Title
Water Quality and the Demand for Recreation

Author
Hanemann, W. Michael

Publication Date
1981-03-01

Division of Agricultural Sciences
UNIVERSITY OF CALIFORNIA

Working Paper No. 164

WATER QUALITY AND THE DEMAND FOR RECREATION

W. Michael Hanemann

California Agricultural Experiment Station
Giannini Foundation of Agricultural Economics
March 1981

WATER QUALITY AND THE DEMAND FOR RECREATION

W. Michael Hanemann

Department of Agricultural and Resource Economics,
University of California, Berkeley, CA 94720

Proposed running head: Recreation Demand

Please mail all correspondence to:

Professor Michael Hanemann
207 Giannini Hall
University of California
Berkeley, CA 94720

WATER QUALITY AND THE DEMAND FOR RECREATION[1]

I. INTRODUCTION

In several recent assessments of the benefits of water pollution abatement programs, the single largest component has been the benefits to recreationists who engage in activities such as swimming, boating, fishing, or water fowl hunting. For example, Freeman [15] examined the benefits of removing conventional water pollutants from U.S. waterways under the provisions of the 1972 Amendments to the Federal Water Pollution Control Act and concluded that the most likely point estimate of the annual benefits in 1985 was $12.3 billion (in 1978 dollars), of which $6.7 billion was attributed to recreation benefits. Tihansky [38] surveyed ten studies on the benefits from the removal of pollutants in various water bodies in the U.S. and abroad, conducted over the period 1960-1972; on the average, recreation benefits accounted for about 44 per cent of the total benefits. A more striking example is provided by Ackerman et al. [1]; in their appraisal of the Delaware Estuary Comprehensive Study, recreation benefits accounted for approximately 95 per cent of the total benefits from the pollution abatement program. Because of the strategic importance of recreation benefits in justifying water pollution control projects, it is worth devoting some attention to the methodology by which they are calculated. Recent studies have generally employed one of three methods: the survey approach, in which respondents are directly questioned about their willingness to pay for (sometimes hypothetical) changes in environmental quality; the demand function approach, in which the recreationists' willingness to pay is inferred from their actual choice behavior, as reflected in fitted demand functions for recreation sites; and the unit day value method, in which a physical measure of attendance is obtained, possibly from a regression model of recreation participation of site demand, and this is converted into monetary benefits through multiplication by a measure of the value of a unit of recreation activity, such as the unit day values promulgated by the Water Resources Council. This paper will focus on the demand function approach. I will discuss some methodological issues and then present an application involving water-based recreation in the Boston area.

The demand function approach has been used in two ways to measure the benefits of changes in water quality. In one case, the question is what would happen if a beach were closed because of pollution. The situation is depicted in Figure 1, where AB is an individual's demand function for visiting a site, OP is the current cost of visiting the site, and OX is his current number of visits. If the site were shut down, the individual's welfare loss is typically measured by the Marshallian triangle, PAB. In the second case the question is one of a variation in quality rather than a beach closing. The situation is depicted in Figure 2, where BC is an individual's demand for visiting a site at its current level of quality, and AD is his demand function when the quality is lowered. The cost of visiting the site is OP and the individual currently makes OX_1 visits; if the quality fell, he would make only OX_2 visits. Following Stevens' [36] suggestion, the common practice is to measure the welfare loss from the quality change by the area ABCD. This paper will be concerned with the problem of measuring quantities corresponding to the areas PAB in Figure 1 and ABCD in Figure 2 in a manner which is consistent with utility theory.

The quantity PAB can be related to the recreation unit value concept mentioned above; the consumer's surplus <u>per trip</u>, given by the length (AP)/2 in Figure 1, can be thought of as a measure of the benefit per

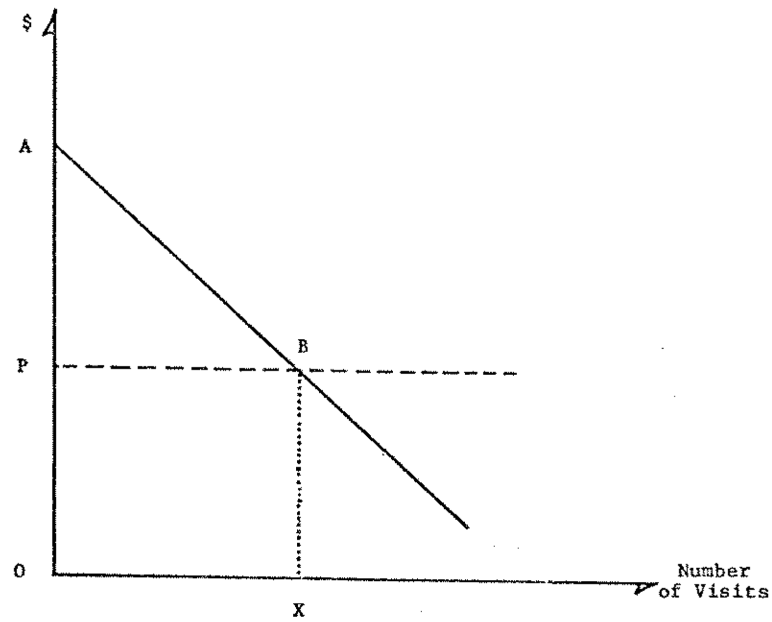

FIG. 1. The welfare loss from the closing of a recreation site.

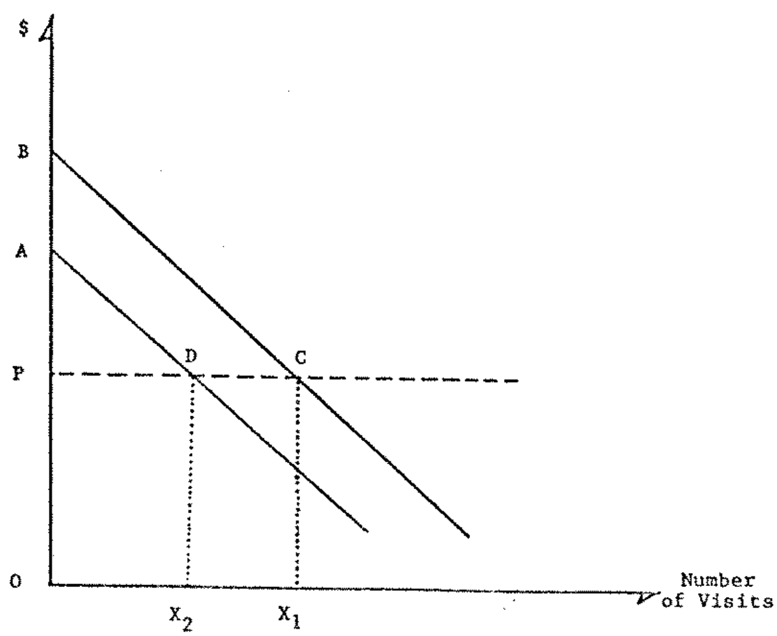

FIG. 2. The welfare loss from a reduction in the quality of a recreation site.

visitor-day. Some estimates of consumer's surplus per visitor-day taken from the recent literature are presented in Table I. These pertain to day trips for general recreation activities—swimming, picnicking, etc. The estimates were made at various times, in various locations, and by various methods, not all of them strictly consistent with utility theory. In 1973 the Water Resources Council issued project guidelines which specified unit day values of 75¢ to $2.25 per visitor day for general water-based recreation [41]. In 1980 these values were raised by 43 per cent to allow for inflation [42]; the new range of values is $1.07 to $3.22 per visitor-day. These sets of values are not incompatible with the consumer's surplus estimates in Table I. However, some recent studies have employed unit values which are considerably higher. In 1974 Tihansky [37] assessed the value of cleaning up ocean beaches at $2.00-$5.00 for swimmers currently using the beaches, and $1.00-$2.00 for new swimmers attracted to the beaches. Heintz, Hershaft and Horak [26] valued the loss of utility to recreationists who continued to visit a site when it changed from being clean to being polluted at $5.75 per visitor-day, in 1973 dollars. The Battelle Memorial Institute conducted a study for the National Commission on Water Quality on the increase in swimming activity at public beaches which had been closed because of poor water quality and which could be expected to reopen as a result of improvements arising from the 1972 Amendments to the Federal Water Pollution Control Act. The increase in swimming activity was divided into two categories: visits diverted from pre-existing sites, and entirely new participation in swimming recreation. In its analysis of Battelle's findings, the NCWQ Staff Report valued a visitor-day of diverted activity at $2.50 in 1976 dollars, and a visitor-day of new activity at $7.50 [33]. I shall argue below that these unit values are

Table I. Estimates of Consumer Surplus Per Visitor Day for General Water-Based Recreation.

Reference	Year of Publication	Site	Consumer's Surplus per Visitor Day ($)
[40]	1958	Feather River, CA	2.09
[41]	1962	Lake Michigan	0.59
[41]	1962	Reservoir 50 miles from St. Louis	1.00
[27]	1964	Kerr Reservoir, NC	3.10[a]
[32]	1968	Lake Cachuma, CA	1.49
[32]	1968	Puddingstone Reservoir, CA	0.25
[32]	1968	Yosemite Lake, CA	0.59
[32]	1968	Lake of the Ozarks, MI	0.83
[18]	1968	Texas Reservoirs	0.43
[7]	1971	3 hypothetical reservoirs 35 to 55 miles from St. Louis	2.43
[34]	1974	Corps of Engineers Lakes	1.19
[34]	1974	Corps of Engineers Lakes	3.33
[8]	1976	11 State Parks, PA	0.06 to 3.39
[6]	1979	8 Wisconsin Lakes	2.72

[a]Consumer's surplus per travel party.

probably much too high.

The purpose of this paper is both methodological and empirical. Among the methodological issues to be discussed are these: How does one estimate a set of demand functions for a group of alternative recreation sites which vary in quality? Moreover, can one develop a set of demand functions which is consistent with a utility maximization model, so that exact measures of consumer's surplus can be derived—it is well known that the areas PAB and ABCD in Figures 1 and 2 are at best approximations? Many previous studies of recreation demand have focused on demand functions for an individual site, thereby ignoring the possible interactions between the visitation of one site and the availability of other competing sites.[2] Intuitively one would expect that, if alternative sites were available, the consumer's surplus from the opening or improvement of any one site would be lower. This, I believe, explains why some of the previous estimates of consumer's surplus are so high. On the empirical side, some of the questions are: How important are alternative indicators of water quality in explaining actual recreation choices? Do perceptions of water quality influence recreation choices more powerfully than objective measures of water quality? How much would a change in water quality actually increase recreation demand? Finally, are the unit day values cited above reasonable when applied to coastal regions where many alternative sites are available, or are they exaggerated?

II. SITE DEMAND MODELS

The most common approach to modeling the demand for recreation sites is to write down a formula for an ordinary demand function containing the conventional arguments—the "price" of the site, perhaps some measure of the quality of the recreation experience at the site, income and possibly other characteristics of the recreationist—and to fit the demand function to observations on these variables. As noted above, from the fitted demand function one calculates quantities corresponding to the areas PAB or ABCD in Figures 1 and 2. However, there are well known problems with using areas under ordinary demand functions to measure compensating or equivalent variations. In the case of the area PAB, the error resulting from the use of ordinary demand functions is discussed in [20]. In [21] I examine the validity of the area ABCD as a welfare measure; I show that unless certain restrictive conditions are satisfied—more restrictive than those generally recognized in the literature (for example, [14, pp. 73-4])—this area does not provide an accurate indication of either the direction or the magnitude of the true welfare change associated with the shift in site quality.

An alternative approach, which avoids these problems and yields exact measures of the compensating or equivalent variations, is to start by writing down an explicit utility function and derive the corresponding ordinary demand functions, which are then estimated from the data. In this case, from the coefficients of the fitted demand functions one can obtain estimates of the coefficients of the utility function and with this one can calculate the true welfare measures. Conceptually this approach is the more appealing but, for reasons described below, it is sometimes more difficult to apply empirically. The set-up in the utility-theoretic approach is as follows. A consumer—an individual or a household—is faced with N alternative recreation sites, and chooses a recreation consumption plan (i.e., which sites to visit, and how often) for, say, the summer recreation season. Let x_i be the number of visits to the i^{th} site, and let p_i be the cost of visiting the site; let $x = (x_1, . . ., x_N)$ and $p = (p_1, . . ., p_N)$. Since the data available in most recreation studies does not cover nonrecreation consumption activities, it is convenient to assume that all other goods can be aggregated into a single commodity whose consumption is denoted by z, which will be taken as the numeraire.[3] The recreation sites are differentiated in quality, and the consumer's welfare depends not only on his consumption levels but also on the quality characteristics of the available sites. Let b_{ik} be the amount of the k^{th} quality characteristic associated with a unit of good i—in the present context, b_{ik} could be, say, the water temperature at site i—and let $b_i = (b_{i1}, . . ., b_{ik})$ be the vector of quality characteristics associated with good i; let $b = (b_1, . . ., b_N)$. Given the consumer's income, y, he chooses his recreation plan by solving the constrained maximization problem:[4]

$$\underset{x \geq 0,\ z \geq 0}{\text{maximize}} \quad u\ (x, z, b) \text{ s.t. } \Sigma\, p_i x_i + z = y. \qquad (1)$$

If one specifies a particular functional form for the utility function u (x, z, b), the specific demand functions for the N sites can be obtained by solving (1).

It is important to distinguish between two types of solution to (1). One is an interior solution; in this case, the consumer makes some visits to <u>every</u> site (i.e., at the optimum $x_i > 0$ for all i). Here the ordinary demand functions have the general form $x_i = h_i\ (p, b, y)$, $i = 1, . . ., N$; i.e., the demand for each site depends on income, the prices of all sites, and the quality characteristics of all sites. On substituting the ordinary demand functions into the direct utility function, one obtains the indirect utility function $v(p, b, y) \equiv u[h(p, b, y), y - \Sigma p_i h_i(p, b, y), b]$, in terms of which the compensating and equivalent variations for a change in quality from b^0 to b^1 can be defined. The coefficients of the indirect utility function can be obtained from those of the fitted ordinary demand functions. In [23], I propose some functional forms for $u(x, z, b)$ which are suitable for empirical applications, I derive the formulas for the resulting systems of ordinary demand functions, and I present the corresponding formulas for the compensating and equivalent variations.

If one is working with aggregate data on the visitation of recreation sites (e.g., data grouped by county of origin), it is likely that the assumption of an interior solution can be invoked—there will probably be visits to all sites from each unit of observation. However, if one is working with microdata on individual households, this assumption will probably not hold—each household will <u>not</u> visit every site in its choice set. This is certainly true of the data set employed below. In that case one is dealing with a corner solution to (1)—i.e., at the optimum $x_i = 0$ for at least one i. One can distinguish between two types of corner solution. One type might be called the "extreme" corner solution, in which $x_i = 0$ for all but one i—the household visits only one of the N sites. This may be compared with the logit and other types of quantal choice model, which focus on the qualitative choice of one out of N items. In the situation envisaged here, there is both a qualitative choice—which site to select?—and a quantitative choice—how many visits are made to the chosen site? In [22], I develop two utility-theoretic demand models applicable to this type of qualitative-quantitative choice situation, and I apply one of them to the recreation demands of the subset of households in my sample who visit only one site; I also calculate exact welfare measures corresponding to the areas PAB and ABCD in Figures 1 and 2. Nevertheless, most households in the sample visit more than one site, although none visits more than eleven out of the thirty or more water-based recreation sites available in the Boston area (see Appendix Table IIA). Therefore the situation is one of a "general" corner solution: the household may visit any subset of the recreation sites, not necessarily one of them and not necessarily all of them. This situation, which is the most realistic, is the most difficult to represent empirically in a manner consistent with an explicit utility maximization model. In [19] and [23] I have developed some utility-theoretic demand models applicable to the general corner solution case, but since they involve (N-1) dimensional multiple integrals they can be implemented only when the number of sites is small—on the order of four or five. Therefore, they cannot be applied here.

This is the crux of the problem of modeling the demand for differentiated commodities (such as recreation sites) with microdata. On the one hand, it would be desirable to employ a demand model which is consistent with a utility maximization hypothesis, rather than a purely statistical model, in order to make use of the utility-theoretic measures of the welfare effects of a quality change. On the other hand, constructing a utility-theoretic demand model severely limits one's flexibility in modeling observed behavior and, in fact, becomes impractical for the complicated types of corner solution which occur in real world data. In the remainder of this section I will describe a demand model which can be implemented empirically, and which is partly consistent with a utility maximization hypothesis. In the next section this model will be applied to microdata on recreation demand in the Boston area.

In order to develop a practical substitute for (1), it is necessary to change the characterization of the choice situation. Suppose that the consumer's decision is made in two separate steps. In one step he decides on the total number of recreation trips, and in the other he decides how to apportion those trips among the available recreation sites. The demand for any site can thus be written

$$x_i = \pi_i X \qquad i = 1, \ldots, N, \tag{2}$$

where X is the predetermined total number of recreation trips, and π_i is the share of these trips allocated to the i^{th} site. The determination of X may depend on the income and other socio-economic attributes of the consumer, and it may also depend on some index of the overal costliness and quality of the available sites in his choice set, but it does not depend on the prices and qualities of the particular sites which he chooses to visit; because of this, the model is not consistent with (1). The allocation of visits among sites can be cast in a utility maximization framework by assuming that each time the consumer embarks on a recreation trip he makes what is now a purely qualitative choice of which site to visit. This can be represented by a quantal choice model such as multinomial logit or probit, which can be interpreted as arising from the following utility maximization problem:

$$\underset{w,\, z \geq 0}{\text{maximize}} \quad u(w, b, z) = \Sigma \phi_i (b_i) w_i + hz \qquad \text{s.t. } \Sigma\, p_i w_i + z = y$$
$$\text{and} \qquad w_i = 0 \text{ or } 1, \text{ all } i$$

Here $\phi_i (b_i)$ is an index of the attractiveness of the i^{th} site, h is a positive constant,[5] and the choice variables $w = (w_1, \ldots, w_N)$ represent a purely qualitative choice, namely the selection of a site to visit; w_i equals 1 if site i is selected, and 0 otherwise. Substituting the budget constraint into the maximand, the problem can be written

$$\underset{w}{\text{maximize}} \quad \Sigma\, \theta_i w_i + hy, \; w_i = 0 \text{ or } 1, \text{ all } i \tag{3}$$

where $\theta_i \equiv \phi_i (b_i) - hp_i$. In order to provide a statistical framework for estimating the quantal choice model it is necessary to introduce a stochastic element into the utility function in (3). This can be done by making the $\phi_i(\cdot)$'s random functions of their arguments; hence, the θ_i's also are random.[6] The allocation shares in (2) can be equated with the qualitative choice probabilities arising from (3)

$$\begin{aligned} \pi_i &= \Pr\{\text{site } i \text{ is selected}\} \\ &= \Pr\{\tilde{\theta}_i > \tilde{\theta}_j, \text{ all } j \neq i\}. \end{aligned} \tag{4}$$

The stochastic element in the $\phi_i(\cdot)$'s can be thought of as representing random differences in tastes across individuals and alternatives in the sample. Because of the large number of sites in the application presented below, it is not practical to employ multinomial probit for the qualitative choice model (3), and therefore multinomial logit will be used.[7] In that case the stochastic element is introduced by writing

$$\tilde{\phi}_i = \gamma_i(b_i) + \tilde{\varepsilon}_i, \; i = 1, \ldots, N \tag{5}$$

where $\gamma_i(b_i)$ is the nonstochastic component of the attractiveness index, and the $\tilde{\varepsilon}_i$'s are assumed to be independently distributed according to the extreme value (EV) distribution with parameters $(1, \alpha_i)$.[8] Equivalently, one can write $\tilde{\phi}_i = \alpha_i + \gamma_i(b_i) + \tilde{\varepsilon}$, where $\tilde{\varepsilon}$ is EV(1, 0). The independence assumption is a strong one since it implies that there is <u>no</u> correlation between an individual's choices on different occasions—the fact that he chose to visit site number 1 last week tells you nothing about which site he will visit this week. In principle, it would be more plausible to allow the $\tilde{\varepsilon}_i$'s to be correlated across individuals; this can be built into the multinomial probit model—for example, see [25]—but there are too many sites to apply this model here. Given the stochastic specification of (5), the qualitative choice probabilities take the form

$$\Pr\{\text{site } i \text{ is selected}\} = \frac{e^{\alpha_i + \gamma_i(b_i) - hp_i}}{\sum_{j=1}^{N} e^{\alpha_j + \gamma_j(b_j) - hp_j}} \quad . \tag{6}$$

As shown in [30], the parameters of the model—the α_i's, h, and the coefficient of the $\gamma_i(b_i)$ functions—can be estimated from (6) by the maximum likelihood method.

In order to perform welfare evaluations for changes in prices or qualities, it is necessary to obtain the indirect utility function corresponding to the direct utility function in (3). Suppose that the consumer selects site i—i.e., the solution to (3) is $w_i = 1$, $w_j = 0$, $j \neq i$; substituting these values into (3), one obtains the conditional indirect utility function $\tilde{u}_i = \tilde{\theta}_i + hy$. The consumer actually selects that site which yields the highest utility. Denote the value of the maximized indirect utility function by $\tilde{u}^*$—i.e., $\tilde{u}^* = \max\{\tilde{u}^*, \ldots, \tilde{u}^*_N\} \equiv \tilde{v}^*(p, b, y)$. Let $\bar{v}^*(p, b, y) \equiv E\{v^*(p, b, y)\}$ be the expected value of the maximized indirect utility function. This function will be used to construct a monetary measure of the welfare effects of a change in the prices or qualities of the available sites. Suppose that prices and qualities change from (p^0, b^0) to (p^1, b^1). The compensating variation measure of the welfare effects of this change is defined to be the quantity C which satisfies

$$\bar{v}^*(p^1, b^1, y - C) = \bar{v}^*(p^0, b^0, y). \tag{7a}$$

Alternatively, the equivalent variation measure of this change is defined to be the quantity E which satisfies

$$\bar{v}^*(p^1, b^1, y) = \bar{v}^*(p^0, b^0, y + E). \tag{7b}$$

In [22], it is shown that for the logit model given by (3) and (5), the expected indirect utility function takes the form

$$\bar{v}^*(p, b, y) = hy + \sum_{j=1}^{N} (e^{\alpha_j + \gamma_j (b_j) - hp_j}) + 0.522 \; . \; . \; \text{(Euler's constant)}$$

and, therefore, for the change from (p^0, b^0) to (p^1, b^1)

$$C = E = \frac{1}{h} \ln (\lambda^0 / \lambda^1)$$

where (8)

$$\lambda^t = \sum_{j=1}^{N} (e^{\alpha_j + \gamma_j (b_j^t) - hp_j^t}) \qquad t = 0, 1.$$

In the case of a pure quality change, one obtains the true welfare measure corresponding to the area ABCD in Figure 2 by setting $p_j^1 = p_j^0$ all j in (8). To obtain the true welfare measure corresponding to the area PAB in Figure 1 for site i, say, one sets $b^1 = b^0$, $p_j^1 = p_j^0$ all $j \neq i$, and $p_i^1 = \infty$ in (8), since this is the price required to drive the consumer's demand for site i to zero (i.e., $\pi_i = 0$) in the logit model. Note that (8) is a measure of the welfare gain (or loss) per choice occasion—i.e., each time the consumer embarks on a recreation trip and makes a qualitative choice of which site to visit he reaps this benefit. Therefore, the total benefit from the change in prices or qualities is obtained by multiplying this benefit per site choice by the consumer's total number of recreation trips over the season, X.

III. ESTIMATION OF THE SITE DEMAND MODEL

In this section the demand model described above will be applied to data on the visitation of water-based recreation sites in the Boston area. The data come from two surveys, both conducted in 1974 and described in more detail in [5] and [19]: a survey in their homes of a stratified random sample of 462 households living in the Boston SMSA to ascertain which beaches they had visited during the 15-week period between Memorial Day and Labor Day, 1974, and how frequent the visits had been; and a survey of 30 major recreation sites in the area to inventory their facilities and collect water samples for chemical analysis.[9] The sites involved in the survey, which include all of the important beaches in the Boston area, are listed in Appendix Table IA; some of the main features of the households' recreation patterns are summarized in Appendix Table IIA. The survey dealt only with trips that lasted one day or less, and it covered several recreation activities. This paper will focus only on swimming and general beach recreation. The households surveyed represent about 0.07 per cent of the total 1970 Boston SMSA population, and have roughly the same incomes, educations and ethnic backgrounds as the larger population.

The demand model to be estimated here is given by (2). As noted above, it has two components: the determination of the total number of visits by the members of a household to all recreation sites, $\tilde{X}$, and the allocation of this total among individual sites, $\pi_1, \ldots, \pi_N$. The estimation of the determinants of total recreation activity involves some statistical problems which will now be examined. The distribution of $\tilde{X}$ across the households in the sample is summarized in the second panel of Table IIA. Three features of this distribution stand out: $\tilde{X}$ is obviously a non-negative random variable; it has a discrete probability mass at zero—56 of the 462 households did not participate in any recreation activity; and it is highly skewed, with a long right tail—the median number of visits to recreation sites by a household is about 10, the mean is 21, and the maximum for any household is 276. These features cause problems for the choice of a statistical model. For example, one might be tempted to postulate the standard regression model

$$\tilde{X} = Z\beta + \tilde{\varepsilon}$$

where Z is a set of explanatory variables, β is a vector of coefficients to be estimated, and $\tilde{\varepsilon}$ is a normally distributed error term. However, this specification is not consistent with the non-negativity, the probability mass at zero, or the skewness of the distribution of $\tilde{X}$. One could handle the non-negativity by postulating that $\tilde{X}$ is a poisson variable with mean $Z\beta$, but this could not simultaneously accommodate the probability mass at zero and the skewness. Another possibility is to assume that $\tilde{X}$ is lognormally distributed; this incorporates the non-negativity and the skewness, but not the probability mass at zero. One could employ the Tobit model:

$$\tilde{X} = \begin{cases} 0 & \text{if } Z\beta - \tilde{\varepsilon} \leq 0 \\ Z\beta - \tilde{\varepsilon} & \text{otherwise} \end{cases}$$

where $\tilde{\varepsilon}$ is normally distributed; this provides for the non-negativity and the probability mass at zero, but not the skewness. One would ideally like to have a lognormal version of Tobit. In [2] and [3] such a model is presented although not in a regression context; a regression version of the model is described in [19], but it does not appear to be computationally feasible. Instead, I will employ a model which is based on Goldberger's [16] two-stage approximation to the Tobit model:

$$\tilde{X} = \begin{cases} 0 & \text{if } Z\beta - \tilde{\varepsilon} \leq 0 \\ g(W, \gamma; \tilde{\eta}) & \text{if } Z\beta - \tilde{\varepsilon} > 0 \end{cases} \tag{9}$$

The logic of this construction is that it decomposes the determination of $\tilde{X}$ into two stages, a qualitative choice of whether or not to participate in recreation and, given a decision to participate, a quantitative choice of how much to participate. Moreover, it permits different factors to influence the two choices; the qualitative choice depends on the variables Z, with coefficients β, while the quantitative choice depends on the variables W, with coefficients γ. Combining the two stages of (9), one has

$$E\{X\} = E\{\tilde{X}|\tilde{X} > 0\} \cdot Pr\{\tilde{X} > 0\} . \tag{10}$$

In the first stage of (9), using all the households in the sample, one estimates the probability that a household participates in recreation

$$Pr\{\tilde{X} > 0\} = F_{\varepsilon}(Z\beta) \tag{11}$$

where $F_{\varepsilon}(\cdot)$ is the c.d.f. of the random term $\tilde{\varepsilon}$ in (9). This is a binary probit model if one uses the normal distribution, and a binary logit model if one uses the logistic distribution.[10] In the second stage, using the subset of households who do participate in recreation, one employs a regression model to estimate the conditional mean number of visits, $E\{\tilde{X}|\tilde{X} > 0\}$. It would be appropriate here to assume that the error term in the second stage, $\tilde{\eta}$, is a skewed, positive random variable, such as a lognormal variable. The lognormal regression can be set up in two different ways. One way is to write the conditional distribution of $\tilde{X}$ given that $\tilde{X} > 0$ as

$$g(w, \gamma; \tilde{\eta}) = \phi(W, \gamma) \cdot \tilde{\eta}, \qquad \tilde{\eta} = e^{\tilde{U}} \tag{12}$$

where $\tilde{U}$ is $N(0, \sigma^2)$. Taking logarithms, one obtains the regression model

Table II. Household Characteristic Variables.

Variable	Definition	Mean Value
INC	Household income ($)	14,317
EDUC	Highest household education (categories: 1 = elementary school, . . ., 4 = some college, . . . 7 = postgraduate education)	4.71
#ADULTS	Number of persons aged 18 and older in the household	2.57
#KIDS	Number of persons under 18 in the household	1.60
IRISH	1 if the household is predominantly Irish, 0 otherwise	0.210
ITALIAN	1 if the household is predominantly Italian, 0 otherwise	0.180
MINORITY	1 if the household is predominantly of a minority group, 0 otherwise	0.104
AUTO	1 if the household owns one or more cars, 0 otherwise	0.825
DAYSWORKED	Number of days per week worked by the head of the household	5.0
SWIMPOOL	1 if the household makes frequent use of a private swimming pool, 0 otherwise	0.160
BOFISH	1 if members of the household engage in fishing or boating, 0 otherwise	0.496
AVDIST	Average of the straight-line distances from the household's home to each of sites 1 through 29	10.196

$$\ell n\ \tilde{X} = \ell n\ \phi(W, \gamma) + \tilde{U} \tag{13}$$

which can be estimated by least squares. Note that in this set-up the conditional mean $E\{X|X > 0\}$ is $\phi(W, \gamma)e^{\sigma^2/2}$, and not $\phi(W, \gamma)$. Therefore $\phi(W, \hat{\gamma})$, where $\hat{\gamma}$ are the least squares coefficient estimates, is a biased estimator of $E\{\tilde{X}|\tilde{X} > 0\}$; in [17], Goldberger describes an adjustment procedure which eliminates the bias. The second approach to lognormal regression, proposed by Amemiya [4], is to assume directly that the conditional mean number of visits is some specific function

$$E\{\tilde{X}|\tilde{X} > 0\} = \psi(W, \gamma) \tag{14a}$$

and the conditional variance is

$$V\{\tilde{X}|\tilde{X} > 0\} = \kappa^2 \psi(W, \gamma)^2 \tag{14b}$$

where κ and γ are the parameters to be estimated.[11] Suppose that $\psi(W, \gamma) = W\gamma$. Amemiya suggests the generalized least squares (GLS) estimator

$$\gamma^{GLS} = (W'\hat{\Omega}^{-1}W)^{-1}W'\hat{\Omega}^{-1}X \tag{15}$$

where $\hat{\Omega} = \text{diag}[(W\hat{\gamma}^{OLS})^2]$ is an estimate of the conditional variance-covariance matrix (14b), based on a first-round OLS estimator $\hat{\gamma}^{OLS} = (W'W)^{-1}W'X$.

To summarize, the estimation of the demand model (2) is conducted in three steps. The first two steps are the estimation of the recreation participation model (11), and the regression model for the total number of recreation trips, based on (12) or (14). The third step is the estimation of the utility-theoretic logit model of the allocation of visits among individual sites, (6). Once estimated, the models may be combined to predict a household's visits to any site using the formula

$$E\left\{\begin{matrix}\text{\# visits}\\ \text{by house-}\\ \text{hold to}\\ \text{site } i\end{matrix}\right\} = Pr\left\{\begin{matrix}\text{house-}\\ \text{hold}\\ \text{selects}\\ \text{site } i\end{matrix}\right\}\cdot E\left\{\begin{matrix}\text{\# visits}\\ \text{by house-}\\ \text{hold to}\\ \text{all sites}\end{matrix}\middle|\begin{matrix}\text{household}\\ \text{partici-}\\ \text{pates in}\\ \text{recreation}\end{matrix}\right\}\cdot Pr\left\{\begin{matrix}\text{household}\\ \text{partici-}\\ \text{pates in}\\ \text{recreation}\end{matrix}\right\}. \quad (16)$$

The variables used in the first two steps are defined in Table II. They include household income, education, size and age composition, ethnic background, automobile ownership, and access to a private swimming pool. All of these variables measure characteristics of the household. In addition, as an index of the availability of recreation opportunities, I included the average distance from the household's home to the main recreation sites in the area.[12] For the recreation participation model I used the binary logit model, estimated by the maximum likelihood method. For the number of recreation trips, I used both the lognormal OLS regression model, (13), and Amemiya's GLS estimator (15); although both yielded similar results, the latter produced a slightly better fit and is reported here. The resulting coefficient estimates are shown in Table III.

These results may be summarized as follows. Household income does not appear to have a significant influence on recreation activity, but household education does—households with a higher education are significantly more likely to participate in water-based recreation; but given that they participate, they do not make significantly more trips. The age composition of the household has some influence. The probability of participation in recreation depends significantly on the number of adults in the household, but not on the number of children. However, given that a household participates, the number of trips which it makes depends on its size but not its age composition: the hypothesis that the coefficients of #KIDS and #ADULTS are the same cannot be rejected at the .05 level. As with other aspects of life in Boston, ethnic background has some effect on recreation behavior. Both Irish and Italian households are somewhat

TABLE III

Two-Stage Model of the Total Number of Visits to All Sites

Explanatory Variable	First-Stage Coefficients	Second-Stage Coefficients	
Equation No.	(1)	(2)	(3)
ℓn(INC)	.1747 (0.58)[a]	-1.846 (0.65)	-1.3663 (0.49)
ℓn(EDUC)	.7023 (1.76)	1.247 (0.31)	1.6572 (0.42)
SWIMPOOL	.6174 (1.23)	-2.5233 (0.70)	-2.9324 (0.81)
AUTO	.1048 (0.24)	4.2562 (1.09)	4.0315 (1.03)
DAYS WORKED	-.1229 (0.80)	0.2089 (0.14)	0.0969 (0.07)
AVDIST	-.1779 (2.74)	-2.0427 (3.69)	-2.0783 (3.76)
#KIDS	.0047 (0.05)	0.1786 (0.20)	
#ADULTS	.4919 (2.64)	1.8473 (1.48)	
PEOPLE			0.7473 (0.98)
IRISH		8.4637 (2.01)	8.2672 (1.97)
ITALIAN	.3017 (0.85)	-3.8338 (1.08)	-3.5341 (1.00)
MINORITY	-.6493 (1.44)	-3.9417 (0.81)	-4.2105 (0.87)
BOFISH		7.3613 (2.55)	7.2292 (2.50)
CONSTANT	.2611 (0.11)	46.569 (1.97)	44.335 (1.88)
SSR		14076	14120
R^2	.100[b]	0.071	0.068
$\ell n\ L^{est}$	-153.554		

[a]The absolute value of the t-statistic.

[b]The Pseudo-R^2 statistic [12, p. 123].

more likely than other white households to participate in recreation at public beaches; given that they participate, Irish households make significantly more visits, whereas Italian households make somewhat fewer visits. Minority households are less likely to participate in recreation at public beaches and, given that they participate, they make somewhat fewer visits.[13] Possession of an automobile has no influence on the probability of participating in recreation, but in households which participate it leads to somewhat more visits being made. Households with access to a private swimming pool also are more likely to participate in recreation at public beaches but, given that they participate, they make somewhat fewer visits. Households which participate in boating and fishing also make significantly more trips for general beach recreation. Lastly, households which are located at a greater distance from the main beaches in the Boston area are significantly less likely to participate in recreation and, if they do participate, they make fewer trips.

The variables used in the logit model for the allocation of visits among sites are listed in Table IV. Six are measures of water quality at each site: AMO, COD, FBACT, PHOS, TEMP and TURB.[14] The variable SITE TYPE, a dummy for freshwater as opposed to ocean sites, is included because there may be distinct preferences for the two types of site. The eighth variable, NUISANCE, measures nonwater aspects of site quality. The variable MINORITY ATT is included in order to allow for racial segmentation in recreation behavior: at certain sites in the Boston area, an unusually high percentage of visitors are from certain ethnic or racial groups. This phenomenon is handled here by creating a dummy and the site is one of those identified as having a special attractiveness to minorities, and 0 otherwise. The variable IRISH ATT is defined similarly for sites traditionally attractive to Irish households. These ten variables constitute the

Table IV. Site Characteristic Variables.

Variable	Definition	Type[a]	Mean Value
AMO	Ammonia content of water at the site (mg/l)	S	0.45
COD	Chemical oxygen demand of water at the site (mg/l)	S	34.2
FBACT	Fecal coliform bacteria in water at the site (#/100 ml)	S	1742.4
PHOS	Total phosphorus content of water at the site (mg/l)	S	0.07
TEMP	Water temperature at the site (° F)	S	67.5
TURB	Turbidity of water at the site (JTU)	S	5.61
SITE TYPE	1 if freshwater, 0 if saltwater site	S	0.26
NUISANCE	1 if the site has heavily urban or noisy setting, 0 otherwise	S	0.30
MINORITY ATT	1 if the household is a minority group and the site is identified as having a special attractiveness to minorities, 0 otherwise	H & S	0.03
IRISH ATT	1 if the household is Irish and the site is identified as having a special attractiveness to Irish, 0 otherwise	H & S	0.06
TRAVEL COST	Cost (in dollars) of traveling by automobile from home to site	H & S	1.62
ENTRY COST	Parking and entrance fees at the site (in dollars)	H & S	0.52
PRICE	Sum of TRAVEL COST and ENTRY COST	H & S	2.12

[a] S = varies over sites; H = varies over households.

components of b_i in the utility model (3). I assume that the function $\gamma_i(b_i)$ in (5) is linear in these variables, and its coefficients are the same across all sites--i.e., $\gamma_i\ (b) = \gamma(b)$ for all i. Moreover I assume that the constant terms are the same across all sites; together with the normalization requirement, this implies that $\alpha_i = 0$ all i. In the terminology of Domencich and McFadden [12, p. 53], this is an entirely "generic" specification of the $\tilde{\phi}_i$'s in (5).

The last three variables in Table IV are price variables. The construction of these prices requires some explanation. Ideally one would like to use the actual cost to households of visiting the sites, which depends partly on the size of the party, the method of transportation, and the route followed. The survey data on actual recreation expenditures were unsatisfactory, and it was therefore necessary to reconstruct the cost of visiting sites. There are two components of this cost. At eleven sites, parking and entrance fees are changed. Only two of these sites are accessible by public transportation; for these two sites I multiplied the parking fee by the percentage of the sample households visiting them which used private automobiles for access. For the other nine sites, I assumed that all access was by private automobile. The resulting values of the ENTRY COST variable are shown in Appendix Table IA.[15] The TRAVEL COST variable was constructed by computing the straight-line distance from each household's home to each recreation site, multiplying this by a conversion factor to obtain an estimate of actual road distance, and multiplying this in turn by an estimated travel cost of 7¢ per road mile (in 1974 dollars). Wilman [43] and others have argued that one should also include an estimate of the time component of travel costs. However, the survey data do not indicate how many adults or children were in the party visiting the site, nor whether the trips were made on weekdays or the weekend. It seems impossible, therefore, to derive any reasonable estimate of the shadow price of the time spent traveling to the sties. The final variable, PRICE, is the sum of TRAVEL COST and ENTRY COST. In principle, site choices should be based on this sum—this should correspond to the variable p_i in (3). This hypothesis is tested below.

There are in fact more than 200 water-based recreation sites available within a one-day trip's distance from the Boston SMSA. The 406 households in the sample who participated in water-based recreation reported that they had visited some 116 sites. However, as noted above, water quality data were available for only the 30 sites listed in Table IA; 372 households visited these sites, and their visits to these sites accounted for about 70% of all the visits reported by all households. Therefore, not too much is lost by concentrating on the demand for these sites. A second problem with the number of sites was less easily overcome. The multinomial logit maximum likelihood estimation program to which I had access could accommodate no more than 20 sites. Therefore, it was necessary to eliminate ten sites. This was done by forming two different subsets of 20 sites: Sample A (Saltwater Sites) includes sites 1-21 and 30, with sites 2 and 3, and 8 and 9 aggregated into two composite "sites" (354 households visited the sites in this sample);[16] and Sample B (Saltwater and Freshwater sites) includes site 2 and 3, and 8 and 9 aggregated into composite sites, plus sites 4, 5, 10-14, 16-19, 21-23, 26-28and 30 (352 households visited the sites in this sample).[17] The multinomial logit model was applied to each sample separately; the resulting coefficient estimates are presented in Table V.

The two sets of results are qualitatively similar and the coefficient estimates are of a similar order of magnitude, although they are by no means

TABLE V

Logit Model of Visit Allocation

EXPLANATORY VARIABLE	SAMPLE A			SAMPLE B		
EQUATION NO.	(1)	(2)	(3)	(4)	(5)	(6)
COD	-0.0392 (31.11)	-0.0304 (25.14)	-0.0391 (34.21)	-0.0444 (25.6)	-0.0358 (20.65)	-0.0427 (27.50)
TURB				-0.0338 (5.39)	-0.0304 (5.32)	-0.0174 (3.07)
PHOS	-7.1719 (12.50)	-11.6268 (19.09)	-0.383 (15.82)	-4.9145 (8.05)	-9.1986 (13.31)	-8.3613 (12.40)
ℓn(FBACT)	-0.1033 (8.36)	-0.0265 (2.22)	-0.0767 (7.28)	-0.2789 (17.91)	-0.1514 (9.78)	-0.2264 (17.93)
SITE TYPE				-2.0497 (27.23)	-1.7204 (22.78)	-1.9923 (29.13)
NUISANCE	-0.7326 (14.28)	-0.2775 (5.68)	-0.7054 (16.48)	-0.7627 (11.33)	-0.4177 (6.38)	-0.7118 (12.88)
PRICE	-0.4304 (31.95)			-0.5283 (34.81)		
TRAVELCOST		-1.3497 (42.30)	-1.1422 (42.67)		-1.3808 (46.68)	-1.3113 (47.40)
ENTRYCOST		0.3697 (15.85)			0.1888 (8.16)	
MINORITY-ATT	1.1177 (10.98)	1.2667 (12.18)	1.0471 (10.23)	0.6463 (6.56)	0.7701 (7.65)	0.6383 (6.45)
$\ell n\ L^{est}$	-14367	-13533	-13658	-13523	-12712	-12745
R^2	.096	.149	.141	.126	.179	.176

identical. In both cases it is clear that water quality conditions have a significant influence on site choice. Higher levels of COD, PHOS, FBACT, and, in sample B but not sample A, TURB significantly reduce the probability that a site is visited. Moreover, in sample B there is evidence of a distinct preference against freshwater sites, even when other water quality variables are taken into account. Two water quality variables have coefficients which are significant but of the "wrong" sign: AMO has a significant positive coefficient, and TEMP has a significant negative coefficient. There is no obvious explanation for this. A lower level of non-water quality at a site, as measured by NUISANCE, significantly reduces the probability that the site is visited. The coefficients of MINORITY ATT are significant and positive, which supports the suggestion of racial segmentation in recreation behavior. Lastly, the coefficient of PRICE (which corresponds to -h in (3)) is significant and negative. As noted above, PRICE is the sum of two components, TRAVEL COST and ENTRY COST. In equations (1) and (2), and (4) and (5) of Table V, I test the hypothesis that these two components have the same coefficient. In each case the hypothesis is decisively rejected.[18] Moreover, the coefficient of ENTRY COST is actually positive and significant. The explanation seems to be that most of the sites at which an entrance fee is charged are higher-quality saltwater or freshwater sites; therefore, the ENTRY COST variable may be serving as a surrogate for site quality. Accordingly, this variable is omitted in the final version of each model—equations (3) and (6) of Table V.

The goodness of fit measures reported in the table are Domencich and McFadden's pseudo-R^2 statistic [12, p. 123]. Other _ad hoc_ measures of goodness of fit are possible. For example, one can calculate the matrix of predicted probabilities that household h selects site i, and take the simple correlation coefficient between the elements of this matrix and those of the matrix of observed site selection probabilities. Using equations (3) and (6) of Table V, the squares of these correlation coefficients are 0.126 and 0.189, respectively. Alternatively one can multiply each household's vector of predicted site selection probabilities by its actual total number of visits to all the sites in the appropriate sample, and sum over all the households in the sample. This yields a prediction of the total number of visits to each site, which may be compared with the actual totals; the square of the simple correlation coefficient between the predicted and actual total number of visits to each site is 0.407 for the sites in sample A, and 0.493 for the sites in sample B. This tests the goodness of fit of the allocation model, (6). In order to test the overall demand model, the above calculation was repeated using the predicted total number of visits by each household to all sites, based on equation (3) of Table III. The calculation was based on the allocation model for sample B sites, and was restricted to the subset of 352 households which visited these sites (i.e., the third term on the right-hand side of (16) was set equal to one). This yielded a prediction of each of these household's total number of visits to all sites; this was adjusted for the sites in sample B by multiplying it by the actual proportion of the household's visits which were made to these sites. The square of the simple correlation coefficient between the total number of visits to each site predicted in this way and the actual total is 0.466.

IV. AN AD HOC MODEL OF SITE DEMAND

It was pointed out in Section II that by insisting on a utility-theoretic approach to the demand for recreation sites one surrenders some flexibility in modeling observed choice behavior. In this section I will adopt the alternative approach in which one postulates a purely statistical model of recreation demand, not derivable from any utility maximization hypothesis. The implications of this ad hoc demand model will be compared with those of the utility-theoretic demand model estimated in the preceding section. In the notation used above, the general form of an ad hoc demand model might be $\tilde{x}_i = F_i(p, b, y)$ $i = 1, \ldots, N$ or, more simply,

$$\tilde{x}_i = F(p_i, b_i, y) \qquad i = 1, \ldots, N, \tag{17}$$

which is the form used in [36], or

$$x_i = F(p_i, p^*, b_i, b_1^*, \ldots, b_K^*, y) \qquad i = 1, \ldots, N \tag{18}$$

where

$$p^* = \phi(p_1, \ldots, p_N) \text{ and } b_k^* = \psi_k(b_{1k}, \ldots, b_{Nk}) \qquad k = 1, \ldots, K.$$

The point of the formulation in (18) is that the demand for a site is influenced by the prices and quality characteristics of other sites only through the general indices p^* and $b_1^*, \ldots, b_K^*$. This is a compromise between the formulation in (17), where the demand for a site is independent of the attributes of the other sites, and the general formulation, where it depends on all of them explicitly. The formulation in (18) was originally suggested by Quandt and Baumol [35] for the purpose of modeling the demand for alternative modes of transportation; they called it the "abstract mode" demand model, because the demand functions are entirely generic. In their formulation the indices p^* and b_k^* were defined as $p^* = \min\{p_1, \ldots, p_N\}$ and $b_k^* = \max\{b_{1k}, \ldots, b_{Nk}\}$— i.e., p^* is the cost of visiting the cheapest site, and b_k^* is the best value of characteristic k attainable at any site. As an alternative, one might define p^* and b_k^* to be the means of the p_i's and b_{ik}'s, or the medians. All three versions will be tested below.

The demand model (18) will be applied to the 30 sites listed in Table IA; as noted above, these sites were visited by 372 households in the sample. The dependent variable in (18) is the number of visits by each household to each site; thus, there are 11,160 (= 372 X 30) observations on this variable. Their distribution is summarized in the third panel of Table IIA. This distribution has the same features as the distribution of $\tilde{X}$, discussed in the preceding section: $\tilde{x}_i$ is a non-negative variable; it has a discrete probability mass at zero—of the 11,160 possible household/site interactions, only 923 (= 8.3%) actually occurred; and its distribution has a long tail. Therefore, I will use the same type of two-stage statistical model for $\tilde{x}_i$ as was used for $\tilde{X}$[19]

$$\tilde{x}_i = \begin{cases} 0 & \text{if } h(p_i,\ p^*,\ b_i,\ b^*,\ y) - \tilde{\varepsilon} \leq 0 \qquad (19a) \\ g(p_i,\ p^*,\ b_i,\ b^*,\ y;\ \tilde{\eta}) & \text{otherwise.} \qquad (19b) \end{cases}$$

Note that this demand model ignores the question of recreation participation —it assumes that the household has already decided to participate. This decision can be incorporated by combining (19) with the participation model (11). Thus, the full model may be written

$$E\left\{\begin{matrix}\text{\# visits by}\\ \text{household}\\ \text{to site i}\end{matrix}\right\} = E\left\{\begin{matrix}\text{\# visits by}\\ \text{household}\\ \text{to site i}\end{matrix}\ \middle|\ \begin{matrix}\text{household partici-}\\ \text{pates in recreation}\\ \text{and visits site i}\end{matrix}\right\} \cdot$$

$$\cdot \Pr\left\{\begin{matrix}\text{household}\\ \text{visits site i}\\ \text{at all}\end{matrix}\ \middle|\ \begin{matrix}\text{household}\\ \text{participates}\\ \text{in recreation}\end{matrix}\right\} \cdot \Pr\left\{\begin{matrix}\text{household}\\ \text{participates}\\ \text{in recreation}\end{matrix}\right\} . \qquad (20)$$

The second term on the right-hand side of (2) corresponds to the first stage of the abstract mode demand model, (19a); the first term corresponds to the second stage, (19b); and the third term corresponds to (11).

Several different formulations of the first stage model, (19a), were estimated using logit; the final formulation is presented in Table VI. The results are qualitatively similar to those obtained for the logit model of the allocation of visits among sites, presented in Table V, although the two models are conceptually different—(19a) involves a binary choice of whether or not to visit each site taken separately, whereas the visit allocation model (6) involves a multinomial choice of one site out of N alternatives. Among the water quality variables, COD, PHOS and FBACT have a significant negative effect on site visitation; as in the visit allocation model, the coefficients of AMO and TEMP are positive, whereas the coefficient of TURB is highly unstable. There is a distinct preference against freshwater and urban sites. The hypothesis that the two components of PRICE have the same coefficient can be rejected and the coefficient of ENTRY COST is positive. Household size has a significant positive effect on site visitation,[20] but other household characteristics (income, education, access to a private swimming pool) do not have significant coefficients. The abstract mode formulation was tested with COD, PHOS, FBACT and TRAVEL COST by using the ratios of each of these variables to their mean, median and maximum or minimum; for all of these variables the abstract mode formulation provided a poorer fit than the formulation presented in the table.

The second stage of the demand model, (19b), consists of the regression of the number of visits to a site on various attributes of the site and the household using the data for the 923 cases in which a site was visited by a household. To estimate the model I used both lognormal OLS regression, as in (13), and Amemiya's GLS estimator, (15). The OLS coefficients provided a better fit, and are reported in Table VII.[21] These results may be compared with those in Table III, which deal with the determinants of the households' overall recreation activity, and with the first-stage

TABLE VI

Ad Hoc Site Demand Model: First Stage

Explanatory Variable	Coefficient Estimate
COD	-0.0384 (13.39)
PHOS	-16.137 (10.48)
ℓn(FBACT)	-0.0709 (2.73)
SITE TYPE	-1.6862 (12.76)
NUISANCE	-0.7363 (7.55)
ℓn(TRAVELCOST)	-1.1808 (22.74)
IRISH-ATT	0.0348 (0.13)
MINORITY-ATT	0.8116 (2.60)
# PEOPLE	0.0508 (3.15)
CONSTANT	0.1945 (1.30)
R^2	.142
$\ell n\ L^{est}$	-2732.77

TABLE VII

Ad Hoc Site Demand Model: Second Stage

Explanatory Variable	Coefficient Estimates				
Equation No.	(1)	(2)	(3)	(4)	(5)
INC[a]	7.206 (1.87)	7.238 (1.89)	7.172 (1.89)	6.996 (1.85)	6.987 (1.85)
# PEOPLE	0.0166 (1.17)	0.0187 (1.33)	0.0187 (1.33)	0.0185 (1.32)	0.0185 (1.33)
SWIMPOOL	-0.2052 (2.45)	-0.2065 (2.47)	-0.2105 (2.55)	-0.21 (2.55)	-0.2109 (2.56)
IRISH	0.0932 (1.23)	0.1016 (1.34)	0.0825 (1.10)	0.0839 (1.12)	0.0837 (1.12)
MINORITY	0.1853 (1.55)	0.2004 (1.69)	0.1976 (1.68)	0.2027 (1.73)	0.2046 (1.74)
ℓn(TRAVELCOST)[b]	-0.4217 (9.86)	-0.4341 (10.93)	-0.4248 (10.93)	-0.4248 (10.96)	-0.4324 (10.66)
SITE TYPE	-0.6414 (3.73)	-0.5158 (3.50)	-0.4416 (4.98)	-0.4428 (5.01)	-0.5259 (3.57)
NUISANCE	-0.2755 (2.77)	-0.262 (2.97)	-0.163 (2.32)	-0.1606 (2.30)	-0.2043 (2.29)
IRISH-ATT	0.572 (2.38)	0.4878 (2.09)	0.4333 (1.90)	0.433 (1.90)	0.4462 (1.94)
MINORITY-ATT	0.2842 (1.01)	0.2246 (0.81)	0.1787 (0.68)	0.1793 (0.69)	0.1555 (0.57)
AMO	0.178 (1.53)				
COD	-0.004 (1.63)	-0.0034 (1.64)			-0.0016 (0.76)
TURB	0.0104 (0.74)				
PHOS	-1.5872 (1.14)				
ℓn(FBACT)	0.0204 (0.76)				
TEMP	0.0116 (0.67)	0.0075 (0.45)			0.0053 (0.32)
RATING-W. QUAL.			0.066 (2.03)	0.0536 (1.94)	0.0487 (1.71)
RATING-CROWDING			-0.0462 (1.72)	-0.0504 (1.93)	-0.0475 (1.80)
RATING-B. FACIL.			-0.0097 (0.34)		
RATING-B. QUAL.			-0.0183 (0.51)		
RATING-TEMP			0.1436 (5.11)	0.1415 (5.07)	0.1407 (5.03)
CONSTANT	1.0228 (0.54)	1.2381 (0.77)	1.1645 (5.16)	1.1301 (5.19)	0.883 (0.43)
SSR	782.8	786.02	758.52	758.99	758.46
R^2	.096	.099	.134	.133	.132
F	10.34	13.48	13.33	15.36	13.33

[a] The coefficient should be multiplied by 10^{-6}.

[b] The coefficient of the log of the ratio of travel cost to the median travel cost to all sites.

results in Table VI. It was shown above that household income does not significantly influence either the total number of trips by a household to all sites or the probability that a household selects an individual site. The results in Table VII imply that household income does have a significant influence on the number of visits made by households to the sites which they do visit—given that a household visits a site, it makes more visits if it has a higher income. However, household education does not have a significant effect. Not surprisingly, larger households make more visits to each site, while households with access to a private swimming pool make fewer visits.[22] Irish households make more visits to each site, and so do minorities. It was shown above that minority households make a smaller total number of visits and there is also evidence (not reported above) that they visit fewer sites than white households; evidently their visits are more concentrated on a small number of sites. As for the effect of site characteristics on demand, there is still a distinct preference against freshwater sites: not only are households less likely to select a freshwater site but, if they do select one, they make significantly fewer visits to it. The same is true of heavily urbanized sites. As before, the hypothesis that the two components of PRICE have equal coefficients can be rejected, and the coefficient of ENTRY COST is positive. The abstract mode formulation was tested by using the ratio of TRAVEL COST to its mean, median and minimum values over all sites; the versions with the mean and, especially, the median provided a better fit.

In addition to the objective measures of water quality used previously —AMO, COD, etc.—it was possible in these regressions to employ indices of the households' subjective assessments of site quality. These assessments were obtained during the household survey as responses to the following question:

For each site with which you are familiar, would you please rate each of the following characteristics of the site on a scale from 1 to 5. For this rating, 1 means bad, 2 means moderately bad, 3 is fair, 4 is moderately good, and 5 is good:

A. Water temperature

B. Water quality (clarity, color, odor, weeds, etc.)

C. Beach facilities (availability, etc.)

D. Beach quality (setting, maintenance)

E. Crowding.

The resulting ratings will be referred to as RWTEMP, RWQUAL, RBFAC, RBQUAL and RCROWD, respectively; they are all integers between 1 and 5. In practice it turned out that most households provided ratings only for the sites which they actually visited plus, possibly, one or two other sites. These ratings could not be used as explanatory variables in the utility-theoretic trip allocation model (6), or in the first stage of the ad hoc demand model, (19a), because there one needs a complete set of observations on site characteristics: for each variable and each household, one needs to know the values of that variable across all sites. This is not needed for the second stage, (19b). Therefore, it is here possible to examine the issue of whether there is a closer relationship between site visitation and either objective or perceived levels of site quality.

The results can be found in Table VII.[23] Of the objective measures of water quality—AMO, COD, FBACT, PHOS, and TURB—only COD consistently has a negative coefficient. Thus, factors such as PHOS and FBACT, which significantly lower the probability that a household selects a site, do not significantly affect the frequency with which it visits the site. Moreover, the inclusion of these variable has virtually no impact on the model's goodness of fit (compare equations (1) and (2), or (3) and (4) in Table VII. By contrast, several of the subjective ratings are significantly correlated with site visitation, and their inclusion substantially improves the model's fit. In particular, there is a significant positive relationship between the frequency of visitation and RWQUAL and RTEMP, and a significant negative relationship with RCROWD. Thus, households make more visits to sites which they judge to be more crowded, and whose water quality and temperature they rate more highly.[24] This raises a question about the direction of causation: do households visit sites more frequently because they perceive them to be of higher quality, or do they rate the sites' quality more highly because they visit them more frequently (i.e., the ratings are offered as a rationalization of their behavior)? On the basis of the present data it is not possible to answer the question definitively, because of the incompleteness of the available ratings. However, in a separate paper [24] I have performed a statistical analysis of the relationship between these ratings and objective measures of site quality. The evidence there suggests that for water temperature and, to a lesser extent, for water quality, households do in fact give a higher rating to sites which they visit more frequently. It is conceivable, therefore, that the causation could run in either direction.

In order to assess the overall fit of the demand model estimated in this section, I used the model to predict the total attendance at each of the thirty sites. This prediction is confined to the 372 households which actually visited these sites (i.e., the third term on the right-hand side of (20) is set equal to one). The coefficients of the first stage of the model are those given in Table VI; the coefficients of the second stage are alternately those given in equations (2) and (4) of Table VII. The square of the simple correlation coefficient between the actual and predicted total number of visits to each site is 0.145, using objective measures of site quality in the second stage (i.e., equation (2) of Table VII), and 0.148 using the subjective ratings of site quality (i.e., equation (4)). By this criterion, the ad hoc model fits the data somewhat less well than the utility-theoretic model estimated in the previous section. However it sheds light on the relationship between recreation choices and objective and subjective measures of site quality, which could not be investigated within the framework of the utility-theoretic model.

V. IMPLICATIONS AND CONCLUSIONS

In this paper two different models of the demand for recreation sites have been estimated. The first is derived from an explicit model of household utility maximization in allocating visits among recreation sites. The second is a purely statistical model and cannot be derived from a utility maximization model; it involves a distinction between whether or not a site is visted at all, and how many trips are made if the site is visited. This dichotomy permits one to investigate whether the two decisions are influenced by different factors and to incorporate in the second-stage data on subjective perceptions, as opposed to objective measures, of site quality. The evidence suggests that the first model fits the data somewhat better in the aggregate. In some respects the empirical implications of the two models are similar; in others they differ. Both sets of results underscore the importance of socio-economic factors, including ethnic and racial background, as determinants of recreation activity. Both also indicate the negative effect of distance on site visitation, and the existence of a distinct preference against freshwater sites. The two models have somewhat different implications for the impact of water quality on site demand. The first model suggests that COD, phosphorus, fecal coliform bacteria and turbidity have a significant negative effect on the choice of a recreation site. The second model suggests that, whereas the first three of these variables have a significant negative effect on whether or not a household visits a site at all, they have relatively little impact on the number of visits made to the sites which are visited; the latter decision is influenced far more strongly by households' perceptions of site quality. This finding is intriguing but, for the reasons noted above, it is not possible to test the influence of perceived versus objective measures of site quality more systematically with the data at hand. Nor is it possible to resolve the direction of causation: it may be that households visit sites more often because they perceive them to be better, or that households rate sites more highly because they visit them more frequently.

In order to illustrate the relative influence of water quality variables and travle cost on site visitation in the two demand models, I simulated the effects of a 10% and 50% reduction in these variables. In each case I assumed that there was a 10% or 50% reduction in the variable at one site alone, with no change in any other variable at that or any other site, and I calculated the consequent change in the total visitation of the site. The results are summarized in elasticity form in Table VIII; these figures are the median elasticities for the 20 sites used in the estimation of the first model, and the 30 sites used in the second model. The results in the first row of the table are based on the coefficients of equation (6) of Table V; following the logic of that model, I assume that these changes would not affect the total number of recreation trips by a household, but only the allocation of this total among the alternative sites. The results in the second row pertain to the ad hoc demand model and are based on the first-stage coefficients in Table VI and the second-stage coefficients in equation (4) of Table VII; they also employ a statistical model of the relationship between changes in the objective water quality variables and the resulting changes in households' subjective ratings of water quality which is developed in [24]. In each case the elasticity of demand with a 50% change is larger than the elasticity with a 10% change: this reflects the nonlinearity of the underlying logit models. In both models the elasticities of demand with respect to TRAVEL COST and COD are considerably larger than those with respect to PHOS, FBACT and TURB. However, the utility-theoretic model suggests that COD has the largest single impact on demand, whereas the ad hoc model suggests that TRAVEL COST has the largest impact.

TABLE VIII

Aggregate Demand Elasticities

	Median Elastitity of:				
	COD	PHOS	FBACT	TURB	TRAVELCOST
Utility-Theoretic Demand Model					
10% Reduction	-1.5	-0.4	-0.2	-0.1	-1.0
50% Reduction	-2.0	-0.4	-0.3	-0.1	-1.3
Ad Hoc Demand Model					
10% Reduction	-1.2	-0.7	-0.1		-1.6
50% Reduction	-1.5	-0.8	-0.1		-3.3

As noted in Section II, the first demand model can be used to calculate exact measures of consumer's surplus corresponding to the areas PAB and ABCD in Figures 1 and 2. I will start with the latter. The idea of this calculation is that by estimating a demand model derived from a specific utility maximization model one can obtain an estimate of the underlying utility function, which in turn can be used to calculate the monetary compensation needed to offset the effects of a change in site quality. In fact, although several water pollution abatement projects have been undertaken in the Boston area to meet the requirements of the 1972 Amendments to the Water Pollution Control Act, according to a recent study there will be no significant improvement in water quality in the lower Charles River and Boston Harbor, where most of the sites considered here are located.[25] Therefore, as an illustration of the methodology, I have computed the benefits from the hypothetical water quality changes mentioned above, involving a 10% or 50% reduction in various pollutants at each site separately. These benefit estimates are shown in Table IX. The logic of the demand model is that quality changes affect site selection probabilities—the first term on the right-hand side of (16)—but not the total household recreation activity—the second and third terms in (16). Each time a household selects a recreation site (i.e., each time it makes a recreation trip) it reaps some benefit from the quality change; this benefit is given by the formula in (8).[26] Over the entire summer recreation season, the total benefit to the household is obtained by multiplying this benefit per site choice by the household's actual total number of recreation trips for the season. Since households vary in their location and socioeconomic characteristics, they do not receive the same benefit from a given quality change at a particular site. Therefore, I have tabulated the average benefit over the 352 households in the sample. Thus, the first entry in Table IX states that the average benefit per household

Table IX. Average Benefit Per Household from a 10% and 50% Reduction in Various Water Pollutants.

Site	Benefit (¢) From 10% Reduction in				Benefit (¢) From 50% Reduction in			
	COD	PHOS	FBACT	TURB	COD	PHOS	FBACT	TURB
2, 3	19.1	3.2	2.3	0.5	134.1	17.3	16.1	2.5
4	21.7	6.4	3.6	0.3	137.3	34.3	25.2	1.3
5	13.7	3.7	1.7	0 [a]	96.1	20.3	12.2	0 [a]
8, 9	18.0	5.8	2.7	1.0	118.3	31.7	19.0	5.0
10	6.3	5.2	0.9	0.5	42.6	33.6	6.6	2.8
11	3.9	0.7	0.2	0.2	42.8	4.3	1.6	0.9
12	8.1	2.3	0.7	0.3	71.6	13.9	4.6	1.4
13	26.4	5.8	2.7	0.6	192.5	31.5	18.9	3.0
14	3.8	1.3	0.5	0.5	27.4	7.5	3.2	2.8
16	0 [a]	5.6	4.0	1.1	0 [a]	29.3	27.4	5.8
17	7.0	2.0	1.4	0 [a]	43.1	10.8	10.1	0 [a]
18	10.5	1.3	1.8	0 [a]	66.2	6.5	12.5	0 [a]
19	5.0	0.3	0.4	0 [a]	41.2	1.6	3.1	0 [a]
21	1.3	0.2	0.2	0.1	9.6	1.1	1.1	0.3
22	1.6	1.8	1.3	0 [a]	8.7	9.8	9.2	0 [a]
23	0.7	1.2	0.1	0.1	4.1	9.3	1.0	0.3
26	2.0	1.0	1.4	0.2	10.7	5.0	9.8	1.0
27	0.6	0.7	0.3	0.3	3.0	4.0	2.4	1.8
28	0 [a]	1.0	0.6	0.6	0 [a]	5.6	4.1	0.3
30	0.2	0.1	0.0	0.0	1.3	0.4	0.2	0.1

[a] No change in pollution level—existing level is zero.

over the summer recreation season from a 10% reduction in COD at sites 2 and 3 is 19.1 cents. In fact, the minimum benefit to any household is 7.6 cents, and the maximum is 43 cents; the standard deviation is 8.0 cents. A similar variation underlies the other entries in the table. It should be noted that the benefits from a 50% reduction in pollutants are generally more than five times the benefits from a 10% reduction; this reflects the nonlinearity of the underlying demand model.

The fitted demand model can also be used to estimate the consumer's surplus from each site, corresponding to the area PAB in Figure 1. As explained in Section II, this is equivalent to the welfare loss from an increase in price from the current level to a price of infinity, at which the household would have a zero probability of visiting the site. The loss per site choice occasion is given by (8), and multiplication by the household's total number of recreation trips yields the total welfare loss (consumer's surplus) over the summer recreation season. The average household's consumer's surplus for each site is tabulated in the first column of Table X. The first entry in the column shows that the average household's consumer's surplus from sites 2 and 3 is about $1. The average over <u>all</u> sites is about 60 cents. These figures are the consumer's surplus for the whole recreation season—i.e., the area PAB. As noted above, most previous estimates of the consumer's surplus from recreation sites have been couched in terms of the benefit per visitor-day—i.e., the length (AP)/2 in Figure 1. In order to compare my results with these estimates, I divided the total consumer's surplus for each site by the predicted total number of visitor-days at the site, using the total number of household visits at each site predicted by the model and survey data on the average number of persons in each household's party when it visits a site. The results are shown in the second column of Table X. The average consumer's surplus per visitor-day over all sites is about 42¢.

Table X. Household's Consumer's Surplus from Boston Area Recreation Sites.

Site	Average Consumer's Surplus (¢)	
	Per Household	Per Visitor-Day
2, 3	100.8	10.5
4	162.4	11.0
5	74.6	13.0
8, 9	119.9	76.6
10	40.0	91.5
11	9.5	3.9
12	27.3	24.7
13	120.4	255.3
14	19.1	167.8
16	188.4	18.1
17	63.1	21.9
18	80.2	22.1
19	18.8	11.1
21	6.3	3.2
22	56.0	20.7
23	5.8	2.1
26	61.2	17.0
27	14.4	60.4
28	25.5	13.6
30	1.3	0.9

These estimates of consumer's surplus per visitor-day, which are of the same order of magnitude as those obtained in [22] using a utility-theoretic demand model applied to the subset of households who visit only one site, are considerably lower than the unit day values promulgated by the Water Resources Council and employed in the recent benefit assessment studies summarized in Section I. If these results can be extrapolated to other coastal areas with abundant recreation sites, they would suggest that the unit day values now widely used in recreation benefit assessment are excessive.

Table IA. Boston Area Recreation Sites

SITE ID	SITE NAME
1	Kings Beach (Swampscott)
2	Lynn Beach (Lynn)
3	Nahant Beach (Nahant)
4	Revere Beach (Revere)
5	Short (Revere)
6	Winthrop Beach (Winthrop)
7	Constitution Beach/Orient Heights (Boston)
8	Castle Island (Boston)
9	Pleasure Bay (Boston)
10	City Point (Boston)
11	L & M Street Beaches (Boston)
12	Carson Beach (Boston)
13	Malibu Beach/Savin Hill (Boston)
14	Tenean Beach (Boston)
15	Wollaston Beach (Quincy)
16	Nantasket Beach (Hull)
17	Wingaersheek Beach (Gloucester)
18	Crane's Beach (Ipswich)
19	Plum Island (Newbury)
20	Duxbury Beach (Duxbury)
21	White Horse Beach (Plymouth)
22	Breakheart Reservation (Saugus)
23	Sandy Beach/Upper Mystic Lake (Winchester)
24	Houghton's Pond/Blue Hills Reservation (Milton)
25	Wright's Pond (Medford)
26	Walden Pond (Concord)
27	Stearns Pond/Harold Parker State Forest (Andover)
28	Cochituate State Park (Natick)
29	Hopkinton State Park (Hopkinton)
30	Cape Code Beaches

Table IIA. Household Patterns of Site Visitation

Number of sites visited	Number of occurrences	Number of visits to all sites	Number of occurrences	Number of visits to each site	Number of occurrences
0	56	0	56	0	10,237
1	106	1 - 10	186	1 - 5	662
2	114	11 - 20	84	6 - 10	129
3	69	21 - 30	36	11 - 15	57
4	54	31 - 40	28	16 - 20	10
5	21	41 - 50	23	21 - 30	36
6	17	51 - 60	13	31 - 40	7
7	10	61 - 100	21	41 - 50	7
8	3	101 - 140	11	51 - 70	8
9	3	141 - 200	2	81 - 100	7
10	2	201-280	2		
11	1				
TOTAL	462	TOTAL	462	TOTAL	11,160

FOOTNOTES

[1]The author would like to thank Professors Dale Jorgenson and Robert Dorfman for their generous assistance in supervising the dissertation on which this paper is based. He is grateful to the U.S. Environmental Protection Agency for funding the collection of the data used in this study.

[2]Notable exceptions are [7], [9], and [8]. Only the last explicitly includes site quality variables; however the demand functions used there are not consistent with a utility maximization hypothesis.

[3]This assumption implies that the prices of all other goods move in the same proportion. Alternatively, one could assume that the utility function is recursively separable, $u = u[f(x, b), z]$, and interpret the resulting demand functions for x_i as conditional demand functions; in this case the argument y becomes total expenditure on beach recreation rather than total income.

[4]This utility model can be regarded as a generalization of Lancaster's model [28], in which the utility function takes the special form

$$u(x, z, b) = u(\sum_i x_i b_{i1}, \ldots, \sum_i x_i b_{iK}, z).$$

It should also be noted that the b_{iK}'s can represent either objectively measured quality characteristics or the consumer's subjective perceptions of quality. The difference between objective and subjective measures of quality will be explored below.

[5]The coefficient h may be interpreted as the marginal utility of income. More general formulations of the utility function may be designed in which h is itself a function of income or other characteristics of the individual.

[6]Hereafter a tilde will be used to denote a random variable or function.

[7]Most applications of multinomial probit involve $N \leq 4$; even if the Clark approximation were used—see [10]—it is unlikely that one could satisfactorily handle cases where N is much larger than, say, 10. For an exposition of multinomial logit see [12] or [30].

[8]That is, $\Pr\{\tilde{\varepsilon}_i \leq s\} = \exp[-\exp(\alpha_i - s)]$.

[9]The data presented in [5] contain a number of errors and anomalies. The corrected data set is described in [19] and is employed here.

[10]This logit or probit model should be regarded as a purely statistical model; there is no utility-theoretic interpretation which is meaningful in the present context.

[11]The specification (14a, b) may be compared with (12) by noting that it implies the following conditional distribution of $\tilde{X}$ given that $\tilde{X} > 0$

$$g(W,\gamma;\tilde{\eta}) = [\psi(W,\gamma)(\kappa^2 + 1)^{-\frac{1}{2}}]\tilde{\eta}$$

where $\tilde{\eta} = e^{\tilde{\tau}}$ and $\tilde{\tau}$ is $N[0, \ell n(\kappa^2 + 1)]$.

[12]It is not practical to include an index of the overall quality of the available recreation sites because this does not vary across the households in the sample: they all have essentially the same choice set. One would have to perform studies like this for a variety of cities in order to obtain variation in site quality and study its effect on the overall level of recreation activity. An alternative procedure, which does not really meet this objection, is to use the estimated value of $\ell n\, \lambda$, where λ is defined in (8), instead of AVDIST. However, the only variation in λ across households comes from differences in p, which are also picked up by AVDIST.

[13]It is worth mentioning that in the two summers after this survey was conducted there were serious incidents of racial violence at some Boston area beaches.

[14]These were selected from a larger set of water quality variables after a cluster analysis had been performed.

[15]At a few sites charges are levied for the use of on-site facilities such as bath houses, but there charges are very small; therefore they are ignored here.

[16]For each composite site, I took the average of the prices and water quality variables of the two component sites.

[17]An alternative procedure for "collapsing" the number of sites would have been to postulate a nested structure based on the Generalized Extreme Value Distribution—see [31]—with all the saltwater sites and all the freshwater sites aggregated into two groups. This was not tried here because of the greater computational complexity.

[18]The likelihood ratio test statistics are 1668 and 1662; both values greatly exceed the .99 percentile of the $X^2(1)$ distribution.

[19]If there were information on the temporal sequence in which households visited sites—which I do not have—an entirely different statistical model of site demand could be developed based on the Markov chain approach.

[20]The logic of including this variable is that, for each site, there is a greater likelihood that it will be visited if the household is larger. I also used #ADULTS and #KIDS as separate variables; the hypothesis that their coefficients are equal could not be rejected at the .01 level.

[21]Recall that the regressand in the OLS model is the logarithm of the number of visits to a site. However, the R^2 statistics reported in the table are the square of the simple correlation coefficients between the actual and predicted numbers of bisits, not the logarithms of these numbers. The constant terms reported in the table have been adjusted in the manner suggested by Goldberger [17].

22 In preliminary tests the hypothesis that the coefficients of #ADULTS and #KIDS are the same could not be rejected at the .01 level.

23 In a few cases some ratings were missing—e.g., a respondent rated the water quality but not the beach quality of a site. In these cases, the missing rating was set equal to the overall mean.

24 The finding that households make more visits to sites which they perceive to be more crowded requires some comment. In principle these are two different ways in which crowding can be related to site visitation. On the one hand crowding renders the recreation experience less pleasant; in this case there should be a negative correlation between crowding and visitation, as has been observed in such studies as [11] and [29]. On the other hand, if a site is more crowded, this may be because something about the site makes it especially attractive to recreationists. To the extent that the source of this attraction is not picked up by the other site quality variables included in the regression, the crowding variable serves as a proxy for site quality and will be positively correlated with visitation; evidently this is happening here.

25 See [13]; this study concludes that ". . . general recreational attractiveness will not be improved in the lower Charles and Boston Harbor due to [continuing] combined sewer overflows and high residual color and turbidity from upstream" (op. cit., p. 13.3).

26 This calculation is based on the coefficients in equation (6) of Table V. It should be noted that, in practice, a water quality improvement project would probably affect several water quality parameters and several sites simultaneously, although the calculations reported here are performed for each parameter and each site separately.

REFERENCES

1. B. Ackerman, S. Rose-Ackerman, J. Sawyer, and D. Henderson, "The Uncertain Search for Water Quality," Free Press, New York (1974).

2. J. Aitchison, On the distribution of a positive random variable having a discrete probability mass at the origin, J. Amer. Stat. Assoc. 50, 901-908 (1955).

3. J. Aitchison, and J.A.C. Brown, "The Lognormal Distribution," Cambridge Univ. Press, London (1966).

4. T. Amemiya, Regression analysis when the variance of the dependent variable is proportional to the square of its expectation, J. Amer. Stat. Assoc. 68, 928-934 (1973).

5. C.S. Binkley, and W.M. Hanemann,"The Recreation Benefits of Water Quality Improvement: Analysis of Day Trips in an Urban Setting" (EPA-600/5-78-010), U.S. Environmental Protection Agency, Washington, D.C. (1978).

6. N.W. Bouwes, and R. Schneider, Procedures in estimating benefits of water quality change, Amer. J. Ag. Econ. 61, 535-539 (1979).

7. O.R. Burt, and D. Brewer, Estimation of net social benefits from outdoor recreation, Econometrica 39 (1971).

8. F.J. Cesario, and J.L. Knetsch, A recreation site demand and benefit estimation model, Reg. Studies 10 (1976).

9. C.J. Cicchetti, A.C. Fisher, and V.K. Smith, An econometric evaluation of a generalized consumer surplus measure: The Mineral King controversy, Econometrica 44, 1259-1276 (1976).

10. C. Daganzo, "Multinomial Probit," Academic Press, New York (1979).

11. T.A. Deyak, and V.K. Smith, Congestion and participation in outdoor recreation: A household production function approach, J. Environ. Econ. Manag. 5, 63-80 (1978).

12. T.A. Domencich, and D. McFadden, "Urban Travel Demand: A Behavioral Analysis," North-Holland Publishing Company, Amsterdam (1975).

13. Environmental Research and Technology, Inc., "Environmental Impact Assessment, Water Quality Analysis, Charles River/Boston Harbor," Concord, Mass. (1976).

14. A.M. Freeman III, "The Benefits of Environmental Improvement: Theory and Practice," Johns Hopkins Press, Baltimore (1979).

15. A.M. Freeman III, "The Benefits of Air and Water Pollution Control: A Review and Synthesis of Recent Estimates," Report prepared for the Council on Environmental Quality, Bowdoin College, Brunswick, Maine, December (1979).

16. A.S. Goldberger, "Econometric Theory," Wiley, New York (1964).

17. A.S. Goldberger, The interpretation and estimation of Cobb-Douglas functions, Econometrica 35, 464-472 (1968).

18. H.W. Grubb, and J.T. Goodwin, "Economic Evaluation of Water-Oriented Recreation in the Preliminary Texas Water Plan," Report 84, Texas Water Development Board, Austin, Texas (1968).

19. W.M. Hanemann, "A Methodological and Empirical Study of the Recreation Benefits from Water Quality Improvement," Ph.D. dissertation, Department of Economics, Harvard University (1978).

20. W.M. Hanemann, Measuring the worth of natural resource facilities: Comment, Land. Econ. 56, 482-486 (1980).

21. W.M. Hanemann, "Quality Changes, Consumer's Surplus, and Hedonic Price Indices," Economics Working Paper No. 116, University of California—Berkeley, Revised November (1980).

22. W.M. Hanemann, "A Qualitative-Quantitative Model of Consumer Choice with an Application to Recreation Demand," Dept. Ag. and Resource Econ., University of California-Berkeley, September (1980).

23. W.M. Hanemann, Quality and demand analysis, in "New Directions in Econometric Modeling and Forecasting in U.S. Agriculture" (G.C. Rausser, Ed.), Elsevier/North-Holland Publishing Co., New York, forthcoming (1981).

24. W.M. Hanemann, "The Perception of Recreation Site Quality," Dept. Ag. and Resource Econ., University of California—Berkeley, April (1981).

25. J.A. Hausman and D.A. Wise, A conditional probit model for qualitative choice: Discrete decisions recognizing interdependence and heterogeneous preferences, Econometrica 46, 403-426 (1978).

26. H.T. Heintz, A. Hershaft, and G. Horak, "National Damages of Air and Water Pollution," Report submitted to the U.S. Environmental Protection Agency (1976).

27. J.L. Knetsch, Economics of including recreation as a purpose of Eastern water projects, J. Farm Econ. 46, 1148-1157 (1964).

28. K. Lancaster, A new approach to consumer theory, J. Polit. Econ. 74, 132-157 (1966).

29. K.E. McConnell, Congestion and willingness to pay: A study of beach use, Land Econ. 53, 185-195 (1977)

30. D. McFadden, Quantal choice analysis: A survey, Annals of Economic and Social Measurement 5, 363-390 (1976).

31. D. McFadden, Modelling the choice of residential location, in "Spatial Interaction Theory and Planning Models (A. Karlquist, L. Lundqvist, F. Snickars, and J.W. Weibull, Eds.), North-Holland Publishing Company, New York (1978).

32. L.A. Merewitz, "Estimation of Recreational Benefit at Some Selected Water Development Sites in California," Planmetrics, Inc., Berkeley (1968).

33. National Commission on Water Quality, "Staff Report," Washington, D.C. (1976).

34. National Economic Research Associates, "Estimates of the Value of Changes in Aquatic Life Resulting from Reductions in Thermal Effluent," National Economic Research Associates, New York (1974).

35. R.W. Quandt, and W.J. Baumol, The demand for abstract transport modes: Theory and measurement, J. Reg. Sci. 6, 13-26 (1966).

36. J.B. Stevens, Recreation benefits from water pollution control, Water Resour. Res. 2, 167-182 (1966).

37. D.P. Tihansky, Recreational welfare losses from water pollution, J. Envir. Quality 3, 335-342 (1974).

38. D.P. Tihansky, A survey of empirical benefit studies, in "Cost Benefit Analysis and Water Pollution Policy" (H.R. Peskin and E.P. Sesking, Eds.), 127-144 (1975).

39. A.H. Trice and S.E. Wood, Measurement of recreation benefits, Land Econ. 34, 195-207 (1958).

40. E.L. Ullman and D.J. Volk, An operational model for predicting reservoir attendance and benefits: Implications of a location approach to water recreation, in "Papers of the Michigan Academy of Arts and Letters" 47, 473-484 (1962).

41. Water Resources Council, Principles and standards for planning water and related land resources, Federal Register 38 (1973).

42. Water Resources Council, Proposed rules: Principles, standards and procedures for planning water and related land resources, Federal Register 45, (1980).

43. E.A. Wilman, The value of time in recreation benefit studies, J. Environ. Econ. Manag. 7, 272-286 (1980).