Estimating the value of water quality improvements in a recreational demand framework

Nancy E. Bockstael, University of Maryland
Catherine Kling, University of Maryland
W. Michael Hanemann, University of California - Berkeley

Available at: https://works.bepress.com/catherine_kling/7/
Estimating the Value of Water Quality Improvements in a Recreational Demand Framework

NANCY E. BOCKSTAEL

Department of Agricultural and Resource Economics, University of Maryland, College Park

W. MICHAEL HANEMANN

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

CATHERINE L. KLING

Department of Agricultural and Resource Economics, University of Maryland, College Park

With the advent of Executive Order 12291, policymakers involved in water quality regulation are increasingly interested in assessing the benefits of their programs. Several methods for valuing water quality improvements using recreational demand models have been developed by economists, most of which depend on observing recreationists visiting an array of sites with varying water quality and costs of access. In this paper, three general types of models are described: systems of demands, discrete choice models, and the hedonic travel cost approach; the latter two models are demonstrated using a common data set on water quality and swimming behavior in the Boston area. The models are contrasted and their relative usefulness in answering policy questions explored.

INTRODUCTION

In the United States the connection between water resources and recreation is a strong one. A report by Freeman [1979] to the Council of Environmental Quality estimated that over 50% of the returns from air and water quality improvements would accrue through recreational uses of the environment. When considering water quality improvements alone, the percentage was even higher. One of the earliest studies attempting to quantify such effects [Federal Water Pollution Control, 1966] estimated that recreationists would receive more than 95% of the benefits derived from water quality improvements in the Delaware estuary. These sentiments were further supported by the National Commission on Water Quality (unpublished report, 1975) which maintained that water based recreationists would be the major beneficiaries of the 1972 Federal Water Pollution Control Act.

The link between recreation and the environment is not limited to environmental quality changes. Studies evaluating the losses from destruction of natural environments or the gains from preserving lands and water bodies in their current state have frequently focused on recreational uses of the resources in question. Examples of such studies are classic articles by Burt and Brewer [1971], Cicchetti et al. [1976], and Fisher et al. [1972], to name only a few. As a consequence, environmental economists have a particular interest in the large and growing literature on recreational demand modeling.

The ultimate purpose of this paper is to discuss some issues which arise in the application of recreation demand models to the valuation of environmental quality changes such as water quality improvements. The next two sections of the paper review the types of recreational demand models which have been used for valuation; this is followed by an empirical demonstration of two contrasting methods. In the applications, a common data set on water quality and swimming behavior in the Boston area is employed and hypothetical changes in objective indices of water quality valued.

A REVIEW OF APPROACHES WITH TRIP ALLOCATION AND SITE VALUATION MOTIVATIONS

Many of the first multiple-site models developed can be used to explain the allocation of visits among alternative sites or to value the introduction or elimination of a site. These models sometimes include site characteristics as explanatory variables but do not always facilitate the valuation of characteristics.

One of the first treatments of multiple sites was in the context of zonal trip allocation models. In 1973 Cesario [1973] suggested the use of these gravity models for the specific purpose of explaining the allocation of trips from each zone to alternative sites. In these models visits between a zone and a site were explained on the basis of zonal and site characteristics and distance, with one set of parameters estimated for all combinations of zones and origins. Usually, such models have been used simply to estimate demand and predict use rates. Freund and Wilson [1974] provided one of the most careful applications of this approach in a study of recreation travel and participation in Texas.

In their paper Cesario and Knetsch [1976] extended the gravity model so that the trips equation for zone i visitors to site j included a factor reflecting "competing opportunities" provided by all other sites. This was intended to make more explicit the substitutability among sites. These authors also introduced the possibility of using gravity models for benefit measurement. Including travel cost (time and money costs) instead of distance, Cesario and Knetsch proceeded by treating the zonal visits equations as demand curves and taking areas behind these curves as measures of consumer surplus.

1 Now at the University of California, Davis.

Copyright 1987 by the American Geophysical Union.

Paper number 6W0667.
0043-1397/87/006W-0667S05.00
The use of gravity models for benefit estimation has been limited, culminating in a rather complex paper by Sutherland [1982]. Unlike his predecessors, Sutherland obtained predictions of individual's behavior rather than simply zonal aggregates. The model had four components which, while inextricably linked, were estimated independently. Each zone's demand for trips to all sites (trip production models) \( T_j \) and each site's aggregate demand from all zones (attractiveness models) \( T_j \) were estimated. Predicted values for these variables were combined with variables reflecting distances in a trip distribution (gravity) model to predict each zone's allocation of visits among all sites \( T_{ij} \). Results from this gravity model were then used to estimate a demand function where predicted trips by zone to each origin were regressed on travel cost (constructed from the distance data).

The most disturbing aspect of the gravity models is that they are statistical allocation models based on no particular arguments about economic behavior. Consequently, when a gravity model is used to "allocate trips from zones to sites," the model does not include the requisite economic behavior to allow estimation of benefits. The relationship between trips and cost must be reestimated to capture the economic behavior implicit in a demand function. It is difficult to understand how gravity models could be useful for benefit estimation purposes, if one does not believe the gravity model is a demand function or any other behavioral function derivable from utility maximization (or cost minimization).

Burt and Brewer [1971] were perhaps the first explicitly to specify multiple-site demand models. Their motivation for going beyond the single-site model was that they were interested in measuring the value of introducing a new water-based recreational site. In deducing the value of the new site, Burt and Brewer set off to estimate how patterns of demand for existing sites would change with its addition. The Burt and Brewer model was a straightforward extension of the singlesite travel cost model to a system of such demands, but the unit of observation was the household rather than the zone. They specified

\[
q_k = f(p_1, p_2, \cdots, p_N, y) \quad k = 1, \cdots, N \tag{1}
\]

where \( q_k \) is the number of trips taken to site \( k \), \( p_k \) is the travel cost to site \( k \), \( y \) is income, and \( N \) is the number of sites in the system considered. Any differences due to the quality characteristics of sites simply show up in the estimated coefficients of the different demand functions.

A similar model (with the omission of income and based on zonal data) was employed by Cicchetti et al. [1976] in their analysis of the Mineral King project in California. Once again the motivation was the valuation of a proposed new site. Similar to Burt and Brewer [1971], the authors estimated a system of demands for alternative sites or site groups as functions of prices (i.e., the costs of traveling to each site), and again, site characteristics were excluded from the model. Similar to the Burt and Brewer approach, the introduction of the new site was assessed by considering the benefits of a price change for the existing site most similar to the proposed site.

Hof and King [1982] asked the very pertinent question, Why do we need to estimate the system of demands in these cases? Why not just estimate the demand for the similar site (as a function of all prices) and evaluate the benefits in that market? In the context of the Burt and Brewer [1971] and the Cicchetti et al. [1976] papers, their arguments seem cogent. If there is only one price change, its effect can be measured in one market [Just et al., 1982]. Even if one expects seemingly unrelated regression problems, ordinary least squares will achieve the same results as generalized least squares when all equations include the same variables.

Hof and King [1982] further argued that Willig's results provide bounds on compensating variation as functions of Marshallian consumer surplus. Thus it is not necessary to estimate the entire demand system so as to impose cross-price symmetry and ensure path independence. In retrospect, this procedure of imposing symmetry (followed by both the Burt and Brewer [1971] and the Cicchetti et al. [1976] papers) seems inappropriate, since there is no reason for the Marshallian demands to exhibit such characteristics. Additionally, the path independence property is not worth worrying about, since the particular functional forms chosen for the systems of demand functions in these papers do not meet integrability conditions [LaFrance and Hanemann, 1984]. In any event, if we are interested in the effect of a single price change, there would seem no especially compelling reason to estimate an entire system of demands if they are to take the form suggested by Burt and Brewer or Cicchetti et al.

While Hof and King's [1982] comments seem justified in this context, the debate leaves unaddressed the broader issue of substitution among sites. The above model treats the new site as an exact replication of an existing site, provided at a lower cost to some individuals. Quality differences across existing sites or between existing and new sites can not be treated in this framework.

THREE APPROACHES TO MULTIPLE-SITE MODELING

Of burgeoning interest in environmental economics is the valuation of improvements in air and water quality. While valuation exercises have frequently taken place in the context of contingent valuation models, economists have concurrently tried to adapt recreational demand models to this task. This has given a new and more insistent motivation for multiplesite modeling. It was quickly realized that in order to value characteristics such as changes in water quality, one needed to estimate demand as a function of these characteristics, and this required observing variation in water quality over observations. This variation could, presumably, be found only by looking across recreational sites. It seemed intuitively plausible that one could deduce willingness to pay for improved water quality from observations on recreationists actual trade-offs between travel costs and quality.

Three distinct modeling approaches have developed to accommodate the valuation of site characteristics drawing on this implicit trade-off, and each can be used to value water quality improvements. One is a modification of the models presented in the previous section which estimate systems of demands for sites, a second employs discrete choice models of site demand. The third type of approach is conceptually quite different in that it estimates demand functions for the site characteristics directly. While each of these approaches will be discussed briefly, only the second (discrete choice) and the third (characteristics demands) will be demonstrated empirically. As we shall see, these are polar cases and highlight the different ways in which the problem can be viewed. Empirical treatment and comparison of variants of the first two approaches can be found in the works by Kling [1986] and Strand et al. [1986].

SYSTEMS OF DEMANDS

As the last section made clear, one of the difficulties with the system of demands approach is that it can not account
explicitly for quality characteristics. While the demand for each site is expected to be a function of price and quality (at both the site in question and substitute sites), no variation in any given site’s quality will be observed over individuals.

While site characteristics cannot be incorporated as separate variables in site demand equations, they can be introduced by means of varying parameters [Freeman, 1979; Vaughan and Russell, 1982; Smith et al., 1983; Smith and Desvousges, 1985]. The varying parameter model was first used in recreational modeling by Vaughan and Russell [1982] to determine the average value of a freshwater fishing day at fee fishing sites. The authors argued that a system of demand equations, where the number of visits was specified as a function of own price and income, could be estimated in the following way:

\[ x_{mi} = \beta_{0i} + \beta_{1i}p_{mi} + \beta_{2i}y_{mi} + \epsilon_{mi} \]  

for each \( i = 1, \ldots, N \)

where \( m \) denotes the individual observation, and \( i = 1, \ldots, N \) denotes the site. The \( 3 \times N \) parameter values from these demand equations could then be regressed against the two observed characteristics of each site \( (b_{1i}, b_{2i}) \):

\[ \beta_{0i} = \alpha_{00} + \alpha_{01}b_{1i} + \alpha_{02}b_{2i} + \mu_{0i} \]  

\[ \beta_{1i} = \alpha_{10} + \alpha_{11}b_{1i} + \alpha_{12}b_{2i} + \mu_{1i} \]  

\[ \beta_{2i} = \alpha_{20} + \alpha_{21}b_{1i} + \alpha_{22}b_{2i} + \mu_{2i} \]

Smith et al. [1983], in a study of recreational benefits from improved water quality, provided a theoretical basis for the varying parameters model based on a household production framework. They estimated the two steps separately using ordinary least squares estimates and in a later paper [Smith and Desvousges, 1985] proposed an alternative model for the first stage which adjusted for the truncation bias. Ordinary consumer surplus measures were derived for changes in quality by determining the effect of a quality change on the predicted coefficients in the system. In a more recent application, McConnell et al. [1984] used a similar model to value water quality changes to Chesapeake Bay boaters.

**Allocation Models**

Rather than estimate conventional demand functions for each recreational site, the second approach models the decision process by which the total number of recreational trips are allocated among alternative sites. One way of doing this is to estimate share models which explain the proportion of total trips taken to each alternative. Several techniques for statistically estimating shares which are consistent with demand functions have been proffered by economists and applied to a variety of economic problems (see, for example, Woodland [1979] and Hanemann’s cataloguing in the work by Bockstael et al. [1986]).

A related approach, which can be found in several applications [Caulkins, 1982; Hanemann, 1978; Feenberg and Mills, 1980; Bockstael et al., 1984; Rowe et al., 1985], retains a similar model but interprets the probabilities not as shares per se, but as choice probabilities arising from some structural economic model. The multinomial probability is used to reflect the probability that alternative \( j \) is chosen on a given choice occasion; the “count” equals 1 if \( j \) is chosen, 0 otherwise, and the likelihood function takes the form

\[ L = \prod_{m=1}^{M} \prod_{g=1}^{G} \prod_{j=1}^{N} \pi_{ym}^n \]  

where \( m \) denotes the individual, \( j \) denotes the alternative, and \( g \) denotes the individual’s choice occasion. The probabilities \( \pi_{ym} \) are themselves functions of the costs and characteristics of the alternatives and the characteristics of the individuals and are derived from a utility maximizing framework.

The consumer’s utility maximizing choice is expressed in terms of the conditional indirect utility function \( v_j(b_j, y, p_j; \epsilon_j) \), where \( v_j \) is the indirect utility function conditioned on the choice of site \( j \), \( y \) is the income available per choice occasion, and \( \epsilon_j \) is a random variable. Then the probability of choosing site \( j \) (with the individual and choice occasion subscripts suppressed) is

\[ \pi_j = Pr(v_j(b_j, y, p_j; \epsilon_j) \geq v_i(b_i, y, p_i; \epsilon_i) \) for all \( k \)

A common assumption is that the random variables \( \epsilon_j, \ldots, \epsilon_k \) are independently and identically distributed extreme value variates and that they are additive in the indirect utility function. This yields the logit model of discrete choices

\[ \pi_j = e^{\epsilon_j}/\sum_{k=1}^{N} e^{\epsilon_k} \quad j = 1, \ldots, N \]

In the work by Bockstael et al. [1984] the generalized extreme value (GEV) distribution [McFadden, 1978] is employed which, together with the assumption of additive errors, yields discrete choice probabilities of the form

\[ \pi_j = e^{\epsilon_j}G(e^{\epsilon_0}, \ldots, e^{\epsilon_N})/G(e^{\epsilon_0}, \ldots, e^{\epsilon_N}) \]

where \( G \) is a positive linear homogeneous function, and \( G( \cdot ) \) is the partial derivative of \( G( \cdot ) \) with respect to its \( j \)-th argument. In either case, the formulas for the choice probabilities may be substituted into the multinomial density (4) for maximum likelihood estimation of the parameters in the \( v( \cdot ) \) functions.

While all of the allocation models (both the share models and the discrete choice models) handle the problem differently, they all have one characteristic in common: the total number of trips taken in a season is not directly determined by the model. One way of redressing this deficiency is to append a separate participation decision. This approach should explain the number of trips (choice occasions) in the season including the possibility of zero trips. A “macro” decision of how many days in the season to recreate was estimated by Bockstael et al. [1984], using a limited dependent variable model which takes account of the fact that decisions will be nonnegative but may be zero for a number of people. The decision is estimated as a function of the characteristics of the individual and the characteristics of the recreational opportunities available as captured through an inclusive value index constructed from the results of the discrete choice (microallocation) model. An example is provided in a subsequent section of this paper.

A comparison of this approach with the Feenberg and Mills [1980] and Caulkins [1982] models exposes an important difference. In the above model the probability that an individual is not a recreationalist, i.e., he does not participate at all in the recreational activity, is estimated directly by either Tobit or
Heckman procedures [see Mandala, 1984]. The latter procedure is particularly appropriate if factors such as old age, ill health, or preferences for other activities cause an individual never to recreate. In the other approaches, total visits are determined by the summation of independent decisions on sequential choice occasions. Here, nonparticipants happen, in a sense, by accident. They are predicted to be those individuals who happen to have a string of zero predicted responses to a sequence of \( N \) independent microdecisions. Modeling the macroallocation decision separately seems to be a more realistic and useful description of individual behavior. However, it does not offer a consistent way to link independent site choice decisions and the demand for total trips with a common underlying utility maximization framework.

**Hedonic Travel Cost**

The final approach to be outlined here has as its sole focus the valuation of site characteristics. The hedonic travel cost model [Brown and Mendelsohn, 1984; Mendelsohn, 1984] attempts to reveal shadow values for characteristics by estimating individuals’ demands for the characteristics. The approach uses information on the extra costs of accessing a site with higher quality characteristics to estimate the demand for quality.

Since chance and not markets provides the array of sites and their qualities, the costs of accessing all possible sites for all individuals will not be an increasing function of characteristics. However, it is a logical result of constrained utility maximization that an individual will only incur greater costs to visit a more distant site if the benefits derived from the visit exceed those from a closer site. Thus observations on costs and site characteristics are included only for those sites which are actually visited by individuals in the regression subsample. The approach depends on costs being a single-valued, increasing function of each element of a vector of site characteristics.

The hedonic travel cost approach consist of two separate procedures. The first step entails regressing individuals’ total costs \( C_{mj} \) of visiting a site \( (j) \) on the characteristics of the site \( b_j \):

\[
C_{mj} = f(b_j) \quad \text{for all} \quad i = 1, \cdots, N \quad (8)
\]

Separate regressions are run for individuals \((m = 1, \cdots, M)\) from each origin \( i \), where the costs of visiting any given site and the characteristics of the site are identical for all individuals visiting the site from the same origin and variation in the data comes from variation in the sites visited by those individuals from the same origin. The partial derivative of cost with respect to a characteristic \( \partial C/\partial b \) is then interpreted as each origin’s hedonic price for that characteristic. These hedonic prices are used as prices in a second stage where demand functions for characteristics are estimated:

\[
\partial C/\partial b = g(b) \quad (9)
\]

The demand function for a characteristic is assumed to reflect the marginal willingness to pay per recreation day for an increase in the quality of the characteristic.

**Some Perspective**

Each of the above modeling approaches depends on observed recreational behavior to extract welfare measures of environmental quality changes. However, the three types of models differ in the way they characterize the nature of the recreational decision. The conventional demand function approach (e.g., varying parameters model) treats each site as a different good with a different demand function, where demand is a function of the quality at the site. This approach implicitly assumes individuals choose interior (nonzero) solutions for each good. Perhaps the greatest drawback of this approach is that in this framework, it is difficult to accommodate the effects of substitute sites.

The discrete choice models take a slightly different perspective. The emphasis here is, in a sense, on the substitutability (or trade-offs) among quality differentiated goods. The decision process modeled is the choice among this finite set of quality differentiated sites. Because of the construction of this model, it is more amenable to the corner solution phenomenon. Since probabilities or shares of total trips are predicted, the model is consistent with the observable phenomenon that recreationists usually visit more than one site but fewer than all sites. While suited to explaining allocations of trips across substitute sites, this class of models is less amenable to the estimation and prediction of total trips taken in a season. The final discrete choice model suggested above attempts to mitigate this problem by appending a macro decision model, but does so in a way which is not completely consistent with a utility maximization framework.

The hedonic travel cost model is based on a completely different view of the problem from the discrete choice-share models. Where the latter view the recreationist’s decision problem as a choice among a finite number of discrete quality differentiated sites, hedonic travel cost presumes that individuals can choose along a continuum of quality. That is, an array of sites exist where increasing quality can be purchased at higher travel costs and the individual can freely choose where to be along that array. Thus one can skip the step of estimating trips demand and move directly to characteristics demands. This approach does not predict behavior. There is no place in the model for the total number of trips taken by an individual. In fact, this dimension does not figure into welfare measurement.

In the next section two models are applied to an actual data set. One is the discrete choice model which takes the recreationists decision as a discrete choice among quality differentiated sites. The other represents the opposing view of the recreationist’s decision as a choice of quality along a continuous array. In each case, the models are estimated followed by a demonstration of how the results are used for valuing water quality improvements. The final section of the paper discusses the nature of benefit measurement in the context of the models.

**Some Estimation Examples**

In this section we demonstrate the application of two methodological approaches to valuing water quality improvements. The two methods, the discrete choice model and the hedonic travel cost, were selected because of their fundamental conceptual differences (for an empirical comparison of other methods, see Kring [1986]). The comparison includes a demonstration of how each approach produces estimates, the data requirements, necessary estimation techniques, and the practical problems, which arise in the estimation process.

The application employs a data set collected by The Environmental Protection Agency in 1975 which includes information on recreational swimming at Boston area beaches. The data set contains information on both participants and nonparticipants, as it is based on random household interviews in the Boston Standard Metropolitan Statistical Area. For each
participant, a complete season's beach use pattern is reported, including the number of trips to each beach in the Boston area. There are objective measures of water quality for 30 beach sites. It should be noted that participants in this data set, like other data sets of this sort, tend to visit more than one site but far less than all sites available.

**Discrete Choice Model**

The multiple-site recreational demand model estimated by Bockstael et al. [1984] has two components; one is the macro-decision: does an individual participate in the activity of interest (swimming at beaches in the Boston-Cape Cod area) and, if so, how many trips does he take in a season? The other component is a site-allocation decision: on each choice occasion, which site does he visit?

In the microallocation model the indirect utility associated with choosing the ith site on any choice occasion is some function of \( z_i \), a vector of attributes of the ith alternative, so that

\[ v_i^* = v_i(z_i) + \epsilon_i. \]

The random component is additive and attributed to the unmeasurable variation in tastes and omitted variables. If the \( \epsilon \)'s are independently and identically distributed with type I extreme value distribution (Weibull), then we have a multinomial logit model. However, the multinomial logit (MNL) implicitly assumes independence of irrelevant alternatives; i.e., the relative odds of choosing any pair of alternatives remains constant no matter what happens in the remainder of the choice set. Thus this model allows for no specific pattern of correlation among the errors associated with the alternatives; it denies, and, in fact, is violated by, any particular similarities within groups of alternatives.

A more general nested logit model [McFadden, 1978] specifically incorporating varying correlations among the errors associated with the alternatives can also be derived from a stochastic utility maximization framework. If the \( \epsilon \)'s have a generalized extreme value distribution then a pattern of correlation among the choices can be allowed. The probabilistic choice model is given by

\[ P_i = \frac{e^{G_i(e^1, \ldots, e^n)}}{G_i(e^1, \ldots, e^n)} \]  

where \( G_i \) is the partial derivative of \( G \) with respect to the ith argument and \( G(Y) \) is

\[ G(Y) = \sum_{m=1}^{M} \beta_m \left( \sum_{i \in S_m} e^{\alpha(Y - a_m)} \right)^{1-\sigma_m} \]

where there are \( M \) subsets of the \( N \) alternatives, \( a_m \) is a parameter which could vary over subsets, and \( 0 \leq \sigma_m < 1 \). This form allows a general pattern of dependence among the alternatives, where the parameters \( \sigma_m \) can be interpreted as indices of the similarity within groups.

The Boston data is particularly amenable to GEV estimation. Among the 30 sites, 8 are beaches at freshwater lakes and 22 are saltwater beaches. It would seem reasonable to suppose that the odds of choosing fresh water site A over salt water site B will be disrupted by the addition of another freshwater lake site. Stated another way, freshwater and saltwater sites are probably viewed as closer substitutes within groups than across groups; while this distinction may seem arbitrary, it is difficult a priori to determine all possible appropriate groupings except on the basis of common sense. However, as we shall see, a posteriori one can test whether the hypothesized groupings are justified.

The GEV model allows us to view individuals (1) as choosing between fresh and saltwater and (2) as choosing among freshwater sites conditioned on the freshwater choice and choosing among saltwater sites conditioned on the saltwater choice. In actuality, the problem is set up so that the individual chooses the "best" within the group of saltwater sites and the best within the group of freshwater sites and then chooses between these two best alternatives on each choice occasion.

To make the estimation process explicit, let us consider the following form of \( v_{im} \):

\[ v_{im} = \theta Z_{im} + \psi W_m \]  

where the \( Z \)'s denote attributes associated with all sites and the \( W \)'s are associated solely with the saltwater-freshwater choice, \( i \) indexes the site, and \( m \) indexes the salt or freshwater alternative. Also let us assume that \( \sigma_m \) is identical across all groups and equal to \( \sigma \). Define the "inclusive value" of group \( m \) as

\[ I_m = \ln \left( \sum_{k \in S_m} e^{\theta Z_{km} + (1-\sigma)} \right) \]

Now, the probability of choosing site \( i \) conditioned on the saltwater/freshwater choice is

\[ P_{im} = \frac{e^{\theta Z_{im} + (1-\sigma)I_m}}{\sum_{k \in S_m} e^{\theta Z_{km} + (1-\sigma)I_m}} \]

and the probability of making the saltwater (or freshwater) choice is

\[ P_m = \frac{e^{\theta W_m + (1-\sigma)I_m}}{\sum_{j=1}^{M} e^{\theta W_j + (1-\sigma)I_j}} \]

These probabilities can be estimated using MNL procedures. First, the \( P_{im} \) are estimated with \( M \) independent applications of the multinomial logit (where \( M = 2 \), one for saltwater beaches and two for freshwater beaches). Note that at this stage \( \theta \) is not recoverable but can be estimated only up to a scale factor of \( 1 - \sigma \). From the results of (14), the inclusive prices (equation (13)) are calculated and incorporated as variables in the second level of estimation (15) where the \( \psi \)'s and the \( \sigma \) are estimated.

In choosing among sites, the determinants of most interest are the site characteristics which vary over alternatives and the costs of gaining access to sites. The quality variables chosen for this model include environmental indicators such as oil, turbidity, fecal coliform, chemical oxygen demand, and temperature. Three other variables reflecting noisy and congested sites, ethnic priorities, and public transportation access are identified as potentially valuable in the site choice model.

Because of the nature of the logit model, variables which are present in the indirect utility function but do not change across alternatives cancel out upon estimation, that is, their coefficients cannot be recovered; income has this property. However, we know from utility theory that income and price must enter the indirect utility function in the form \( Y - p \). Thus the coefficient on price will be income's implicit coefficient as well, an important fact which will be drawn upon in calculating benefits.

Estimation of the second stage of the model requires the calculation of inclusive values, defined in (13), which capture the information about each group of sites in stage 1. Thus if
water quality were to change at some sites, the inclusive values would change. Additionally, other variables may enter at this stage, variables which affect the salt-freshwater decision but do not vary over alternatives within each group, including the size of the household, the proportion of children, and whether or not the household has access to a swimming pool.

Table 1 presents the estimated coefficients and test statistics for the first stage of the GEV model, and Table 2 presents the second stage results. In the first stage the estimated coefficients on quality characteristics all are significant at the 5% significance level and of the expected sign (with the possible exception of temperature and turbidity in the freshwater equation). Additionally, individuals (ceteris paribus) visit closer beaches, avoid noisy areas, and are discouraged by beaches heavily populated by ethnic groups different from their own. Individuals who do not own cars are less likely to visit beaches not serviced by public transportation.

From the first stage results the "inclusive" values are calculated and introduced in stage 2. A value of 0.854 is estimated for $1 - \sigma$ implying a $\sigma$ of 0.146, which is significantly different from both 0 and 1. This suggests that fresh and saltwater sites are considered significantly different, but all freshwater sites are not viewed as perfect substitutes for one another and neither are all saltwater sites. Thus we can expect that there are gains from using the GEV specification and that partitioning of the alternatives into these subgroups was justified. Additionally, the results suggest that larger families tend to go to lakes but families with a larger portion of children tend to go to saltwater beaches. Those who have access to a swimming pool are more likely to visit saltwater beaches. The constant term suggests that ceteris paribus, people prefer saltwater beaches.

The second part of the model is a single-activity model of swimming participation. We use the Tobit model which presumes that individual's decisions can be described as

$$x_i = h(z_i) + \delta_i \quad \text{if} \quad h(z_i) + \delta_i > 0 \quad (16a)$$

$$x_i = 0 \quad \text{if} \quad h(z_i) + \delta_i \leq 0 \quad (16b)$$

and that the decision of whether or not to participate and how much to participate are dictated by the same forces. Income, size, and composition of household and ownership of water sports equipment are included as explanatory variables.

Additionally, we would wish to include variables reflecting the cost and quality of the swimming activities available. Since it is appropriate to capture the quality and costs of the best alternative for each individual, not necessarily the characteristics of the closest site or the average characteristics over sites, the inclusive value concept is appealing. It represents the value of different alternatives weighted by their probabilities of being chosen. Defining an inclusive value from both stages of the GEV estimation gives us

$$I_p = \ln \left( \sum_{j \in J} e^{v_j} + \sum_{j \in J} e^{v_j} \right) \quad (17)$$

where $J_s$ is the set of saltwater sites, $J_p$ is the set of freshwater sites, and $v_j = \theta Z_j + \psi W_j$, where the $Z$'s are explanatory variables in the first stage and the $W$'s are explanatory variables in the second stage.

The results of the macro decision are presented in Table 3. The estimated coefficients for all other variables except income are statistically significant from zero at the 2.5% level (for one-tail test). The "inclusive value" variable, included to reflect the value of recreational opportunities, is significant and positive as expected.

The estimated model allows us to capture three types of changes in beach use. The discrete-continuous macroallocation model (estimated as a Tobit) permits the prediction of two aspects of the beach use decision: whether or not to participate and, if so, how many trips to take. Both aspects of the decision are functions of water quality variables as re-

<table>
<thead>
<tr>
<th>Table 2. Second-Stage GEV Model Estimates of Choice Between Saltwater and Freshwater Beaches, Boston and Cape Cod, 1975</th>
</tr>
</thead>
<tbody>
<tr>
<td>Estimated coefficient</td>
</tr>
<tr>
<td>------------------------</td>
</tr>
<tr>
<td>Estimated coefficient</td>
</tr>
<tr>
<td>$t$ Ratio</td>
</tr>
<tr>
<td>Likelihood</td>
</tr>
<tr>
<td>Chi-squared with 5</td>
</tr>
<tr>
<td>degrees of freedom</td>
</tr>
</tbody>
</table>

*This $t$ ratio tests significant difference from zero. A more appropriate test is significant difference from 1; the relevant $t$ ratio is - 4.044.
flected in the inclusive value term. The GEV model permits the prediction of the allocation of trips among sites as functions of costs, etc. The predicted probabilities which the model produces can be interpreted as shares of the household's total trips.

The ultimate purpose of the modeling effort however is to estimate the benefits associated with improvements in water quality. Given the GEV model, the expected value of the indirect utility function for a choice occasion can be shown to equal

\[ V(p^0, b^0, y) = \ln G(e^{\alpha_1 y}, \cdots, e^{\alpha_n y}) + k \]  

where \( k \) is a constant. A compensating variation measure of a change from \( b^0 \) to \( b^1 \) can be defined as \( C \) in the following equation [see Hanemann 1982, 1984]

\[ V(p^0, b^0, y) = V(p^0, b^1, y - C) \]

or

\[ \ln G(e^{\alpha_1(y - C, z^1, w^1)}, \cdots, e^{\alpha_n(y - C, z^1, w^1)}) = \ln G(e^{\alpha_1 y}, \cdots, e^{\alpha_n y}) + k \]  

There is no closed-form solution for compensating variation in this case, but we can approximate it by

\[ CV \approx \left[ \sum_{m=1}^{2} e^{\alpha_m W^m_0 + (1 - \sigma) I_m 0} - \sum_{m=1}^{2} e^{\alpha_m W^m_1 + (1 - \sigma) I_m 1} \right] \cdot \left[ \sum_{m=1}^{2} \gamma_1 e^{\alpha_m W^m_0 + (1 - \sigma) I_m 0} + (1 - \sigma) I_m 1 \right]^{-1} \]  

where \( m = 1, 2 \) denotes the saltwater and freshwater alternatives \((W^0_1, I^0_1)\), and \((W^1_1, I^1_1)\) represent values of variables before and after the water quality change, respectively, and \( \gamma_1 \) and \( \gamma_2 \) are the implicit income coefficients in the saltwater and freshwater models.

The calculation of \( CV \) according to (20) yields an estimate of the expected compensating variation per choice occasion for the household. To obtain annual or seasonal benefit estimates this number must be multiplied by the predicted number of trips the individual takes. One should note that even if the individual takes no more trips in response to the quality change (either because he is constrained or because a more substantial quality change is necessary to increment the number of trips), the benefits of improvements are still measureable.

**TABLE 3.** Estimates of Tobit Model of Boston Swimming Participation and Intensity

<table>
<thead>
<tr>
<th>Variable</th>
<th>Tobit Estimates</th>
<th>Initial Value (OLS Estimates)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>26.01</td>
<td>35.98</td>
</tr>
<tr>
<td>(2.57)</td>
<td>(4.59)</td>
<td></td>
</tr>
<tr>
<td>Inclusive value</td>
<td>0.897</td>
<td>1.02</td>
</tr>
<tr>
<td>(1.86)</td>
<td>(2.74)</td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>-1.19</td>
<td>-0.07</td>
</tr>
<tr>
<td>(-0.56)</td>
<td>(1.79)</td>
<td></td>
</tr>
<tr>
<td>Size of household</td>
<td>-24.10</td>
<td>-8.1</td>
</tr>
<tr>
<td>(-2.76)</td>
<td>(-2.08)</td>
<td></td>
</tr>
<tr>
<td>Percent children</td>
<td>-6.18</td>
<td>-14.71</td>
</tr>
<tr>
<td>(-1.23)</td>
<td>(-2.03)</td>
<td></td>
</tr>
<tr>
<td>Water sports equipment</td>
<td>13.05</td>
<td>6.42</td>
</tr>
<tr>
<td>(3.44)</td>
<td>(2.05)</td>
<td></td>
</tr>
</tbody>
</table>

Chi-squared statistic 262; t ratios are in parentheses.

**TABLE 4.** Average Compensating Variation Estimates of Specific Reductions in Pollutants at Boston Area Beaches

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>10% Reduction at All Sites</th>
<th>30% Reduction at All Sites</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Per Choice Occasion</td>
<td>Per Season</td>
</tr>
<tr>
<td>Oil</td>
<td>$0.05</td>
<td>$0.96</td>
</tr>
<tr>
<td>COD</td>
<td>$0.12</td>
<td>$2.65</td>
</tr>
<tr>
<td>Fecal coliform</td>
<td>$0.02</td>
<td>$0.19</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pollutant</th>
<th>30% Reduction at All Sites</th>
<th>30% Reduction at Downtown Boston Beaches</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Per Choice Occasion</td>
<td>Per Season</td>
</tr>
<tr>
<td>Oil, turbidity, COD, and fecal coliform</td>
<td>$0.50</td>
<td>$12.04</td>
</tr>
</tbody>
</table>

Estimates are given in 1974 dollars.

In Table 4 the estimated benefits (in 1974 dollars) of a series of hypothetical water quality changes are reported. The hypothetical water quality changes introduced include a 10 and a 30% reduction in each of the following water quality parameters individually: oil, chemical oxygen demand, (COD) and fecal coliform; these reductions were introduced uniformly across all sites. Also reported in Table 4 are the results of a 30% reduction at all sites in oil, turbidity, COD, and fecal coliform simultaneously. This figure can be compared to the same sort of pollutant reductions if they affect only beaches in Boston harbor. Reductions in pollutants at downtown Boston beaches (8 of the 30 sites) generate more than half the benefits reported when all sites are uniformly improved. These examples are offered to demonstrate the sorts of questions which can be answered with this model.

**Hedonic Travel Cost Model**

The above application of the discrete choice multiple-site demand model is a relatively detailed one, the basic model having been expanded and developed to fit the particular problem and information at hand. As such, it seems somewhat unfair to apply any other estimation approach to this same data set unless equal care can be given in the model development. Additionally, the application of two or more methods to the same data set suggests a comparison which, under the circumstances, is unwarranted. A comparison of benefit estimates across models is ultimately meaningless, since there is no way of proving one set of answers superior (i.e., closer to the truth) than another.

Bearing this in mind, we present in this section an application of the hedonic travel cost model with the belief that the nature of an estimation approach can best be understood through an empirical application. The story which underlies the hedonic travel cost is substantially different from that of either the discrete choice-share models or the demand systems models. This suggests that different methodological approaches may answer different questions or be better suited to different types of valuation problems. In any event, an application may help to underscore the types of problems which are best and least suited to this estimation approach.

The goal of the hedonic travel cost model is to value characteristics (such as water quality dimensions) directly from
(recreationally related) demands for characteristics, rather than through the demand for recreational trips to sites. The premise is that the extra costs necessary to travel to a better site reflect the value of the superior quality of that site.

The model was estimated for the subset of saltwater sites. Including the freshwater sites, which were substantially different for reasons other than water quality, would seem to confuse the issue. In fact, a preliminary estimation including both freshwater and saltwater sites produced poor results. The two most important environmental quality indices (oil and COD) were included in the model. Adding more quality indices introduced serious multicollinearity and prevented the estimation of the hedonic prices. The site characteristics were indexed here such that increasing values imply improving water quality to facilitate interpretation.

The first stage of the hedonic travel cost model requires the regression of each origin with the necessary variation in dependent and independent variables coming from the variation across (destination) sites. The data set contained information on households from 93 different origin zones. Unfortunately, there was so little variation in site choice within these zones that regressions were infeasible. As a result, the origins were aggregated into larger zones (25 of them) to obtain sufficient variation. This aggregation may have violated the premise of the model which requires regressions being run on homogeneous groups with identical travel costs to any given site. It is likely that expanding origin zones produced more variation in site choice because there was more heterogeneity in travel cost.

It should be noted that about three fourths of the participants sampled visited more than one site. This problem was handled in a manner similar to that used in the discrete-continuous choice model. Different site choices by the same individual were included in the regressions as additional observations (in effect as though they were site choices by different individuals). This provided the added benefit of more variation in sites visited by individuals from the same origin.

Given the linear functional form of the Brown and Mendelsohn [1984] application, the hedonic prices of oil and COD are the estimated coefficients from the first stage regression of the following form:

\[
C_{ij} = \beta_{0j} + \beta_{1j}O_{i} + \beta_{2j}D_{i} + \epsilon_{ij} \tag{21}
\]

where \(C_{ij}\) is travel costs from origin \(j\) to site \(i\), \(O_{i}\) is the transformed index of oil at site \(i\), and \(D_{i}\) is the transformed index of COD at site \(i\). (The indices are transformed so that increasing values represent water quality improvements.)

Once the first stage has been estimated, marginal value functions for quality characteristics are then estimated by regressing the derived hedonic prices for individuals from each origin to each site on the level of the quality characteristics at the relevant sites (together with other individual related variables). In this application, these variables included income and the ethnic dummy variable which had turned out to be important in the discrete choice model. We also included an instrumental variable for the number of trips the individual took, since this variable was included in the Brown and Mendelsohn [1984] application. As in that paper, trips were initially regressed on the other individual-specific variables (ethnic dummy and income) as well as dummy variables for origins. Then the predicted values were included in the following marginal value functions for each characteristic:

\[
PO_{i} = x_{0} + x_{1}O_{i} + x_{2}D_{i} + x_{3}Y_{i} + x_{4}E_{i} + x_{5}X_{i} + u_{i} \tag{22}
\]

\[
PD_{i} = y_{0} + y_{1}O_{i} + y_{2}D_{i} + y_{3}Y_{i} + y_{4}E_{i} + y_{5}X_{i} + w_{i} \tag{23}
\]

where \(PO_{i}\) and \(PD_{i}\) are the derived prices of improvements in oil and COD levels (\(\beta_{1j}\) and \(\beta_{2j}\) from equation (22)); \(Y_{i}\) is income; \(X_{i}\) is predicted visits; and \(E_{i}\) is the ethnic dummy.

The results of the first stage produced 50 "hedonic prices" (25 coefficients for each quality index). Unfortunately, only seven of the 50 were positive and significantly different from zero. In contrast, 23 of the 50 were negative and significantly different from zero. Since not all hedonic prices from the first stage were positive, it is not clear as to whether observations on all prices should be included in the final stage demand function. The results of two separate approaches are reported in Appendix A. In the appendix, oil and COD denote indices which increase with declining levels of these pollutants and as a consequence are indices of "goods," not "bads"; \(t\) statistics are in parentheses. The first set of characteristics demand functions includes only those observations for which there are positive prices; the second set includes all observations. The functions estimated on reasonable prices (the positive ones) did not produce negative coefficients on own prices (that is, they failed to generate downward sloping demand functions). In both cases, both price coefficients were significantly different from zero and positive. Only when negative prices were included was a negative slope estimated, and then only for COD.

Given these problems, the alternative estimation procedures presented in Mendelsohn [1984] were applied. Here the first-stage regressions (i.e., the hedonic price equations) were non-linear Box-Cox transformations, where the following was estimated:

\[
\frac{C_{ij}^{x_{j}} - 1}{\lambda_{j}} = \beta_{0j} + \beta_{1j} \left( O_{i}^{x_{j}} - 1 \right) \frac{\lambda_{j}}{\lambda_{j}} + \beta_{2j} \left( D_{i}^{x_{j}} - 1 \right) + e_{ij} \tag{24}
\]

for all \(j\)

which allowed some flexibility in form as well as a hedonic price which was a function of characteristic levels. Characteristics' prices can not be determined directly from these results, but must be constructed as derivatives of (24). There are now 25 price gradients for each characteristic, each of which can be evaluated at the observed levels of \(O_{i}\) and \(D_{i}\). Better results were obtained with this approach than the linear specification, since 11 of the 25 price gradients for COD produced positive prices and 16 of the 25 price gradients for oil produced positive prices.

The next step of this procedure required estimating instrumental variables for characteristic prices (in addition to visits) before including these prices in the characteristics demand functions. Following Mendelsohn [1984], the constructed prices were regressed on income, the ethnic dummy, and site dummies producing predicted prices; this procedure did not appreciably increase the number of positive prices, however.

The final step of the procedure involved the estimation of characteristics demand functions with quantity on the left-hand side as opposed to price. The forms of these functions were

\[
O_{i} = \alpha_{0} + \alpha_{1}P_{0i} + \alpha_{2}PD_{i} + \alpha_{3}Y_{i} + \alpha_{4}E_{i} + \alpha_{5}X_{i} + u_{i} \tag{25a}
\]
\[ D_i = \gamma_0 + \gamma_1 P_i + \gamma_2 P_i^2 + \gamma_3 Y_i + \gamma_4 E_i + \gamma_5 X_i + w_i \]  

Again one set of regressions was estimated from observations which had only positive predicted prices and a second set

Once again, those regressions which included only positive prices were unsatisfactory. The own price coefficient was positive
and significant for oil and insignificant (although negative) for COD. When both positive and negative prices were included,
however, both the oil and COD regressions behaved, respectably. For example, in the oil equation, the demand for
cleaner water (less oil) decreased with the “price” of cleaner water (in terms of oil), increased with the price of cleaner
water (in terms of COD), increased with income, and decreased with total number of trips taken. In the COD equation,
death for less COD decreased with the price of less COD and increased with the price of less oil. However, the
signs on income and total (predicted) visits were reversed from the previous results.

The characteristics demand functions can presumably be used to value water quality characteristics. In the present ex-
ample the only viable characteristics demand is the COD equation from the second procedure, because this was the only
demand function for which a negative slope was obtained from using only positive hedonic prices. Using this demand
function, we could calculate the value of a characteristic at the margin. However, we are normally interested in discrete
changes, such as those which would come about as the result of an environmental policy change. Brown and Mendelsohn
[1984] and Mendelsohn [1984] do not use their models in this way and there is some question as to whether the character-
istics demands can appropriately be used to value discrete changes. If we were to take the area behind the characteristics
demand for a 10% change in COD and call it the consumer surplus associated with this change as some practitioners have
done we would obtain a consumer surplus of $450/visit.

Concluding Comments About Welfare Measurement

The types of benefit measures which can be obtained from models of recreational behavior depend very much on the way
in which the recreationists decision problem is perceived. The discrete choice model described above models each step of a
recreationists choice given that he can choose whether to par-
ticipate in the recreational activity (in this case swimming), and choose how many recreational trips to take and how to
allocate those trips among quality-differentiated sites. The model can be used to predict how his decisions at all of these
levels could change as a result of a policy which would change (water) quality at any one site or all sites. Finally, the model
can be used to value any of these changes.

In contrast, the hedonic travel cost model treats quality as a decision variable, where quality is purchased at higher costs.
Presumably one can deduce, given the current configuration of sites, what the marginal value of a quality characteristic is.
However, since quality in this model is a choice variable which can be bought at a higher travel cost, a question such as
“what is a public action worth which improves water quality?” makes little sense.

Additionally, the demand functions are associated with characteristics and not sites, and thus it seems particularly
difficult to assess the value of a site specific change in quality (such as might be brought about by a regulation, etc.). An
environmental policy change which altered water quality at one or all sites would alter the cost functions for each origin
and thus would have the effect of changing the hedonic prices, but it is not clear how the model, once estimated, could be
used to predict new hedonic prices after the change.

A second problem with the hedonic travel cost approach is that these functions do not capture any information about
how individuals’ behavior would change with a change in quality. How do we account for the fact that a change in a
quality characteristic will alter behavior, affecting the demands for characteristics and the use rates of the sites? How
does this approach capture the benefits accruing to new participants who might be attracted into the recreational activity
by improved environmental quality?

In contrast, the discrete choice model, with all its shortcomings and cumbersome stages, tells a story, and there is
solace in stories. A model based on a comprehensive and meaningful story about decisionmaking allows us to ask relevant
welfare questions and lends credence to the benefit estimates obtained from them.

APPENDIX A: DEMAND FOR CHARACTERISTICS USING THE
HEDONIC TRAVEL COST APPROACH (LINEAR HEDONIC
EQUATION, INVERSE DEMAND FUNCTION)

Regressions Include Only Positive Prices

| Price oil = -0.06 + 0.0006 oil + 0.007 COD - 0.04 ethnic | (1.18) | (1.18) | (1.18) |
| -2.3 x 10^{-9} inc - 0.006 visits | (0.323) | (0.323) |

Price COD = 0.017 + 0.00007 COD + 0.003 oil

| (4.2) | (4.2) |
| -0.0006 ethnic + 4.7 x 10^{-9} inc - 0.0004 visits | (4.2) | (4.2) |

Regressions Include All Prices

| Price oil = 0.06 + 0.0015 oil + 0.005 COD + 0.024 ethnic | (1.77) | (1.77) | (1.77) |
| -2.19 x 10^{-8} inc - 0.0035 visits | (3.05) | (3.05) |

Price COD = -1624 - 89.4 COD + 485.1 oil

| (0.97) | (0.97) |
| -731.7 ethnic + 0.002 inc - 375.5 visits | (11.75) | (11.75) |

APPENDIX B: DEMAND FOR CHARACTERISTICS USING THE
HEDONIC TRAVEL COST APPROACH (NONLINEAR HEDONIC
EQUATION, QUANTITY-DEPENDENT DEMANDS)

Regressions Include Only Positive Prices

| Oil = 36.30 + 7.93 pricé oil + 3.89 pricé COD | (43.94) | (43.94) |
| + 1.6 x 10^{-6} inc + 1.3 ethnic + 0.08 visits | (5.26) | (5.26) |

|
COD = 64.39 - 0.98 price COD + 6.03 price oil
   (19.27)(-0.12) (1.56)
   - 4.8 x 10^-6 inc + 3.4 ethnic - 2.06 viisits
   (-3.91) (1.32) (-3.98)

Regressions Include All Prices
Oil  = 44.54 - 1.17 price oil + 8.79 price COD
   (110.9) (-2.57) (6.68)
   + 1.9 x 10^-6 inc - 1.17 ethnic - 0.11 viisits
   (9.2) (-5.19) (-17.34)
COD  = 24.05 -17.06 price COD + 4.33 price oil
   (15.11) (-3.27) (2.40)
   - 4.1 x 10^6 inc + 34.25 ethnic + 0.37 viisits
   (-5.01) (25.96) (14.62)

Acknowledgments. This research was funded through EPA Coop­erative Agreement CR-811043-01-0 and the Maryland Agricultural Experiment Station.

REFERENCES
Federal Water Pollution Control, Delaware Estuary comprehensive study: Preliminary report and findings, technical report, Dep. of Inter, Philadelphia, Pa., 1966.

N. E. Bockstael, Department of Agricultural and Resource Economics, University of Maryland, College Park, MD 20742.
W. M. Hanemann, Department of Agricultural and Resource Economics, University of California, Berkeley, CA 94720.
C. L. Kling, Department of Agricultural Economics, University of California, Davis, CA 95616.