ are the cumulative distribution function (CDF) and probability density function (PDF), respectively, of the residuals ϵ.

What matters here is the difference $u_{(m,k)} - u_{(l,j)}$ between utility levels offered by (m, k) and (l, j), not their absolute magnitudes. Therefore, if the conditional utility u is additively separable between the choice-specific and non-choice-specific attributes, leading to the following form

$$\tilde{u} = v\left(y^i - p^t, \ q^t\right) - h(S_i) + \epsilon,$$

the non-choice-specific personal attributes will drop out of the micro choice decision because they are constant across all PL-site alternatives.

The Multinomial Logit Model

If the residuals ϵ are independently and identically distributed with *type I extreme value* distribution for which the CDF is

$$F(\epsilon) = \exp(-e^{-\epsilon})$$

and PDF is

$$f(\epsilon) = \exp(-\epsilon - e^{-\epsilon}).$$

then it can be shown that

$$\pi_{(m,k)} = \frac{e^{u_{(m,k)}}}{\sum_{(l,j)\in\Omega_t} e^{u_{(l,j)}}} = \frac{e^{u_{(m,k)}}}{\sum_{l=1}^{M}\sum_{j=1}^{J_l} e^{u_{(l,j)}}}$$

which is the multinomial logit model.[4]

The type I extreme value distribution is in fact a special case of the Gumbel distribution[5] that has the CDF

$$F(\epsilon) = \exp\left(-\exp\left[-\mu\,(\epsilon-\eta)\right]\right),\ \mu > 0$$

and PDF

$$f(\epsilon) = \mu\ \exp\left[-\mu(\epsilon-\eta) - \exp(-\mu\,(\epsilon-\eta))\right]$$

where η is a location parameter and μ is a scale parameter. The type I extreme value distribution simply assumes that $\eta = 0$ and $\mu = 1$. The Gumbel distributed residuals ϵ all have the same mean

$$\eta + \frac{\gamma}{\mu}$$

and variance

$$\frac{\pi^2}{6\mu^2}$$

where $\gamma(\approx 0.5772)$ is the Euler constant, and result in the probability of PL m-site k alternative:

$$\pi_{(m,k)} = \frac{\exp(\mu\, u_{(m,k)})}{\sum_{(l,j)\in\Omega_t} \exp(\mu\, u_{(l,j)})}.$$

Since the parameter μ is not econometrically identifiable, it is common practice to set it arbitrarily to 1, yielding the same probabilities as the type I extreme value distribution. As pointed out by Ben-Akiva and Lerman (1985, p. 104), the assumption

[4] It is called conditional logit by McFadden (1974).

[5] See Ben-Akiva and Lerman (1985, pp. 104-107) for a discussion.

of a constant η for all alternatives is not restrictive as long as each systematic utility has a constant term. Though the Gumbel distribution is used for analytic convenience, its choice can be defended as an approximation to the normal density.

Note that the probability (III.5) can also be expressed as the product of a conditional probability and a marginal probability

$$\pi_{(m,k)} = \pi_{k|m} \cdot \pi_m \tag{III.6}$$

where

$$\pi_{k m} = \frac{e^{u_{(m,k)}}}{e^{I_m}} \quad \text{and} \quad \pi_m = \frac{e^{I_m}}{\sum_{l=1}^{M} e^{I_l}} \tag{III.7}$$

and

$$I_m \equiv \log\left(\sum_{j=1}^{J_m} e^{u_{(m,j)}}\right).$$

With the Gumbel assumption, it can be shown that

$$E\left[\max_j\{\tilde{u}_{(m,j)}\}\right] = \log\left(\sum_{j=1}^{J_m} e^{u_{(m,j)}}\right) + \gamma = I_m + \gamma.$$

Hence the inclusive value index I_m reflects weighted information about the alternatives in PL m and is a measure of the expected maximum utility one can get from choosing PL m.[6]

We assume that the systematic part u of the random utility $\tilde{u}$ can be separated into the part that varies only with PLs and the part that varies with both PLs and sites as follows:

$$u_{(m,k)} = \gamma Z_m + \beta X_{(m,k)}.$$

In this case the probabilities (III.7) become

$$\pi_{k|m} = \frac{e^{\beta X_{(m,k)}}}{\sum_j e^{\beta X_{(m,j)}}} \quad \text{and} \quad \pi_m = \frac{e^{\gamma Z_m + I_m}}{\sum_{l=1}^{M} e^{\gamma Z_l + I_l}}.$$

---

[6] A discussion of the Gumbel properties is in Ben-Akiva and Lerman (1985, p. 105).

Now the inclusive value index for product line m becomes:

$$I_m = \log\left(\sum_{j=1}^{J_m} e^{\beta X_{(m,j)}}\right)$$

where the "inclusive value” is the expected utility of an individual for the site-specific attributes X, net of the integrating constant γ.

Estimation can thus be carried out by sequentially applying MNL to each PL m, calculating I_m for all m PLs, and then calculating $\pi_{(m,k)}$ using formula (III.6). Obtaining the maximum likelihood estimates of a multinomial logit model in general poses no computational difficulty since it has been proved by McFadden (1974) that the log likelihood function

$$LL = \sum_{i=1}^{N} \log P^i_{(m,k)} = \sum_{i=1}^{N}\left(u_{i,(m,k)} - \log \sum_{(l,j)\in\Omega_i} e^{u_{(l,j)}}\right)$$

is globally concave under relatively weak conditions. The Newton-Raphson algorithm will therefore always converge within finite steps, often in just a few iterations, to a unique solution.

The way a simple MNL models the PL-site decision is to treat each PL-site combination as a feasible choice. Given that we have M product lines and J_m potential sites for each product line m (= 1, 2, . . . , M), the total number of alternatives one faces is $\mathcal{J} = \sum_{m=1}^{M} J_m$, as illustrated in figure III.1. A restrictive feature of this modeling approach is the aforementioned *Independence from Irrelevant Alternatives.* It is highly implausible that the odds ratio $\frac{\pi_{(m,k)}}{\pi_{(l,j)}}$ of any two PL-site choices (m,k) and (l,j) will be independent of the conditions of other available alternatives, as implied by the MNL specification where

$$\frac{\pi_{(m,k)}}{\pi_{(l,j)}} = \frac{\exp(u_{(m,k)})}{\exp(u_{(l,j)})}.$$

Consider a situation where an individual i can only choose between site *A* for PL 1 and site *B* for PL 2. With the addition of a site *C* for PL 2 that has exactly

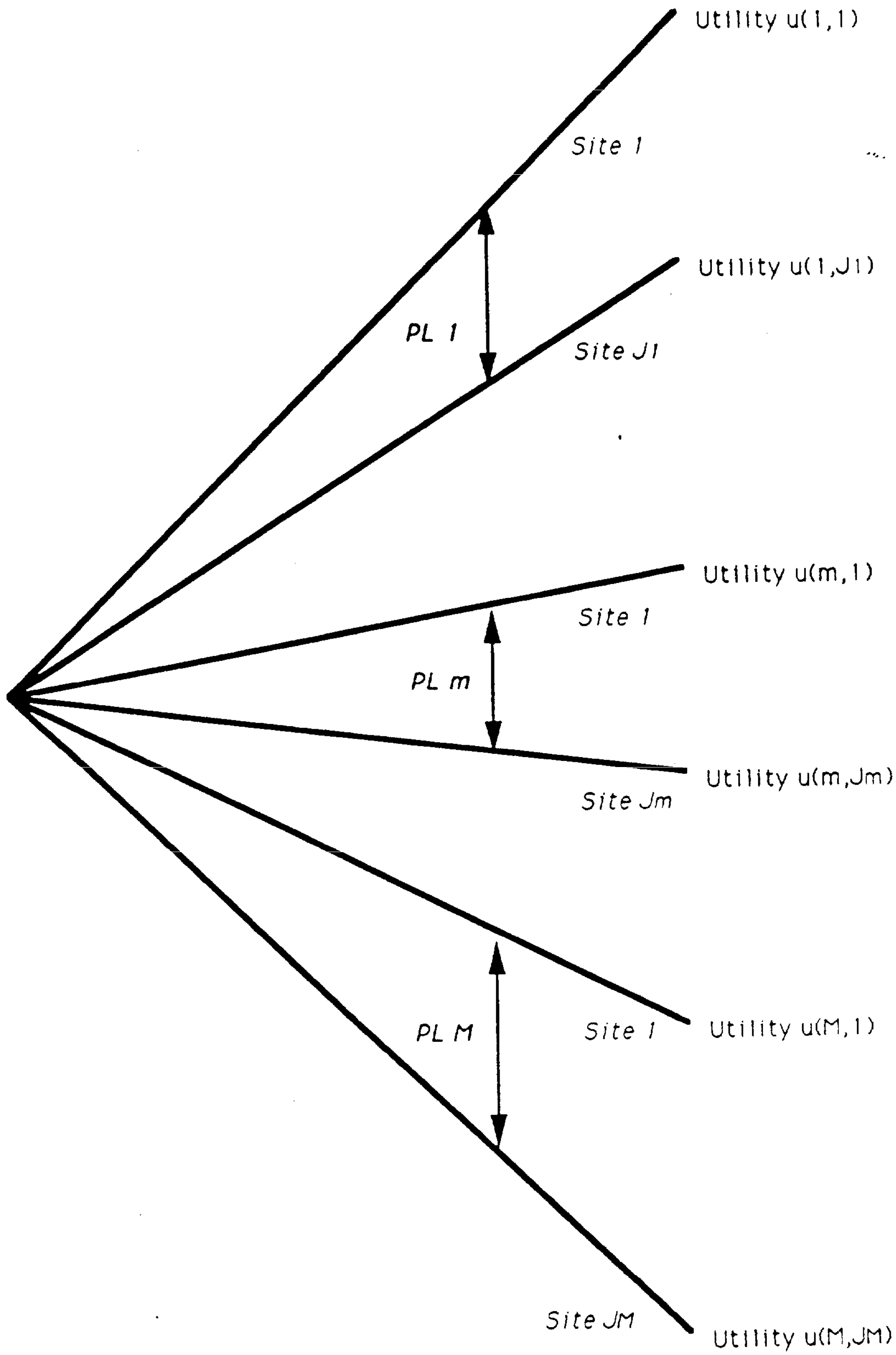

Figure III.1: The flat micro PL-site decision structure.

the same attributes [7] as site *B*, the probability $\pi_{(2,B)}$ would probably be only half its original level, while $\pi_{(1,A)}$ will, most likely, not change. This is just an analog of the famous red-bus/blue-bus problem in the transportation literature. Therefore, where there are obvious differences in patterns of substitution and complementarity across alternatives, the IIA assumption, and hence the MNL, is not appropriate.

Nested Multinomial Logit Model

To avoid the IIA restriction, the *nested multinomial logit* (NMNL) model is better suited for our study. Individuals are hypothesized to adopt a two-level tree-like decision process on any choice occasion. They first determine the target product line, and then choose a site conditional upon the product line decision. This is illustrated in figure III.2. The result is that the IIA property is imposed on sites within a product line, but not across product lines.

Assume that the random utility $\bar{u}$ *(mj),* an individual can receive from first choosing PL *l* arid then site *j* is

$$\tilde{u}_{(m,j)} = \alpha Z_m - \beta_m X_{(m,j)} + \epsilon_{(m,j)}$$

where the attribute vector *X(m,j)* and random terms $\epsilon_{(m,j)}$ are specific to the PL-site choice *(m,j),* while variables in vector Z_m vary only with PLs. The PL characteristics Z_m are shared by all sites available to PL *m.* Also assume that the random terms ϵ follow the generalized extreme value (GEV) distribution defined below.

Definition III.2 The generalized extreme value distribution is defined as

$$F(\epsilon_1, \epsilon_2, \ldots, \epsilon_N) = \exp\left[-G(e^{-\epsilon_1}, e^{-\epsilon_2}, \ldots, e^{-\epsilon_N})\right]$$

where $G(y_1, y_2, \ldots, y_N)$ satisfies the following conditions

1. G is a nonnegative function of $y_i \geq 0$, $i = 1, 2, \ldots, N$..

[7] And consequently C is a perfect substitute for B.

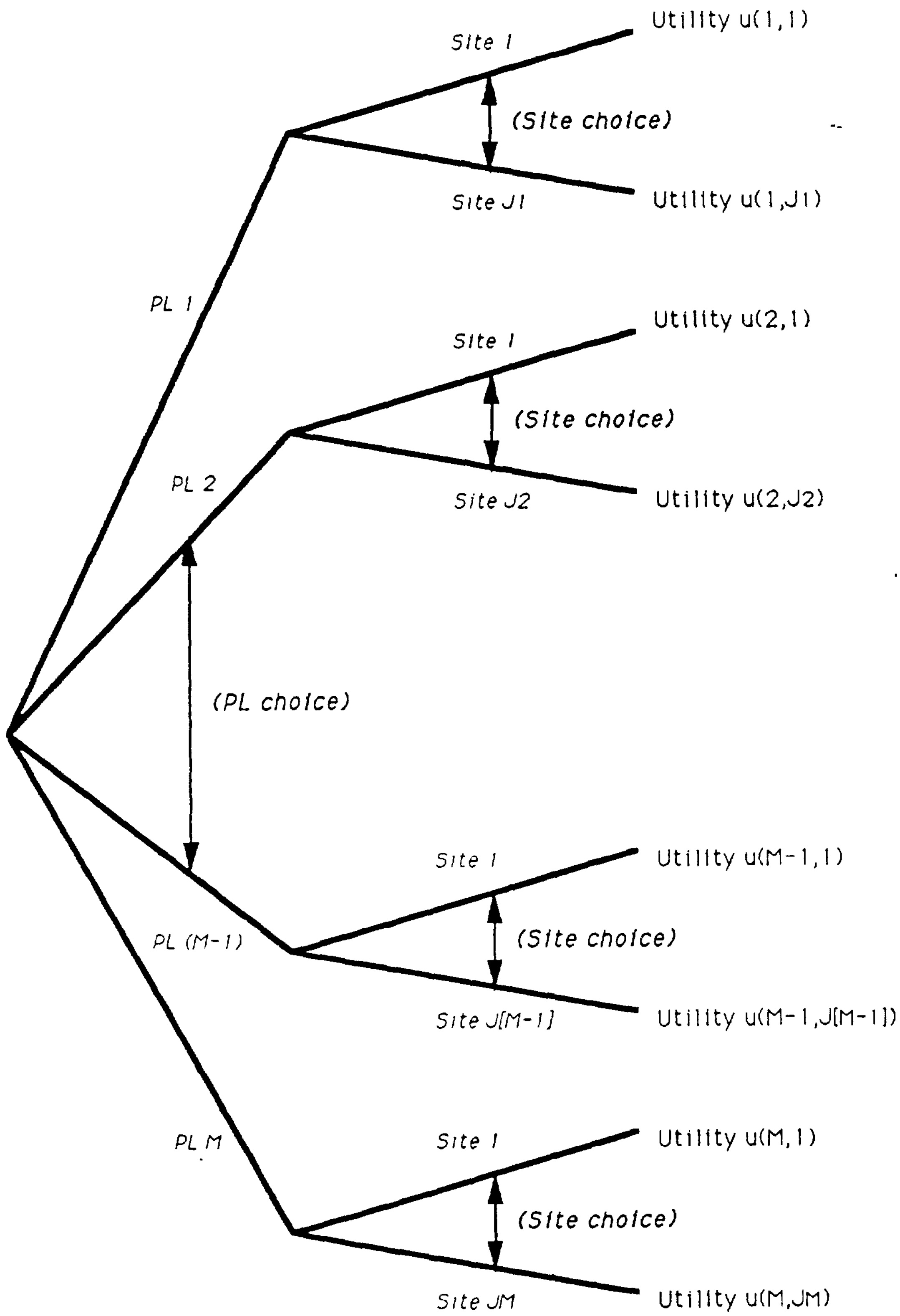

Figure III.2: The two-stage micro PL-site decision structure.

2. G is homogeneous of degree $\mu > 0$
3. $\lim_{y_i \to \infty} G(y_1, y_2, \ldots, y_N) = +\infty$ for $i = 1, 2, \ldots, N$
4. The sth derivative of G with respect to any combination of s distinct y_i's, $i = 1, 2, \ldots, N$, is non-negative if s is odd, and non-positive if s is even. ∎

McFadden (1978) proves the GEV distribution implies that the probabilistic-choice model consistent with utility maximization gives choice probability of the form

$$\pi_i = \frac{e^{u_i} G_i(e^{u_1}, e^{u_2}, \ldots, e^{u_N})}{\mu\, G(e^{u_1}, e^{u_2}, \ldots, e^{u_N})}$$

where G_i is the first derivative of G with respect to y_i.

Now assume that the function G is homogeneous of degree 1 and has the form

$$G(y_{(m,j)},\ \forall (m,j)) = \sum_m \left[\sum_j \left(y_{(m,j)} \right)^{1/\theta} \right]^{\theta}.$$

Therefore, the disturbances ϵ have the joint distribution

$$F(\epsilon_{(m,j)},\ \forall (m,j) \in \Omega_i) = \exp\left(-\sum_{m=1}^{M} \left[\sum_{j=1}^{J_m} \exp\left(\frac{-\epsilon_{(m,j)}}{\theta} \right) \right]^{\theta} \right). \tag{III.8}$$

The probability of (m,k) being chosen is then

$$\pi_{(m,k)} = \frac{\exp(\beta_m X_{(m,k)}/\theta)}{\sum_{j=1}^{J_m} \exp(\beta_m X_{(m,j)}/\theta)} \cdot \frac{\exp(\alpha Z_m + \theta I_m)}{\sum_{l=1}^{M} \exp(\alpha Z_l + \theta I_l)} \tag{III.9}$$

$$= \frac{\exp(\beta_m X_{(m,k)}/\theta)}{\exp(I_m)} \cdot \frac{\exp(\alpha Z_m + \theta I_m)}{\sum_{l=1}^{M} \exp(\alpha Z_l + \theta I_l)} \tag{III.10}$$

$$= \pi_{k|m} \cdot \pi_m \tag{III.11}$$

where

$$I_m \equiv \log\left(\sum_{j=1}^{J_m} \exp(\beta_m X_{(m,j)}/\theta) \right) = E\left[\max_j \{\tilde{u}_{j|m}\} \right] - \text{constant} \tag{III.12}$$

is the inclusive value of the sites in PL m. Note that

$$\frac{\pi_{(m,k1)}}{\pi_{(m,k2)}} = \frac{\exp(\beta_m X_{(m,k1)}/\theta)}{\exp(\beta_m X_{(m,k2)}/\theta)}.$$

Consequently, the IIA property continues to hold for sites in the same product line, but not across product lines.

The inclusive value I_m is an index of the overall quality of fishing opportunities of the sites in PL m, or the expected maximum utility the site, of PL m can offer, excluding the utility one can get from PL attributes Z that do not vary across sites. Similarly we can calculate the inclusive value

$$I^* = \log\left(\sum_{m=1}^{M} e^{\alpha Z_m + \theta I_m}\right), \tag{III.13}$$

as an index for the desirability of participation in recreation. The value I^* represents the expected utility of taking a fishing trip[8] and will be used in the participation decision modeling.

Estimates of the parameters can be obtained by employing a two-step procedure: First, the estimates for $\beta'_m\ (\equiv \beta_m/\theta)$ are obtained by repeatedly applying MNL to each product line m. The inclusive values I_m can then be calculated and used, along with the PL-specific variables Z_m, in the second-stage MNL estimation of α and θ. The original parameters β can then be recovered as $\beta_m = \theta\,\beta'_m$. Note that in the simple logit setup (III.7). the parameter θ is exogenously set to 1, thus excluding the case where different PLs have inherently different utility effects.

The way the NMNL avoids the IIA property is to allow a general pattern of dependence among the choices. This is embodied by the GEV distribution assumption, as opposed to the independent residuals assumption of the simple logit model. This can be more intuitively seen from an alternative derivation of the NMNL by Ben-Akiva and Lerman (1985, pp. 285-91) or Cardell and Steinberg (1988).

They start by splitting the PL-site alternative residuals into two parts and assuming that

$$\tilde{u}_{(m,k)} = u_m + u_{(m,k)} + \epsilon_m + \theta\,\epsilon_{(m,k)}$$

where ϵ_m is the random component attributable to the product line m and common to all sites in PL group m. This generates a correlated structure for the errors across

[8] Net of the constant terms from both the site and product line levels of analysis.

alternatives in the same PL group When $\epsilon_{(m,k)}$ is Weibull distributed, it can be shown that there exist some $\theta \in [0, 1]$ and a random variable ϵ_m, independent of $\epsilon_{(m,k)}$, such that $[\epsilon_m + \theta \epsilon_{(m,k)}]$ is also Weibull distributed. The probability statements derived from this specification are exactly identical to (III.9) and (III.11).

The parameter θ is called the *dissimilarity index* because it indicates the share of the common components in the error variance. The smaller θ is, the more similar the sites under PL m are.[9] Therefore, as θ approaches 0, the nested MNL becomes more appropriate. The flat MNL can only be justified when θ is close to 1.

McFadden (1978) shows that a sufficient condition for a NMNL model to be consistent with random utility maximization is that the coefficient θ of the inclusive value I lies in the [0, 1] unit interval. An estimated θ outside the unit interval range hence raises questions about a potential mis-specification of the model.

Estimation of the Sequential Multinomial Logit

Due to the nonlinearity and complexity of the model (III.9), estimation by maximum likelihood is practically infeasible. Instead, the sequential estimation method described above is employed.

One complication arising from this stepwise procedure needs to be addressed. Because the inclusive value index variables used in all stages above the lowest level are in fact estimated from the lower stages, not actually observed variables, the covariance matrices calculated by MNL will be biased and have to be corrected.[10]

A more serious problem with the sequential estimation method that can not be

---

[9] Many authors use $\rho = 1 - \theta$ instead of θ as a measure to indicate the correlation among alternatives in the same group. For example, Maddala (1983), Bockstael et al. (1986, 1988), and Greene (1989).

[10] The process of correcting the estimated covariance matrix to generate consistently estimated standard errors is described in the Appendix in McFadden (1981). Schmalensee and Joskow (1986) also discuss techniques of using estimated parameters as independent variables.

so easily corrected is the loss of efficiency in the estimation process. Note that since parameters β and θ appear in both $\pi_{k|m}$ and π_m of the probability statement (III.11), the full information maximum likelihood (FIML) estimates can only be obtained by taking derivatives of the complete log likelihood function

$$LL = \sum_{i=1}^{N} \log \pi^i_{(m,k)} = \sum_{i=1}^{N} \log \pi^i_{k|m} + \sum_{i=1}^{N} \log \pi^i_m$$

with respect to the parameters β, θ and α and setting the first derivatives to zero.

Alternatively, the multi-stage procedure estimates $\beta' = \beta/\theta$ in the first stage by maximizing only

$$LL_1 = \sum_{i=1}^{N} \log \pi^i_{k|m},$$

and then estimates α and θ in the second stage by maximizing

$$LL_2 = \sum_{i=1}^{N} \log \pi^i_m$$

These estimates are thus only limited information maximum likelihood. (LIML) estimates since not all available information in the data is utilized.

The Nested Multinomial Logit Specification

We adopt a linear utility function for this study, which implies there are no income effects from quality changes. Consequently, the compensating variation and equivalent variations for a quality change will be equal.[11] The utility an individual i receives from choosing PL m and site k is assumed to be

$$\tilde{u}^i_{(m,k)} = \alpha_i + [\xi_m + \gamma Z_m] + [\lambda_{(m,k)} + \beta_m X_{(m,k)}] + \epsilon_{(m,k)} \qquad \text{(III.14)}$$

---

[11] Most empirical MNL studies. from the early mathematical psychology work of Luce (1959) to recent recreation demand studies of Bockstael et al. (1988) and Morey et al. (1988), adopt a linear form for the conditional utility function. Feenberg and Mills (1980) argue that, though linearity can hardly be literally true, it “may be a good approximation if the utility function is smooth and the sample variances of the parameters are not too large” (p. 112). The most important reason for our adopting a linear utility function, however, is the ease of consumer surplus computation. When the conditional utility is nonlinear formula for the consumer surplus per choice occasion is very hard to obtain.

where α_i is the individual constant, ξ_m is the PL constant, $\lambda_{(m,k)}$ is the PL-site constant, Z_m is the characteristics vector that varies only with PLs, and $X_{(m,k)}$ are the attributes specific to both PLs and sites. The random elements ϵ are assumed to follow the GEV distribution defined by (III.8).

Given the PL choice m, individual i will select a site k that offers the highest "site" utility

$$\tilde{v}_{(m,k)} = \lambda_{(m,k)} + \beta_m X_{(m,k)} + \epsilon_{(m,k)}.$$

The variables X we use in the estimation include site quality Q and the choice occasion income net of travel costs to site $k\,(Y_i - P_{ik})$, which is available to spend on consumption of market goods, where P_{ik} is the travel cost of an angler i visiting site k. Therefore, the conditional utility function $\tilde{v}_{(m,k)}$ becomes

$$\tilde{v}_{(m,k)} = \lambda_{(m,k)} - \eta\,[Y_i - P_{ik}] + \beta_m Q_{(m,k)} + \epsilon_{(m,k)}.$$

As explained above, we cannot obtain estimates for λ, η and β at this stage; instead we get only $\lambda' = \lambda/\theta, \eta' = \eta/\theta$, and $\beta'_m = \beta_m/\theta$ for each PL m using the sample of individuals who we observed choosing PL m.

In our analysis (as is generally the case) we do not have data on choice occasion income. These missing data are not a problem in the site-choice level of analysis, because the income is constant across sites and so drops out in the estimation (which employs differences between the conditional utility functions.) However, when the marginal utility of income is not constant across alternatives in a nested MNL model, the lack of choice occasion income will affect the higher-level estimation – in our analysis, the choice of product line – and the welfare calculations.

Consequently, we derive in some detail below the value of the lower-level inclusive value index for sites in product line m, I_m, and of the higher-level inclusive value index across product lines. The goal is to explicitly identify the role of the choice occasion income variable. All inclusive value calculations are performed separately for each trip duration category.

The inclusive value I_m for sites in product line m, as defined in (III.12), is

$$
\begin{aligned}
I_m &= \log\left(\sum_j \exp(v_{(m,j)}/\theta)\right) \\
&= \log\left(\sum_j \exp(\lambda'_{(m,j)} + \eta'_m[Y_i - P_{ij}] + \beta'_m Q_{(m,j)})\right) \\
&= \log\left(\sum_j \left[\exp(\eta'_m Y_i)\ \exp(\lambda'_{(m,j)} - \eta'_m P_{ij} + \beta'_m Q_{(m,j)})\right]\right) \\
&= \log\left(\exp(\eta'_m, Y_i) \sum_j \exp\left[\lambda'_{(m,j)} - \eta'_m P_{ij} + \beta'_m Q_{(m,j)}\right]\right) \\
&= \eta'_m Y_i + \log\left(\sum_j \exp\left[\lambda'_{(m,j)} - \eta'_m P_{ij} + \beta'_m Q_{(m,j)}\right]\right) \\
&\equiv \eta'_m Y_i + \log\left(\sum_j \exp(v'_{(m,j)})\right) \\
&\equiv \eta'_m Y_i + \bar{I}_m
\end{aligned}
$$

The estimated parameter η of the travel cost variable P_{ij} is individual i's constant marginal utility of income for product line m. Because of missing data on choice occasion income Y_i, we can only calculate the *pseudo*-inclusive value $\bar{I}_m$ from the estimates λ', η'_m and β'_m.

In the upper-level PL-choice modeling in the NMNL we estimate the parameters of the PL conditional indirect utility function

$$
\begin{aligned}
v^i_m &= \alpha_i + \xi_m + \theta I_m + \gamma Z_m \\
&= \alpha_i + \xi_m + \theta[\eta'_m Y_i + \bar{I}_m] + \gamma Z_m \\
&= \alpha_i + \xi_m + \eta_m Y_i + \theta \bar{I}_m + \gamma Z_m
\end{aligned}
$$

where I_m is the inclusive value index calculated above from the lower-level site-choice MNL estimation. As defined by (III.13), the inclusive value of taking a trip is

$$
I_i^* = \log\left(\sum_m \exp(v_m)\right)
$$

$$= \log\left(\sum_m \exp(\alpha_i + \xi_m + \theta I_m + \gamma Z_m)\right)$$

If $\eta_m = \eta$ for all m, then we can further simplify the formula:

$$\begin{aligned}
I_i^* &= \log\left(\sum_m \left[\exp(\alpha_i + \eta Y_i^*)\exp(\xi_m + \theta \bar{I}_m + \gamma Z_m)\right]\right) \\
&= \log\left(\exp(\alpha_i + \eta Y_i^*)\sum_m \exp(\xi_m + \theta \bar{I}_m + \gamma Z_m)\right) \\
&= \alpha_i + \eta Y_i^* + \log\left(\sum_m \exp(\xi_m + \theta \bar{I}_m + \gamma Z_m)\right) \\
&= \alpha_i + \eta Y_i^* + \log\left(\sum_m \exp(v'_m)\right) \\
&\equiv \alpha_i + \eta Y_i^* + \bar{I}_i^* \qquad \text{(III.15)}
\end{aligned}$$

The individual-specific constants α_i are not identifiable, and as stated above, we do not know the choice occasion income Y_i^*. Therefore, we cannot calculate the real level I_i^*, and so will use the pseudo-inclusive value $\bar{I}_i^*$ in chapter V for the calculation of consumer surplus.

Because $\alpha_i + \eta Y_i^*$ is constant across product lines (assuming $\eta_m = \eta$), estimation with $\bar{I}_i^*$ is equivalent to estimation with I_i^*. In Chapter V, we further show that if the marginal utility of income is constant across alternatives, the lack of choice occasion income does not pose problems for the welfare analysis. However, if MUI is *not* constant across product lines, then we cannot calculate I_i^* as the three separable components in (III.15): the choice occasion income term remains an integral component of the calculation. As a consequence, estimation using $\bar{I}_i^*$ in place of I_i^* yields a mis-specification.

In our NMNL estimation procedure below, we do not impose the constraint of constant MUI across product lines within a duration group,[12] due to the computational

[12] We would not expect constant MUI across duration groups, but this issue does not pose any difficulties because we model trip-duration choice within the participation model.

complexity it would cause, though such a restriction seems conceptually appropriate. We have developed a procedure for handling the problems posed by not constraining the MUI to be constant. In the process of specifying the correct consumer surplus formulas for the case of varying marginal utility of income, we derive in Chapter V a “weighted” marginal utility of income, where the weights represent the ex-post probabilities of choosing the alternative. The weighted MUI serves the same role in the consumer surplus formula that the (constant) MUI serves in the simpler context. We will substitute the weighted MUI in the calculation of I_i^*, in lieu of performing the NMNL estimation with cross-estimation constraints. Employing this conceptual framework, use of $\bar{I}_i^*$ in the estimation procedure is not a mis-specification.

The Valuation of Time

The Analysis Framework

A critical component of the conditional indirect utility function for site alternatives specified above in equation III.14 is the travel cost P_{ik} per trip by individual i to site k. Conceptually. travel cost consists of two components – distance costs and time costs. As discussed above in Chapter II the time cost component is controversial. We derive in some detail several models which support several different treatments of the opportunity costs of trip time in the discrete choice literature. We show how the models imply not only different measures of travel cost but also different definitions of the choice set of feasible sites.

Anglers are assumed to maximize utility subject to a full-income constraint for the choice occasion, where full income refers to the money budget plus the value of time, following the household production function literature. Note that, among the sites within the angler’s choice set, we only observe the amount of time the individual allocated for a trip to the *chosen* site. We must make assumptions about how much time an individual would allocate for trips to other sites.

The first two frameworks are based on an assumption that *total trip time* will be the same for all sites (as for the chosen site); we label this the "exogenous total trip time" framework. The two variants we develop employ alternative measures of trip time. The third framework is based on the assumption that *on-site time* will be the same for all sites (as for the chosen site); we label this the "exogenous on-site-time" framework. We show below in the Appendix the potentially large effect these differences may have on model results.

Variable Definitions

We first define the following variables for our discussion.

X = market goods (set price=1 as numeriare

D = number of days in trip

w = post-tax wage rate

c = vehicle operating cost per mile

s = driving speed (miles per hour)

q_j = quality of site j, where $j = 1, \ldots, J$

D_j = round-trip distance in miles to site j

$P_j = c\,D_j$ = round trip travel cost to site j

$R_j = D_j/s$ = round trip travel time in hours to site j

S_j = time spent on site j in hours

$T_j = R_j + S_j$ = total trip time to site j in hours

C = choice occasion time in hours

y = money allocated to the recreational choice occasion

Y^* = full income allocated to the choice occasion

Exogenous Trip Time

Within the exogenous trip-time framework, we impose the assumption that an individual allocates all of the choice occasion time C either to visiting a site j, thereby incurring round-trip travel time R_j plus on- site time S_j, or to other activities (working, other recreation). We can write the generalized time budget, conditional upon participation, for the choice occasion:

$$\sum_j \delta_j (R_j + S_j) = \sum_j \delta_j T_j = C.$$

As defined previously, the indicator $\delta_j = 1$ if site j is chosen for the visit, and $\delta_j = 0$, otherwise. If $\delta_k = 1$, then $\delta_j = 0$ for all $j \neq k$.

The *sources* of full income Y^* include wC, time during the choice occasion valued at the post-tax wage rate w,[13] and y, the money income allocated for expenditure during the period:

$$Y^* = wC + y \tag{III.16}$$

The *uses* of full income during the period are expenditures during the trip on market

---

[13] Recognizing that recreational activities take up time, much of the recreation demand literature relates the opportunity time costs to the wage income forgone when a trip is taken. However, the labor supply literature now recognizes that work, time may not be a continuous choice variable over which individuals can freely trade-off income and recreation at the wage rate. Only individuals with flexible work hours can adjust their marginal rate of substitution between work and recreation and make it equal to their marginal wage rates. These individuals are said to be at interior solutions. Others, who either have to work fixed hours or do not work at all, are at corner solutions, and their wage rates cannot serve as the value of their leisure time.

Some authors in the recreation demand literature have also adopted this view and have treated interior solutions and corner solutions differently. For people at interior solutions, work time is at their discretion, and trip time can be traded for income at their marginal wage rate. For others without such freedom, no opportunity, wage cost exists since they cannot increase their work effort even if no trip is taken.

Unfortunately no wage rate was directly measured in our data, so the above treatment is impossible. The budget frontier therefore has to be assumed a straight line, and people are assumed to be at interior solutions. We assume that people value their time at their wage rate, calculated as annual personal income divided by projected working hours per year.

goods X plus distance costs and time costs to the chosen site:

$$Y = X + \sum_j \delta_j [cD_j + wR_j + wS_j]$$

Setting *sources* equal to *uses,* we have

$$wC + y = X + \sum_j \delta_j [P_j + wR_j + wS_j]. \tag{III.17}$$

Conditional upon participation in recreation at site j (i.e., $\delta_j = 1$), we can solve for $X = y - P_j$. The indirect utility function conditional upon participation at site j is

$$V_j = V(y - P_j, q_j).$$

Assuming a linear functional form, the conditional indirect utility function becomes

$$V_j = \beta_1 [y - P_j] + \beta_2 q_j.$$

The important point to note is that time costs have completely dropped out of the travel cost measure for site choice, because the amount of time allocated to "producing" recreation (the choice occasion) equals the trip duration. The use of the standard travel cost variable incorporating both time and distance costs cannot be supported in this framework. [14] With a conditional direct utility function of the form

---

[14] Besides monetary vehicle costs P_j, the amount of driving could conceivably have at least two other effects on an angler's utility: a reduction in available fishing time S_j and the (dis)utility of driving R_j itself. A more elaborate and complete utility function will, therefore, be

$$\begin{aligned} V_j &= V(y - P_j, R_j, S_j, q_j) \\ &= V(y - cD_j, R_j, T_j - R_j, q_j) \end{aligned}$$

The linear estimating function is then

$$\begin{aligned} V_j &= -\beta_1 [cD_j] + \beta_2 R_j + \beta_3 [T_j - R_j] + \beta_4 q_j \\ &= -\beta_1 [cD_j] + \beta_2 [D_j / s] + \beta_3 [T_j - D_j / s] + \beta_4 q_j \\ &= \left(-\beta_1 c + \frac{\beta_2 - \beta_3}{s} \right) D_j + \beta_3 T_j + \beta_4 q_j \\ &= \left(-\beta_1 + \frac{\beta_2 - \beta_3}{cs} \right) P_j + \beta_3 T_j + \beta_4 q_j \end{aligned}$$

$V_j(X, q_j)$, the time costs affect angler decision-making only at the higher level where the participation choices for each trip duration are made.

We suggest two alternative methods for implementing the exogenous trip duration model: the first measures trip duration in hours, based on the self-reported hours (and days) for the beginning and end of the trip; the second measures trip duration based on the number of days the individual reported being away. Though the conditional indirect utility specification is the same, the definition of the choice set for each individual, the consistency checks for selecting individuals into the sample, and the value of the time cost variable in the participation model differ.

Hypothesis 1: Exogenous Trip Time (Using Self-Reported Trip Hours)

In this case, total trip time *T* is based on self-reported trip duration. The amount of income allocated to the choice occasion is calculated as (III.16). but has no practical significance in the site choice analysis.

Let constant *h* be the presumed number of hours people are awake and active during a day. Its complement h' $(\equiv 24 - h)$ is then the time in hours people rest each day. For a trip of *D* days, people necessarily must rest for $(D - 1)$ nights, a total of $(D - 1)\,h'$ hours, at their fishing site or on the road. Therefore. we impose two time constraints for a *D-day* trip to any potential site *j:*

$$\textbf{A1}: T_j \leq h^* \equiv D\,h + (D-1)\,h'$$

$$\textbf{A2}: \quad S_j = T_j - R_j \geq (D-1)\,h'$$

The real marginal utility of income $\eta = \beta_1$ cannot be identified. The marginal utility of income measure we obtain without including the R_j and S_j terms in V_j is $\eta' = \beta_1 - \frac{\beta_2 - \beta_3}{cs}$, which is actually a combination of the various driving effects. Anticipating negative utility from driving and reduced fishing time, we have $\beta_2 < 0$ and $\beta_3 > 0$. Therefore,

$$\eta' > \eta > 0.$$

The resulting consumer surplus we derive using η' will consequently be an under-estimate of its real value.

Constraint A1 asserts that the total trip duration cannot exceed the limit of *h'* hours, which is the sum of active and resting time during the day. Constraint A2 enforces that people still have time left after accounting for driving and resting to fish and enjoy site amenities. Combined: they imply that $R_j \leq D\,h$. In words; this posits that round-trip driving time (R_j) cannot take up all the time people are awake during the trip.

Any site *j* that violates either A1 or A2 is considered infeasible for a visit on the choice occasion in question. The time constraints A1 and A2 thus define the individual choice set of feasible sites. People whose reported destination sites of the observed trips violate these time constraints are thus treated as outliers and deleted from the MNL sample.[15] The value selected for *h* hence plays a critical role in the site choice analysis.[16] When a larger *h* is used: more people will be included in the sample, and more sites will be included in each individual's choice set.

To be included in the MNL sample: each angler's destination site *k* must satisfy time constraints A1 and A2 specified above. That is,

$$\mathbf{S_1}: \begin{cases} T_k & \leq h^* \\ S_k = T_k - R_k & \geq (D-1)\,h' \end{cases}$$

For an individual in the sample S_1 who chooses site *k* (and hence has a pre-determined total trip duration T_k). site *j* will be included in his or her choice set if and only if

$$\mathbf{C_1}: \begin{cases} T_j \equiv T_k & \leq h^* \\ S_j = T_k - R_j & \geq (D-1)\,h' \end{cases}$$

---

[15] In the actual implementation of the model, an individual can be excluded from entering the MNL sample due to three reasons: (1) the trip duration data T is missing, (2) the observed chosen site itself violates the time constraint A1 and/or A2, or (3) the chosen site is the only feasible site, so no other site is contained in his or her choice set. People who are left out from the site MNL estimation for the third reason may be included in the upper level PL MNL estimation because the inclusive value can still be calculated even though there is only one feasible site for the chosen PL.

[16] The choice set definition in turn has direct effects on the MNL and total trip estimation, as well as the final welfare calculation.

The satisfaction of the first constraint of C_1 is guaranteed since it is exactly the same as the first restriction of S_1 for inclusion in the sample. The restrictions of S_1 are applied to the selected sites, while the restrictions of C_1 are applied to all other sites.

Hypothesis 2: Exogenous Trip Time (Using Total Trip-Days)

In this case, we eschew using self-reported measures of total trip time and alternatively impose the assumption that the total trip duration T equals the maximum number of trip hours allowed in a *D-day* trip, h^*. Therefore, $T = h' = Dh - (D - 1)\, h'$. This procedure eliminates the measurement error and missing data problems posed by the first procedure: but incorporates probably greater measurement error in the trip time (cost) variable by imposing the assumption that each trip uses all waking hours of the day.[17]

Under this hypothesis, the conditions A1 and A2 combined are equivalent to

$$\mathbf{S_2} : \; R_k \leq D\,h.$$

The choice set of an individual can also be computed by including in it any site j that satisfies

$$\mathbf{C_2} : \; R_j \leq D\,h.$$

Hypothesis 3: Exogenous On-Site Time

The final hypothesis imposes an alternative assumption that on-site time is fixed across site choices: rather than trip duration. To ensure an exogenously defined choice occasion, we must include in the model the possibility of 'slack time' during the choice occasion Even using this device. there is some question for day trips as to whether

[17] We can calculate the difference from the site-hours variable, but we do not know how much measurement error there is in site-hours.

one of the conditions for welfare analysis necessarily holds: that only one trip could be taken during the choice occasion[18] We rewrite the time-budget from above to acknowledge that choice occasion time C and trip time T are no longer assumed to be the same duration:

$$\sum_j [R_j + S_j + \theta_j] = \sum_j [T_j + \theta_j] = C,$$

where θ_j is slack, and $C = h^* = Dh - (D - 1)\, h'$.

The solution to equation (III.17) in which we set sources equal to uses is:

$$X = [y - P_j] + w[C - R_j - S_j].$$

By assumption, y, C, and S_j are constant across sites, and so are not relevant to the site choice decision-making. The measure of travel cost, however, is $(-P_j - wR_j)$: the time cost of travel is included along with the distance cost! unlike for the models above. in which only the distance cost $(-P_j)$ is included.

The conditions an individual must meet to get into sample S_3 are the same constraints A1 and A2 that define S_1 above. Therefore, $S_3 = S_1$. For an individual i in the sample S_3 whose actual destination is site k, the constant on-site time is calculated to be

$$S^* \equiv S_k = T_k - R_k.$$

The choice set C_3 of i is defined by including any site j that satisfies

$$\mathbf{C_3} : \begin{cases} T_j = S_k + R_j & \leq h^* \\ S_j \equiv S_k & \geq (D - 1)\, h' \end{cases}$$

Now the second constraint of C_3 is always satisfied.

Though the two samples of anglers S_1 and S_3, defined respectively for the exogenous trip duration and the exogenous on-site time hypotheses, are identical, there is

[18] This hypothesis is adopted by McConnell et al. (1990) and Parsons and Kealy (1990). The latter assumes that all anglers spend a fixed amount of four hours on site.

no relationship between the choice sets C_1 and C_3 defined for an individual under the alternative hypotheses. [19]

It is also obvious that

$$S_1 \subseteq S_2 \text{ and } S_3 \subseteq S_2$$

since for S_2 (1) the deletions necessitated by S_1 and S_3 due to to missing data on self-reported trip-lengths are avoided, and (2) the trip durations in S_2 are assumed to be their maximum value h^*. It can further be shown that, for people in both samples S_1 and S_2,

$$C_1 \subseteq C_2$$

and, for people in both samples S_3 and S_2,

$$C_3 \subseteq C_2.$$

We will show the sensitivity of the model to the different trip time hypotheses in the Appendix.

---

[19] It is easy to show this with examples. Suppose h = 15 and D = 1. For an angler with T_k = 5 and R_k = 2, a site j with R_j = 6 will be in set C_3, whereas not in C_1. But if T_k = 14. R_k = 1 and R, = 4, constraints for C_3 will be violated, while those for C_1 are satisfied. The first example is a short day trip (e.g., an afternoon trip) to a nearby site. In this case faraway sites might be included in C_3 and excluded from C_1. The second example is a long day trip (e.g.. a whole-day fishing excursion) to a nearby site. In this case farther sites will probably be left out from C_3, but still incorporated in C_1. Conceivably this will mainly happen to people taking day trips, as most sites will be available to all anglers taking long trips.