AN ABSTRACT OF THE THESIS OF

Kenneth Charles Gibbs for the Doctor of Philosophy
(Name) (Degree)

in Agricultural Economics presented on May 7, 1969
(Major) (Date)

Title: The Estimation of Recreational Benefits Resulting from an Improvement of Water Quality in Upper Klamath Lake: An Application of a Method for Evaluating the Demand for Outdoor Recreation.

Abstract approved: Signature redacted for privacy.
H. H. Stoevener

The objectives of this thesis include the development and testing of a methodology that would be appropriate to determine the economic benefits from a recreational facility which is to be developed, determining the relationship between water quality and recreational use, and determining the economic benefits accruing to society from a postulated improvement in water quality.

The value of recreation can be computed with the use of a recreational demand curve. The estimation of the value of outdoor recreation has proceeded in two general directions: the "direct" and the "indirect" method to determine consumer's willingness to pay.

This thesis presents a new form of the indirect approach in the evaluation of recreational resources, which avoids some limiting assumptions of the previous methods. The focal point of the theory

was the individual recreationist. Travel costs were treated independent of variable on-site costs. A demand model for the number of days a recreationist will stay at a particular site per visit was derived. The statistical model forthcoming from the theoretical framework was:

$$q_1 = q_1 [(k^* - k), (p_1^* - p_1)] \qquad \text{for} \quad (k^* - k), (p_1^* - p_1) \geq 0$$

$$k^* = k^* (p_1, y, Sw, Ws, B, F, C, Si)$$

$$p_1^* = p_1^* (k, y, Sw, Ws, B, F, C, Si)$$

where q_1 is the number of days a recreationist spends at a site per visit, k is the travel cost, k^* refers to the critical, or maximum, travel cost, p_1 is the variable on-site cost, p_1^* reflects the recreationist's willingness to pay on-site costs, and Sw, Ws, B, F, C, Si represent the recreational characteristics of a site.

An equation was developed to express the relationship between the number of visits at a site, and travel cost, income, and site characteristics:

$$V = V(k, y, Sw, Ws, B, F, C, Si)$$

where V refers to the number of visits from a county, and the remaining variables are as previously defined. This relationship was used to derive the aggregate demand function.

The study was based principally on Upper Klamath Lake, but three other lakes--Lake of the Woods, Odell, and Willow--were involved. All lakes are located in the southwestern section of Oregon. A sample of recreationists was chosen and personally interviewed to obtain information concerning the length of stay, the income, and a detailed account of the recreationist's expenditures.

The equations, for the four lakes, were estimated as:

$$k^* = -36.711 + \underset{(2.322)}{6.248\,W} + \underset{(7.800)}{3.779\,F} + \underset{(.0002)}{.0003\,Si} + \underset{(.0018)}{.0020\,y} + \underset{(3.349)}{10.435\,p_1} \qquad R^2 = .616$$

$$p_1^* = -7.263 + \underset{(.197)}{7.80\,W} + \underset{(.815)}{2.630\,F} + \underset{(.000025)}{.000067\,Si} - \underset{(.002)}{.004\,k^2} + \underset{(.143)}{.269\,k} \qquad R^2 = .684$$

$$V = -71{,}166.121 + \underset{(460.199)}{7{,}141.764\,W} + \underset{(1{,}651.425)}{19{,}825.384\,F} + \underset{(.059)}{.641\,Si} - \underset{(115.473)}{379.786\,k} \qquad R^2 = .868$$

$$\ln q_1 = .759 - \underset{(.0018)}{.0064}\,(k^* - k) + \underset{(.0189)}{.0637}\,(p_1^* - p_1) \qquad R^2 = .113$$

where the site characteristics are represented by the W, F, and Si variables (the use-intensities of water skiing, swimming, boating, camping, and fishing, and the size of the lake).

In order to analyze a particular lake's demand model, it is necessary to hold the site characteristic variables constant in the k^*, p_1^*, and V equations. By increasing the value of the site characteristic variables, to represent an improvement in water quality, the incremental value of the lake can be determined.

Estimates of the economic values of all four lakes were derived. Two water quality changes were postulated for Upper Klamath Lake. Estimates of the economic benefits accruing to society from these hypothetical improvements were also made.

The Estimation of Recreational Benefits Resulting from an Improvement of Water Quality in Upper Klamath Lake: An Application of a Method for Evaluating the Demand for Outdoor Recreation

by

Kenneth Charles Gibbs

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

June 1969

APPROVED:

Signature redacted for privacy.

Associate Professor of Agricultural Economics

in charge of major

Signature redacted for privacy.

Head of Department of Agricultural Economics

Signature redacted for privacy.

Dean of Graduate School

Date thesis is presented May 7, 1969

Typed by Nancy Kerley for Kenneth Charles Gibbs

ACKNOWLEDGMENT

The author is indebted to many persons for advice and reassurance in the preparation of this manuscript. In particular, debts of gratitude are due to:

Dr. Herbert H. Stoevener, as my major professor, for the many hours of consultation devoted to this work;

My wife, Jacquie, for her help in typing, for the mental stamina she developed listening to all of my ideas, and for her faithful support during my years in graduate school;

Dr. John A. Edwards, without whom the theoretical framework in this thesis would not have existed;

Federal Water Pollution Control Administration, Project number WP-01477, entitled "Economic Benefits from Improvements in Water Quality," for financial assistance;

Dr. G. Burton Wood, and the Department of Agricultural Economics without whose help and understanding the opportunity to attend graduate school would not have been available;

The staff members, and my fellow graduate students for their enlightening ideas;

F. W. P. C. A. Graduate Training Grant number STI-WP-71 entitled "Economics of Water Quality and Supply," for financial assistance throughout my graduate work.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

THE ESTIMATION OF RECREATIONAL BENEFITS RESULTING FROM AN IMPROVEMENT OF WATER QUALITY IN UPPER KLAMATH LAKE: AN APPLICATION OF A METHOD FOR EVALUATING THE DEMAND FOR OUTDOOR RECREATION

CHAPTER I

INTRODUCTION

Statement of the Problem

Growth of biological organisms in Upper Klamath Lake, located in Southern Oregon, has resulted in levels of water quality which restricts the use of this water resource for recreational purposes. Extensive research has been carried out by the staff at the Pacific Northwest Laboratory of the Federal Water Pollution Control Administration, on the physical and biological aspects of the eutrophication process of the Lake.

In general, as a lake ages it undergoes changes and a natural maturation process takes place. Precipitation and natural drainage contribute nutrients which support and facilitate the growth of vegetation within a lake. The extensive activities of man, however, can increase the amounts of nutrients deposited in a lake in several ways: by a more intensive use of the agricultural land; by urbanization; and by the discharges of industrial wastes, and waste treatment plant effluents. The process of enrichment of waters with nutrients, referred to as eutrophication, that occurs naturally is often accelerated by man's activities. The resulting quality of the water may thus change significantly and often at a relatively rapid pace. It has not been established, however, that man's activities have been the cause of the water quality problem in Upper Klamath Lake. Eutrophication can have beneficial effects, as in Rohlick's words (1968, p. 7):

> An obvious benefit of eutrophication is the increase in the biomass which can be supported in a body of water. Additional nutrients usually result in increased growth of microscopic organisms and a consequent increase in fish production. In terrestrial situations, the increase in the yield of a crop after fertilization is desirable; . . .

The effects of eutrophication, when in conflict with man's alternative uses of a body of water, are often undesirable.

> Aesthetic values may be lowered because of increase in algal growth and production of floating algal scums which are a nuisance to those who wish to use the water for recreational purposes. . . . Because of the development of anaerobic conditions in the hypolimnion water quality is impaired (Rohlich, 1968, p. 7).

The research, being done by the Pacific Northwest Laboratory, is directed towards finding technically feasible methods of improving water quality in Klamath Lake. Such a method, when found, will have to stand the test of economic feasibility in order to be implemented. This thesis will seek to estimate the economic benefits resulting from such a possible improvement in water quality. Such benefits would arise largely from increased recreational use of the Lake. The likely recreational benefits were evaluated from the viewpoint of developing one of the Nation's natural resources.

Objectives

This thesis was directed toward three principal objectives: (1) To develop and empirically test a methodology that would be appropriate to determine the economic benefits accruing to society from a recreational facility which is to be developed. (2) To determine the relationship between water quality and recreational use, using the new methodology developed. A prediction of the change in recreational use with a substantial improvement in water quality, in a large body of water, will then be feasible. (3) To determine the economic benefits accruing to society, in general, from the postulated improvement in water quality and the associated increase in recreational use.

Considerable scientific effort is being devoted to the understanding of the eutrophication process of Upper Klamath Lake. Aside from the scientific value of this work, it is directed toward finding techniques which would permit alteration of the biological processes in such a way as to result in a substantial improvement in water quality. Ultimately, however, the question of the economics of such a water quality improvement will have to be considered. Given the scarcity of available resources for water quality improvement, it is imperative that they be devoted to projects where the payoff, in terms of benefits, is greatest. The economic evaluation of water pollution control is often difficult, especially if, as in the case of Upper Klamath Lake, the benefits are in the nature of "extramarket goods," such as outdoor recreation.

The research in this thesis is directed toward evaluating the economic benefits resulting from increased utilization of water resources for outdoor recreation. This is important for two reasons: First, it provides a guideline for decision-makers concerned with the allocation of public funds for water quality improvement, in the special case of Upper Klamath Lake. Secondly, it is anticipated that the methodology developed in this study will be useful in the evaluation of recreational benefits resulting from water quality improvements in other cases. In regard to the latter point, it should be noted that some recent developments in economic analysis have provided for the estimation of the demand for outdoor recreation. These, however, have been restricted by their assumptions and the theoretical aspects. The theory of recreational demand needs to be developed further to permit an application to a more diversified range of problems. Further, none of the recent studies has dealt with the problem of deriving the demand relation for a recreational facility which is to be developed. This type of information could be very useful to decision-makers.

In order to achieve the aim of this study it was necessary to consider the problem in three major sections. The first was to examine the theoretical concepts involved with the estimation of demand relationships for outdoor recreation. In Chapter II the previous work done in the area of economic analysis of outdoor recreation will briefly be reviewed. Chapter III is devoted to the development of a theoretical model to determine the relationships involved in a recreational demand function.

The second section draws upon the theoretical model, to develop the statistical model for recreational demand. Chapters IV and V will be devoted to this purpose. Measurement problems and a discussion of the variables in the theoretical model are included in Chapter IV. In Chapter V the selection of the sample is discussed, which includes the procedures for determining the sample size and drawing the sample. The study area will also be defined.

In the third section information from the second section was used to estimate the magnitude of the direct economic benefits accruing to society from an improvement in water quality in Upper Klamath Lake. This section consists of Chapters VI and VII of this thesis. The empirical results in general, and the statistical problems encountered, are discussed in Chapter VI. In Chapter VII the empirical demand model for each of the four lakes, and the estimated economic value of an improvement in water quality is derived.

In the remaining chapter, Chapter VIII, the main categories of this study, and the conclusions are summarized. Also discussed in Chapter VIII are limitations of the study and recommendations for future research in this area.

CHAPTER II

METHODS TO DETERMINE ECONOMIC VALUE OF OUTDOOR RECREATION

Public decision-making in the management of land and water resources is concerned with the allocation of public funds toward multiple-use natural resources. Proper management of our public natural resources requires the satisfaction of several social objectives. One such objective that decision-makers need to be concerned with is economic efficiency. That is, efficient management requires some knowledge of the possible values of all the alternative uses of the natural resources. This does not necessarily imply that the economic values of the alternative uses should be the only source of information to the policy-makers. It does imply, however, that if economic efficiency is deemed an important goal, decision-makers need to know the value of each alternative use of a resource. Outdoor recreation is one of the prominent uses of the land and water resources today.

Of the various competing uses for natural resources the evaluation of the economic benefits accruing to outdoor recreation is especially difficult to quantify. Unlike any of the other uses of natural resources, an adequate market has not been developed for outdoor recreation: there may be two general explanations for this phenomenon.

First, due to the technical production and consumption relationships an efficient market cannot be employed. Second, even if an efficient market could be developed for outdoor recreation, the American public would perhaps not be in favor of it. Traditionally the out-of-doors has been a "free" commodity.[1]

It is thus important to be able to quantify the value of outdoor recreation without an existing market. The value of recreation can be computed with the use of a recreational demand curve. The estimation of the value of outdoor recreation has proceeded in two general directions: both are directed toward determining the amount recreationists are willing to pay for the privilege of being able to use a certain facility. One direction is the "direct" method of estimating the consumer's willingness to pay (Knetsch and Davis, 1965). The recreationists are asked, by means of a personal interview, to state how much they would be willing to pay for the use of the recreational facility rather than be excluded. The demand estimates obtained in this fashion are defensible on theoretical grounds, but the real crux lies in the measurement problem. The degree of reliability placed on the respondent's answers is the real issue to be determined. Many biases are likely to exist especially when questions are asked

[1]See Stoevener and Brown, 1967; also McKean, 1968, for a comprehensive discussion of this problem.

dealing with matters of opinion concerning a person's activity that has been regarded as "free." One type of bias is the consumers understatement of his preference for a commodity in the hopes that if a charge were made he would still be able to enjoy the activity at the present level of use, and not have to pay as much as he would be willing. This type of bias can be expected since the recreationist can see uniformed park officials at most national park facilities, as well as at many other sites, thus visualizing the power to exclude them.

Another source of bias, the counterpart of the possibility that the recreationist understates his willingness to pay, is the chance that he may overstate his willingness to pay. He may do so in the belief that the area might then have a greater chance to be improved and preserved as a recreational site.

Biases are likely to be significant and this can affect the results by reducing the accuracy and thus the reliability of predictions. In the estimation of recreational benefits high accuracy is extremely important since any results will need to be projected to future dates and policy recommendations will be based on these projections.

The direct method of estimating recreational demand is not restricted to estimating only the effective demand for outdoor recreation. That is, it can be concerned with those persons now enjoying the services from outdoor recreation, and also persons that may decide to do so at a later date. The option demand may be evaluated by this procedure. However, the procedure has the disadvantage of it being difficult to make inferences about the action of recreationists due to the hypothetical nature of the questions posed.

The second type of methodology that has developed to determine the amount recreationists are willing to pay for the use of recreational facilities has been by observing the reaction of recreationists to changes in costs of travel to the recreational site. Willingness to pay can be computed from this "indirect" evidence. The indirect procedure, while not posing hypothetical questions to recreationists, is limited to the extent inferences may be made to whole populations of people consisting of recreationists and non-recreationists. This procedure is restricted to those persons now enjoying the services of outdoor recreation and not those that may do so in the future. The effective demand can be analyzed, but not the option demand, by using this measurement technique. It is not, however, involved with hypothetical questions since the actions of the individuals are observed rather than posing questions as to what actions would be, given certain circumstances.

Hotelling, in a letter to the U. S. National Park Service, 1949, is credited with the original idea of defining concentric zones around the recreational site in such a way that the cost of traveling to the site from a given zone would be approximately constant. The idea was then to use the travel cost that existed within each zone as the price variable to be compared to the number of visitors from each zone, to obtain a demand function for recreation.

Clawson (1959) recognized problems in Hotelling's formulation but used the basic idea underlying his approach to give it an interpretation that further facilitates the measurement of recreational values. Clawson envisioned deriving two demand relations. The first demand relation is expressed as a relationship between the level of travel cost and the rate of participation, derived in the manner suggested by Hotelling. From the first demand relation he derives a second demand schedule. The rate of participation from a given population group, for a certain fee increase, can be predicted by referring to the observed participation rate of another population with travel costs equal to the travel cost of the group in question, plus the fee increase. A demand schedule for a population group can then be obtained by relating various fee increases to the resulting participation rates, which then need to be applied to the number in the population to get an estimate of the quantity of use from that population zone. This procedure can be applied to each zone, and the resulting demand schedules added horizontally to obtain the aggregate demand schedule for the recreational resources.[2]

[2] For a more complete presentation of this procedure, see Clawson (1959).

Brown et al. (1964) have expanded upon the Clawson model to include family income as an explanatory variable. They also expanded upon the concept of travel cost to include not only the cost of driving to and from the recreational site, but also the costs of food, lodging, and any other costs the recreationist incurs in transit. They refer to these costs as transfer costs.

Recently Stevens (1966) further refined the Clawson model by discussing and incorporating the quality of the recreational experience by using the angling success per unit of angling effort as an independent variable. Work was also done on the inclusion of distance as a separate explanatory variable, but was found to be highly correlated with transfer cost due to the manner in which transfer cost was defined.

Many important and restrictive assumptions are involved in the indirect approach discussed up to this point. First, it is implied that the reaction of recreationists to a fee would be identical to an equal addition in the cash cost of travel. No distinction is made between a recreationist's reaction to variable cost increases (daily costs while at the recreational site) and fixed cost increases (costs of travel). Secondly, in predicting different groups' response to costs by observing other groups, they assume that all of the groups, stratified by distance, face identical alternatives to the recreational resource in question. Thirdly, recreationists in all distance zones will react identically to increases in the cost of travel. This forces the same slope on each of the groups' demand schedule, the only difference being in the level of the costs and quantity of participation. Fourthly, the assumption is made that the population groups are homogeneous with respect to the recreationists' characteristics and preferences, i. e., their utility functions are identical. Finally, it assumes (since it cannot be proven in any way) that the first demand relation is for the "entire recreational experience," including anticipation, travel to the site, experience on the site, travel from the site, and recollection; while the second demand schedule, derived from the first, is only for the recreational site "itself" excluding the other phases of the experience.

This thesis presents a new form of the indirect approach to the evaluation of recreational resources which avoids some of the limiting assumptions of the previous methods.[3] In the next chapter the theoretical development of this procedure will be examined with the focal point being the individual recreationist, instead of a population group. This avoids the necessary assumptions about the characteristics and homogeneity of the populations used to derive the second demand

---

[3]It should be noted that independent of the construction of the theory found in the next chapter of this thesis, Peter Pearse (1968) developed a very similar approach. His approach differs slightly in the technical development and his application is entirely different.

curve. Aggregation of the data within each distance zone, and merely using averages create statistical problems. High coefficients of determination (R^2) can be obtained by this procedure, since most of the variation is eliminated. The use of averages, as observations, indicates the assumption that little variation exists within each group, the only variation that is being tested is that between groups. This assumption has little validity. When only a few observations are used for regression (for example, 5) the degrees of freedom are substantially reduced. The assumptions made in the new approach, being presented in this thesis, are not different from those Clawson made in deriving his first demand relationship, except that of aggregation. Both procedures need to make assumptions about the preferences and characteristics of the recreationists. Clawson's approach derives a second demand curve which upon doing makes many limiting assumptions, as was enumerated above. The procedure developed herein derives only one demand function and divides the price variable, used by Clawson, into two components.

CHAPTER III

A NEW CONCEPTUAL FRAMEWORK TO ESTIMATE RECREATION DEMAND

In this chapter a conceptual framework for analyzing the economic behavior of recreationists will be discussed.[4] In order to analyze the economic aspects of recreationists it is necessary to have a model that will account for the number of visitor-days[5] recreationists will take at various expenditures. This chapter will deal with a model to analyze the number of days a recreationist will stay at a particular site per visit. The approach is of the indirect nature in the evaluation of recreational benefits.

---

[4]The basic idea behind the theory conceptualized in this Chapter came from an unpublished paper by Dr. John A. Edwards, Department of Agricultural Economics, Oregon State University.

[5]A recreation visitor-day consists of 12 visitor-hours, which may be aggregated continuously, intermittently, or simultaneously by one or more persons. The visitor-hours contained therein must be spent by persons in any activities, except those which are a part of or incidental to the pursuit of a gainful occupation.

A recreation visit, on the other hand, is the entry of any person upon a site, or area of land or water, generally recognized as an element in the recreation population. Visits must be made in order to engage in any activities, except those which are a part of, or incidental to, the pursuit of a gainful occupation. These are the definitions adopted by the U. S. Forest Service in their Forest Service Handbook 2309. 11.

Theory of Consumer Behavior

The rational consumer strives to maximize his utility, U; a function of the amounts of the products purchased, say q_1 and q_2:

$$U = U(q_1, q_2) \qquad \text{for} \qquad q_i \geq 0, \ i = 1, 2$$

subject to the fixed budget constraint

$$y = p_1 q_1 + p_2 q_2 \qquad y, p_i, q_i \geq 0, \ i = 1, 2$$

where y represents the consumers' fixed income available for the purchasing of the two commodities, and p_i represents the unit price of the i^{th} commodity.

That is to say, the consumer will maximize the Lagrangian function

$$V = U(q_1, q_2) + \lambda \ (y - p_1 q_1 - p_2 q_2) \tag{1}$$

where λ is the Lagrangian multiplier. The conditions for maximization of utility are fulfilled when the partial derivatives of (1) with respect to q_1, q_2, and λ are set equal to zero

$$\partial V/\partial q_1 = \partial U/\partial q_1 - \lambda \, p_1 = 0 \tag{2}$$

$$\partial V/\partial q_2 = \partial U/\partial q_2 - \lambda \, p_2 = 0 \tag{3}$$

$$\partial V/\partial \lambda = y - p_1 q_1 - p_2 q_2 = 0 \tag{4}$$

Solving equations (2) and (3) simultaneously, the necessary condition for constrained maximization of utility can be obtained as

$$\frac{\partial U}{\partial q_1} \Big/ \frac{\partial U}{\partial q_2} = p_1/p_2 \tag{5}$$

The consumer will allocate his time and income up to the point where the ratio of his marginal utilities associated with consuming Q_1 and Q_2 just equals the ratio of their respective prices.[6]

Divergence in Theory of Consumer Behavior

Consider a slight divergence in the basic micro-approach to consumer demand. In the presentation above it was assumed that in order to purchase, or consume, a unit of the i^{th} commodity the consumer needed to pay the price, p_i. Now consider the case where the purchaser needs to pay a certain charge, or cost, referred to as k, in order to be able to buy any units of Q_i. The value of k does not depend on the amount of Q_i purchased.

Recreation is a good example where k is relevant. In order to enjoy any amount of recreation, Q_1, the recreationist must incur a certain price per day, p_1, while recreating, but in addition he must travel to the recreational site. The travel costs, including the transportation cost, food, lodging, camping fees, etc. that occur while enroute to and from the recreational site will be referred to as k.

The recreationist will allocate his income in order to

---

[6]For the purposes of this thesis the second-order condition will be assumed to be satisfied. This will be indicated by the convexity of the indifference curves.

maximize his utility, U, which is a function of the quantity of the commodities purchased:

$$U = U(q_1, q_2) \qquad q_i \geq 0, \; i = 1, 2$$

where q_1 indicates the number of recreation-days the recreationist enjoys at a particular site per visit, and q_2 represents all other goods and services the recreationist could purchase with his income. The recreationist, as previously discussed, is limited to a fixed budget. The budget constraint maintains that the income allocated to the consumption of the two commodities, y_o, must just equal the total amount spent for the recreation-days commodity, $p_1q_1 + k$, plus the total expenditures for the nonrecreation-days commodity, p_2q_2, i.e.:

$$y_o = p_1q_1 + k + p_2q_2 \qquad y_o, p_i, q_i \geq 0, i = 1, 2 \qquad k > 0$$ [7]

where p_i is the unit price of the i^{th} commodity.

Thus the recreationist will maximize the Lagrangian function

$$V = U(q_1, q_2) + \lambda\ (y_o - p_1q_1 - k - p_2q_2) \tag{6}$$

by solving the partial derivatives of (6) simultaneously to obtain the first-order condition for maximization of constrained utility:

$$\frac{\partial U}{\partial q_1} \Big/ \frac{\partial U}{\partial q_2} = p_1/p_2$$

The consumer will consume recreation and non-recreation up to the

---

[7] The case when $k = 0$ will be treated separately on the next page of this thesis.

point where the ratio of the marginal utilities associated with recreation and non-recreation just equal the ratio of their respective prices.

The budget constraint, or the iso-income line, can be written as

$$y_o - k = p_1q_1 + p_2q_2 \quad . \tag{8}$$

This illustrates better the role played by k, in reducing the income available to the consumer for purchasing the commodities in his budget. It is further assumed that k can be zero if and only if no recreation is consumed, i.e.,

$$k = 0 \Longleftrightarrow q_1 = 0$$

This assumption points out a unique characteristic of the budget constraint, namely that if $k = 0$ the constraint takes on the new form of

$$y_o = p_2q_2 \qquad \text{for} \quad y_o, p_2, q_2 \geq 0 \qquad k = 0 \tag{9}$$

If faced with the budget constraint in (9) the consumer would maximize his utility by taking as many non-recreation units as his income would allow, i.e., he would consume $q_2 = y_o/p_2$ units of Q_2 and no units of Q_1.

Variations in Travel Cost (k)

The indifference map and budget constraints of a typical consumer[8] are presented in Figure 1. The two prices, p_1 and p_2 and

[8]The conclusions drawn in this analysis will hold for any system of indifference curves in which both commodities have positive

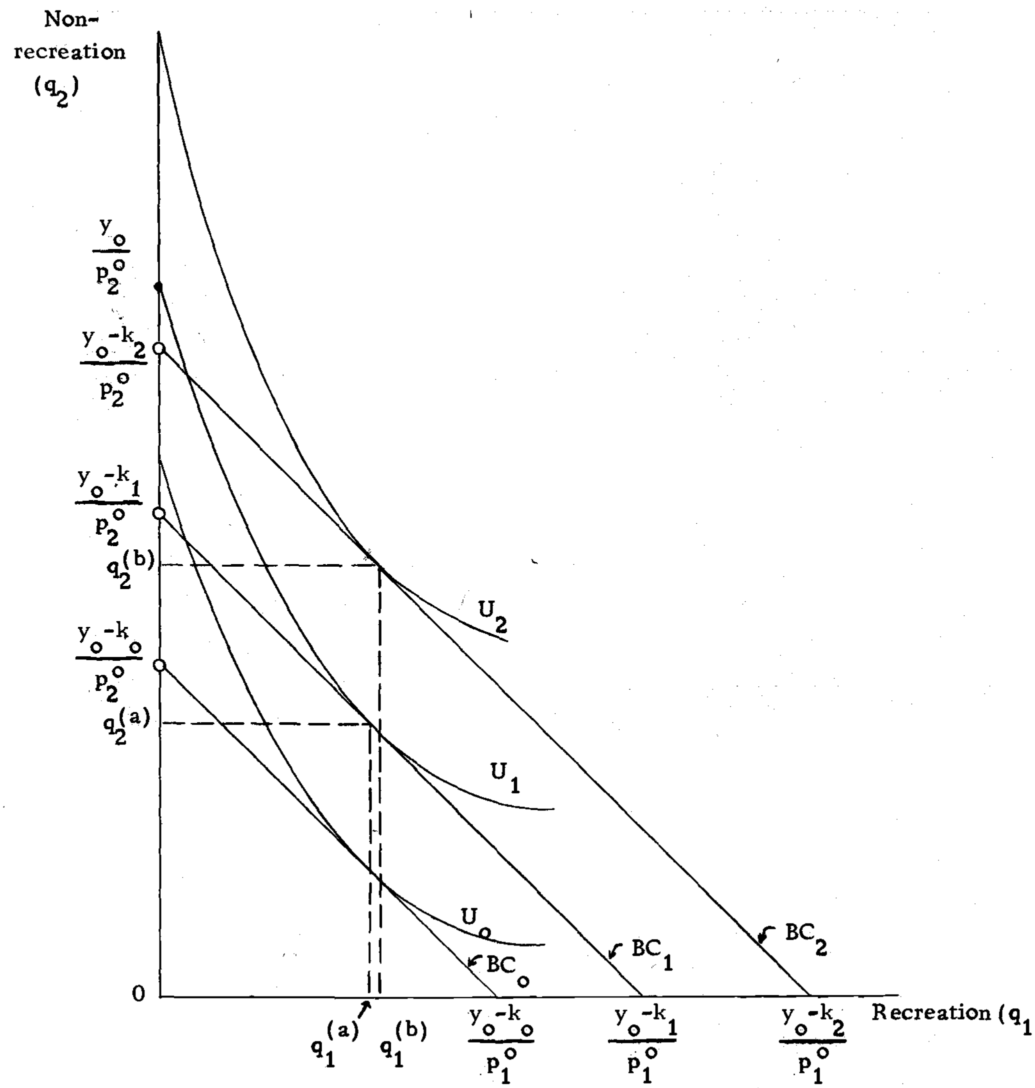

Figure 1. The optimal combinations of recreation and non-recreation that a typical consumer would take, if faced with travel costs of $k = k_o$, k_1 and k_2 (where $k_o > k_1 > k_2$), for given prices p_1^o and p_2^o, and fixed income y_o. $k = k_1$ is the critical travel cost for the conditions given.

the level of income are held constant. The only variation in the budget constraints is k. The budget constraint BC_o, is one for which $k = k_o$, a large positive value. The point on the vertical axis at $q_2 = y_o/p_2^o$ illustrates the uniqueness of the budget constraint--the discontinuity involved. If the consumer allocated all of his income to the consumption of non-recreation, he will not only be able to reach the level $\frac{y_o - k_o}{p_2^o}$, but it will be within his budget to attain the position y_o/p_2^o, since he will not need to withstand the additional travel cost required to enjoy recreation, as at any other point on the budget line. With the consumption of no recreation $(q_1 = 0)$ the consumer has y_o dollars at his disposal to purchase non-recreation commodities, i.e., he has k_o more dollars to purchase Q_2 than he would have if he chose to recreate.[9]

Given the set of indifference curves and the budget constraint BC_o in Figure 1, the consumer will prefer the combination $\{q_1 = 0, q_2 = y_o/p_2^o\}$ over any other attainable combination. He will acquire a level of utility of U_1, while if he chose any other set of the commodities he could at most attain a level of utility of U_o, where

---

marginal utilities and that posses the general shape consistent with maximization of utility, i.e., they satisfy the first-order conditions. The commodities Q_1 and Q_2 must be defined in such a way that the indifference curves intersect the q_2 axis, i.e., it must be possible to have positive utility while consuming no Q_1, the entire budget being allocated to the consumption of Q_2

[9]See equation (9) above.

$U_o < U_1$.

If the travel cost, in order to recreate, were less than previously considered, say $k = k_1$, where $k_1 < k_o$, then the consumer would be faced with the budget line BC_1 in Figure 1. The budget line shifting to the right indicates the availability of more income to allocate to the two commodities, namely $k_o - k_1$ more dollars available. Now the consumer, in maximizing his utility, has two alternatives: (a) he can take the combination $\{ q_1 = 0, q_2 = y_o/p_2^o \}$, or (b) he can consume the set $\{ q_1 = q_1^{(a)}, q_2 = q_2^{(a)} \}$. This choice exists since either position provides a U_1 level of utility, i. e., they lie on the same indifference curve.

A further reduction in the travel costs would shift the iso-income line further to the right as BC_2 in Figure 1. The consumer now prefers to consume a combination of recreation and non-recreation. He will purchase $q_1^{(b)}$ units of recreation and $q_2^{(b)}$ units of non-recreation. He will attain a level of utility of U_2, where $U_2 > U_1$.

In general, more recreation-days per visit will be demanded by the consumer as the travel costs decrease, ceteris paribus. This phenomenon is due to the fixed budget available. If more was spent in travel cost, then less would be available for the variable on-site costs. A decrease in travel costs is analogous to an increase in the recreationists' income, since the income that affects the decision-making framework as to how many days he will recreate per visit, is $y - k$.

An alternative hypothesis could state that as the travel costs increase the amount of time spent at the recreational site per visit would increase. The recreationists that are forced to drive a considerable distance, according to this argument, will generally be willing to spend more than a short time at the site when they arrive. The persons, on the other hand, that live near a recreational site would be more likely to use the site a short period of time per visit.

Based on the theoretical concepts developed in this chapter, the hypothesis that recreationists will tend to spend fewer days at a site per visit as his costs of travel increase will be tested. There is, however, a limit as to how high the travel costs can become, beyond which the consumer will not recreate. From Figure 1, when $k = k_1$ the consumer was indifferent between recreating and not recreating, while if $k < k_1$, he preferred to recreate a certain amount per visit depending on how much smaller k was than k_1. If $k > k_1$ the consumer could maximize his utility by not recreating, as when $k = k_o$. The travel cost of $k = k_1$ will be referred to as the "critical" travel cost and will be denoted k^*, i.e., in this case

$$k_1 = k^* \qquad \text{for} \quad p_1 = p_1^{\,o}$$

The effect on the number of days demanded of recreation, per visit, is zero for decreases in k when

$$k - \Delta k > k^*$$

i.e., when, even after a decrease in k, the resulting travel cost is still larger than the critical travel cost, k^*. For any decrease in k when

$$k - \Delta k < k^*$$

the effect on the quantity demanded for recreation will be the same as an increase in income. The effect on the amount of recreation consumed depends on how much smaller k is than k^*, i.e., the size of $(k^* - k)$.

The value of the critical travel cost can be expressed as a function of three independent variables. k^* depends on the amount of the variable costs at the site relative to the cost of one unit of q_2, the level of income, and the utility function:

$$k^* = k^*(p_1, p_2, y, U) \quad [10] \tag{10}$$

<u>Change in On-Site Costs (p_1)</u>

Changes in the relative prices of recreation and non-recreation commodities have an effect on the demand for recreation, q_1,

---

[10]As p_1 increases, the value of k^* will have to decrease due to the fixed budget available; $\partial k^*/\partial p_1 < 0$. As the level of income increases, <u>ceteris paribus</u>, so will k^*; $\partial k^*/\partial y > 0$, due to the nature of recreation, i.e., it is a normal good. The effect of U on k^* will be discussed later in this Chapter when the utility variable is expanded upon.

however in a different way than did the travel cost considered above. A change in either of the prices, p_1 or p_2, will change the slope of the iso-income line, since the slope is equal to the ratio of the two prices. A change in the slope will directly affect the optimal budget allocation of the consumer. The indifference map and budget constraints of an individual similar to the one presented in Figure 1 are presented in Figure 2. The travel cost is fixed at $k = k_o$, and the level of income and p_2 are held constant. The only variable in this diagram is p_1. In this way, holding the other variables fixed, the effect on the quantity of recreation-days per visit demanded as a function of the variable costs of a recreational day, can be illustrated.[11]

The budget line BC_o is identical to the one found in Figure 1, and illustrates that the consumer faced with this situation would prefer not to recreate. A decrease in the price of recreation from p_1^o to p_1' is represented by the iso-income line BC_o'. After the price decrease the maximum utility attained is U_1 and may be obtained in either of two ways. The consumer can allocate his entire budget to the consumption of non-recreation, y_o/p_2^o units, or he could involve the consumption of $q_1^{(c)}$ units of recreation and $q_2^{(c)}$ units of non-recreation. The consumer would be indifferent between the two

---

[11]The analysis could proceed analogously by observing the resulting changes due to variations in p_2, instead of p_1.

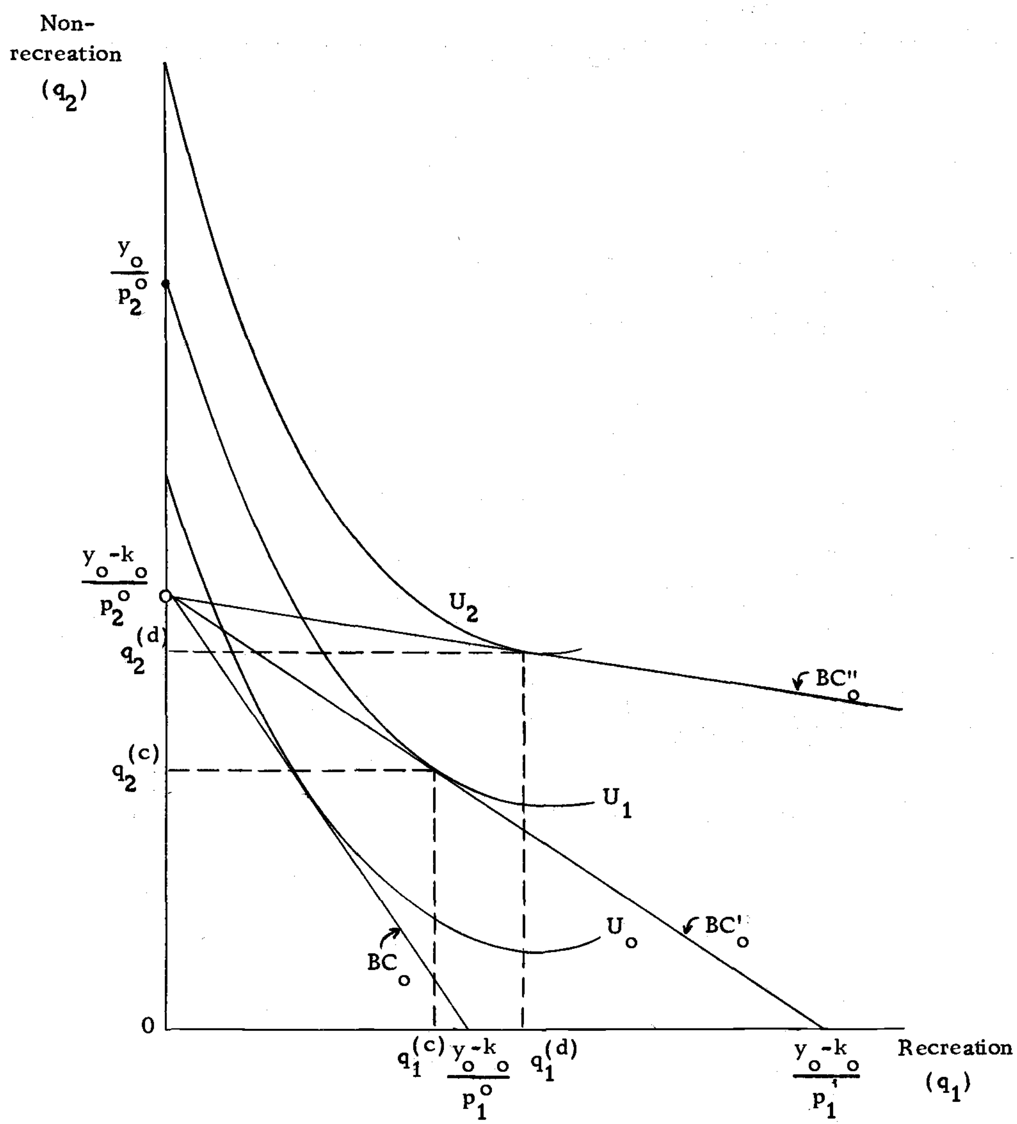

Figure 2. The optimal combinations of recreation and non-recreation type commodities that a typical consumer would take if faced with variable costs at the site of $p_1 = p_1^o, p_1'$, and p_1'', (where $p_1^o > p_1' > p_1''$), for given travel cost of $k = k_o$, and fixed price of q_2 of p_2^o, and fixed income y_o. $p_1 = p_1'$ is the critical price given these conditions.

choices, however for any value of p_1 such that $p_1 > p_1'$, recreation would be completely excluded from the optimal budget. A further decrease in p_1, indicated by the budget line BC_o'' in Figure 2, will change the optimal budget to $q_1^{(d)}$ units of recreation, and $q_2^{(d)}$ units of non-recreation. Note that, as the unit costs of recreation decreased the quantity of recreation consumed increased.[12] It can be seen in Figure 2 that any value of p_1 such that $p_1 < p_1'$ will indicate the consumer prefers a combination of recreation and non-recreation instead of consuming just non-recreation. In accordance with the terminology introduced earlier, the price of a unit of recreation of p_1' will be referred to as the "critical" price of recreation, denoted p_1^*, as

$$p_1' = p_1^* \qquad \text{for} \quad k = k_o$$

It will be instructive at this time to look at one more case, when $k = k_1$. The question that needs to be answered is whether the critical price of recreation will change when the travel cost changes. In Figure 3 the travel costs are fixed at $k = k_1$, and p_2 and the level of income are again held constant. As in Figure 2, the only variation in the budget lines is due to a change in the price of recreation. When $p_1 = p_1^o$, as in Figure 1, the consumer was indifferent between recreating and not recreating, illustrated by budget line BC_1. If the

---

[12]This is not true, of course, for inferior goods. Recreation is assumed to be a normal good, so the above conclusions follow.

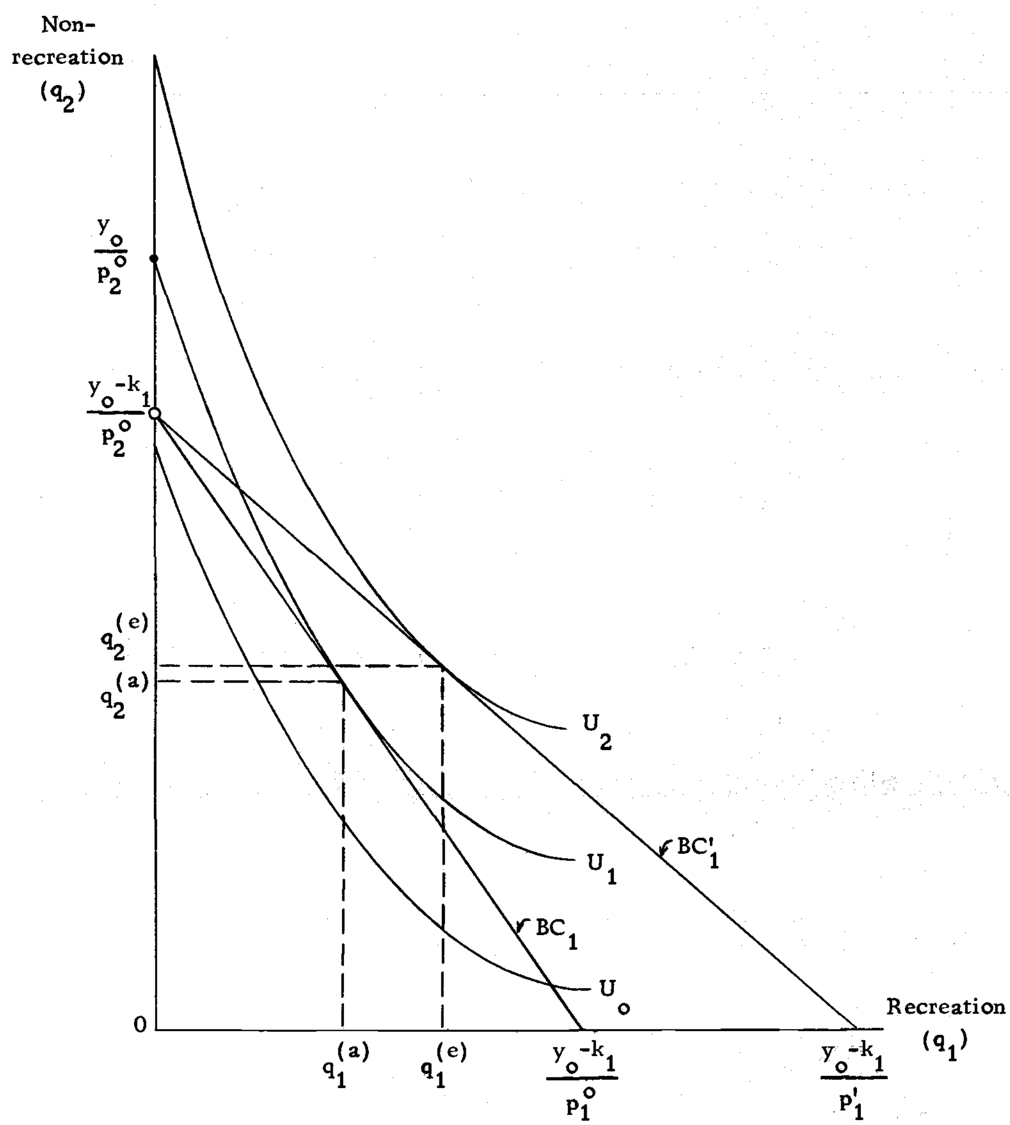

Figure 3. The optimal combinations of recreation and non-recreation type commodities that a typical consumer would take faced with variable costs at the site of $p_1 = p_1^o$, p'_1 (where $p_1^o > p'_1$), for given travel cost of $k = k_1$, ceteris paribus. $p_1 = p_1^o$ is the critical price given these conditions.

price of recreation were to increase the consumer would maximize his utility by not recreating. On the other hand if p_1 were to decrease to p_1', as in the iso-income line BC_1', the rational consumer would recreate $q_1^{(e)}$ units and take $q_2^{(e)}$ units of non-recreation. That is to say, if p_1 were such that $p_1 > p_1^o$, no recreation will be forthcoming. But if p_1 were less than p_1^o, the consumer would maximize his utility by including recreation in his budget. The price of p_1^o can then be termed the "critical" price of recreation when $k = k_1$:

$$p_1^o = p_1^* \qquad \text{for} \qquad k = k_1$$

In general, for given conditions relating to preferences, income, and prices of other commodities, there exists a critical value of p_1 for every value of k and a critical value of k for every value of p_1.

The effect on the quantity demanded of recreation is zero for decreases in the price of recreation when

$$p_1 - \Delta p_1 > p_1^*$$

that is, even after a decrease in p_1 of $\Delta p_1, p_1$ is still larger than the critical value, P_1^*. There will be an effect on q_1 for decreases in p_1 such that

$$p_1 - \Delta p_1 < p_1^*$$

The effect on the amount of recreation consumed depends on how much smaller p_1 is than p_1^*.

The value of the critical price of recreation depends upon the travel costs, the level of income, the price of other commodities, and the utility function; i.e.,

$$p_1^* = p_1^*(k, y, p_2, U) \quad [13] \tag{11}$$

Effect of Utility (U)

A change in the shape of the indifference curves, through a change in the utility function, can have an effect on the amount of recreation demanded. The utility variable can be represented, at least in part, by several other variables. That is, there are variables that can be observed that may give an indication as to how the recreationist will react to a change in his utility function. Some of the variables thought important in this respect are: the value of a recreationist's equipment, the characteristics of a site, the characteristics of the recreationist, and the leisure-time available to the recreationist.

Fixed investment. The amount of fixed investment is sometimes not considered a variable to be concerned with in decision-making. It is hypothesized, however, that there is some correlation between the demand for recreation and the amount of money tied up in fixed equipment, such as boats, campers, trailers, etc. The

---

[13] The effect of k on p_1^* was illustrated; as k increases the critical price of recreation decreases, $\partial p_1^*/\partial k < 0$. On the other hand as the level of income increases it is expected that the value of the critical price would also increase, $\partial p_1^*/\partial y > 0$, due to the fact that recreation is a normal good. The effect of U will be examined next.

rationale for this hypothesis is based on the assumption that persons with a large investment in recreational equipment, ceteris paribus, will have a different utility function, with respect to recreation, than would persons with less invested in equipment. This will directly affect the decision as to how many days the recreationist will spend at the site per visit, as well as how many visits he should make, due to the nature of the indifference curves faced by the different individuals. It is thought that the person with the larger investment in equipment will stay more days at the recreational site than the recreationist with less equipment. For example a person that owns a camper and large motor boat will probably tend to recreate more than the person who merely owns a tent, other things held constant. A qualification needs to be made, however, this phenomenon would be expected to occur only when a comparison is made within the same type of sites. That is, a wilderness-type lake cannot be compared to a highly-developed site, since each requires a unique combination of equipment.

Investment may be able to replace income as an explanatory variable. It is believed useful to use the amount of investment as a replacement for income for two reasons. First, it is correlated more with the permanent income of the recreationist than is current annual income. This is important since it is thought that the permanent income is more closely associated with the quantity of recreation taken than is current income. A case in point is a retired person. He has a low current annual income, but the amount of equipment he owns is related to his previous permanent income. It is hypothesized that the retired person will probably spend more time recreating[14] than a non-retired person. On the other hand, a young person beginning his career will probably purchase equipment with his future in mind. Even though he may have a low current income he plans on future increases in income. It is his future permanent income, then, that he correlates with his purchases of equipment, and the number of days he recreates. Thus, the current income of the recreationist may have little correlation with the amount of recreation demanded.

Secondly, the amount of fixed equipment reflects the recreational group[15] characteristics more than does current income. Recreation equipment is used primarily for recreation. The family income of a recreationist, however, is not all available for recreation, since the cost of living must be deducted. Thus, in most cases the value of the recreational equipment is more closely associated with the group's recreational budget. It is the group that makes the decision about the quantity of recreation demanded.

---

[14] See definition of visit and visitor-day at the beginning of this Chapter. This implies a very narrow definition of recreation. Recreation is not referring, in this thesis, to merely "not working."

[15] The recreational group is composed of those persons who recreate together. It may consist of either family members, non-family members, or both.

Site characteristics. The characteristics of a site help to influence the consumer's decision concerning the number of days he should recreate. Examples are the quality of the water, the size of the lake, the location of the site, the facilities available, the climate, etc. One reason the recreationist chooses a site at which to recreate is the "desirability" of the characteristics. The amount of time spent recreating per visit depends on the degree to which the characteristics are thought desirable, or undesirable. These characteristics are fixed as far as the site is concerned, but differ when considering several sites.

Characteristics of recreationist. The characteristics of the recreationist can, of course, influence his decision on his consumption of recreation. The background of the person, his stage in his family cycle, his age, etc. influence his "utility" function for recreation.

Leisure-time. The available leisure-time of a recreationist seems to be very important in determining the recreational demand of an individual. As the amount of leisure-time increases, ceteris paribus, the recreationist would tend to recreate more. The time available for recreation can be considered a constraint as much as, or more than, the income available to the recreationist.

The Theoretical Model

The theoretical model for an individual can now be written in terms of three structural equations: the quantity demanded of recreation per visit equation, the critical travel cost equation, and the critical variable cost of recreation equation.

$$q_1 = q_1 [(k^* - k), (p_1^* - p_1)]^{16} \quad \text{for} \quad (k^* - k) \geq 0, \; (p_1^* - p_1) \geq 0 \tag{12}$$

$$k^* = k^* (p_1, p_2, y, U) \tag{13}$$

$$p_1^* = p_1^* (k, y, p_2, U) \tag{14}$$

The variables in the above theoretical model will be analyzed in terms of empirical considerations, in Chapter IV. The statistical model will then be stated.

---

[16] It is hypothesized that the quantity demanded of recreation per visit will increase as the travel cost becomes greater than the critical travel cost, and as the on-site costs become larger than the critical on-site costs, i. e.,

$$\frac{\partial q}{\partial(k^* - k)} > 0, \quad \text{and} \quad \frac{\partial q}{\partial(p_1^* - p_1)} > 0.$$

CHAPTER IV

SPECIFYING THE STATISTICAL RECREATION DEMAND MODEL

In order to test the hypotheses derived in the previous chapter, information needs to be obtained from recreationists. This was done by personal interviews at four sites while persons were involved in the recreational activities. The variables that were used in the analysis discussed in Chapter III, will now be elaborated on and the empirical considerations discussed.

The theoretical discussion in Chapter III focuses on a single decision-making unit, whether it be a family, a group of unrelated individuals, or an individual. The sampling unit was the recreation unit, i.e., the group that was recreating together, regardless of the type of group. A problem can arise when discussing the measurement of the variables. The value of many expenditures increases as the size of the recreational group increases. Examples are food, lodging, and all other expenses that have prices based on individual consumption. The travel costs vary slightly, up to a point, as the number in the group changes. After the size increases to a certain point the travel cost will substantially increase, e.g., when more cars are required to transport the group.

In order to reduce the variation in the variables that is due to the size of the group, each variable in this study was expressed in terms of individuals. The expenses that a group incurred were divided by the number in the party. Another argument for expressing the variables in terms of individuals, and not groups, is that the family income obtained from the questionnaire represents the family income of the person interviewed, not the group. The appropriate income, when using the group as a focal point, would be the average family income of each member of the group--for parties composed of unrelated individuals. There are, of course, disadvantages of expressing everything on an individual basis. When the number in the group increases, ceteris paribus, economies of size are realized in some of the expenses. For example, the travel cost, especially when only transportation is involved, would not change very much with variations in the group size. But, the average cost per person would decrease as the size increased. In family groups it is difficult to perceive treating children in the same category as adults. In non-family groups, however, these economies of size can be easily accepted, since often the lower-per-person costs are being considered in the decision-making.

The biases that are forthcoming in this study as a result of expressing the variables on a per person basis are not serious. It is believed that the economies of size bias is less serious than the biases introduced by using group observations.

The theoretical demand model, conceptualized in Chapter III is as follows:

$$q_1 = q_1 [(k^* - k), (p_1^* - p_1)] \qquad \text{for} \quad (k^* - k), (p_1^* - p_1) \geq 0 \qquad (15)$$

$$k^* = k^* (p_1, p_2, y, U) \qquad (16)$$

$$p_1^* = p_1^* (k, y, p_2, U) \qquad (17)$$

The variables will now be defined and discussed.

Days of Recreation Per Visit (q_1)

The questionnaire determined when the recreationist arrived and departed from the site. From this information, the number of days that the recreationist visited the site for that trip was determined, within a half-day interval.

Travel Cost (k)

The travel cost the recreationist incurs in order to recreate consists of the cost of transportation to and from the site, food expenditures while traveling to and from the site, lodging, camping fees, and any other expenses while enroute to, or from, the recreational site.

Cost of Transportation

The transportation cost is the amount it costs the recreationist to drive[17] to and from the recreational site. This includes gasoline, oil, depreciation, insurance, maintenance and repairs and any other miscellaneous items involved with maintaining an automobile. This type of information was not specified on the questionnaire. The transportation cost was figured by multiplying the total number of miles traveled, in both directions, by a cost per mile of five cents. The cost per mile figure was determined by using information from previous studies, as well as current research, in which recreationists were asked to enumerate their transportation cost.[18]

Food

The recreationist that spends time away from home needs to make arrangements for his food consumption. He may prepare food at home to take with him, or purchase the ingredients at grocery stores to prepare at the recreational site, or he may patronize restaurants and cafes and purchase prepared food. In any case the recreationist may end up spending more, less, or the same for food while traveling as he would if he had stayed home.

When determining the value of a recreational site, the

---

[17] This could include other forms of transportation, but in this study no other means was encountered.

[18] See Guedry, and also Stevens (1966).

relevant food expenditure, made by the recreationist, is not the total amount spent for meals while traveling. The amount that he would have spent at home, had he chosen not to recreate, should be deducted from the amount of his expenditures while traveling.

In this analysis the food expenditure that was considered appropriate was the amount consumed while traveling, less the amount the person would have consumed at home. The amount a recreationist would spend for food at home was not directly available since this information was not collected on the questionnaire. Secondary data were available, however, which indicated the average daily food cost per person at home for various income groups.[19] This average figure was subtracted from the recreationists' food expenditure in the following way.

If the recreationist made expenditures for meals while enroute to the recreational site, account was made for what he would have spent at home.[20] The average daily expenditure per person for food at home, times the number of days traveled to and from the site,[21] was subtracted from the actual amount spent while traveling.

---

[19]See U. S. Dept. of Agriculture, 1965, page 7, Table 2. The 1964 expenditure figures were inflated by the consumer price index until equivalent to 1967 information.

[20]If no expenditure was made for meals while enroute, no adjustment for what he would have spent at home was necessary.

[21]The number of days spent traveling was not obtained in the questionnaire. It had to be assumed that the recreationist averaged

This difference, in some cases negative and some cases positive, was added to the calculations of the k variable.

Other Costs of Travel

The cost of lodging, camping fees, and any miscellaneous expenses while traveling were recorded in the personal interview.

The value of k was positive in every case, which is reasonable. Even though it is possible for a recreationist to have less expenses for food while traveling than at home, the transportation cost is a more significant portion of the travel cost. Since persons usually recreate in groups, the travel cost was divided by the number in the group to get the desired travel cost per person.

On-Site Costs (p_1)

p_1 is the total cost incurred by the recreationist per day while visiting the recreational site. Involved in p_1 are the costs of lodging, camping fees, equipment rentals, meals, and other miscellaneous expenses, incurred at the site. The total expenses that the recreational group incurred per day was divided by the number of persons in the group in order to obtain the daily on-site cost per

approximately 40 miles per hour in driving. This was then divided into the distance traveled, both ways, to give the number of hours spent traveling. The number of days was then figured, to the nearest tenth, by assuming 12 hours per day.

person.

For the reasons mentioned above, the value of meals the recreationist would have consumed at home was subtracted from the on-site food expenditures. The daily food cost per person while recreating at the site was computed from information obtained by the questionnaire. From this figure the average daily expenditure per person for food while at home, within the corresponding income groups, was subtracted. The difference in the two figures, whether positive or negative, was added to the calculations to determine the p_1 variable. In some cases, especially those persons recreating a short time per trip, the daily cost per person while at the recreational site, or p_1, was negative. That is to say that it was less expensive to visit the recreational site than to stay at home--not considering the cost of travel to and from the site.[22] It does not seem unreasonable to expect this type of phenomenon to occur, in a few cases. All other components of the on-site cost variable were accounted for by the questionnaire.

---

[22]Negative values of p_1 can also arise due to the fact that the average expenditure of food at home was subtracted from the individual recreationists' on-site food expenditure. That is, if the recreationist would have spent less than the average amount for food at home, it could have been just as expensive for him to stay at the site. But, since a larger value was subtracted, it appears as though it was less expensive for him to recreate.

Income (y)

There is often confusion as to which income is of interest in the economic analyses of recreation. It is seldom specified if personal income, family income, disposable income, etc., is being considered. In this analysis the family income of the recreationist, after taxes, was used. It was important to subtract income taxes from the family income since the income after taxes is all the recreationist has to use in his decision-making. Information was recorded as to the size of the recreationist's family, the age of the household head, and the gross family income. It was then possible, using a Federal Income Tax Table, to estimate the Federal Taxes paid, for the year of 1967. It was not feasible to estimate the State Income Taxes, since recreationists came to Oregon's lakes from several different states. Thus, the income used in this analysis is biased upward. However the error, it is believed, is not significant. The amount of the Federal Income Tax is usually much larger than the State Income Tax. The majority of the relevant taxes that should be subtracted have been, so exclusion of the State Income Taxes is not serious.

Price of Other Commodities (p_2)

The price of other commodities, in theory, is equal to the weighted average of the prices of the appropriate commodities:

$$p_2 = \alpha_1 p_3 + \alpha_2 p_4 + \ldots + \alpha_{n-2} p_n \tag{18}$$

where the p_i (where $i = 1, 2, \ldots, n$) is the price of the i^{th} commodity. In practice, however, the commodities that comprise the alternatives to recreation are so diverse between individuals to make it impossible to specify one or two commodities that would represent the alternatives to recreationists. The inclusion of the p_2 variable in the statistical model, was impractical. Thus, the p_2 variable was deleted in the statistical demand function for recreation.

Utility (U)

The actual utility function of the recreationist cannot, of course, be measured. Knowledge of several variables can be used as an alternative to knowledge of the utility function, however. The use of variables such as the characteristics of the site, the characteristics of the recreationist, the value of the recreationist's equipment, and the amount of leisure-time available, may help to indicate a change in the utility function.

Characteristics

It is beyond the scope of this thesis to try to handle both the characteristics of the individual and the characteristics of the site.[23]

[23] A study being conducted concurrently with this study has as it's major objective to specify the relationship between the

This study will concentrate its efforts on the use of the characteristics of a site to help estimate recreational demand relationships.

If a study were conducted to determine the recreational value of a particular site, and information gathered only at that site, the characteristics of the site are fixed and do not need to be stated explicitly. However, if more than one site is used these characteristics can be correlated to the quantity of recreation-days observed. This would be appropriate if contemplating the introduction of a new recreation site. The proposed site would possess certain characteristics[24] that could be introduced into the demand relationship in order to make estimates of recreational value. The specific characteristics that were of interest in this study are: the size of the lake, in acres; and the activities provided for at the site. When combining more than one site, the size of the lake and the use-intensities of the activities were used as independent variables. The activities considered were swimming, boating, water skiing, fishing, and camping. Discrete values were used to rank the use at each lake by activities:

---

characteristics of the individual and the demand for recreation. See Guedry.

[24] In fact, a look at the characteristics that have a substantial effect on the use, can help decision-makers to provide a new site with these characteristics. For example boat ramps, camp sites, etc.

no use of an activity was given a value of zero, low use a value of one, medium use a two, and high use a three.[25] The size, in acres, of each lake was used as a continuous variable.

Amount of Recreational Equipment

The value of a recreationist's equipment could be considered as an independent variable in the model, possibly replacing income. This study, due to the large number of problems to be solved, did not use investment as an explanatory variable.

Leisure-Time Available

The recreationist's leisure-time available seems to be very important in determining the recreational demand of an individual. Further examination, however, will reveal an important ambiguity involved with the definition, and hence the quantification, of leisure-time. Leisure-time can not only be defined in numerous ways, but in fact hinges around a person's value judgement as to what is considered leisure-time, i.e., it is closely associated with a person's utility function. An example might be cutting the lawn. One person

---

[25]The use-intensities were estimated by personal interview with U. S. National Forest Service personnel, and employees of the F. W. P. C. A. who were knowledgeable of the lakes' characteristics. It is recognized that the values are subjective. See Appendix Table I for the estimated site characteristic values.

considers that as an enjoyable activity to look forward to, while another may consider it as much a necessary and laborious job as his profession.

Another problem in trying to quantify the leisure-time available of an individual, in order to regress this on the amount of recreation, is the small variability in the total amount of leisure-time, compared to the wide dispersion of time spent recreating. In other words, leisure-time as an independent variable would statistically explain little variation in the dependent variable beyond that accounted for by the variables previously discussed. In a time-series analysis, however, the total amount of leisure-time a society has is very important in determining the total quantity demanded for recreation. This relationship can be seen intuitively over the past several decades.

Critical On-Site Cost (p_1^*)

Another variable that needs to be discussed is p_1^*, the recreationist's maximum willingness to pay variable costs at the site.[26] p_1^* is related to the amount of travel costs, the level of income, and the utility function, both through the characteristics of the individual and the characteristics of the site. The measurement of the critical

[26] This Chapter henceforth will focus on p_1^* rather than k^*. The procedure for determining k^* is analogous to the determination of p_1^*, so the reader can extrapolate this analysis to the use of k^*, if so desired.

price variable is difficult due to it's unique nature. The measurement technique that this thesis proposes is the following. If the estimation procedure is performed on more than one recreational site, then the observations should be categorized by sites, and then placed in homogeneous groups determined by the income level and the travel costs.[27] The groups were determined as follows: the incomes of all recreationists were listed in descending order of magnitude. Consideration was given to the "natural" breaks in the list, but due to the small number of observations only two cut-off points could be chosen. The cut-off points were chosen to ensure that each group would have approximately the same number of observations in them. That is, for all lakes, the three income groups were selected as: (1) less than or equal to $8, 000, (2) between $8, 000 and $10, 000, (3) greater than or equal to $10, 000. The distribution of travel costs, due to their unique nature, was different for each lake. For this reason the groups were

---

[27] If only one site is considered, arrange the observations into groups according to the income and travel costs incurred.

Note should also be made of the fact that only a part of the utility variable (characteristics of the site) is accounted for by this procedure. This can be justified by the fact that there would not be sufficient data to use all of the possible classifications inherent in the utility variable. It would be valuable to experiment with certain of the characteristics in which the observations could be grouped. It might be that some of the characteristics of the individual are more important than some of the characteristics of the site. These are testable hypotheses that might be examined.

determined separately for each lake. The critical points for Klamath Lake were \$1.00 and \$19.00. That is, the three groups were: (1) $\leq$ \$1.00, (2) between \$1.00 and \$19.00, (3) $\geq$ \$19.00. For Lake of the Woods the break-off points of k were \$2.00 and \$20.00; for Odell Lake \$5.00 and \$10.00 were used; and for Willow Lake the values of k were \$2.00 and \$10.00.[28]

The statistical problem arises as to the estimation of the maximum p_1 of the distribution within each group, i.e., p_1^*. Within each of the groups the variable costs, p_1, should be arranged in ascending order of magnitude. The n^{th} p_1 is the maximum p_1 observed in each group, and is referred to as the n^{th} order statistic.[29] The n^{th} order statistic can be used as a reliable estimate of the maximum p_1 in the population. The use of the maximum observed p_1 as an estimator of p_1^* does not imply that all other observed p_1's are ignored. The maximum p_1 is chosen only after it has been compared to all other p_1's in the group. Thus, use is being made of all the sample information.

One situation that can arise is if there are too few observations

---

[28]In computing the groups for the k^* equation, the distribution of p_1 and income were considered. The same income groups were utilized as above, and the groups for p_1, used on all lakes, were: (1) $\leq$ \$1.50, (2) between \$1.50 and \$2.50, and (3) $\geq$ \$2.50.

[29]For a theoretical treatment of order statistics see Hogg and Craig (1965, ch. 6).

in a group. The estimate of p_1^* is thus less reliable than if more observations were encountered. Reliability refers to the size of the variance; the less reliable the estimators are, the higher the variance. In utilizing regression analysis it is assumed that the diagonal elements in the variance-covariance matrix are constant, or nearly so. In this case some groups have higher variances than others, thus the constancy of variances assumption is violated. The coefficients estimated, by ignoring this type of problem, will be unbiased but will not have minimum variances. It is necessary, then, to make adjustments in the analysis to ensure unbiased minimum variance estimators.

It is hypothesized that the variance of the error term, ϵ_i, in the regression equation, is inversely related to the number of observations in the group. That is, $V(\epsilon_i) = \sigma^2/n_i$, where σ^2 is a constant variance. If ϵ_i was multiplied by $(n_i)^{\frac{1}{2}}$, then the variances would be constant: $V[(n_i)^{\frac{1}{2}}\epsilon_i] = n_i V(\epsilon_i) = \sigma^2$. Multiplying ϵ_i by $(n_i)^{\frac{1}{2}}$ is equivalent to multiplying each variable by $(n_i)^{\frac{1}{2}}$. The coefficients estimated by this procedure would be unbiased and have minimum variance. This procedure is referred to as weighted least squares. More weight is given to the reliable estimates than to the unreliable ones. This procedure was used to obtain the p_1^* and k^* relationships.

By utilizing the above procedure an estimate can be obtained for the critical variable cost, in each subgroup. The level of income, and the travel cost, are then regressed on the resulting p_1^*'s to obtain an estimate of the relationship of the independent variables and p_1^*.[30]

The Statistical Demand Model

The statistical model forthcoming from the preceding discussion, when more than one site is analyzed is:

$$q_1 = q_1 [(k^* - k), (p_1^* - p_1)] \quad \text{for} \quad (k^* - k), (p_1^* - p_1) \geq 0 \tag{19}$$

$$k^* = k^* (p_1, y, Sw, Ws, B, F, C, Si) \tag{20}$$

$$p_1^* = p_1^* (k, y, Sw, Ws, B, F, C, Si) \tag{21}$$

where Sw indicates the intensity of swimming use, Ws the intensity of water skiing, B the boating use-intensity of the site, C the camping use-intensity, and Si the size of the lake, in acres. All other variables are as defined above.

Equations (19), (20), and (21) need to be solved simultaneously to obtain estimates of all the coefficients. Two-stage least squares analysis was employed on the three structural equations in order to obtain the required estimates. The system of equations is

---

[30] If more than one site is being combined, then the characteristic of the site variables need to be included with the other independent variables.

overidentified--thus no estimates of the coefficients in equation (19) can be obtained by solving the system simultaneously with the use of a reduced form equation.

Equations (20) and (21) were estimated by the method of weighted least squares. Then the predicted values of p_1^* and k^* were utilized in equation (19). Then equation (19) could be regressed by ordinary least squares, since all the variables had observations.

The demand function of interest is the aggregate demand function for the total number of visitor-days per season. The preceding model is appropriate for explaining the number of visitor-days per visit, i.e., it is for an individual, not the population of users. The total number of visitor-days can be obtained by multiplying equation (19) by the number of visits, V, that the site would experience during the season. The appropriate aggregate model is as follows:

$$Vq_1 = f\,[\,(k^* - k),\ (p_1^* - p_1)] \qquad \text{for } (k^* - k),\ (p_1^* - p_1) \geq 0 \qquad (22)$$

$$k^* = k^*\,(p_1, y, Sw, Ws, B, F, C, Si) \qquad (23)$$

$$p_1^* = p_1^*\,(k, y, Sw, Ws, B, F, C, Si) \quad . \qquad (24)$$

Where V indicates the number of visits at a lake for the season. In order to analyze a particular lake, the site characteristics of that lake have to be substituted into the equations. The total number of visits, at each site during the relevant season, was obtained from the

U. S. Forest Service, since the lakes studied had facilities maintained by the Forest Service.

The Number of Visits Relationship

Two possible lines of endeavor exist at this point: The first would be to use the estimate of the number of visits and have it remain fixed with respect to any possible changes in the independent variables. The second, and possibly more desirable, method would be to use the estimated number of visits, as above, but then express V as a function of some appropriate variables. In this way when a postulated change would occur in the independent variables the effect on V, as well as q_1, could be observed. This would give a more reliable estimate of the total number of visitor-days. This procedure was adopted in this study.

The number of visits forthcoming at a site for an individual can be expressed as a function of the travel cost, and the income of the recreationist.[31] In this study, information was not obtained from recreationists as to how many visits they would make during the season. Thus it was impossible to focus attention on the individual recreationist, when concerned with the number of visits. An alternative does exist, however. The total estimated visits, V, can be

---

[31]If more than one site is considered the characteristics of the site would also be appropriate variables.

proportioned into several groups, representing regions, and then the relationship between the number of visits from the i^{th} region and the explanatory variables can be computed. It was hypothesized that the number of visits forthcoming from an area, V, is functionally related to the average travel cost from persons residing in the area, k, the average income of persons that recreate from the area, y, the total number of persons living in the area, pop, and the characteristics of the site, i.e.,

$$V = V\,(k, y, pop, Sw, Ws, B, F, C, Si) \quad . \tag{25}$$

Counties were used to represent the areas on which the variables were defined. Indication was made on the questionnaire as to the county from which the recreationists reside. Record was made of the number of recreationists that originated from each county. The percent of persons sampled from each county was computed, and then multiplied by the estimated total visits for the season, to assign an estimated number of visits, for the season, for each county from which recreationists were observed.

The travel cost and income of all persons from each county had to be averaged to obtain an average value for the county. The population of each county was determined from the U. S. Bureau of the Census, Population Estimates, 1966. The relevant counties, the population of the counties, as well as the information on visits, k and y are contained in Table I. Each lake is kept separate. Note should

TABLE I. THE POPULATION, NUMBER OF SAMPLE VISITS, PERCENT OF SAMPLE VISITS, VISITS PER SEASON, TRAVEL COST, AND INCOME, BY COUNTIES, FOR EACH OF THE FOUR LAKES.

County	Population of County (pop)	Sample Visits	Percent of Sample in County	Visits (V)	Travel Cost (k)	Income (y)
			KLAMATH LAKE			
Oregon:						
Deschutes	27,600	1	.021	3,070	4.67	18,372
Jackson	91,300	5	.104	15,205	2.16	7,547
Josephine	37,000	2	.042	6,140	3.83	10,329
Klamath	49,600	29	.604	88,305	0.26	8,385
Lane	200,700	1	.021	3,070	4.70	3,956
Multnomah	534,900	2	.042	6,140	51.33	5,342
Washington	126,100	1	.021	3,070	80.32	3,860
California:						
Orange	1,171,400	1	.021	3,070	28.63	13,440
Los Angeles	6,814,500	3	.063	9,211	80.26	13,355
San Bernardino	628,900	1	.021	3,070	119.39	5,228
Santa Clara	929,800	1	.021	3,070	19.29	15,315
Trinity	8,200	1	.021	3,070	63.17	7,414
Total		48	1.002	146,491		

TABLE I. (continued)

County	Population of County (pop)	Sample Visits	Percent of Sample in County	Visits (V)	Travel Cost (k)	Income (y)
			LAKE OF THE WOODS			
Oregon:						
Coos	54, 100	1	. 008	2, 130	14. 36	7, 528
Deschutes	27, 600	1	. 008	2, 130	9. 17	12, 048
Douglas	72, 600	2	. 016	4, 260	11. 51	7, 131
Jackson	91, 300	46	. 362	96, 388	1. 80	10, 471
Josephine	37, 000	6	. 047	12, 515	0. 78	10, 589
Klamath	49, 600	19	. 150	39, 940	2. 84	8, 327
Lane	200, 700	2	. 016	4, 260	8. 66	10, 335
Lincoln	25, 500	1	. 008	2, 130	91. 30	10, 746
Linn	65, 600	2	. 016	4, 260	9. 96	11, 850
Multnomah	534, 900	2	. 016	4, 260	4. 58	12, 506
California:						
San Francisco	714, 600	2	. 016	4, 260	58. 63	13, 106
Monterey	229, 900	3	. 024	6, 390	50. 86	7, 057
Contra Costa	514, 400	1	. 008	2, 130	79. 38	13, 740
Orange	1, 171, 400	3	. 024	6, 390	33. 85	10, 790
Alameda	1, 030, 400	4	. 031	8, 254	36. 71	15, 570
Los Angeles	6, 814, 500	10	. 079	21, 035	54. 78	14, 942
Humboldt	101, 300	1	. 008	2, 130	6. 48	10, 215
Santa Barbara	253, 400	1	. 008	2, 130	19. 20	5, 646
Siskiyou	35, 000	3	. 024	6, 390	3. 92	11, 893
San Mateo	519, 100	1	. 008	2, 130	17. 01	11, 010
San Bernardino	628, 900	1	. 008	2, 130	26. 76	13, 440

TABLE I. (continued)

County	Population of County (pop)	Sample Visits	Percent of Sample in County	Visits (V)	Travel Cost (k)	Income (y)
			LAKE OF THE WOODS			
Santa Clara	929, 800	9	.071	18, 905	28.92	15, 412
San Joaquin	283, 500	1	.008	2, 130	35.28	11, 010
Sacramento	597, 700	2	.016	4, 260	26.20	6, 492
Del Norte	16, 700	1	.008	2, 130	9.10	7, 414
Nevada:						
Clark	233, 700	1	.008	2, 130	46.22	13, 590
Idaho:						
Canyon	60, 400	1	.008	2, 130	8.68	10, 620
Total		127	1.004	267, 327		

TABLE I. (continued)

County	Population of County (pop)	Sample Visits	ODELL LAKE Percent of Sample in County	Visits (V)	Travel Cost (k)	Income (y)
Oregon:						
Benton	49, 100	3	. 031	5, 595	6. 45	8, 318
Clackamas	146, 100	4	. 041	7, 400	7. 22	5, 259
Coos	54, 100	2	. 021	3, 790	11. 29	5, 280
Crook	10, 100	1	. 010	1, 805	3. 87	7, 414
Deschutes	27, 600	2	. 021	3, 790	12. 14	5, 373
Douglas	72, 600	2	. 021	3, 790	2. 40	7, 748
Jackson	91, 300	3	. 031	5, 595	8. 17	11, 194
Klamath	49, 600	4	. 041	7, 400	2. 32	6, 822
Lane	200, 700	30	. 309	55, 769	4. 66	7, 256
Lincoln	25, 500	3	. 031	5, 595	6. 45	11, 804
Linn	65, 600	1	. 010	1, 805	5. 00	8, 472
Marion	141, 700	5	. 052	9, 385	7. 67	6, 688
Multnomah	534, 900	7	. 072	12, 995	9. 86	11, 747
Polk	30, 900	2	. 021	3, 790	10. 58	9, 005
Washington	126, 100	4	. 041	7, 400	11. 77	12, 567
California:						
Contra Costa	514, 400	2	. 021	3, 790	23. 11	8, 669
Orange	1, 171, 400	1	. 010	1, 805	44. 53	18, 540
Alameda	1, 030, 400	2	. 021	3, 790	44. 04	6, 940
Los Angeles	6, 814, 500	6	. 062	11, 190	28. 79	15, 502
Shasta	75, 500	1	. 010	1, 805	12. 37	11, 784
Siskiyou	35, 000	1	. 010	1, 805	15. 09	7, 114

TABLE I. (continued)

County	Population of County (pop)	Sample Visits	ODELL LAKE Percent of Sample in County	Visits (V)	Travel Cost (k)	Income (y)
San Mateo	519, 100	1	. 010	1, 805	12. 87	12, 048
Santa Clara	929, 800	3	. 031	5, 595	29. 44	11, 826
Modoc	7, 500	1	. 010	1, 805	10. 05	9, 519
Kern	325, 200	1	. 010	1, 805	24. 97	10, 746
Marin	188, 600	1	. 010	1, 805	11. 72	21, 804
San Diego	1, 188, 000	2	. 021	3, 790	40. 36	8, 330
San Joaquin	283, 500	1	. 010	1, 805	8. 39	15, 615
Sonoma	193, 700	1	. 010	1, 805	26. 59	1, 750
Total		97	. 998	180, 304		

TABLE I. (continued)

County	Population of County (pop)	Sample Visits	WILLOW LAKE Percent of Sample in County	Visits (V)	Travel Cost (k)	Income (y)
Oregon:						
Douglas	72,600	1	.031	3,397	2.77	6,700
Jackson	91,300	17	.531	58,187	1.90	8,127
Josephine	37,000	3	.094	10,301	1.94	9,589
Klamath	49,600	1	.031	3,397	2.30	8,586
Lane	200,700	1	.031	3,397	8.65	7,870
Lincoln	25,500	1	.031	3,397	28.53	10,614
Marion	141,700	1	.031	3,397	14.53	1,750
California:						
Los Angeles	6,814,500	5	.156	17,094	78.60	6,299
Siskiyou	35,000	2	.063	6,904	9.83	3,580
Total		32	.999	109,471		

be made that observations were made in only four states: Oregon, California, Nevada, and Idaho. This is due partly to the reputations of the sampled lakes--none is very well known outside the state of Oregon. Another explanation is due to the increased number of alternatives available at further distances from these lakes.

Since more than one site was being analyzed, the data from all four lakes were utilized to estimate the visits relationship. For every observation from Klamath Lake, for example, the values of the characteristics of Klamath Lake were used, along with the appropriate values of the remaining independent variables. There were data from 12 counties for Klamath Lake, 27 counties for Lake of the Woods, 29 from Odell Lake, and 9 from Willow Lake. Thus 77 observations were used, i.e., 12 + 27 + 29 + 9.

Since the independent variables k and y were aggregated within each county, some were more reliable than others, due to the number of observations within each county. Weighted least squares, similar to that discussed for determining the p_1^* and k^* relationships, was used to correct this problem.

Multiple regression was then used to express the relationship in equation (25). It should be clarified that equation (25) is redundant if one is not interested in changing some of the independent variables, and predicting a change in total visitor-days. If no changes were considered, equation (25) will merely predict the total visits for the site--

the information that the data for equation (25) was derived from. The total visits were estimated prior to the estimation of this equation. However, if the study was based on changing the existing situation and predicting the resulting change in the number of visitor-days, then the relationship is very important. It adds much information since visitor-days is made up of two components--visits and the length of stay per visit.

A word of caution is in order here. The interpretation of the visits relationship should be clarified. The population figure is not the population of recreationists, but the population of a county. The population of recreationists in a county is a subset of the total population in a county. Inferences are being made to the population of recreationists, even though county data is being used. It is recognized that this procedure is not entirely without criticism, but it is believed to be the best method of obtaining a relationship for the number of visits, with the existing information on hand.

In Chapter V the sampling scheme, the determination of the sample size, and the drawing of the sample are discussed.

CHAPTER V

SELECTION OF THE SAMPLE

Area of Study

The study area under consideration is located in the southwestern section of Oregon, predominately in Klamath County. The largest population center in the County is Klamath Falls which has a population of about 35, 000 within five miles of its downtown area. Many other towns are in Klamath County but are much smaller in size being either predominately lumbering or recreation-oriented. The availability of over 100 lakes and 80 streams and rivers in the County illustrate the abundance of water-related recreational resources. The present study was based principally on Upper Klamath Lake, located just outside Klamath Falls, but will also involve other lakes in the vicinity. Upper Klamath Lake is the largest body of fresh water in Oregon, being over 30 miles in length.

The Sampling Scheme

"Sampling is the taking of a part of a whole or total number of individuals from which to draw inferences or conclusions in regard to the characteristics of the group from which the sample was taken" (Cochran, 1960, p. 1). Sampling techniques were used in this study, as was mentioned earlier. A "scheme" can be characterized as a systematic plan for attaining some objective or purpose. It is thus important not to lose sight of the objective of the current research when discussing the sampling scheme. The principal objective, to which sampling is relevant, is the estimation of the recreational use of Upper Klamath Lake, were it cleaned of the water pollution problem. In order to accomplish this objective it was necessary to select other lakes in the area in which to obtain the necessary interviews with recreationists. The characteristics that a lake possesses are important, and were considered explicitly in this study. The selection of the lakes to be sampled was done with the purpose in mind of choosing them such that the levels of the characteristics were different. That is, it would not be rational, statistically, to choose a sample of lakes that possessed the same, or nearly the same, level of each of the characteristics thought to be important, e.g., the size of the lake. For this reason, then, a sample of lakes was chosen--not randomly--but purposively.

Three lakes, besides Upper Klamath Lake, were chosen to represent the sample of lakes. The lakes chosen were Odell Lake, Lake of the Woods, and Willow Lake. They were chosen because many of the recreationists living in the Klamath Falls area are now using either these three lakes, or lakes in the close vicinity of these lakes, in which to recreate. Another reason these particular lakes were chosen was due to the degree to which they possessed the characteristics chosen for this study: the size of the lake, and the use-intensities of the activities. For example, Klamath Lake is a very large lake, and thus represents a high potential of use. Only lakes that were easily accessible, that had relatively high water quality, and that had overnight facilities available were considered, since these attributes were deemed a necessity in order for an average recreationist to substitute that lake for Klamath Lake, due to the characteristics that Klamath Lake possesses. By choosing the sample of lakes for the reasons mentioned, it was then possible to correlate the recreational use with the characteristics, and in this way to make inferences about the projected use of Upper Klamath Lake were it free of it's water quality problem.

The population for this study was the total number of recreationists that recreated at the four lakes studied, during 1968. The sampling unit was the recreation unit, i.e., the group that recreates together, whether it is a family, a group of relatives, or a group of unrelated individuals.

It might seem that the individual recreationist should be the sampling unit. It is this individual who consumes the services that recreation offers as "inputs" in order to produce satisfaction. In order to use the individual as the sampling unit one must conclude that he is the primary decision-making unit as far as consuming the services from recreation is concerned. But, in reality, very few individuals are completely free to choose as they please among the various alternative ways of producing satisfaction. Children clearly are not without some form of supervision, and their parents make decisions about the family as a whole.

Similarly, the family unit is not the appropriate sampling unit, since there can be more than one family in a decision-making unit. This would add much confusion to the analysis if several unrelated individuals were to recreate as a group. Because of this problem--the family unit and the decision-making unit not always coinciding--the recreation unit was used as the unit of inquiry for this study.

The appropriate frame for this study would be a list of all the recreation units found in the area of study for 1968. A list of this sort, that would be adequate for the objective of this study, would be impossible to obtain. Since only four lakes were sampled it was adequate to take a stratified random sample at each lake. This was done by first dividing the area immediately surrounding the lake into several geographic blocks. All of the camping areas, boat ramps, picnic areas, lodges, cabins, etc. were placed in a separate block. A random sample of recreationists was then drawn from each of the blocks in a proportion consistent with the use of each block. That is if one block had ten times as much use than another block, the number of recreationists sampled in the first was ten times that in the lower use block. This was done for each of the four lakes. In this manner a good coverage of the lake was obtained and a representative sample selected.

Sample Size

The number of recreationists to be interviewed, in order to provide adequate information for the empirical basis of this study, was derived from statistical theory. The main steps to follow when determining sample size are as follows:

(1) A determination of the precision that is expected of the sample must be made.

(2) An equation needs to be found that connects the sample size with the desired precision.

(3) All unknown parameters in the equation found in (2) above need to be estimated.

(4) If more than one item is being measured in the survey, and a desired degree of precision is prescribed for each item, a method must be found for reconciling any conflicting values of the sample size.

(5) Finally, the sample size chosen from steps (1) through (4) must be appraised in light of an economic framework, i.e., to see whether the size of the sample is consistent with the available resources. Further, an analysis of the

responsiveness of the precision with respect to changes in the sample size needs to be viewed. Then the equating of the marginal cost of an additional interview with the marginal value of the additional precision, can be made.

The primary objective of this research is to estimate the total number of recreation-days forthcoming on Upper Klamath Lake if a substantial improvement in water quality were made. The total recreation-days can be defined as:

$$N\bar{q}_1$$

where $\bar{q}_1$ is the average number of recreation-days taken per visit, and N is the total number of visits during the season for all four lakes.[32] For the reason that estimating the total number of recreation-days is the most important item to be computed, the sample size was chosen in order to satisfy a desired precision computed on the random variable $N\bar{q}_1$.

In compliance to the steps outlined above to determine the sample size, this author first must state the degree of precision he was willing to accept. In stating the desired precision, a person is in essence, stating the amount of error which he is willing to tolerate in the sample estimates. This amount is determined in light of what is

[32] N, for purposes of this chapter is the sum of the V's, the number of visits from each site. Due to the nature of the sampling problem it was believed to be preferrable to use N instead of using V and performing the operation four times.

being estimated and the actual use that is planned of the estimate. To a degree it is arbitrary, but can represent a very close approximation as to what is acceptable. The author deems an error limit of approximately ten percent in the estimate of $N\bar{q}_1$ to be small enough to allow a proper use of the estimator. In addition to the margin of error of ten percent a small risk, α, which we are willing to incur that the actual error is larger than ten percent, has been agreed upon. That is, we want

$$P\left[\,\left|N\bar{q}_1 - \widehat{N\bar{q}_1}\right| \geq t_\alpha \sqrt{V(N\bar{q}_1)}\,\right] = \alpha \tag{26}$$

where $\widehat{N\bar{q}_1}$ is the estimate of $N\bar{q}_1$, and $t_\alpha \sqrt{V(N\bar{q}_1)}$ is equal to ten percent of the value of $N\bar{q}_1$. Equation (26) can be written as

$$P\left[-t_\alpha \leq \frac{\widehat{N\bar{q}_1} - N\bar{q}_1}{\sqrt{V(N\bar{q}_1)}} \leq t_\alpha\right] = (1 - \alpha) \tag{27}$$

in order to arrive at the confidence interval for $N\bar{q}_1$ as:

$$\widehat{N\bar{q}_1} \pm t_\alpha \sqrt{V(N\bar{q}_1)} \tag{28}$$

where $t_\alpha \sqrt{V(N\bar{q}_1)}$ will be set equal to ten percent of the value of $\widehat{N\bar{q}_1}$. The confidence interval in equation (28) forms a connection between the desired precision and the sample size. There are several parameters that need to be estimated before any computation is possible. The study that is concurrently underway at Oregon State

University, referred to in footnote 23 of Chapter IV, was useful in the acquisition of data to estimate the necessary parameters, since the data for that study had been collected. The parameters were estimated as follows: the average number of days that a recreationist visited a site, $\bar{q}_1$, was five, with an estimated variance, σ^2, of approximately 25. The total population in the four lakes being sampled, N, was estimated from United States Forest Service Data at approximately 700, 000.

The confidence interval, equation (28), can now be written as

$$\widehat{N\bar{q}_1} \pm (2) \sqrt{V(N\bar{q}_1)} \tag{29}$$

if we use $\alpha = .05$, i. e., we use a 95 percent confidence interval. Further manipulation yields

$$\widehat{N\bar{q}_1} \pm (2)\ N \sqrt{V(\bar{q}_1)} \quad , \tag{30}$$

since $V(N\bar{q}_1) = N^2V(\bar{q}_1)$; and is equal to

$$\widehat{N\bar{q}_1} \pm (2)\ N \sqrt{\hat{\sigma}^2/n} \quad ^{33} \tag{31}$$

---

[33] The variance of the predicted mean value of q_1, denoted $\bar{q}_1$ at a specific value of x (where x denotes, in general, the independent variable in the regression equation), say x_k, is

$$V(\bar{q}_1) = \frac{\sigma^2}{n} + \frac{(x_k - \bar{x})^2\ \sigma^2}{\Sigma\,(x_i - \bar{x})^2} .$$

For the particular problem in this study, however, the specific x_k would be the mean value of the independent variable, i. e., $x_k = \bar{x}$.

or

$$\widehat{N\bar{q}_1} \pm (2)\ N \frac{\hat{\sigma}}{\sqrt{n}} \quad . \tag{32}$$

Substituting the appropriate values into (32) above will give the confidence interval of

$$3,500,000 \pm (2)\ (700,000)\ \frac{5}{\sqrt{n}} \quad . \tag{33}$$

The right-hand portion of the above confidence interval should then be set equal to ten percent of the left-hand side, since this is the stated precision, as

$$(2)\ (700,000)\ \frac{5}{\sqrt{n}} = (.10)\ (3,500,000) \tag{34}$$

$$(7,000,000)\ (\frac{1}{\sqrt{n}}) = 350,000$$

$$\sqrt{n} = 20 \quad \rightarrow \quad n = 400$$

which implies that the sample size should be about 400 for a ten percent chance of error and an α of .05.

A responsiveness of the precision with respect to changes in the sample size should be considered to make sure the sample size chosen satisfies an economic criterion. If the sample size was equal to 300, what sort of variability might be expected with a 95 percent confidence interval? Using the same formulation and estimates of the parameters as before, we have

---

Thus the variance of $\bar{q}_1$ will be a minimum and equal to σ^2/n, as was used in the text above. For a more thorough discussion of this topic see Draper and Smith (1966, p. 22).

$$\widehat{N\bar{q}_1} \pm (2)\ N \frac{\hat{\sigma}}{\sqrt{n}} \tag{35}$$

or

$$3,500,000 \pm (2)(700,000) \frac{5}{\sqrt{300}} \tag{36}$$

$$= 3,500,000 \pm (700,000) \frac{1}{\sqrt{3}}$$

$$= 3,500,000 \pm 404,624 \quad .$$

This indicates a 11.6 percent limit in the error expected in the sample estimates. If, on the other hand, a sample size of 350 was substituted into the confidence interval used above, a limit of error of approximately 10.5 percent could be expected.

In viewing the high cost of each interview taken, due to it's complex nature, it seemed more efficient to use a sample size of 300, instead of 400, since little precision was sacrificed for the relatively high cost savings involved. A smaller sized sample, on the other hand, would not seem wise due to the number of lakes to be sampled, and the group breakdowns needed at each lake. Thus, this study used a sample of 300 recreationists interviewed while they were at the four lakes recreating.[34]

---

[34]It should be noted that the methodology used in this section to determine the sample size is invariant to the value of the total number of recreation-days variable, $N\bar{q}_1$. This can easily be shown if the random variable $CN\bar{q}_1$ were substituted into the confidence interval for $N\bar{q}_1$. Here C is a constant greater than zero. In the computation the C will cancel and the sample size as computed above will solve the equation.

Drawing the Sample

The first step necessary in order to draw the sample of recreationists, was to make a complete list of the campsites, picnic sites, boat ramps, etc., including all places available for recreationists to use the service of the water resource. A list was made for each of the four lakes that were sampled. Most of the sites on the lakes were United States Forest Service Campgrounds, some were County owned boat ramps and campsites, while still others were privately owned resorts leased from the federal government.

The second step was to obtain estimates as to how much use each site had in 1967. The number of recreational-days that the site was used in 1967 was recorded. A recreational-day, as defined by the United States Forest Service, is a 12 hour period in which one person actively enjoys a recreational facility. For example, if two persons stayed in a campground for six hours, this would be equivalent to one recreational-day; if they stayed 24 hours it would be equivalent to four recreational-days, and so forth.

When all of the use-data were recorded, the sample size of 300 was divided according to the amount of use at each lake, in direct proportion to the number of recreational-days enjoyed at each lake in 1967. For example if Odell Lake had three times as many recreational-days recorded as Willow Lake did, then the number of interviews taken at Odell Lake was three times as large as those taken at Willow Lake. The number of interviews taken at each site depends, again, on the amount of recreation-use each received in 1967. The number allotted for each lake was then distributed to the various compsites on the lake. The total number of interviews taken at each site is recorded in Table II.

The recreationists, within each campsite at each lake, were then drawn at random to be interviewed. Care was taken that the sample was representative of the use the site received. That is, if approximately 2/3 of the people at a site were picnicking and 1/3 water skiing, then the sample should contain twice as many picnickers as water skiiers. Consideration was also given that the recreation-units were not interviewed because of the type or size of the equipment being used.

One more question had to be answered: when should the interviews be taken? It was deemed important that the interviews were distributed between the week days and the weekends, since the characteristics of the recreationists could be quite different. Estimates were obtained from the agencies as to the relative use of the lake during the week and throughout the weekend. The sample was then proportioned by this factor, e.g., if for any given day during the week there would be one-tenth as many people recreating as would occur on a given day on the weekend, then 80 percent of the sample

TABLE II. NUMBER OF INTERVIEWS TAKEN AT EACH LAKE, BY CAMPSITE, FOR THE FOUR PERIODS UTILIZED.

Lake	Campsite	Week Days	Week-end	First Week Total	Week Days	Week-end	Second Week Total	Total
Klamath	Recreation Creek	1	2	3	1	2	3	6
	Odessa	0	1	1	0	0	0	1
	Rocky Point Resort	3	11	14	3	11	14	28
	Moore Park	2	7	9	2	6	8	17
	Pelican Marina	0	1	1	0	1	1	2
	Yacht Club	0	1	1	0	1	1	2
	Public Boat Launch and Bank Fishermen	0	0	0	0	1	1	1
Total Klamath Lake		6	23	29	6	22	28	57
Lake of the Woods	Rainbow Bay	5	17	22	5	17	22	44
	Aspen Point	3	13	16	3	12	15	31
	White Pine	0	2	2	0	2	2	4
	Lake of the Woods Resort	3	11	14	3	10	13	27
Total Lake of the Woods		11	43	54	11	41	52	106

TABLE II. (continued)

Lake	Campsite	Week Days	Week-end	First Week Total	Week Days	Week-end	Second Week Total	Total
Odell	Princess Creek	6	7	13	6	7	13	26
	Sunset Cove	3	3	6	2	3	5	11
	Trapper Creek	10	11	21	11	10	21	42
	Pebble Bay	0	0	0	0	0	0	0
	Odell Creek	1	2	3	1	1	2	5
	Odell Summit Lodge	1	2	3	1	1	2	5
	Odell Lake Resort	1	1	2	1	1	1	4
	Shelter Cove Marina	0	1	1	0	1	1	2
Total Odell Lake		22	27	49	22	24	46	95
Willow	County Campground	5	16	21	5	16	21	42

size attributed to the particular site should be taken on weekends, and 20 percent during the week.

All possible weeks in the season of use for the four lakes were numbered, and two weeks were randomly chosen. The total number of interviews needed was then divided into two equal parts, one-half taken on each of the two weeks chosen. The number of interviews taken during each time period, broken down by weekly and weekends, is indicated in Table II.

In order to avoid drawing substitutes and replacing them in the sampling scheme, due to refusals, oversampling was employed. That is, a 90 percent completion rate was expected to be obtained,[35] so the quota would be completed if 11 percent more interviews were attempted. Thus 333 interviews were attempted, allowing for a ten percent refusal rate, including questionnaires that were incomplete and unusable.

The empirical results are enumerated in Chapter VI, including some of the statistical problems encountered.

---

[35] The 90 percent figure was obtained by personal communication with Mr. Roy Bardsley, co-owner of a Marketing Research Analysis firm. His experience in this type of study led him to recommend this figure.

CHAPTER VI

ESTIMATING THE STATISTICAL RECREATION DEMAND MODEL

The empirical results consist mainly of four predicted equations: The critical travel cost relationship, k^*; the critical on-site cost relationship, p_1^*; the visits relationship, V; and the demand relationship, q_1. Each will be discussed separately in this Chapter. The variables utilized in the relationships are as defined in Chapter IV.

The estimated relationships for the critical travel cost equation, the critical on-site costs equation, and the visits equation were determined by the method of weighted least squares. Multicollinearity was observed between three of the site characteristic variables; swimming, water skiing, and boating use-intensities. The simple correlation coefficient, ranging in value from .957 to .980, was large enough to make it impossible to disentangle the separate influences of these variables and determine their relative effects in explaining the variation in the dependent variable. When swimming intensity had entered the equation, in the step-wise program, neither water skiing nor boating could explain any more of the variation, i.e., all three variables were explaining the same portion of the variation in the dependent variable. This problem was solved by combining the three into one variable. The use-intensity rankings for swimming, water skiing, and boating were summed to represent a single variable, denoted by W.

Another case of multicollinearity was observed between camping intensity and income. This was a statistical problem, having no particular economic significance. It was not feasible to combine the two variables due to their different nature, so camping intensity was deleted from the estimated equation. The amount of camping, being a function of man-made facilities and natural characteristics, does not clearly represent a characteristic of the site--thus the relationship is not altered significantly by the removal of this variable.

Since weighted least squares was utilized, the original equations had to be "deflated." That is, since each observation was multiplied by a positive value, $(n)^{\frac{1}{2}}$,[36] the constant term in the equation, $\hat{\beta}_o$, had to be recalculated. The calculation was done by the original definition of $\hat{\beta}_o$ in linear regression:

$$\hat{\beta}_o = \bar{y} - \hat{\beta}_1 \bar{X} \qquad (37)$$

where $\bar{y}$ is the mean of the dependent variable and $\bar{X}$ represents the mean of the independent variable (or variables). After $\hat{\beta}_o$ was recalculated, the weights were disregarded, and the value of the coefficients on all of the independent variables remained as estimated.

---

[36] Here n refers to the number of observations in the group.

The coefficients will be different if weighted least squares analysis was employed than if it was not, but no adjustment was necessary since weights were utilized in order to affect these coefficients.

Both a linear function and a curvilinear function were fitted to the data. Combinations of squared terms and linear terms were experimented with until the "best" combination was obtained. For example, both income, y, and income-squared, y^2, were analyzed, then only y^2, etc., for each of the independent variables. The criteria for choosing the functional form was based on the t-value of the variables, the simple correlation coefficients associated with the variables, the deviation of the actual minus the predicted values of the dependent variable, and theoretical considerations.

Upon estimating an equation, the deviations were first analyzed to determine if a random distribution existed, or if some relationship existed between residuals and the dependent variable, or between the residuals and the independent variables. If it was determined that the residuals were not random, then a search was made for another functional form, since the one in question would not be appropriate for the particular set of data.[37]

If, however, the residuals were nearly randomly distributed, the t-values associated with the coefficients were observed. If an

---

[37] Draper and Smith (1966, Ch. 3) have an excellent discussion of the investigation of residuals.

additional variable was entered into the equation with a very low t-value, then the reason had to be determined. It may have been due to a high correlation with another independent variable, or to the fact that it merely did not explain any additional variation in the dependent variable. In either case it was not considered appropriate to leave the additional variable in the equation in that form. The t-values of other variables had to be considered to see if they were correlated with the added variable. By removing another variable it may be appropriate to let the added variable remain in the equation. This solution must be determined in light of the theoretical concepts, i.e., there are certain economic variables that should remain in the equation. Many interesting statistical problems became evident when determining the correct functional form of the estimated equations. Each relationship will now be discussed.

k^* Relationship

As was discussed earlier in this thesis, the k^*'s were determined by using the n^{th} order statistic, in the distribution of k, to give a defensible estimate of the maximum k of the distribution. The individual observations were sorted into groups, classified by the independent variables: lake characteristics, income, and on-site costs. When the k^* relationship was estimated, average values of income and on-site costs were used. For this reason, inferences can be made only to the groups, not the individual observations.

The predicted equation is:

$$k^{*} = \underset{}{-36.711} + \underset{(2.322)}{6.248W^{\#\#}} + \underset{(7.800)}{3.779F} + \underset{(.0002)}{.0003Si} + \underset{(.0018)}{.0020y} \tag{38}$$

$$+ \underset{(3.349)}{10.435p_1^{\#}} \qquad R^2 = .616^{\#}$$

where W refers to the swimming, water skiing, and boating variable, F indicates the fishing intensity, Si signifies the size of the lake, y is the average income in the group, and p_1 is the average on-site costs for the group. The standard errors of the coefficients are listed in parentheses immediately under the coefficient. It was tested to determine if the coefficients were significantly different from zero. The level of significance is indicated by a # mark on the variable designation. One mark, #, implies the variable is significantly different from zero at the one percent level, two marks, ##, indicate a five percent level of significance, and three marks, ###, refer to a ten percent level of significance. If no marks are listed the coefficient is significantly different from zero at a level of significance greater than ten percent. In equation (38) only the on-site costs variable is significant at the one percent level, while the "water" variable, W, is significant at the five percent level. All other variables failed to obtain the ten percent level of significance.

The site characteristics, W, F, and Si, were hypothesized to have a positive effect on k^*. This means that as the characteristics of a site, both the asthetic properties, as well as the quality of the water, improve, recreationists will be willing to pay more travel costs in order to utilize the site. The sample evidence supports this hypothesis; that is, with the available information it would not be possible to reject the hypothesis that the site characteristics have a positive effect on k^*. However, due to the low significance of the fishing intensity and the lake size variables, few conclusions can be stated with much reliability concerning the existence of a relationship between these variables and k^*. The evidence does suggest, however, that if a relationship did exist the characteristic variables would have a positive effect on k^*. If the swimming, water skiing, and boating variable were increased by one unit (indicating a higher use in these activities) we would expect persons to be willing to pay an additional \$6.25 in travel cost, to recreate. That is, if a recreationist was faced with the decision of recreating on two identical lakes, except one had a one unit higher value of it's W variable, the person would be willing to pay \$6.25 more in travel cost to recreate at that lake, than at the other. The coefficients on the F and Si variables can be interpreted in similar manner.

The sample data suggests a positive relationship between the level of income and the critical travel cost variable. Income was not very significant in the estimated regression equation, so heavy reliability can not be placed on this relationship. It is hypothesized, however, that as the level of income increases, ceteris paribus, so will k^*. The lack of significance may be due, at least in part, to the procedure in which k^* was estimated and the broad income groups used. The incomes within each group were not homogeneous, thus an extraction of the complete effect of income on k^* was not possible. The groups should be defined for narrower ranges of income, if an adequate amount of data would permit doing so. Another possible explanation for income being slightly significant is due to the definition of the income variable. The income the decision-maker must use is that portion of net family income that can be allocated to recreation. This is the amount of money to which he is limited, not his entire income. The difference between total income and that portion available for recreation is diverse between individuals. Thus the use of total income tends to cover up most of the variation inherent in the income available for recreation variable.

The on-site costs, the most statistically significant variable in the equation, has an estimated positive relationship with k^*. The theory suggested that as a person was required to spend more at a recreational site, he would be willing to pay less in travel costs, due to his fixed budget. This hypothesis is rejected by the sample data. Possible explanations might include the fact that the on-site costs comprise a much smaller portion of the recreationist's budget than do the travel costs. Because of this, other phenomena have influences in the decision-making. For example, the more desirable sites may generally have a higher daily on-site cost, due to the increased demand for the site. Therefore recreationists would be willing to pay higher travel costs to visit the more desirable site.

The coefficient of determination, R^2, for this estimated relationship is .616. That is, the independent variables explain 61.6 percent of the variation in k^*, the remaining being unexplained and absorbed into the error term. The R^2, in this case, is significantly different from zero at the one percent level.

p_1^* Relationship

The data for the p_1^* variable was obtained, as was explained in Chapter IV, by grouping the observations according to lakes and then by the income level and travel cost. Average values of income and travel cost were used in the p_1^* relationship. Inference, then, may only be made toward the groups, not the individual recreationists. Inferences can be made at the individual level when values of p_1^* and k^* are substituted into the demand relationship, to be discussed later in this Chapter.

The estimated relationship is:

$$p_1^* = -7.263 + \underset{(.197)}{7.80W^{\#}} + \underset{(.815)}{2.630F^{\#}} + \underset{(.000025)}{.000067Si^{\#\#}} - \underset{(.002)}{.004k^{2\#\#\#}}$$

$$+ \underset{(.143)}{.269k^{\#\#\#}} + \underset{(.0000000094)}{.000000017y^{2\#\#\#}} \qquad R^2 = .684^{\#} \tag{39}$$

where W, F, Si are the site characteristic variables, as previously defined, and k is the average travel cost of the group, k^2 is the square of the travel cost, and y^2 is the average income of the group, squared. The standard error of the coefficients are written in parentheses below the coefficients. The significance level is indicated, as above, by the # sign. Only two variables, W, and F are significantly different from zero at the one percent level, the size of the lake being significant at the five percent level, while travel cost, travel cost squared, and income squared are all significant at the ten percent level of significance.

The site characteristics should, it was hypothesized, have a positive relation to the critical on-site costs. One would expect, and can conclude with the sample evidence obtained, that a more desirous site would induce recreationists to be willing to pay higher daily costs, than recreationists at a comparable site, with lower site characteristic values. This assumes, of course, that the variables quantified in this study to represent the site characteristics are an adequate measure of the desirability of a site. There are many criteria in which a site is judged by decision-makers, most of which can not be quantified except by an expression of the individual's utility function.

From the estimated equation it is possible to see the predicted effect on p_1^* of changing the characteristics. For example, if fishing intensity were to increase one unit, say due to a change in the water quality, then the recreationist would be willing to increase the maximum he is willing to pay at the site by \$2.63. The average p_1^* for all groups was estimated at \$5.78.

The conceptual framework in this study suggested a negative relationship between travel cost and critical on-site costs. Evidence exists, however, that the relationship is positive. The available data implies that as travel costs increase the critical on-site costs will increase, but at a decreasing rate. It is reasonable to expect that one recreationist that travels a longer distance than another recreationist, i.e., has a higher value of k, would be willing to pay a higher on-site cost than the other. This is partly due to the difference in magnitude of the two types of expenses. The travel cost requires a much larger outlay than the daily on-site costs. After a person spends, say, \$20 in travel costs in order to recreate, he would be willing to pay, say, \$3.00 per day to remain at the site. Another person, on the other hand, who had to pay only \$5.00 for traveling, would be reluctant to spend as much as \$3.00 per day to use the recreational site. Thus a positive relationship between p_1^*

and k is not unreasonable.

Income had, as was hypothesized, a positive influence on p_1^*. The reliability on this relationship is not as high as possible, since the coefficient is significant only at the ten percent level. The evidence indicates, however, that as income increases the critical on-site costs will increase at an increasing rate. This seems reasonable, even though, as discussed above, the income variable used reflects the total income of the family.

The R^2 for the equation was .684, significant at the one percent level, or 68.4 percent of the variation in p_1^* is explained by the six independent variables. The remaining 31.6 percent of the variation is unexplained.

Visits Relationship

Information concerning the number of visits was obtained from aggregating individual observations into county units. Both the number of visits and the population were derived from counties. The travel cost and income of recreationists were averaged within each county in order for these variables to represent county information. Any inference from the sample is limited to a county level, not an individual recreationist. In this study use was made of this predicted equation, not to infer about county reactions, but to sum all the county estimates to explain the total visits at a lake. In the visits relationship, it is assumed that all recreationists react the same to changes in the independent variables, except as is taken into account by the site characteristics. Thus, it is assumed, the estimated number of visits per county would differ, for the four sites, only as the characteristic variables differed. The total number of visits would differ, for the various sites, as the characteristics differed and as the number of counties represented varied.

The estimated visits equation is:

$$V = \underset{}{-67,947.046} + \underset{(842.177)}{7,312.442W^{\#}} + \underset{(2,366.595)}{21,024.198F^{\#}} + \underset{(.068)}{.648Si^{\#}}$$

$$- \underset{(167.595)}{149.953k} - \underset{(.824)}{.118y} - \underset{(.0019)}{.003pop^{\#\#\#}} \qquad R^2 = .874^{\#} \qquad (40)$$

The site characteristics, referred to by W, F, and Si, are as defined above, k refers to the average travel cost of recreationists within the counties, y represents the average income of recreationists in the counties, and pop refers to the total population of persons residing in each county. The standard error of the coefficients are indicated in parentheses beneath the coefficients. The level of significance of each coefficient is indicated by the procedure discussed in the preceding sections.

As hypothesized, the number of visits from a county to a recreational site will increase as the characteristics improve. Based on the sample information, one could expect 648 more visits to one site than another for every 1, 000 acres larger. This relationship seems plausible, however it seems intuitively feasible that this relationship would not remain constant for any size lake. Economies of size would be expected to be reached after the size considered becomes so large. As soon as one considers a lake that is sufficiently large to easily accomodate water skiiers, boat ramps, swimming areas, and all other water related activities, it would do little to consider a larger lake. There seems to be an optimal size, above which little is gained. This points to a criticism of the present study's use of lake size as a characteristic. This problem could better have been handled by a dummy variable indicating a small, medium, or large lake, instead of using actual acres to represent size. This problem will be further illustrated in the next chapter. All of the site characteristics were significant at the one percent level.

As travel costs of recreationists increase one would expect fewer visits from a given county, everything else held constant. This was observed from the sample data, indicated by the negative coefficient on the k variable. If the travel cost of recreationists in an average county were increased one dollar, it is hypothesized that 150 fewer visits would be observed from that county. The coefficient on travel cost was not significant at the ten percent level. The simple correlation coefficient between travel cost and population was .711, thus indicating the fact that as the travel cost increased, or alternatively, as the distance to the site increased, the population of the counties increased. There is no causality indicated in this correlation, however due to the unique location of the recreational sites, i. e., low population counties nearby, and high population counties at further distances, the population was observed to be higher at larger values of k.[38]

When either population or travel cost was included in the estimated equation, the introduction of the remaining one explained little additional variation. This fact points to the reason why neither population nor travel cost was very significant. The negative coefficient on the population variable is further evidence of this intercorrelation, since fewer visits were observed at higher values of k, these also being associated with high values of population. Thus where few visits were observed, large population counties were also observed.

It was hypothesized that as the average income of recreationists from a county increased, so would the number of visits. In

---

[38]Johnston and Pankey (1968), observed this same phenomenon. They derived a correlation coefficient between population and distance of .891. The sign on their population variable was positive. The reservoirs studied were located in California, thus relatively near the sites were counties with high populations, while the counties at further distances were characterized as lower population counties. In their equation that is similar to the visits model in this study (p. 33), income was not a significant variable. This lends further evidence that the problems in the formulation in equation (40) are not unique to this study, i. e., the multicollinearity between income and site characteristics and between population and distance (or travel cost).

equation (40) income (y) has an estimated negative coefficient. This implies a decrease in the number of visits as the income increases. In order to conclude that significant evidence exists to refute the hypothesis, the t-value of the coefficient should be examined to determine at what level it is significantly different from zero. The coefficient of the income variable is not significant at the ten percent level; thus, not enough information exists to reject the hypothesis. A high level of correlation between W (the swimming, water skiing, and boating use-intensity variable) and y (income) was encountered (.694). The indication is that if W is included in the relationship, the existence of income means little, i.e., it fails to reduce the variation in V significantly.

The two statistical problems mentioned above--the high correlation between k and pop, and between y and W--force two possible alternative formulations of the equation. The first solution would be to delete both the income and the population variables, since neither add much to the equation. The resulting equation was estimated as:

$$\begin{aligned} V = -71,166.121 &+ \underset{(460.199)}{7,141.764W^{\#}} + \underset{(1,651.425)}{19,825.384F^{\#}} + \underset{(.059)}{.641Si^{\#}} \\ &- \underset{(115.473)}{379.786k^{\#}} \qquad R^2 = .868^{\#} \end{aligned} \tag{41}$$

The variables retain their original definitions. It should be noticed that all coefficients are significant at the one percent level. Travel cost was greatly improved after removing the population variable. The R^2 was not reduced significantly, indicating that the predictive qualities have not been dampened.

The second possible procedure in which to solve the multicollinearity problem would be the following. The income variable, as before, should be deleted. Instead of deleting the population variable, however, it was combined with the visits variable, i.e., visits was divided by population. The relationship would now predict the number of per capita visits, V/pop, for counties. The definition of the remaining variables were compatable with the new formulation. In this manner the statistical assumption--that of independence--which was violated, may be valid. The estimated per-capita relationship is:

$$V/pop = -.608 + \underset{(.007)}{.091W^{\#}} + \underset{(.025)}{.190F^{\#}} + \underset{(.0000009)}{.00001Si^{\#}} - \underset{(.002)}{.009k^{\#}} \qquad R^2 = .843^{\#} \tag{42}$$

V/pop refers to the per-capita visits from a county, and the remaining variables are equivalent to those defined earlier. All of the independent variables were significant at the one percent level, indicating that the multicollinearity has been eliminated. The signs on all of the coefficients in equation (42) are as hypothesized.

Caution should be given, at this stage, not to confuse the definition of per-capita visits with inferences toward the population of recreationists. In this equation per-capita visits refers to the total population in a county, not the population of recreationists. If it is assumed that the same percentage of the total population in each county were recreationists, then use of the total population would be appropriate in a regression problem, since it would represent a one-to-one transformation on the dependent variable. Since the total number of recreationists in each county was unknown, it had to be assumed that the same percent of each county's population were recreationists.

Both procedures to estimate visits were considered and the one that best explained the total number of visits at each lake was used. Equation (41) was used to estimate the number of visits for each county, and then all estimates were added together to obtain a total for the lake. This procedure was applied for each lake.

Equation (42), on the other hand, is designed to predict per-capita visits. Per-capita visits for each county were estimated and then multiplied by the population in the county to obtain an estimate of the number of visits, i.e., (V/pop)(pop) = V. The estimated visits for each county were summed to obtain an estimate of the total number of visits at each lake.

Days Per Visit Relationship

Values of k^* and p_1^* were predicted for each of the respective groups by assigning values of the independent variables that were pertinent for each group. For example in the k^* equation certain values of the characteristic variables were fixed depending on the lake under question, i.e., when predicting values of k^* for the groups within Klamath Lake the Klamath Lake characteristics were substituted into the equation, leaving only p_1 and y as variables. Then for each group the average on-site costs and average income of the group were held constant and an estimate of k^* was obtained for each group. A similar procedure was adopted for predicting p_1^*.

Individual observations of the length of stay per visit, q_1, travel cost, k, and on-site costs, p_1, were used for the predicted equation in the following manner. Each recreationist's travel cost, k, was subtracted from the maximum k for the group, k^*; similarly each observation's on-site cost, p_1, was subtracted from the maximum p_1 for his group, p_1^*. The two differences, i.e., $(k^* - k)$ and $(p_1^* - p_1)$, were used as independent variables, and regressed with q_1 as the dependent variable. Only 34 unique values of k^* and 34 unique values of p_1^* existed, due to the necessity of classifying the data into groups. That is, there were 34 groups both for determining k^* and p_1^*. There were, however, 304 observations on k and p_1. Thus, 304 unique differences existed for each of the two independent variables. Thus, the regression was performed on 304 observations.

The demand relationship was estimated by ordinary least squares as:

$$q_1 = 16.941 - \underset{(.033)}{.186}\,(k^* - k)^{\#} - \underset{(.361)}{.805}\,(p_1^* - p_1)^{\#\#} \qquad (43)$$

$$R^2 = .101^{\#}$$

Further examination of the data revealed the existence of two distinct groups of recreationists. One group spent less than three weeks at the site per visit and had relatively high values of $(p_1^* - p_1)$. That is, their on-site costs were low compared to what was predicted they would be willing to pay. On the other hand were recreationists that spent a period of time at the site per visit, up to 169 days, and had relatively low values of $(p_1^* - p_1)$, i.e., they were spending near their critical level for daily expenses on the site. Most of the recreationists in this category were retired people spending either the entire season, or most of the season, at a site. Some were individuals spending a month or more, e.g., teachers with no job commitment during the summer.

Due to the extreme divergence in individuals, it was assumed that the recreationists in one category would react quite differently in response to their on-site and travel costs, than persons in the other group. In fact it was hypothesized that the recreationists remaining a considerable length of time per visit would not respond significantly to the price stimulus. That is no relationship was hypothesized between the length of stay per visit and the two independent variables $(k^* - k)$ and $(p_1^* - p_1)$.

The estimated relationship for the recreationists with less than 20 days per visit is:

$$\ln q_1 = .759 - \underset{(.0018)}{.0064} (k^* - k)^{\#} + \underset{(.0189)}{.0637} (p_1^* - p_1)^{\#} \qquad R^2 = .113^{\#} \tag{44}$$

The average length of stay per visit for recreationists in this group was 3.4 days.

It was hypothesized that as the difference between the critical travel cost and the actual travel cost increased so would the days a recreationist would remain at the site per visit. In equation (44) a negative coefficient was estimated, which was significant at the one percent level, for $(k^* - k)$. In other words, evidence suggests that, for a given critical travel cost, as a person's actual travel cost increases he will tend to recreate more days per visit. Two explanations seem apparent. First, as was mentioned earlier, the appropriate income in the budget constraint depicted in the theory was $(y - k)$. As k increased less income was available to spend on the site after arrival--thus fewer days would be forthcoming. The problem is that the income as measured in this study is so much larger than the travel cost, that the travel cost does not significantly reduce the income available for expenses at the site, $(y - k)$. Second, since the majority of a person's expenses for a recreational trip are involved with travel costs, after he has invested a considerable amount he is more likely to stay a longer time at the site, at a relatively inexpensive rate per day, than he is to spend fewer days. If the daily costs at the site became competetive for the recreationist's income, people would probably begin reacting differently to increases in travel costs.

It was conceived that as the daily on-site costs increased, given a fixed p_1^*, fewer days per visit would be observed. This hypothesis was substantiated by the sample data. The $(p_1^* - p_1)$ variable was significant at the one percent level.

The R^2 for equation (44) was equal to .113, which was significant at the one percent level. Explaining only 11.3 percent of the variation in q_1 doesn't seem very useful, at first sight, but when the number of observations is very large much variation is expected. Even if a regression equation predicted quite well, i.e., if all of the points were very close to the estimated line, there would be much more unexplained variation with a large number of observations than with fewer. There were 282 recreationists in the group who recreated less than 20 days per visit.

The remaining recreationists, i.e., those who remained at the site for more than 20 days, were considered in a separate relationship. There were 22 observations involved in the following estimated equation:

$$q_1 = 75.642 - \underset{(.234)}{.132}\,(k^* - k) - \underset{(2.975)}{3.427}\,(p_1^* - p_1) \qquad R^2 = .127 \tag{45}$$

where the variables are as defined above, and the standard error of the coefficients are noted in parentheses.

Neither of the coefficients of the independent variables were significantly different from zero, except at extremely high levels. Both variables indicate a negative relationship, but due to the low level of significance, mean very little. The coefficients, being near zero, could well have been estimated as positive. The recreationists in this category, it was concluded, do not respond to the variables analyzed. It was also noted that each lake contained approximately an equal number of people who stayed long periods of time. To some retired people the amount spent at a site would have been allocated at home had they not recreated, so their actual price of recreating was near zero.

For the above reaons, the latter group of recreationists, i.e., those with $q_1 > 20$ days, was not regarded as important for the purposes of this study. They add value to each lake, but if a change in value is being analyzed it would be irrelevant to consider these recreationists, since no evidence existed to conclude their behavior would vary with a change in water quality.

The demand model for each lake, and an estimate of the recreational value of an improvement in Klamath Lake will be evaluated in Chapter VII.

CHAPTER VII

APPLYING THE STATISTICAL RECREATION DEMAND MODEL

In this chapter a closer inspection is made of the predicted equations of Chapter VI. The estimated relationships will be presented as applied specifically to each of the recreational sites. Assumptions were made, throughout the analysis, regarding the similarity of reactions of individuals, regardless of the site at which they were recreating, except as the differences in lakes were represented by the characteristic variables. If a lake's characteristics were substituted into equations (38) through (42), from Chapter VI, the resulting relationships would be appropriate for that lake. Equation (43) can be transformed into an equation representing a particular lake by holding the k^*, p_1^*, and k variables at their mean values, i.e., the mean values associated with that lake.

In this manner the demand model for each lake can be viewed separately. The demand model for a site, as discussed in Chapters III and IV, consists of the combination of the k^*, p_1^*, and q_1 equations, and then multiplied by the estimated number of visits, V. The result will represent the aggregate demand model for a recreational site. The demand function for an average individual recreationist is attained by holding the independent variables in the q_1

relation at their means, and observing the relationship between daily on-site costs, p_1, and days per visit, q_1. The independent variables in the p_1^* relationship should also be held at their mean values, and an estimated p_1^* for the average individual will be forthcoming. The demand function, in general terminology, for an average individual, per visit, at a particular lake will take the form of the one in Figure 4. In Figure 4, $\bar{p}_1^*$ is the critical on-site costs for the average individual, $\bar{p}_1$ represents the amount of on-site costs the average recreationist spent per day, q_1^* is the length of stay per visit if the daily on-site costs were at the critical level, and $\bar{q}_1$ is the resulting number of days, per visit, the average recreationist would stay if confronted with daily variable costs of $\bar{p}_1$.

Several interpretations of value are possible using the concept of consumer surplus. First, the value per visit is equivalent to the area $\bar{p}_1 \bar{p}_1^* AB$, in Figure 4, or in terms of a mathematical derivation, the area:

$$\int_{\bar{p}_1}^{\bar{p}_1^*} [f(p_1)\,dp_1] = \text{Value per person per visit} \ . \tag{46}$$

To determine the total value of a site it is necessary to multiply the per-visit value by the total number of visits. Since two possible predictive equations exist for expressing the number of visits, a choice had to be made. Each equation was used to estimate the original data, i.e., to acquire an estimate of the number of visits

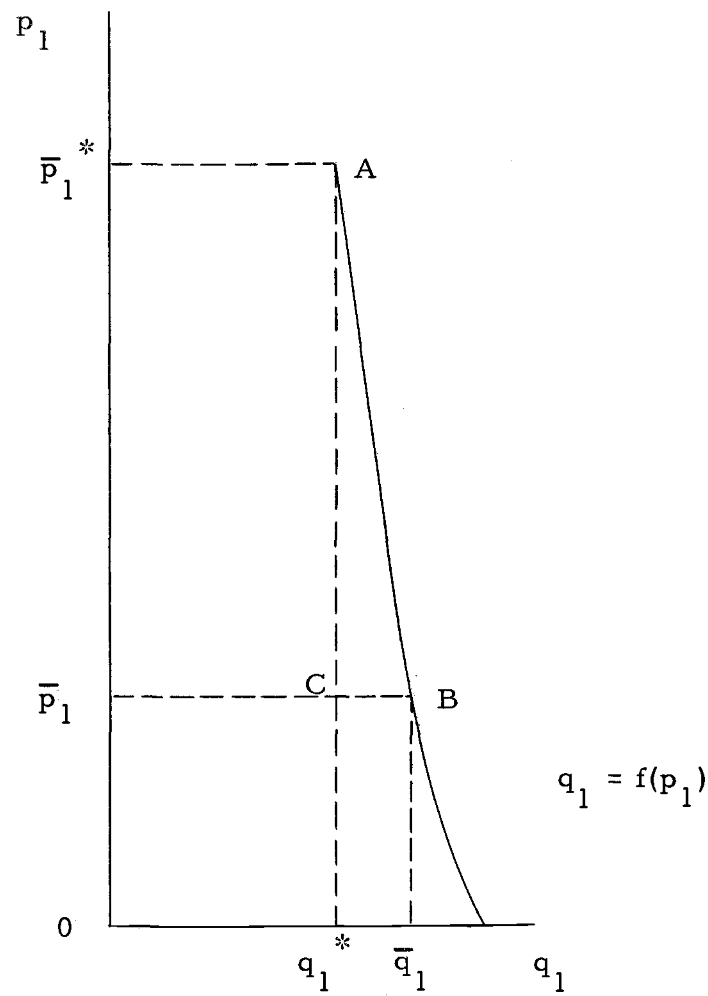

Figure 4. A general illustration of an average individual's demand curve, per visit, for a particular lake.

at each lake. Values, per county, were assigned the independent variables and all county estimates were added, for each site. The total represents the number of visits per lake. The two estimates were compared to the number of visits estimated by the U. S. Forest Service. The results of the comparison are summarized in Table III.

The choice was made abvious by the comparison. It is not certain why the per-capita visits equation does such a poor job explaining the data. One possible explanation would be that it is designed to estimate per-capita visits, not total visits. It does not appear that this would make that much difference, however, it may. Another possibility is that the manner in which the statistical problem was purported to be solved, was not legitimate. That is, dividing population into the dependent variable may not be a procedure to solve a multicollinearity problem. In any case, the per-capita procedure to estimate the number of visits at each lake, is unacceptable.

One problem is still prevalent in the visits equation. The total number of visits for the study area, i.e., the sum of each lake's visits, is accurately explained by the equation. The visits by lakes, however, are not so precisely estimated. Some lake's visits are underestimated while others are over-estimated; on the average an accurate estimate can be obtained. It was of interest in this study to utilize the individual lake's estimated visits. If the predicted visits were used for each lake the relationship between lakes would be lost,

TABLE III. A COMPARISON OF THE PREDICTIVE PROPERTIES OF TWO EQUATIONS EXPLAINING THE NUMBER OF VISITS: RELATED TO THE U. S. FOREST SERVICE ESTIMATES.

Lake	Number of Visits as Estimated by the U. S. Forest Service	Number of Visits as Estimated by V = f(W, F, Si, k)	Number of Visits as Estimated by Per-Capita Model V/pop = f(W, F, Si, k)
Klamath	146, 491	243, 211	1, 579, 859
Lake of the Woods	266, 327	335, 331	451, 758
Odell	180, 304	14, 968	-2, 468, 222
Willow	109, 471	110, 361	-3, 089, 331
Total	702, 593	703, 871	-3, 525, 936

since one figure was substantially under-estimated while the remaining ones were over-estimated. The particular manner in which the characteristics were represented in the equation, it is thought, is responsible for the lack of predictive power at each site. Too large a variation was observed in the values assigned to the characteristics. For example, lake size varied from 320 acres to 98, 560 acres; swimming, water skiing, and boating intensity ranged from 1 to 9; etc. A direct and constant relationship between visits and the characteristics is implied. That is, by this assumption, one could expect size to influence the number of visits proportionately. So, as very high values of the characteristics are observed, a very large estimated number of visits will be attained. If the values of the characteristics are generally low, then a very low estimate of the number of visits will be forthcoming. This phenomenon can be seen to exist in this study.

To avoid this problem, the original U. S. Forest Service estimates were employed to obtain the aggregate demand function, and total value, of each site. When it became necessary to estimate the change in total visits at Klamath Lake due to a change in water quality, an additional step was necessary. The estimated equation was employed to predict the number of visits subsequent to a change in water quality, by assigning new values of the characteristics in the equation. The new values of the characteristics are larger than the original, and since the coefficients of these variables are positive in the visits relationship, an increase in the number of visits will be postulated. The increase in the estimated visits should be expressed as a percent change. This percentage should then be applied to the Forest Service estimate. In this manner an estimated change in the number of visits can be obtained by use of the visits equation. It was assumed that even though the predicted equation could not accurately predict each lake's visits, it could represent a percent change adequately. This was assumed because the total visits for the study area was so closely estimated.

The demand model for each lake and the value accruing to each lake will now be discussed.

Lake of the Woods

The use-intensities of the activities and the size of Lake of the Woods were substituted into the k^*, p_1^*, and q_1 equations. The three estimated relationships became:

$$k^* = 23.617 + .0020\, y + 10.435\, p_1 \tag{47}$$

$$p_1^* = 2.458 + .269\, k - .004\, k^2 + .000000017\, y^2 \tag{48}$$

$$q_1 = e^{.759 - .0064\, k^* + .0064\, k + .0637\, p_1^* - .0637\, p_1} \tag{49}$$

The demand model for Lake of the Woods is represented by the above three equations. Recreationists stayed at Lake of the Woods for an average of approximately four days and spent an average of $2 per day. Travel cost for recreationists at Lake of the Woods averaged about $15. On the other hand, the individuals using this site had incomes averaging just over $10, 500. These recreationists, it was estimated, would be willing to pay up to $6 per day and still prefer to remain at the site; while, they would be indifferent to recreating at a travel cost of about $66.

When the above information was included in the demand relation, a demand function was obtained as follows:

$$q_1 = e^{.821 - .0637 p_1} \tag{50}$$

The following geometric interpretation, Figure 5, helps explain the demand function for an individual recreationist, per visit, at Lake of the Woods. The value per visit can be calculated, as described earlier in this chapter, as:

$$\text{Value per visit} = \int_{2.01}^{6.10} (e^{.821 - .0637 p_1})\, dp_1 \tag{51}$$

$$= \$7.17$$

The value per visit is approximately $ 7. The 1968 seasonal value of Lake of the Woods was obtained by multiplying the per-visit value by the number of visits for 1968:

$$(\$\ 7.17)(266, 327) = \$1, 909, 565 \quad .$$

The proper interpretation of this value should be made. The $1.9

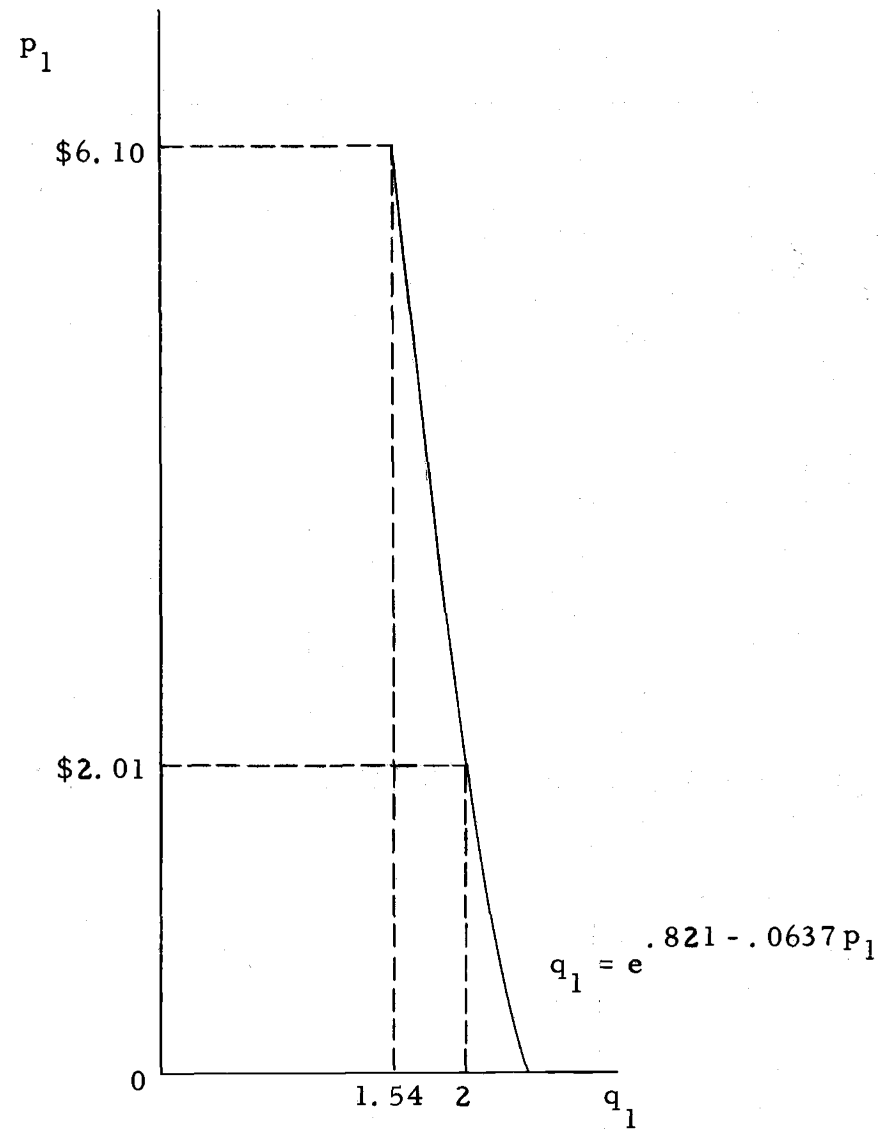

Figure 5. An illustration of the average recreationist's demand curve, per visit, for Lake of the Woods. Values of $\overline{p}_1^*$, $\overline{p}_1$, q_1^*, $\overline{q}_1$, that are applicable to Lake of the Woods, are placed in the diagram.

million value refers only to those recreationists that visited the site for less than 20 days. The relationship between daily cost and number of days for those persons staying more than 20 days could be represented by a vertical line. No reaction to price was observed, thus indicating that at any price they would remain as long as they did. To compute the value attributable to the lake from these recreationists would require multiplying their daily cost by the number of days at the site. Since there were only nine persons sampled in this category, the few dollars added to the $1.9 million would mean very little.

The derived value of Lake of the Woods refers to the 1968 season. Any inferences concerning estimated value for subsequent years requires additional assumptions. It must be assumed that the reactions of recreationists remain constant, and the measured variables either are not altered, or the exact change can be specificied. Any use of the lake made in other seasons, i.e., other than the summer season, was not considered in this value estimate. For example, ice fishing, duck hunting, etc., that is performed at Lake of the Woods in the fall or winter was excluded in this study.

The computed values of the remaining lakes are subject to the above interpretations.

Odell Lake

The demand model for Odell Lake, upon evaluating the appropriate functions at the characteristic values associated with Odell Lake, is:

$$k^* = -18.076 + .0020\,y + 10.435\,p_1 \tag{52}$$

$$p_1^* = 1.641 + .269\,k - .004\,k^2 + .000000017\,y^2 \tag{53}$$

$$q_1 = e^{.759 - .0064\,k^* + .0064\,k + .0637\,p_1^* - .0637\,p_1} \tag{54}$$

At Odell Lake, recreationists spent an average of \$2.34 daily at the site, and approximately \$11 in travel cost. They had an average income of \$9,063, and were estimated to have critical on-site costs, p_1^*, and travel costs, k^*, averaging \$5.57 and \$24.71, respectively. The incorporation of the above information into the demand model enables a closer look to be taken at the demand function:

$$q_1 = e^{1.025 - .0637\,p_1} \tag{55}$$

The geometric view of the demand function for a single visit for Odell Lake is given in Figure 6. The value per visit for 1968, was calculated as:

$$\text{Value per visit} = \int_{2.34}^{5.57} (e^{1.025 - .0637\,p_1})\,dp_1 = \$7.17 \tag{56}$$

Odell Lake is worth approximately \$7 per visit, on the average for

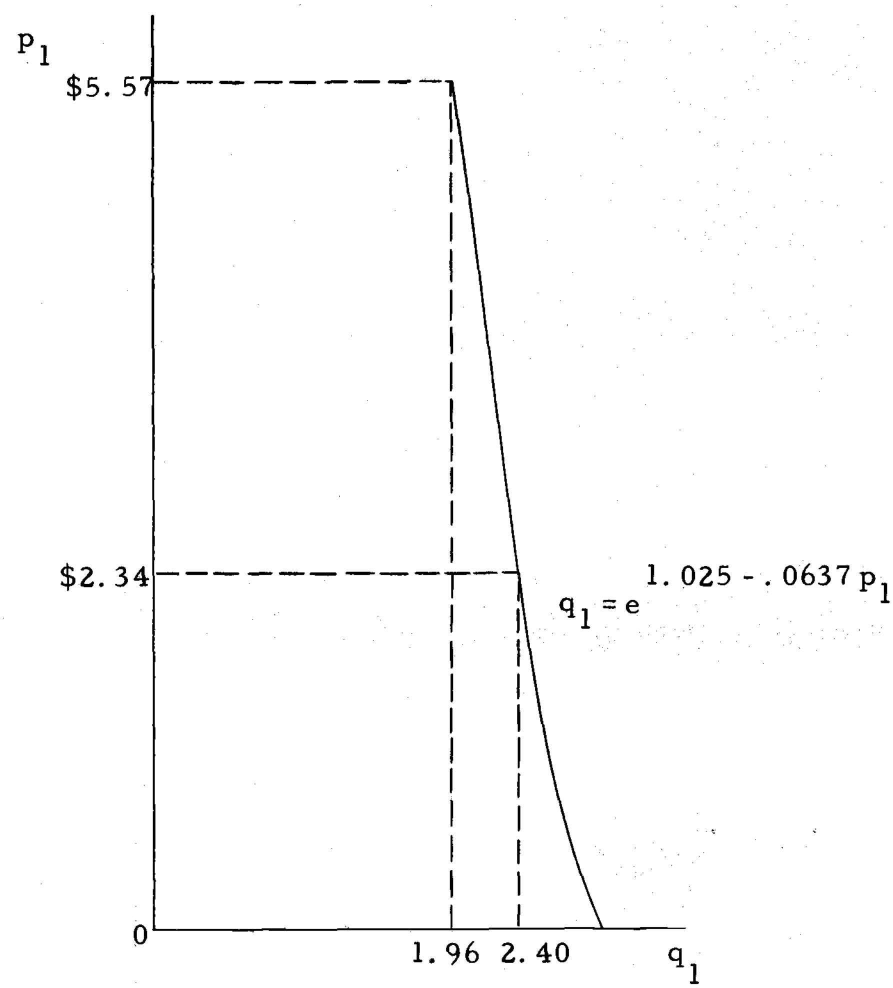

Figure 6. An illustration of the average recreationist's demand curve, per visit, for Odell Lake. Values of $\bar{p}_1^*$, $\bar{p}_1$, q_1^*, $\bar{q}_1$, that are applicable to Odell Lake, are placed in the diagram.

1968. When this figure was multiplied by the number of visits for the 1968 season a total value for 1968 was obtained:

$$(\$7.17)(180,304) = \$1,292,780 \quad .$$

The total value of Odell Lake, for the summer season in 1968 is about \$1.3 million, excluding the recreationists staying more than 20 days. Due to the fact that only five recreationists were observed at Odell Lake staying more than 20 days per visit, the exclusion of these persons will not significantly affect the estimated value.

Willow Lake

Willow Lake is the smallest of the four lakes in this study, but has approximately average values of the use-intensities for the various activities. When these characteristics are held constant in the k^*, p_1^*, and q_1 equations, the demand model applicable to Willow Lake becomes:

$$k^* = -6.534 + .002\, y + 10.435\, p_1 \tag{57}$$

$$p_1^* = 2.988 + .269\, k - .004\, k^2 + .000000017\, y^2 \tag{58}$$

$$q_1 = e^{.759 - .0064\, k^* + .0064\, k + .0637\, p_1^* - .0637\, p_1} \tag{59}$$

The recreationists at Willow Lake recreated an average of nearly two days per visit, spending about \$1.72 per day while visiting. They had average incomes of \$7,790 and incurred approximately \$6 for travel cost per visit. The estimated critical travel cost averaged nearly \$31, while the estimated critical daily cost averaged about \$4.50. When p_1^*, k^*, and k were assigned average values applicable to Willow Lake, the demand function for an individual, per visit, was:

$$q_1 = e^{.888 - .0637\, p_1} . \tag{60}$$

An analysis of this relationship can help decision-makers determine what courses of action would be economically feasible, given the estimated reactions of individuals to on-site costs. A geometric interpretation may enhance the concept of "value" as derived for Willow Lake. The demand function, with p_1^* incorporated, for an average recreationist per visit, is similar to the one in Figure 7. The value per visit, on the average, was calculated as follows:

$$\begin{aligned} \text{Value per visit} &= \int_{1.72}^{4.55} (e^{.888 - .0637\, p_1})\, dp_1 \\ &= \$5.64 \end{aligned} \tag{61}$$

The value of Willow Lake to an average individual visiting the site during 1968 was approximately \$6. Expanded by the total estimated number of visits for the summer of 1968, an estimate of the seasonal economic value of Willow Lake was obtained:

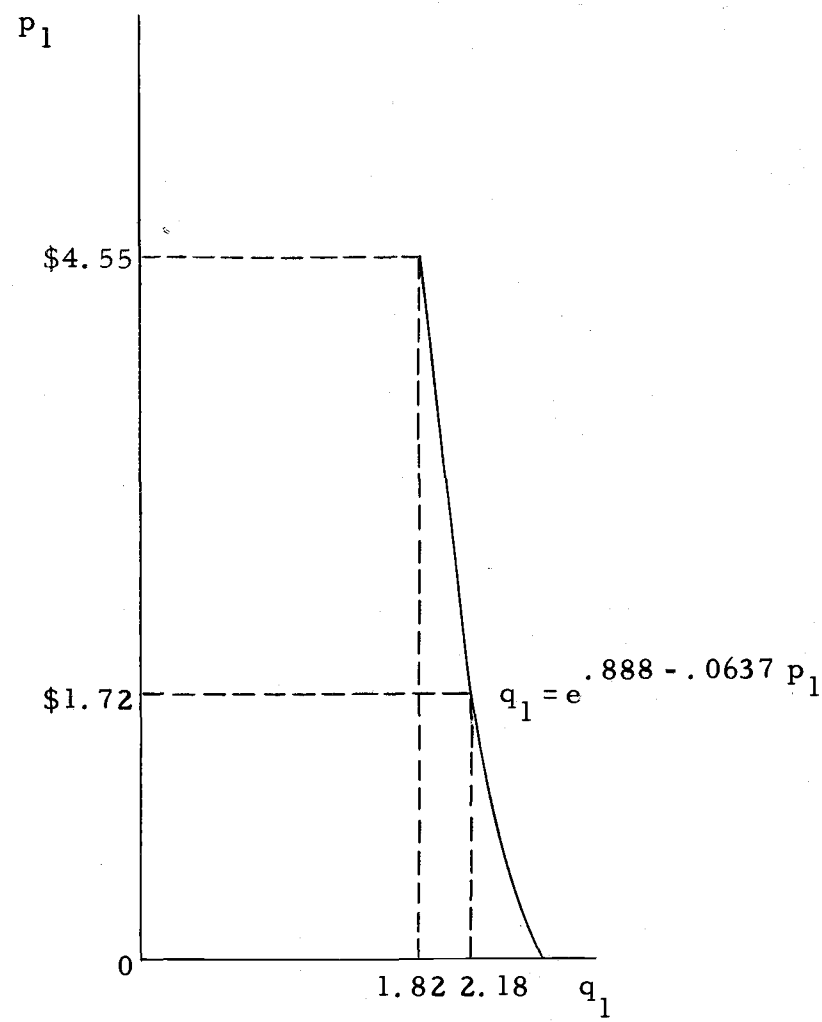

Figure 7. An illustration of the average recreationist's demand curve, per visit, for Willow Lake. Values of $\overline{p}_1^{\,*}$, $\overline{p}_1$, $q_1^{\,*}$, $\overline{q}_1$, that are applicable to Willow Lake, are placed in the diagram.

$$(\$5.64)(109,471) = \$617,416$$

The total economic value of Willow Lake was estimated at \$617,416 for 1968. Only two recreationists were observed in the sample who remained at Willow Lake for more than 20 days. Excluding that group of recreationists, it is thought, does not seriously hamper the predictions.

Klamath Lake

The Present Situation

The demand model appropriate to Klamath Lake as it now exists is:

$$k^* = 9.132 + .002\,y + 10.435\,p_1 \tag{62}$$

$$p_1^* = 3.531 + .269\,k - .004\,k^2 + .000000017\,y^2 \tag{63}$$

$$q_1 = e^{.759 - .0064\,k^* + .0064\,k + .0637\,p_1^* - .0637\,p_1} \tag{64}$$

where the values of the site characteristics were held constant in the k^*, p_1^*, and q_1 equations presented in Chapter VI. An average of 1.6 days per visit was spent at Klamath Lake, with a daily expense per recreationist of \$1.84. The average recreationist had an annual income of \$8,900, and allocated about \$6.80 to travel costs, per visit.

Average critical values of travel cost and on-site costs were estimated at approximately \$55 and \$5.50, respectively. Upon fixing the p_1^*, k^*, and k variables at their means in the q_1 relation, the demand function for Klamath Lake, as it now exists, can be viewed.

$$q_1 = e^{.801 - .0637\, p_1} \tag{65}$$

The demand function, is shown in Figure 8. The economic value of Klamath Lake, as it exists for the summer of 1968, can be computed by the use of the demand curve, as was done in the preceding sections.

$$\text{Value per visit} = \int_{1.84}^{5.54} (e^{.801 - .0637\, p_1})\, dp_1 + \tag{66}$$

$$= \$6.37$$

The economic value of Klamath Lake, per visit for recreationists, was estimated at approximately \$6 There was an estimated 146,491 visits to Klamath Lake during the past season, so the total economic value of the lake for 1968, to recreationists is approximately \$933,148, i.e.,

$$(\$6.37)(146,491) = \$933,148$$

Only six recreationists were observed at Klamath Lake that remained more than 20 days at the site, per visit. The exclusion of

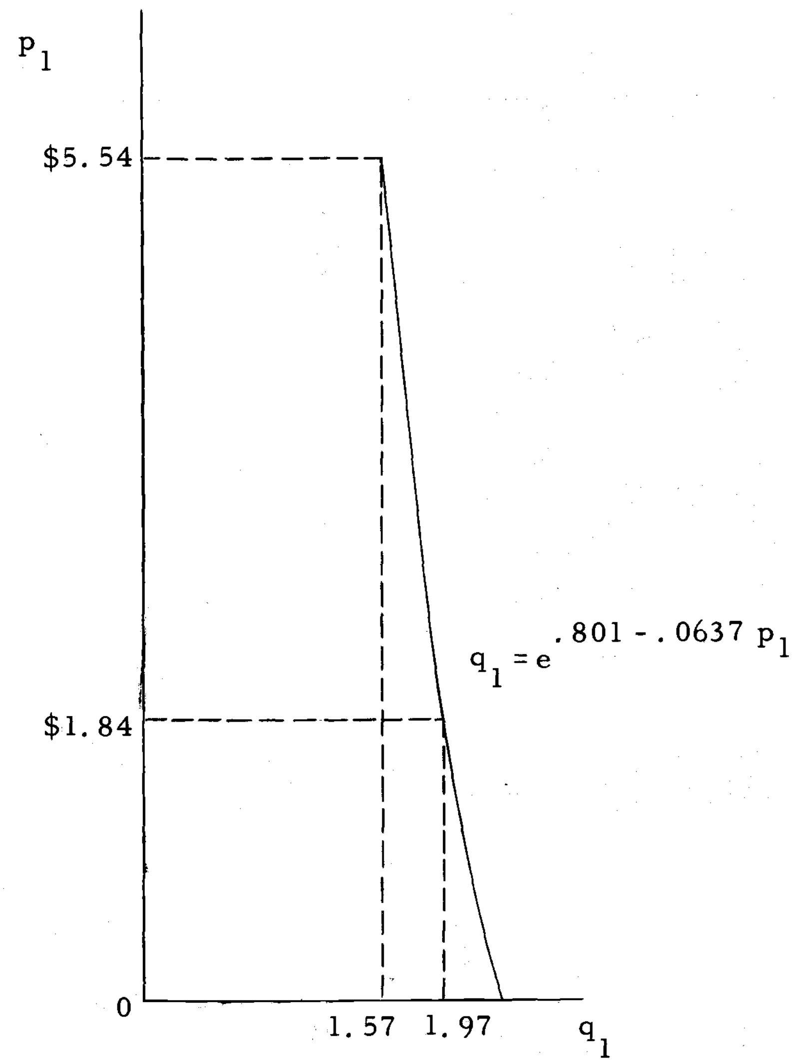

Figure 8. An illustration of the average recreationist's demand curve, per visit, for Klamath Lake, as it now exists. Values of $\bar{p}_1{}^*$, $\bar{p}_1$, $q_1{}^*$, $\bar{q}_1$, that are applicable to Klamath Lake, as it now exists, are placed in the diagram.

these is not considered serious.

A Change in Water Quality

A possible improvement in water quality in Klamath Lake needs to be considered at this point. The proposed change would probably consist of a two-step plan of action, according to the personnel at the F. W. P. C. A. Pacific Northwest Water Laboratory in Corvallis, Oregon. Step 1 would attempt to remove the blue-green algae present in Klamath Lake. It is well beyond the scope of this thesis to be concerned with the manner in which the algae is to be removed. The results presented are invariant with respect to the exact procedures involved in removing the algae.

The second step would attempt to lower the water temperature and improve the beaches. The temperature could be lowered in many ways, but is not considered in this thesis. It would be desirable to lower the water temperature so the blue-green algal growth would not be as stimulated, however the temperature should not be lowered enough to discourage swimming. The beaches have never been used extensively for swimming since they are quite muddy. If, however, an adequate beach could be developed it would greatly enhance swimming participation, as well as other water related activities.

The effect on the economic value of Klamath Lake due to the change in water quality will be discussed in two parts--corresponding to each of the two proposed steps, discussed above.

Klamath Lake (Step 1)

The swimming, water skiing, and boating use-intensity index was projected to change from two to seven, upon the removal of the blue-green algae. To break this variable into its component parts would be helpful in analyzing the changes involved. Swimming would be expected to increase from no use to medium use, while water skiing would increase from low to high use. Boating, on the other hand, would change from low use to high use, while fishing would increase from low to medium use, with a postulated removal of the blue-green algae. These estimates were based on consultations with the personnel at the F. W. P. C. A. Pacific Northwest Water Laboratory in Corvallis, and the U. S. Forest Service in Klamath Falls, Oregon.

An estimate of possible changes in other variables in the demand model also had to be made. Other variables that had to be analyzed were p_1 (on-site costs), k (travel costs), and y (income). Would a significant change in the average daily costs, travel costs, or income of recreationists occur at Klamath Lake due to a change in water quality? This can be evaluated as a testable hypothesis. The methodology, to test this hypothesis, that was proposed by this thesis was as follows. The average value of p_1, k, and y were computed for each lake, these were then utilized as dependent variables. These variables were regressed against the site characteristics. The site characteristics were represented by summing the swimming, water skiing, boating, and fishing use-intensities for each lake. Size was not used as a variable since a postulated change in water quality would not involve a change in the lake size. The results of the regressions were then used to estimate projected changes in the p_1, k, and y variables. By assigning the proposed new values of W and F into the relationships, an estimate of p_1, k, and y that correspond to a higher level of quality would be obtained. The three regression results are:

$$p_1 = 1.638 + \underset{(.046)}{.066\,Q} \qquad R^2 = .507 \qquad (67)$$

where Q indicates the sum of the use-intensities at each lake. The coefficient on Q is not significantly different from zero at the ten percent level. Even though the sign of the coefficient is as hypothesized, not enough observations were available to conclude that a relationship exists between p_1 and Q. The R^2, equal to .507, with only two degrees of freedom, is not significant at the ten percent level.

$$k = 12.075 + \underset{(.323)}{.475\,Q} \qquad R^2 = .519 \qquad (68)$$

The conclusion drawn from the k-Q relation, estimated above, is that not enough evidence has been observed to conclude that a relationship exists between k and Q. Neither the coefficient of Q nor the R^2 were significant at the ten percent level.

$$y = 5,031.295 + \underset{(220.150)}{529.637}\,Q \qquad R^2 = .743 \tag{69}$$

No relationship could be concluded from this estimated equation, since the coefficient on Q and the R^2 were not significant at the ten percent level.

With the available information no possible changes in the p_1, k, or y variables could be postulated, if water quality were improved. The computed relationships do lend confidence that if more observations were available the three variables would increase when water quality was improved. Only changes in W and F were accounted for in Steps 1 and 2.

With larger values of the characteristics, the k^* and p_1^* functions will shift upward due to the positive coefficients on the variables, i.e., k^* and p_1^* will now be larger for the same values of the independent variables. The demand model for Klamath Lake after step 1 is:

$$k^* = 44.151 + .002\, y + 10.435\, p_1 \tag{70}$$

$$p_1^* = 10.060 + .269\, k - .004\, k^2 + .000000017\, y^2 \tag{71}$$

$$q_1 = e^{.759 - .0064\, k^* + .0064\, k + .0637\, p_1^* - .0637\, p_1} \tag{72}$$

The only difference in the k^* and p_1^* functions in this case compared to any other lake is the constant term. The constant term differs only to the extent that the characteristics differ. The q_1 relation, at this point is unaltered. When new average values of p_1^* and k^* are substituted into the q_1 relation, the demand function shifted to the right. The estimated average critical travel cost was \$90.38, and the average critical daily cost was \$12.07. By introducing these averages into the q_1 relation, and by keeping k fixed at the mean value derived prior to step 1, the demand function was estimated as:

$$q_1 = e^{.993 - .0637\, p_1} \tag{73}$$

The difference in this demand function and one previously derived for Klamath Lake is the larger constant term. Thus the demand curve shifted to the right, indicating that the per-visit value has increased. The value per visit was computed, with the aid of a geometric conceptualization of equation (73), Figure 9. The average number of days per visit has increased to approximately two days,

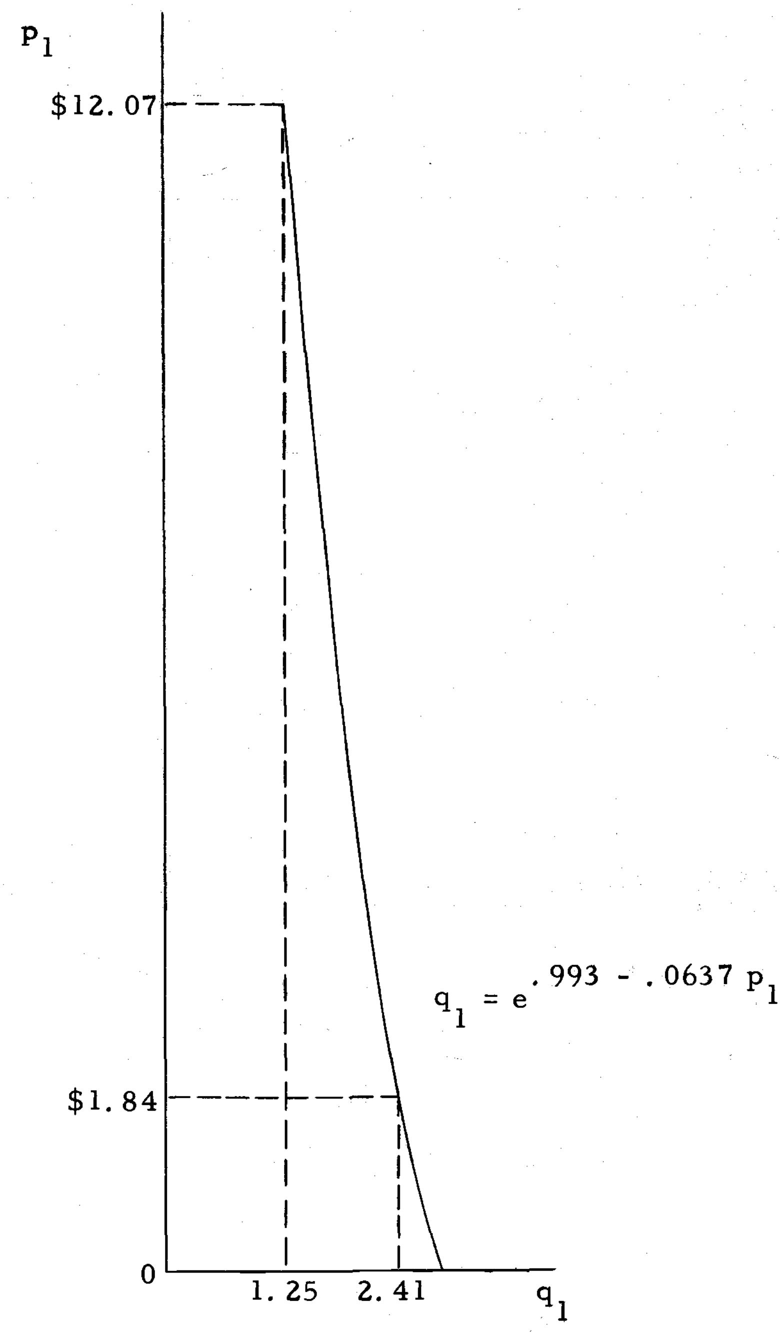

Figure 9. An illustration of the average recreationist's demand curve, per visit, for Klamath Lake (Step 1). Values of $\overline{p}_1^{\,*}$, $\overline{p}_1$, $q_1^{\,*}$, $\overline{q}_1$, that are applicable to Klamath Lake (Step 1), are placed in the diagram.

while the average income, on-site costs, and travel costs remain unchanged.

$$\text{Value per visit} = \int_{1.84}^{12.07} (e^{.993 - .0637 p_1})\, dp_1 \qquad (74)$$

$$= \$18.29$$

The value per visit has increased to \$18.29 an estimated incremental change of \$11.92 after the removal of the blue-green algae. To estimate the new value of Klamath Lake the revised estimated number of visits was multiplied by the value per visit. The revised number of visits was estimated at 234,386, an increase of 160 percent, attributable to the changed water quality. The value of Klamath Lake for 1968, given a removal of algae was estimated as:

$$(\$18.29)(234,386) = \$4,286,920 \quad .$$

The value of Klamath Lake as it now exists was \$933,148, thus the value of removing the algae is represented by the difference in the two estimated values, or \$3,353,772. That is, it would be worth an additional \$3.4 million to society in general, to remove the algae present in Upper Klamath Lake, for the summer season, 1968.

Klamath Lake (Step 2)

The second step of the proposed improvement in the water quality in Klamath Lake is that of lowering the temperature of the water and improving the beaches. The swimming use-intensity was proposed to increase from low, after step 1, to high after step 2. Other changes were hypothesized as: water skiing, no change after step 2; boating, no change after step 2; and fishing from medium to high use after step 2. The swimming, water skiing, and boating use-intensity index would thus rise from seven to nine, and the fishing index from two to three. Other variables, such as daily costs, travel cost, income, and lake size, were unaltered by the proposed second step. The demand model, after reflecting the quality changes, is:

$$k^* = 60.426 + .002\,y + 10.435\,p_1 \tag{75}$$

$$p_1^* = 14.250 + .269\,k - .004\,k^2 + .000000017\,y^2 \tag{76}$$

$$q_1 = e^{.759 - .0064\,k^* + .0064\,k + .0637\,p_1^* - .0637\,p_1} \tag{77}$$

The k^* and p_1^* functions have again shifted, representing an increase in the average k^* and p_1^* for recreationists at Klamath Lake. The revised averages are \$106.65 and \$16.26, respectively. The proposed revised demand function is now:

$$q_1 = e^{1.156 - .0637\,p_1} \tag{78}$$

A shift has been recorded in the demand function, since the constant term is larger than before.

The value, per visit, to recreationists at Klamath Lake after the water quality has been substantially improved is seen more clearly with the use of Figure 10. The per-visit value was calculated as:

$$\text{Value per visit} = \int_{1.84}^{16.26} (e^{1.156 - .0637 p_1}) \, dp_1 = \$26.72 \tag{79}$$

The per-visit value increased another \$8.43 since step 2. That is, it was worth an additional \$8.43 per visit to recreationists to lower the water temperature and improve the beaches. More visits would be projected, attributable to the second step in the procedure. Total visits were projected at 377,947, an increase of 258 percent over the original estimated visits. A new total value of Klamath Lake was computed as:

$$(\$26.72)(377,947) = \$10,098,744 \; .$$

An increase in the economic value of Klamath Lake of \$9,165,596 was estimated allocatable to both steps 1 and 2 for 1968. That allocated to step 2, over and above step 1, was \$5,811,824. That is to say, it would be worth \$9.2 million to society to activate the two-step procedure as estimated by the 1968 data. It would be worth \$3.4 million to undertake the first step and approximately \$5.8 million additional to

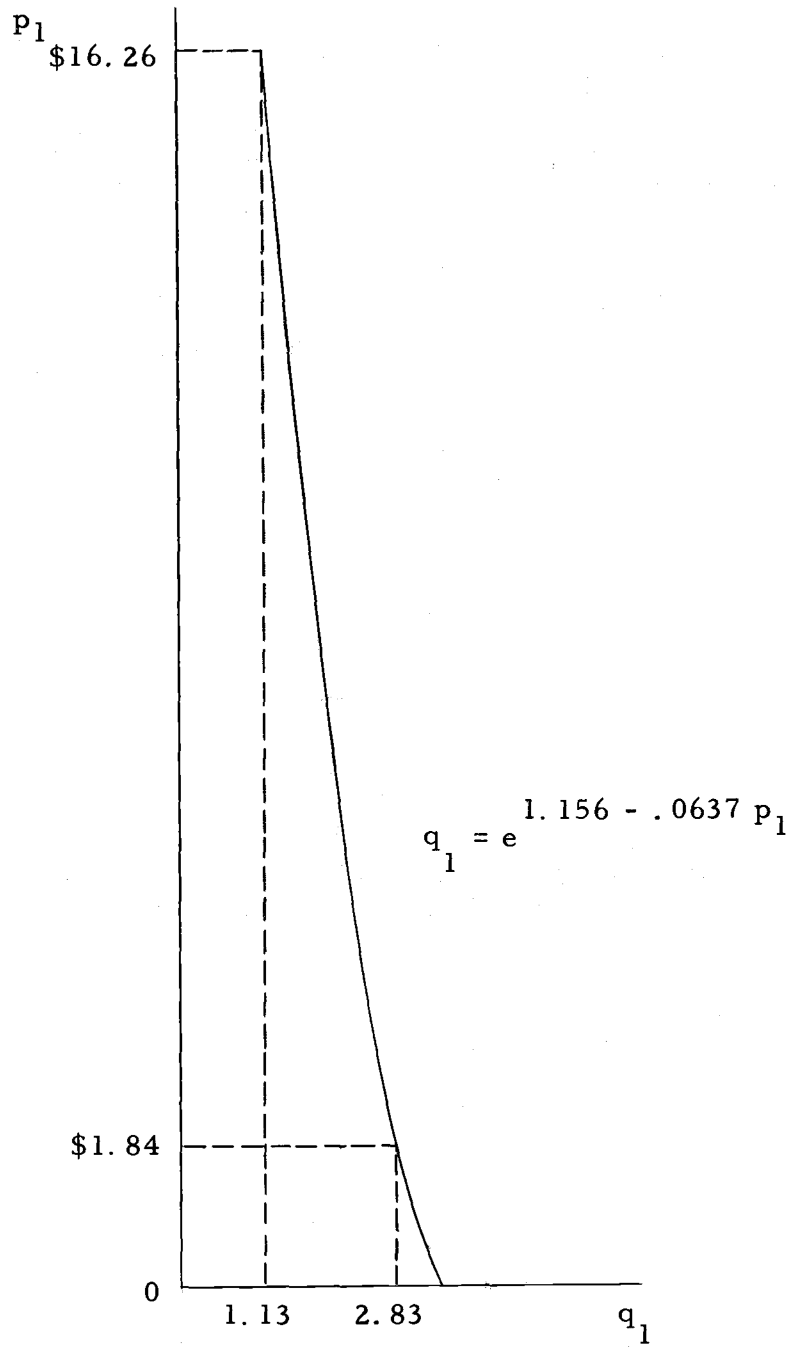

Figure 10. An illustration of the average recreationist's demand curve, per visit, for Klamath Lake (Step 2). Values of $\bar{p}_1^*$, $\bar{p}_1$, q_1^*, $\bar{q}_1$, that are applicable to Klamath Lake (Step 2), are placed in the diagram.

carry out the second proposed step.

A second definition of economic value exists. The amount recreationists spent per visit, $\bar{p}_1\bar{q}_1$, could be added to the consumer surplus to represent a second interpretation of economic value. The estimates of value obtained in this manner would be more appropriate for determining a "gross value" of a recreational site. In the context of a decision-making framework, similar to the one considered in this study, it would be more appropriate to have an estimate of the "net" economic value.

A third interpretation of value that arises with the use of truncated demand curves should be considered. An estimate of the consumer surplus, in a Marshallian context, has been calculated for each recreational site in this Chapter. These estimates represent one interpretation; however, alternative definitions exist when using truncated demand curves. The theoretical framework underlying the truncated demand curve indicates that no levels of q_1 will be observed for lengths of stay less than q_1^*, in Figure 4. The question arises as to whether the demand schedule actually exists in the range of q_1 where $q_1 < q_1^*$. If it does not, the definition of consumer surplus that includes the area bounded by $\bar{p}_1\bar{p}_1^*$AC, in Figure 4, may be questionable. If the portion where $q_1 < q_1^*$ is irrelevant, the estimates of value change considerably (see Table IV). The "true" estimate of value may be represented by an area less than the area bounded by $\bar{p}_1\bar{p}_1^*$ AB, in Figure 4, but larger than the area ABC, Figure 4.

Table IV. A COMPARISON OF TWO ALTERNATIVE DEFINITIONS OF CONSUMER SURPLUS, INCLUDING AND EXCLUDING THE PORTION OF THE DEMAND CURVE WHERE $q_1 < q_1^*$.

Lake	Total Value: Including the Portion Where $q_1 < q_1^*$	Total Value: Excluding the Portion Where $q_1 < q_1^*$
Lake of the Woods	$1,909,565	$ 231,704
Odell	1,292,780	151,455
Willow	617,416	53,641
Klamath	933,148	82,035
Klamath (Step 1)	4,286,920	1,289,123
Klamath (Step 2)	10,098,744	3,941,987

Additional thought needs to be given to the conceptualization and empirical estimation of the relevant consumer surplus applicable to a truncated demand curve. The estimated values of the recreational sites, for the purposes of this study, may be taken as intermediate between the values estimated in Table IV. That is, the estimated value of removing the algae in Upper Klamath Lake, given the water quality parameter assumptions assumed throughout this study, may range from $1.2 to $3.4 million. The predicted value of the second step may lie between \$2.7 and \$5.8 million, while the total value of performing both steps may be limited to the range of \$3.9 and \$9.2 million.

The values of the characteristics for each lake and for Klamath Lake after each proposed step, as well as average values of all the remaining variables, and the estimated number of visits, are contained in Table V.

The next chapter will summarize the procedures and results of this study. The conclusions and limitations of the results will also be discussed.

TABLE V. VALUES OF LAKE CHARACTERISTICS AND AVERAGES OF THE RELEVANT VARIABLES USED IN THE DEMAND MODEL,[1] AND THE NUMBER OF VISITS, LISTED BY LAKES, AND FOR KLAMATH LAKE AFTER EACH STEP IN A PROPOSED CHANGE IN WATER QUALITY.

Item	Lake of the Woods	Odell Lake	Willow Lake	Klamath Lake	Klamath Lake (Step 1)	Klamath Lake (Step 2)
W	9	1	3	2	7	9
F	1	3	3	1	2	3
Si	1,055	3,500	320	98,560	98,560	98,560
p_1^*	6.10 (2.14)	5.57 (2.36)	4.55 (2.05)	5.54 (1.76)	12.07 (1.81)	16.26 (1.77)
p_1	2.01 (2.72)	2.34 (2.86)	1.72 (1.95)	1.84 (2.66)	1.84 (2.66)	1.84 (2.66)
k	15.15 (26.11)	10.86 (11.41)	6.04 (8.05)	6.84 (19.25)	6.84 (19.25)	6.84 (19.25)
y	10,569 (4,576)	9,063 (5,075)	7,790 (3,919)	8,943 (4,583)	8,943 (4,583)	8,943 (4,583)
k^*	66.18 (23.92)	24.71 (28.32)	31.14 (25.93)	55.36 (31.69)	90.38 (31.61)	106.65 (31.61)

TABLE V. (continued)

Item	Lake of the Woods	Odell Lake	Willow Lake	Klamath Lake	Klamath Lake (Step 1)	Klamath Lake (Step 2)
q_1	4.14 (3.91)	3.77 (3.41)	2.03 (1.59)	1.61 (2.27)	2.41 x	2.83 x
V	266, 327	180, 304	109, 471	146, 491	234, 386	377, 947

(Figures in parentheses are standard deviations)

x = No standard deviations available.

[1] The averages computed were based on the 282 observations where the length of stay was less than 20 days.

CHAPTER VIII

SUMMARY AND CONCLUSIONS

Summary

The main objectives of this thesis were: (1) To develop and empirically test a methodology that would be appropriate to determine the economic benefits accruing to society from a recreational facility which is to be developed. (2) To determine the relationship between water quality and recreational use, using the new methodology developed. A prediction of the change in recreational use with a substantial improvement in water quality, in a large body of water, will then be feasible. (3) To determine the economic benefits accruing to society, in general, from the postulated improvement in water quality and the associated increase in recreational use.

The value of recreation can be computed with the use of a recreational demand curve. The estimation of the value of outdoor recreation has proceeded in two general directions: the "direct" method of estimating consumer's willingness to pay, and the "indirect" procedure to determine willingness to pay. Recreationists are asked, directly, to state how much they would be willing to pay for the use of the recreational facility by the "direct" method of estimating recreational value. In the "indirect" approach the reaction of recreationists to changes in costs of travel to a recreational site are observed.

This thesis presented a new form of the indirect approach to the evaluation of recreational resources, which avoids some limiting assumptions of the previous methods. The focal point of the theory was the individual recreationist. The procedure derives one demand function and divides the traditional price variable into two components. Travel costs were treated independent of variable on-site costs. A demand model for the number of days a recreationist will stay at a particular site per visit was derived. The statistical model forthcoming from the theoretical framework was:

$$q_1 = q_1 \left[(k^* - k), (p_1^* - p_1) \right] \qquad \text{for } (k^* - k), (p_1^* - p_1) \geq 0 \tag{80}$$

$$k^* = k^* (p_1, y, Sw, Ws, B, F, C, Si) \tag{81}$$

$$p_1^* = p_1^* (k, y, Sw, Ws, B, F, C, Si) \tag{82}$$

where q_1 is the number of days a recreationist spends at a site per visit, k is the travel cost, k^* refers to the critical, or maximum, travel cost, p_1 is the variable on-site cost, p_1^* reflects the recreationist's willingness to pay on-site costs, and Sw, Ws, B, F, C, Si refer to the recreational characteristics at a site.

The study area under consideration was located in the southwestern section of Oregon, predominately in Klamath County. The study was based principally on Upper Klamath Lake, but three other lakes--Lake of the Woods, Odell, and Willow--were involved.

A sample of recreationists was chosen and personally interviewed at the four recreational sites. Information was sought concerning the total number of days the recreationist visited the particular site, and a detailed account of their expenditures, both in traveling to and from the recreational site, and while visiting the site.

Conclusions

The empirical results consisted mainly of four predicted equations: the critical travel cost relationship, k^*; the critical on-site cost relationship, p_1^*; the visits relationship, V; and the demand relationship, q_1. The predicted equations, for the four lakes, were estimated as:

$$k^* = \underset{}{-36.711} + \underset{(2.322)}{6.248\,W} + \underset{(7.800)}{3.779\,F} + \underset{(.0002)}{.0003\,Si} + \underset{(.0018)}{.0020\,y} + \underset{(3.349)}{10.435\,p_1} \qquad R^2 = .616 \tag{83}$$

$$p_1^* = -7.263 + \underset{(.197)}{7.80\,W} + \underset{(.815)}{2.630\,F} + \underset{(.000025)}{.000067\,Si} - \underset{(.002)}{.004\,k^2} + \underset{(.143)}{.269\,k} \qquad R^2 = .684 \tag{84}$$

$$V = -71,166.121 + \underset{(460.199)}{7,141.764\,W} + \underset{(1,651.425)}{19,825.384\,F} + \underset{(.059)}{.641\,Si}$$

$$- \underset{(115.473)}{379.786\,k} \qquad R^2 = .868 \qquad (85)$$

$$\ln q_1 = .759 - \underset{(.0018)}{.0064}\,(k^* - k) + \underset{(.0189)}{.0637}\,(p_1^* - p_1) \qquad R^2 = .113$$

The variables are as defined above. Each of the four estimated equations applies equally to all four lakes. That is, it was assumed that the only differences in the reactions of recreationists to changes in the independent variables were reflected by the site characteristics. A further assumption was that the site characteristics were adequately depicted by the variables utilized: the use-intensities of water skiing, swimming, boating, camping, and fishing, and the size of the lake.

In order to analyze a particular lake's demand model, it is necessary to hold the site characteristic variables constant in the k^*, p_1^*, and V equations. The days per visit equation does not change at this stage; a change is observed only after values of k^*, p_1^*, and k have been placed in the equation, fixed at their mean values.

The demand model for Klamath Lake, as it now exists, was calculated as:

$$k^* = 9.132 + .002\,y + 10.435\,p_1 \qquad (86)$$

$$p_1^* = 3.531 + .269\,k - .004\,k^2 + .000000017\,y^2 \qquad (87)$$

$$q_1 = e^{.759 - .0064 k^* + .0064 k + .0637 p_1^* - .0637 p_1} \tag{88}$$

The only difference in the demand model for Klamath Lake and the other lakes is in the constant term. The coefficients of the independent variables are the same for each lake, since identical reactions, regardless of the lake, were assumed.

The demand function for Klamath Lake, in it's present situation, when average values were fixed for the k^*, p_1^*, and k variables, was computed as:

$$q_1 = e^{.801 - .0637 p_1} \tag{89}$$

The demand function is also similar, regardless of the particular lake, except for the constant. However, the slope is identical. The demand function for Klamath Lake yields an economic value per visit, as estimated in 1968, ranging from \$.56 to \$6.37. An estimated 146,491 visits were purported to have occurred during the summer of 1968. When the value per visit was combined with the number of visits, an estimate of the total seasonal value was forthcoming. The seasonal value of Klamath Lake, in it's present condition, was estimated to exist between \$82 thousand and \$933 thousdand for 1968.

A proposed change in water quality would occur in two steps. Step one would attempt to remove the blue-green algae present in Klamath Lake, while step two would endeavor to lower the water temperature and improve the beaches. Since the demand models were derived in relation to the site characteristics, a change in water quality can be accounted for by these equations. The variables that would be predicted to change, given an increase in water quality, in the Klamath Lake demand model were W (water skiing, swimming, and boating use-intensities) and F (fishing use-intensity). All other variables were assumed unaltered. When the projected values of W and F were placed in the demand model, a new demand function was obtained, and the determination of the economic value of each proposed step was possible.

It was estimated that it would be worth between \$3.9 and \$9.2 million to society to activate the two-step procedure. It would be worth \$1.2 to \$3.4 million to undertake the first step, and between \$2.7 and \$5.8 million additional to carry out the second step. These estimates are based on the data collected in the summer of 1968. Any inclusion of other seasons, or other years would need to be done to supplement the results in this study.

Average values of the relevant variables, for each lake, and Klamath Lake after each step, are included in Table V.

Limitations and Suggestions for Future Research

Some limitations of the methodology suggested in this thesis are in order at this point. The restrictions encountered because of too few observations should, ideally, be removed. Many observations were necessary in order to divide them into groups to obtain estimates of k^* and p_1^*. The question of determining the appropriate sample size should not only be concerned with statistical precision, but also the cost of obtaining a larger sample. This study utilized 304 observations, however a study that would follow the methodology presented, but concerned with only one site, would need substantially fewer observations.

Biases are likely to occur in many instances and for several reasons. Biases may have occurred in the results of this thesis due to non-responses, refusals, incompleteness of the questionnaire, inability of respondents to estimate the items in the desired detail, etc. Biases could occur for two reasons: due to the construction of the questionnaire, or due to the manner in which the interviewing was done. First, as with any questionnaire, the one used in this study had some weak points. Five categories were excluded: the number of cars that were driven to the site should have been included to enable computation of travel cost more accurately; the investment items that were borrowed should have been specified; the length of time it took to drive to the site should have been obtained, so the number of meals enroute could be established; and a space should have been provided to account for Golden Eagle cards purchased. The place of consumption of the purchased food should have been indicated. Food could be purchased enroute to be eaten at the site. Thus a more reliable estimate of travel costs and variable costs could be calculated.

Secondly, biases could occur due to the manner in which the interviewing was done. This type of bias could occur in this study, as in any other, however, it is thought that their presence is limited. The use of a professional marketing research firm for the collection of the data has kept these biases to a minimum.

Many limiting assumptions were made that would unquestionably cause biases in the results. It had to be assumed that all the recreationists would react identically to economic phenomena, except for allowance of the differences. The characteristic variables were to account for all the possible differing reactions. This assumption was necessary in order to express participation in outdoor recreation as a function of water quality. It was assumed that the major differences in the sampled lakes were due to the degree of water quality possessed. Another vital assumption was that the variables chosen to represent lake characteristics, actually depicted the biological differences in the lakes. The productivity of a lake is determined by several parameters: the average depth of a lake; the shape of the lake's basin; the average water temperature; the transparancy of the water; the dissolved oxygen content; and the pH of the water.[39] Thus, the variables chosen in this study--the use-intensities of swimming, water skiing, boating, fishing, and camping, and the lake size--attempted to reflect the water quality parameters. The amount of use of the activities should indicate, to a degree, the quality of the water. Information was not available, for the lakes included in this study, to include the exact measurements of the parameters listed.

The manner in which the intensities of use were measured could be improved in future research. Rankings of low, medium, and high use were assigned index numbers for each activity. Information should be sought from the U. S. Forest Service to obtain the estimated visitor-days at a site, broken down by activities. This information should give an improved estimate of the relationship of use between lakes--the information was not available for all lakes sampled in this study.

The lake size, it is thought, should be revised as a variable depicting site capacity. Some measure of accessibility should be investigated, e. g., the miles of easily accessible shoreline, etc. Possibly a dummy variable could be used to represent the lake size, i. e., lake size could be defined as small, medium, or large. Biases

---

[39] These "water quality parameters" were determined upon consultation with biologists at the F. W. P. C. A. Pacific Northwest Water Laboratory in Corvallis, Oregon.

can be expected to occur from inadequate specification of the characteristic variables. For example, too large an influence may be attributed to lake size, thus the estimated change in the value per day and visits could be overstated. It is thought that this bias exists in the results of this study.

Future research in this area should, of course, consider additional variables. Possibilities include the use of fixed investment as a replacement for income, the characteristics of the recreationists, fishing success, lake reputation, number of alternatives available, amount of development, average daily temperature, and remoteness of the site.

Other biases are likely to occur from interviewing only recreationists at the sites, instead of also those not presently at the sites. In this manner more defensible estimates could be obtained for k^* and p_1^*, since observations would be available from persons with values of k and p_1 less than the critical levels.

Deleting all recreationists staying more than 20 days could also affect the conclusions. The effect of this was discussed in Chapter VII of this thesis.

Estimates of the coefficients derived for the days per visit relationship are unbiased for predicting the log of q_1. However, when these coefficients are utilized to draw inferences toward q_1, they no longer are unbiased. The error terms in the log form are additive, but when the function is transposed to the exponential form, the error terms are multiplicative. Severe statistical problems can arise due to a violation of the assumption of additive error terms.

One omission may cause serious difficulties in the estimation of the "worth" of substantially improving the water quality in Upper Klamath Lake. The effect of the water fowl migration habits has been ignored. The water fowl hunting, rated excellent at present, may be seriously hampered by an improvement in water quality. The exact effect, however, depends on the chosen procedure to improve the water quality. Some recommendations would eliminate a majority of the hunting grounds, while others may preserve them. Since it was uncertain as to which scheme was most feasible, the water fowl hunting was ignored. The value estimates, however, do represent a value to recreationists using the lake's facilities during the summer season.

Future work in this area should focus on the problem of substitutes. That is, it was assumed for purposes of this study that the relation of attendance between lakes was relatively insignificant. For example, when an increase in the length of stay and number of visits was proposed for Klamath Lake, no decrease in the use of other lakes was assumed. The reduction in the use of alternative sites should be included in an analysis, to determine the "net" effect relevant to the area.

It can be concluded that even though recognizing many limitations in this study, many important and restrictive assumptions were avoided, prevalent in past methodologies. Some of the assumptions involved in other procedures were enumerated in Chapter II of this thesis. It was deemed important to consider the travel costs and on-site costs separately. The differences in the reactions of individuals can be seen from the postulated demand models. If these two influences are combined into one variable the two effects will be counterbalanced, and only the stronger will be observed. Also of importance is to evaluate individual recreationists, and to define the quantity variable carefully--ideally it should be disaggregated into two components: the days per visit, and the number of visits per unit of time.

Future work should strive to obtain information from recreationists regarding the number of visits that they will make per season. By this procedure one will not have to resort to the use of county information, as was done in the visits equation of this thesis. All inferences may then be made toward the population of recreationists.

The estimated economic benefits resulting from an improvement in water quality may not be reliable due to the manner in which the site characteristics were utilized. This, however, does not limit the usefulness of the methodology. The methodology, it is thought, would be adequate if an improved procedure to account for the differences in lakes were available. It would also be a valuable procedure in the estimation of the recreational demand for a single site.

BIBLIOGRAPHY

Bond, Carl E., Charles R. Hazel and David Vincent. 1968. Relations of nuisance algae to fishes in Upper Klamath Lake. Corvallis. 119 numb. leaves. (Oregon State University, Dept. of Fisheries and Wildlife. Terminal progress report on Research Grant WP00625 of the U. S. Federal Water Pollution Control Administration) (Unpublished manuscript)

Brown, William G., Ajmer Singh and Emery N. Castle. 1964. An economic evaluation of the Oregon salmon and steelhead sport fishery. Corvallis. 47 p. (Oregon. Agricultural Experiment Station. Technical Bulletin 78)

California. Franchise Tax Board. 1966. Annual report, 1966 calendar year. Sacramento. 102 p.

Castle, Emery N. 1965. The market mechanism, externalities, and land economics. Journal of Farm Economics 47:542-556.

Clawson, Marion. 1959. Methods of measuring the demand for and value of outdoor recreation. Washington, D. C., Resources for the Future. 36 p. (Reprint no. 10)

Clawson, Marion and Jack L. Knetsch. 1966. Economics of outdoor recreation. Baltimore, Maryland, Johns Hopkins University. 317 p.

Cochran, William. 1960. Sampling techniques. New York, Wiley. 317 p.

Crutchfield, James H. 1962. Valuation of fishery resources. Land Economics 38:145-154.

Draper, N. R. and H. Smith. 1966. Applied regression analysis. New York, John Wiley. 402 p.

Guedry, L. J. The role of selected population and site characteristics in the demand for forest recreation. Ph. D. thesis. Corvallis, Oregon State University. (In preparation)

Harper, F. H. 1930. Elements of practical statistics. New York, Macmillan. 284 p.

Hines, Lawrence G. 1958. Measurement of recreation benefits: a reply. Land Economics 34:365-367.

Hogg, Robert V. and Allen T. Craig. 1965. Introduction to mathematical statistics. New York. Macmillan. 383 p.

Johnston, W. E. and V. S. Pankey. 1968. Use of prediction models for Corps of Engineers reservoirs in California. In: An economic study of the demand for outdoor recreation: Conference proceedings of the Cooperative Regional Research Technical Committee, San Francisco, 1968. p. 15-47. (Report no. 1)

Knetsch, Jack L. 1963. Outdoor recreation demand and benefits. Land Economics 39:387-396.

Knetsch, Jack L. and Robert K. Davis. 1966. Comparisons of methods for recreation evaluation. In: Water research, ed. by Allen V. Kneese and Stephen C. Smith. Baltimore, Maryland. John Hopkins University. p. 125-142.

McKean, Roland N. 1968. Public spending. New York, Mc-Graw-Hill, 179 p.

Oregon. State Tax Commission. Research and Planning Section. Analysis of Oregon's personal income tax returns for the tax year 1966. Salem. 15 numb. leaves. (Mimeographed)

Pearse, Peter H. 1968. A new approach to the evaluation of non-priced recreational resources. Land Economics 44(1):87-99.

Portland State College. Center for Population Research and Census. 1967. Population estimates of counties and incorporated cities of Oregon, July 1, 1967. Portland. 6 p.

Rohlich, Gerard A. 1968. Eutrophication. A paper presented at the 41st Annual Convention of the Soap and Detergent Association, New York City, January 25, 1968. 11 p.

Stevens, Joe B. 1966. Recreation benefits from water pollution control. Water Resources Research 2(2):167-182.

Stoevener, Herbert H. 1963. An economic evaluation of water pollution control alternatives. In: Economics of transferring water; Economics of water quality; the economic base study in river basin planning. Water resources and economic development of the west: Conference proceedings of the Committee on the Economics of Water Resources Development of the Western Agricultural Economics Research Council, Salt Lake City, 1963. p. 47-59. (Report no. 12)

__________ 1965. Water use relationships as affected by water quality on the Yaquina Bay. In: New horizons for resources research: Issues and methodology: Proceedings of the Western Resources Conference, Boulder, University of Colorado. p. 87-99.

Stoevener, Herbert H. and William G. Brown. 1967. Analytical issues in demand analysis for outdoor recreation. Journal of Farm Economics 49:1295-1304.

Stoevener, Herbert H. and Emery N. Castle. 1965. Input-output models and benefit-cost analysis in water resources research. Journal of Farm Economics 47:1572-1579.

Trice, Andrew H. and Samuel E. Wood. 1958. Measurement of recreation benefits. Land Economics 34:196-207.

U. S. Bureau of the Census. 1968. Population Estimates: Estimates of the population of counties: July 1, 1966. (Ser. p-25, no. 401) Washington, D. C., Aug. 28, 1968. 20 p.

__________ 1968. Population Estimates: Estimates of the population of counties: July 1, 1966. (Ser. p-25, no. 404) Washington, D. C., Sept. 20, 1968. 20 p.

__________ 1968. Population Estimates: Estimates of the population of counties: July 1, 1966. (Ser. p-25, no. 407) Washington, D. C., Sept. 27, 1968. 22 p.

__________ 1963. U. S. Census of Population: 1960. Vol. 1. Characteristics of the population. Part 4. Arizona. Washington, D. C. p. 112-113.

__________ 1963. U. S. Census of Population: 1960. Vol. 1. Characteristics of the population. Part 14. Idaho. Washington, D. C. p. 137-140.

U. S. Bureau of the Census. 1963. U. S. Census of Population: 1960. Vol. 1. Characteristics of the population. Part 28. Montana. Washington, D. C. p. 151-155.

__________ 1963. U. S. Census of Population: 1960. Vol. 1. Characteristics of the population. Part 30. Nevada. Washington, D. C. p. 105-106.

__________ 1963. U. S. Census of Population: 1960. Vol. 1. Characteristics of the population. Part 46. Utah. Washington, D. C. p. 130-132.

__________ 1963. U. S. Census of Population: 1960. Vol. 1. Characteristics of the population. Part 49. Washington. Washington, D. C. p. 170-173.

__________ 1963. U. S. Census of Population: 1960. Vol. 1. Characteristics of the population. Part 52. Wyoming. Washington, D. C. p. 108-110.

U. S. Dept. of Agriculture. Agricultural Research Service. 1965. Food consumption of households in the U. S.: Household food consumption survey, 1965-1966. Washington, D. C. 212 p. (Report no. 1)

U. S. National Park Service. 1949. The economics of public recreation: An economic study of the monetary evaluation of recreation in the national parks. Washington, D. C. 36 p.

Warren, Charles E. 1957. The influence of dissolved oxygen upon the survival, development, growth, activity, and movements of fresh water fishes. Corvallis. 49 p. (Oregon State University. Progress report to the U. S. National Institutes of Health on Research Grant no. RG-4352)

Wennergren, Boyd. 1964. Valuing non-market priced recreational resources. Land Economics 40:303-314. Nov.

__________ 1967. Demand estimates and resource values for resident deer hunting in Utah. Logan. 44 p. (Utah. Agricultural Experiment Station. Bulletin 469)

APPENDIX TABLE I. THE ESTIMATED USE-INTENSITIES OF CERTAIN ACTIVITIES, AND LAKE SIZE, BY LAKE.

Lake	Use-Intensities					Lake Size
	Swimming	Water Skiing	Boating	Fishing	Camping, Picnicking	
Lake of the Woods	3	3	3	1	3	1,055
Odell	0	0	1	3	2	3,500
Willow	1	1	1	3	2	320
Klamath	0	1	1	1	0	98,560
Klamath (Step 1)	1	3	3	2	2	98,560
Klamath (Step 2)	3	3	3	3	3	98,560

FIELD SURVEY QUESTIONNAIRE

OREGON STATE UNIVERSITY

July, 1968

Hello, I'm ________________. I'm working on a recreation survey for Oregon State University and would like to ask you a few interesting questions if you don't mind!

1 - 1 Visit lake (continue) 2 Other purpose (DISCONTINUE)	Was the main purpose of your trip to visit this particular lake, or are you taking your trip for some other purpose?
2 - ________________ Date ________________ Time (AM or PM)	May I ask when you arrived at this particular site--the date and the approximate time?
3 - ________________ Date ________________ Time (AM or PM)	Now, when do you plan to leave this particular site--again the date and approximate time of day or night?
4 - ________________ City/Town ________________ County ________________ State	Where do you live at the present time--the city or town, county and state?
4a- 1 In city/town (skip to 5) 2 Suburban area (skip to 5) 3 Rural outside (ask 4b)	Do you live right in the city (town), a suburban area, or a rural area outside of the city (town)?
4b- ________________ Miles 1 Nearer site 2 Away from site	How many miles do you live out of the city (town)? (INT: Mark whether nearer or farther away from site)
5 - ________________ Number	Including yourself, how many persons are there in your party which is stopping at this particular place?
6 - 1 Immediate family 2 Other relatives 3 Unrelated individuals 4 Other (explain below)	Does your party consist mainly of your immediate family, mainly of other relatives, or mainly of unrelated individuals, such as neighbors and friends?
7 - ________________ Number	Including yourself, how many persons are there in your immediate family?

To help the University figure out how valuable recreation is to the state, I'd like to ask you about your party's expenditures from your home to this area.

8 - $ ____________ Enroute	Approximately how much did your party spend for food and liquor in cafes, restaurants or taverns while you were en-route to this particular site? (just your best estimate)

$ ________	Here	About how much will your party probably spend in restaurants, cafes or taverns while you are stopping at this particular site? (just your best estimate)
9 - $ ________	Home	Approximately how much did your party spend for this trip in grocery or liquor stores before you left home? (just your best estimate?)
$ ________	Enroute	About how much did your party spend in grocery or liquor stores while you were enroute to this particular site? (just your best estimate?)
$ ________	Here	What do you think your party will spend in grocery and liquor stores while you are stopping at this particular place? (just your best estimate?)
10 - $ ________	Enroute	While you were enroute to this site, about how much did your party spend for lodging in motels, hotels, or trailer parks? (just your best estimate)
$ ________	Here	What do you think your party will spend for lodging in motels, hotels or trailer parks while you are stopping at this site? (just your best estimate)
11 - $ ________	Enroute	How about camping fees--how much, if any, did your party spend for camping fees while you were enroute to this site? (just your best estimate)
$ ________	Here	What do you think your party will probably pay for camping fees while you are here at this site? (just your best estimate)
12 - ________	Miles	How many miles, if any, did your party drive yesterday while at this site?
12a - ________	Purpose	What was the purpose of your drive yesterday?
12b - ________	Miles	(If not at site yesterday) About how many miles, if any, will your party probably drive today while at this site?
12c - ________	Purpose	For what purpose will today's drive be for?

Now, please think of the gasoline and oil that will be purchased for your party's car and boat for this entire trip.

13 - ________	% before leaving	First, about what percent of gas and oil for the car and boat was purchased for the trip before you left home? (just your best estimate)
________	% both ways	Now, think of all the gas and oil that will be purchased between home and here and between here and back home. Approximately what percentage of the gas and oil will be purchased between home and the time you get back home from here, that is, both ways? (Estimate)
________	% at this site	What percent of all gas and oil purchased for the car and boat will you probably make while you are stopping at this site? (just your best estimate)

14 - 1 Yes (ask 14a) 2 No (skip to no. 15)		Did your party bring a boat with you to this site?
14a - ________________	Gallons Gas	About how many gallons of gasoline does your boat use a day at this particular site?
________________	Quarts Oil	How many quarts of oil does your boat use in a day while here?

(INTERVIEWER: Refer to question no. 6. Ask question 15 series only if code 2, 3, or 4 is circled in no. 6)

15 - 1 Mine (ask 15a) 2 Someone else (skip to 15b)	Whose car did you bring on the trip--yours or someone else's in your party?
15a - $____________________	How much, if any, did other members of your party contribute for gas, oil and automotive expenses thus far on the trip?
15b - $____________________	How much, if any, have you contributed thus far to the owner of the car for gas, oil and automotive expenses?
16 - $____________________	How much money, if any, has your party spent on boat launching fees while on this trip?

17 - Thus far, we have talked about expenses for the automobile, boat, food and liquor, and for lodging and camping fees. Can you think of any other types of expenses you have had coming here, such as camera supplies, souvenirs, etc. (If YES) What type?

1 No
2 Yes Type ____
Total cost of these expenses $ ____________________

17a - What other types of expenses will you have while stopping at this site?
1 No
2 Yes Type ____
Estimated cost of these expenses $ ____________________

18 - (HAND CARD TO RESPONDENT) Here is a list of items which either you or other members of your party may own, which you have brought with you to this site. Looking over the list, will you please tell me which owned items were brought with you? Do not inclue rented items. (INT: Mark X for each item. Then ask remaining questions on your card for each X'd item)

Items	Year Purchased	Amount Paid for Item	Type & Location of Store Where Purchased	Maintenance
____ Boat				
____ Outboard motor				
____ Boat trailer				
____ Fishing tackle (rod, reel, tackle box, etc.)				
____ Camper (van, truck, trailer camper, etc.)				
____ Tent trailer				
____ Tent				
____ Back pack				
____ Sleeping bag				
____ Stove				
____ Lantern				
____ Cooler				
____ Boat equipment not included in price of boat (preserver, fire extinguisher, etc.)				
____ Water skiis, ropes, etc.				
____ Special clothing (such as rubber boots, coats, rainwear, swimming suits, etc.)				
____ Any other items? (If YES) What? ____				
____				
____				

19 - (HAND RENTAL CARD) Looking at this list of items, will you please tell me which, if any, of these items you or other members of your party have rented for this particular trip? (INT: Mark X for each item. Then, for each X'd item, ask the remaining questions on your card for each X'd item)

Item	Rental Rate (Daily, Hourly, Weekly)	Type & Location of Store Where Rented	Total Rent Expected to Pay for Item
____ Boat			
____ Outboard motor			
____ Boat trailer			
____ Fishing tackle (rod, reel, tackle box, etc.)			
____ Camper (van, truck, trailer camper, etc.)			
____ Tent trailer			
____ Tent			
____ Back pack			
____ Sleeping bag			
____ Water skiis			
____ Life vests			
____ Other equipment for boats			
____ Any other items? (If YES) What? ________			
________			
________			
________			
________			

20 - $________ About how much will you spend at this site for various baits--just that amount that will be used at this particular site?

0 None

21 - (HAND RESPONDENT INCOME CARD) Would you please look at this card and tell me which one of these groups best fits your total family income before taxes for last year? Just call your answer by letter, please.

1 (a) Less than $3,500	8 (h) $11,000 - $11,999
2 (b) $3,500 - $4,999	9 (i) $12,000 - $12,999
3 (c) $5,000 - $6,999	_0 (j) $13,000 - $14,999
4 (d) $7,000 - $7,999	1 (k) $15,000 - $16,999
5 (e) $8,000 - $8,999	2 (l) $17,000 - $19,999
6 (f) $9,000 - $9,999	3 (m) $20,000 - $24,999
7 (g) $10,000 - $10,999	4 (n) $25,000 or over

(INT: If $25,000 or over, get range from respondent)

22 - INTERVIEWER: Mark below the type of activity the respondent was doing when you first approached (him) (her), or the type of activity the respondent just finished doing.

________________________________ Activity

23 -	1 Male	1 Under 21 years of age	
	2 Female	2 21 - 29 years	
		3 30 - 39	
		4 40 - 49	
		5 50 - 59	
		6 60 or over	Age and sex of respondent

24 - ______________________________ Site where interview was taken

25 - __________ ____________________ Telephone number of respondent.
Area Code (For verification purposes only)

X I hereby certify this interview was actually taken with the person described above, and represents a true and accurate account of the interview.

______________________________ ____________________, 1968
(Interviewer's Signature) (Date)

COMMENTS ON INTERVIEW (if any):