The Travel Cost Model

Article · January 2003

DOI: 10.1007/978-94-007-0826-6_9 · Source: OAI

1 author:

George R. Parsons
University of Delaware

.

Chapter 9

THE TRAVEL COST MODEL

George R. Parsons
University of Delaware

1. INTRODUCTION

The travel cost model is used to value recreational uses of the environment. For example, it may be used to value the recreation loss associated with a beach closure due an oil spill or to value the recreation gain associated with improved water quality on a river. The model is commonly applied in benefit-cost analyses and in natural resource damage assessments where recreation values play a role. Since the model is based on observed behavior it is used to estimate use values only.

The travel cost model is a demand-based model for use of a recreation site or sites. A site might be a river for fishing, a trail for hiking, a park for wildlife viewing, a beach for swimming, or some other area where outdoor recreation takes place. It is useful to separate travel cost models by those that estimate demand for a single site and those that estimate demand for many sites.

Single site models work like conventional downward sloping demand functions. The 'quantity demanded' for a person is the number of trips taken to a site in a season and the 'price' is the trip cost of reaching the site. Variation in price is generated by observing people living at different distances from the site. Price is low for people near the site and high for those living further away. The demand function slopes downward if trips decline with distance to the site.

Single site models are useful when the goal is to estimate the total use or 'access value' of a site. The elimination of a site is the usual application. The lost value is the total consumer surplus under the single site demand function -- the difference between a person's total willingness to pay for trips and the actual trip cost incurred over a season. Some valuation examples are

- a beach closure due to an oil spill,
- a fish consumption advisory that closes a lake for fishing, or
- a development that eliminates a natural area for wildlife viewing.

It is also possible to use a single site model to estimate the value associated with a change in the cost of access to a site. For example, an increase in an entry fee or the opening of a new entrance could be evaluated using a single site model. More difficult, and really calling for the transfer of a single site model from one site to another, is valuing the addition of a new site such as reservoir created by a dam.

There are some variations of the single site model for valuing changes in site characteristics such as improved water quality on a lake or an increase in the number of hiking trails in an wilderness area. However, this is not the strength of the model. When the goal is to value changes in site characteristics at one or more sites or to value the access to more than one site simultaneously, a multiple site model is preferred.

The random utility maximization (RUM) model is the most widely used multiple site model. A RUM model considers an individual's discrete choice of one recreation site from a set of many possible sites on a single choice occasion in a season. The choice of site is assumed to depend on the characteristics of the sites. For example, an individual making a fishing trip may consider trip cost, catch rate of fish, and site amenities. The choice of site implicitly reveals how an individual trades off one site characteristic for another. Since trip cost is always included as one of the characteristics, the model implicitly captures trade-offs between money and the other characteristics.

Some examples of valuing changes in site characteristics (at one or more sites) using the RUM model are

C improvements in water quality on lakes and rivers
C increases in the catch rate of fish on lakes and rivers
C improving the conditions of access to several local parks
C increases in moose populations in several hunting areas
C increasing the number of mountain biking trails in a state park

The RUM model may also be used to value access to one or more sites simultaneously. For instance, it may be used to value the loss of several beaches closed due swimming advisories or to value several ski slopes opened in a development project.

This chapter is organized into two sections. The first covers the single site model. The second covers the RUM model.[1] Both sections are organized around a table listing the steps required to estimate a basic version of the model. The emphasis is placed how one applies a basic modern version of the model and not on the frontiers of modeling. For more on recent developments see Herriges and Kling (1999) or Phaneuf and Smith (2002). For a good historical account of the model which was first applied over 50 years ago, see Ward and Beal (2000).

1. THE SINGLE SITE MODEL

This section is divided into four parts: basic model, steps in estimation, an example, and variations. I will focus on using the single site model to value site access since that is its strength and most common application.

2.1 Basic Model

The single site model is a demand model for trips to a recreation site by a person over a season. The 'quantity demanded' is the number of trips a person takes to the site. The 'price' is the trip cost of reaching the site which includes a person's travel expenses and time cost necessary to make the trip possible. In its simplest form the single site model is

$$(1) \quad r = f(tc_r)$$

where r is the number of trips taken by a person in a season to the site and tc_r is the trip cost of reaching the site. Like any demand function, one expects a negative relationship between quantity demanded (trips r) and price (trip cost tc_r). People living closer to the site face a lower cost of reaching the site and, all else constant, probably take more trips.

Trip costs alone will not explain an individual's demand for recreation trips.

It will also depend on things like income, age, experience in the recreation activities available at the site, and proximity to other recreation sites. Accounting for these factors gives a more realistic demand function with a set of shifters

$$(2) \quad r = f(tc_r, tc_s, y, z)$$

where tc_s is a vector of trip costs to other recreation sites, y is income, and z is a vector of demographic variables believed to influence the number of trips.

By incorporating trip cost to other sites, the model now accounts for substitutes. The tc_s's are the 'prices' of trips to substitute sites. If a person lives near a substitute site, the number of trips (r) is likely to decline as the person substitutes trips to that site and away from the site of interest in the analysis. A positive coefficient on tc_s would pick up this substitution effect. The other shifters work much as one would expect. A positive coefficient on income, for example, implies that the number of trips taken increases with income and so forth.

Figure 1 shows a linear version of equation (2) corresponding to

$$(3) \quad r = \beta_{tc_r} tc_r + \beta_{tc_s} tc_s + \beta_y y + \beta_z z \, .$$

If a person faces a trip cost of tc_r^0 in this model, he or she takes r^0 trips. The area A is his or her total consumer surplus for trips to the site during the season -- the difference between total willingness to pay for trips (area A+B) and total trip cost (area B). This is also called the individual's access value for the site. If the site were closed for a season, the individual would lose access to the site and consequently the area A. Notice that access values increase the closer a person lives to the site.

Figure 1. Access Value in a Linear Single Site Model

A more general expression for consumer surplus or access value, which applies to any functional form, is the area under the demand curve between an individual's current price and the choke price. The choke price is the price at which trips fall to zero in the model. Mathematically the consumer surplus is

$$(4) \quad \Delta w = \int_{tc_r^{0}}^{tc_r^{choke}} f(tc_r, tc_s, y, z)\, dtc_r$$

where tc_r^{choke} is the choke price and $tc_r^{\,0}$ is the individual's trip cost.

In application then, one seeks to estimate an equation like (3) using visitation data to a site. Table 1 shows the form of a typical single site data set -- trip count, trip cost, and demographic data for a sample of individuals. The data are gathered by survey. Once assembled, equation (3) is estimated by regressing trips (r) on the relevant explanatory variables in the data set (tc_r, tc_s, y, z_1, z_2). After the model is estimated, the parameters are used to compute access value (area A) for each individual in the sample. Means are computed and an estimate is extrapolated to the population. (The model is usually not linear but the principle is the same regardless of functional form.) The next section describes these steps in detail. For a derivation of the basic model from utility theory or from a household production function see Freeman (1993) or Bockstael (1995).

Table 1. Typical Form of a Data Set for Estimating a Single Site Model

ID	Number of Trips to the Site for the Season (r)	Trip Cost to Site *$* ($tc_r$)	Trip Cost to Substitute Site *$* ($tc_s$)	Annual Household Income *$000* ($y$)	Years Engaged in this Form of Recreation (z_1)	Number of Children (z_2)
1	7	45	200	45	17	0
2	1	150	20	100	2	2
3	21	12	65	77	22	1
.	.	.	.	.	.	.
.	.	.	.	.	.	.
N	3	98	25	55	4	1

2.2 Steps in Estimation

The steps are shown in Table 2. I will discuss each of these in a separate subsection. I will assume that a model is being estimated for the purpose of valuing site access, that the analyst is using day trip data only, and that the data are individual-based. This is the most common application of the model. Valuing quality changes at a site, as already noted, is not the strength of the single site model and incorporating overnight trips, while possible, is less common and presents special complications.

Following this section, I will discuss an actual single site application by Brent Sohngen (2000). I will describe how he handled each step of the analysis. He estimated a model of beach use for two Lake Erie beaches.

2.2.1 ***Step 1***: Define the Site to be Valued

The analysis begins with a definition of the site to be valued. The site might be a park, lake, beach, wilderness area, river segment, or some other area used for outdoor recreation. The boundaries are often easy to delineate. This is true in the case of a park, lake, or wildlife reserve. In other cases the delineation is not so clear. A river for fishing or white-water rafting, for example, calls for the analyst to define the length of the segment to be analyzed. A site for hunting or ocean fishing usually requires researcher defined boundaries as well. In these cases, one seeks to define a site broadly enough to include most, if not all, of the area affected by the policy being analyzed.

Government agencies, park services, and tourist bureaus often have literature and maps which help in defining a site. Sometimes government agencies charged with managing natural resource use within a state have their own definitions for sites, such as wildlife management units. These site definitions often work for a travel cost model.

Table 2: Steps in Estimating a Single Site Model

Step 1	Define the Site to be Valued
Step 2	Define the Recreation Uses and the Season
Step 3	Develop a Sampling Strategy
Step 4	Specify the Model
Step 5	Decide on the Treatment of Multiple Purpose Trips
Step 6	Design and Implement the Survey
Step 7	Measure Trip Cost
Step 8	Estimate the Model
Step 9	Calculate Access Value

2.2.2 ***Step 2***: Define the Recreation Uses and the Season

In some cases the site will have a single or dominate recreation use. In other cases there will be multiple uses such as fishing, swimming, boating, and viewing on a lake. Ideally one would like to include all recreation types and estimate separate demand functions for each. Sometimes, policies call for focusing on a single recreation type. Again, government agencies, park services, and tourist bureaus often have information and data to help identify major uses.

If recreation types are similar enough, they may be aggregated. For example, if there are many types of boating, one might treat these as a single group. The more similar the recreation types, the less problematic the aggregation. Aggregating sail-boating and motor-boating is less problematic than aggregating motor-boating and swimming. Most studies have some aggregation. Beach use, which can include sunbathing, swimming, surfing, jogging, and even surf fishing, is often treated as a single recreation type. Aggregation simplifies data collection and analysis. It means less information from each respondent, fewer observations, and less modeling. Again, one must be careful not to bundle recreation types that are too dissimilar.

Since individuals are sometimes observed engaging in more that one type of recreation on a single visit, a common practice is to identify the primary purpose of the recreation trip and classify the use accordingly. For example, one might ask respondents in a survey to report the number of trips taken "*primarily* for the purpose of fishing," and then "*primarily* for the purpose of boating." And, so on.

Along with defining the uses of the site one must also define the season for each use. Hunting and fishing may have a season defined by law. Skiing, rock climbing, and beach use will have a season defined by periods of favorable weather. Others uses, such a viewing and hiking may be year round.

2.2.3 ***Step 3***: Develop a Sampling Strategy

Next, a strategy is developed to sample the users and potential users of the site. There are essentially two approaches: on-site and off-site sampling.

In on-site sampling recreationists are intercepted at the site and asked to complete an oral or written survey. The survey may be completed on the spot or handed to people to mail to a specified address at a later time. It is even conceivable to recruit respondents on-site and later mail a survey to their home. In this case, the recreationist only reports an address at the site.

On-site samples have the advantage of hitting the target population directly. Every person interviewed has visited the site. Compare this to a random survey of the general population where the percent of people who have visited the site in the current season is likely to be quite small for most types of outdoor recreation. In this case, the number of contacts or interviews required to get a reasonable sample of users can be rather large. For this reason, on-site samples are popular for single site studies.

There are a number of issues to be aware of when using on-site samples. First, people who do not visit the site are missed. This implies a sample with no observations taking zero trips. This compromises the accuracy of the estimated intercept, the 'choke price', for the demand function. Think of a scatter of data points being used to estimate the demand function in Figure 1. With on-site sampling the data are truncated at one trip. We have no scatter of data points at zero trips and are forced to estimate the intercept using trip data over individuals having taken one or more trips. This is extrapolating outside the range of the observed data.

Second, on-site samples can be difficult to conduct in such a way that a random sample of users is obtained. Think about randomly drawing a user on a beach. When and where do you sample? A strategy must be devised. For example, randomly drawing several week days and weekends during the season for interviewing and then interviewing every tenth person is a strategy that attempts to approximate a random sample. Clear entry points, such as gates, help in on-site random sampling. Consideration must also be given to how one conducts a survey on-site. Interrupting someone as they sleep on the beach or as they put a boat in the water is hardly advisable. Catching respondents at an opportune time with minimum disruption will help response rates and extend common courtesy. Consideration must also be given as to whether or not one samples a person as they arrive or depart. The latter has the advantage that respondents know more about the actual recreation experience – catch rate of fish, time spent on the site, activities, and so forth.

Third, in estimation one must correct for selection bias inherent in on-site samples. The error term implicit in equation (3) in an empirical analysis will be truncated such that no observations less than one are observable. This will cause the estimated demand function to be too steep giving biased parameter and welfare estimates unless corrected in estimation. See Creel and Loomis (1990). On-site sampling will also over sample more frequent users. This is called endogenous stratification. A person that visits a site 10 times during the season is ten times more likely to be sampled than someone visiting the site only once. This also introduces bias which must be corrected. See Shaw (1988) or Haab and McConnell (2002, p. 174-181).

The alternative sampling strategy is off-site sampling. In off-site sampling individuals from the general population are contacted, usually by mail or phone, and asked to complete a survey. Unlike an on-site sample, an off-site sample will include people who take trips (participants) and people who do not take trips (nonparticipants). This gives the information needed to estimate the intercept, avoids the selection biases, and is simpler to design for random response. It also possible to model the individual's decision of whether or not to visit the site (see Step 8). Unfortunately, as noted above, off-site samples can be costly to assemble because the participation rates from the general population for a given site tend to be low.

Off-site sampling also raises the issue of determining the extent of the market. Is the full day trip population located within 20 minutes of the site, 3 hours of the site, or further? A local fishing pond may have a small geographic market. A popular ocean beach may have a large geographic market. In principle one would want to sample the entire market randomly.

In a day trip model of recreation use, a safe bet for the extent of the market is a maximum days drive to reach the site -- perhaps 3 to 4 hours. However, the further one gets away from the site, the lower the rate of participation and the higher the cost of assembling a sample with participants. Prior knowledge of where users come from may help establish the extent of the market. For example, parks may keep records of where their visitors come from. Stratified sampling is also possible whereby residents living closer to the site are sampled more heavily than those living further away. In this case one must correct for the sampling in estimation and welfare computation.

Using an off-site targeted population is a common and effective way to circumvent the participation rate problem. For many types of recreation, users must obtain a licence or pay a registration fee. There are, for example, fishing, boating, and hunting licenses. Participation rates among the set of license holders is significantly higher than the general population. Hence, if one can sample from the set of license holders, the number of people one must contact will be considerably lower than over the general population. Often times agencies that issue fishing licenses or boating permits will give an analyst access to addresses of license or permit holders. If so, one can randomly sample from this set.

For a good example of studies using on-site surveys see Sohngen (2000) and Siderelis and Moore (1995). See Boyle ( ) for a discussion and analysis using a targeted population.

2.2.4 ***Step 4***: Specify the Model

Before data collection begins the right hand side variables for equation (2) must be idntified. This establishes the information the analyst needs to gather in the survey.

Every model includes an individual's trip cost to the site (tc_r). This serves as own price in the demand model. Step 7 addresses measurement of trip cost at length. Most models also include a measure of trip cost to substitute sites (tc_s). There are usually three or fewer included. The analyst looks for sites frequently visited by the same population of users, sites similar in character to the site of interest in the analysis, and sites nearby. From this mix, a set is formed. Occasionally, proxies for substitute sites are used, such as the number of lakes or acreage of water within a certain distance of one's home.

As shown in equation (2) most models also include a measure of income (y) and a set of demand shifters (z) or demographic variables. The shifters are factors other than trip cost believed to influence the number of trips taken over a season. Some common shifters are

- family size
- age
- gender
- urban/rural residence
- occupation
- level of education
- club membership
- equipment ownership
- attitudinal information
- experience in activity

Family size or family composition may matter for many types of recreation such as beach use or hiking. Families with young children are more likely to use a beach. Club membership might include a fishing or hunting club, environmental group, or some other such association. One might use subscriptions to specialized recreation magazines as a explanatory variable as well. These variables pick up unobserved measures of intensity of interest in the recreation activity.

Urban/rural, occupation, and education are sometimes used as shifters as well. Occupation and education are usually categorical variables such as unemployed (yes/no), student (yes/no), retired (yes/no), and so forth. Education might be high school completed (yes/no) and college completed (yes/no). Attitudinal information is response data based on questions like "Would you consider yourself an advocate for the environment?" Experience includes things like number of years a person has been rock climbing or a self evaluation of level of expertise. This might come from the answer to a question like "How would you rate you level of rock climbing expertise? Novice, intermediate, or expert?"

Although the list is long, most analysts are parsimonious in their selection of shifters. Studies typically include anywhere from one to five variables. Sometimes analysts will report two or three model specifications such as a short model that includes only trip cost and income and a longer model that adds a set of shifters.

2.2.5 ***Step 5***: Decide on the Treatment of Multiple Purpose Trips

Recreation trips may be single or multiple purpose. On a single purpose trip an individual engages only in recreation at the site. On a multiple purpose trip a person does more. For example, they may visit family or friends along the way, take side trips for business, go shopping, go site seeing, visit other recreation sites and so forth.

Single purpose trips fit the travel cost model well. An individual leaves home, travels to a recreation site, engages in some type of recreation, and returns home. All travel expenses can, more or less, be attributed to creating the recreation experience.

Multiple purpose trips are more complicated. Trip expenses no longer purchase a recreation experience alone. Instead, they purchase a package of experiences. The logic of treating trip cost as the price of a recreation trip or recreation experience becomes tenuous. Attempts to apportion trip cost across the different purposes meet with little success. There is no logical way to identify the marginal cost of the recreation portion of the trip unless some restrictions are placed on the model (see Parsons and Wilson (1997)). How then does one handle multiple purpose trips in the analysis?

With day trip data the most common approach is to assume all trips are single purpose. If a case can be made that individuals engage in few other purposes along the way and that these other purposes are mostly incidental, then all trips may be treated as single purpose with little objection. With day trip data this is often a reasonable assumption. With overnight data it is not. Indeed, this is one reason many analysts confine their attention to day trips.

Another approach is to drop multiple purpose trips from the analysis. In this case, the analyst ask respondents to report multiple and single purpose trips separately or to report only single purpose trips. The model is then confined to single purpose trips.

A last approach is to amend to the basic model to accommodate multiple purpose trips. Mendelsohn et. al. (1992) and Parsons and Wilson (1997) are recent examples. Mendelsohn's framework broadens the definition of a site to include multiple purposes and proceeds with the same logic as the basic travel cost model. Their application is to a multiple site model where the other purposes are recreation at other sites but the reasoning would apply to a single site model. Parsons and Wilson (1997) present an empirical model, with theoretical support, that introduces a simple demand shifter for multiple purpose trips in a single site model in which all trips are analyzed together. They also offer an argument for why access values for a site may embody the consumer surplus associated with the other purposes of the trip.

2.2.6 ***Step 6*:** Design and Implement the Survey

The next step is to design and implement the survey. It is useful to start with a survey from a past study as a guide. The survey is usually divided into four parts

- Introductory material
- Trip count questions
- Last trip questions
- Demographic/Household Characteristic questions

The introductory material introduces the interviewer, identifies his or her affiliation, explains the purpose of the study, and provides assurances that keep respondents interested (eg., the survey is short and the responses are confidential). There are usually opening questions as well that are easy to answer, familiarize the respondent with the material, set the stage for the trip count questions, and help define the site and the season.

The next three parts of the survey are used to form a data set like that shown in Table 1. The trip count questions ask the respondent to report the number of trips taken to the site over a designated time period. These questions may be divided by recreation type (number of fishing trips, number of boating trips and so forth), by day and overnight trips, and/or by multiple and single purpose trips.

The last trip questions pertain to the most recent trip taken by the respondent and typically include information such as time spent on-site, number of people sharing travel expenses, other expenses incurred, and information on the trip experience such as number of fish caught. These data are used to construct trip cost and sometimes to create other explanatory variables in the demand model. These are gathered for the last trip only because gathering them for each trip over the season can lengthen a survey considerably and are difficult for respondents to recall for every trip.

The demographic/household characteristic questions correspond to the vector *z* in equation (3). These questions close out the survey and include the respondent's income and location of the respondent's hometown which is required to estimate trip cost. Income is used as a shifter and usually to calculate trip cost. In principle, one needs to know how trip cost will be calculated in this step to know what data to gather. For this reason steps 6 and 7 overlap somewhat.

It is essential to use good survey research methodology in developing the survey. Champ's Chapter 3 is a good reference on survey research methods. They cover issues pertaining to effective question writing, keeping response rates high, sampling, and more. While there is no need or room to repeat the principles of good survey research here, there are a couple of issues of special concern in travel cost surveys that deserve some attention: trip recall and trip categorization.

When people are asked to report their number of trips, the analyst assumes they will remember how many were taken. Since the model calls for a count of trips over a season, respondents are often asked to recall the number of trips taken over many months or even a year. This raises obvious questions. How accurate is an individual's recall of past trips? Will approximations be valid?

There is no evidence I am aware of from controlled experiments as to how serious recall error may be, but it is hard to deny the potential. One approach to improve recall is do the survey at several intervals over the season asking respondents to report trips taken only over the proceeding month or so. While sensible, this approach can raise survey costs considerably and lower responses rates through respondent attrition over the season.

Off-site surveys are usually conducted immediately following a season while trip recall is good. On-site surveys are done within a season, which is also likely to lessen recall problems. However, with on-site surveys the seasonal data are truncated because the respondent can only report the number of trips taken to date. The balance of the season is unknown. Two approaches are used to fill in the data for the end of the season: ask respondents to estimate the number of trips or predict the number trips based on current trip behavior. The later assumes an individual's rate of trips is constant throughout the season. Note that on-site samples that gather addresses only for an end-of-the-season mail survey have a complete trip count but probably have a greater recall problem.

As mentioned in step 2, there is often more than one type of recreation at the site. If so, the survey may be designed for multiple recreation types. The most common survey strategy is to proceed use-by-use in the questioning. For example, in valuing recreation at a local lake one might begin by asking "How many day trips did you take to the site for the primary purpose of fishing?" Then, following the fishing questions, there would be a similar block of questions about swimming, and then boating, and so on.

There are short-cuts that reduce the length of the survey. One might simply ask people to report "How many trips did you take to the site for purposes of recreation?" And then, "What is your primary recreation use of the lake?" This avoids questioning about each recreation use. People are then classified by their primary use. There is a model of lake visitation for people that use the site primarily for fishing, one for people that use the site primarily for boating, and so on. There is some mixing of trips within models in this case. However, if people tend to have a single dominate use for a site, this strategy can be effective.

It is also useful to isolate side trips and trips originating form somewhere other than the resident's hometown. For example, if a person is on a business trip and takes a trip to the beach or goes on a whale watching tour while on the trip, this is a side trip. The trip cost from the individual's home to the site would clearly overstate the real marginal cost of the trip. It is easiest to isolate and delete side trips from the analysis by clearly defining a day trip as being to and from one's primary residence when the trip count is being made. See Parsons and Wilson (1997) for an approach that adjusts trip cost and then incorporates side trips into an analysis.

A similar issue arises when an individual owns a cabin or cottage near the recreation-site of interest. Suppose the person spends most of the summer at the cottage and makes frequent day trips to the site from that second residence. How should these trips be handled in the analysis? As one 3 month long overnight trip from their permanent residence or as it many day trips to the site from their cottage? While the former choice should, in principle, embody all the individual's consumer surplus it is simply not amenable to the single site model. There are two strategies that one might consider. Either drop the observation from the sample and lose a heavy recreation user or include the observation and count the trips as day trips from the cottage. The latter approach understates the full the surplus, but avoids deleting important observations. To identify these persons is it necessary to include a question that identifies whether or not the individual owns property near the recreation-site and makes trips to the site from there.

2.2.7 *Step 7:* Measure Trip Cost

Once the raw data are assembled and organized, the trip costs or 'prices' to the site and any substitutes are computed. Trip cost is the sum of the expenses required to make a trip possible. Typical costs for a day trip include

- travel cost
- access fees
- equipment cost
- time cost

Travel cost includes all transit expenses. In a model of day trips where most or all of the trips are made by car, travel cost is measured using the U.S. Department of Transportation's or the American Automobile Association's estimate of the average cost of operating a vehicle per mile. Current studies are using about 35 cents per mile. These costs include fuel and upkeep. The round trip distance to the sites is usually calculated using a software package such a PC Miler. The per mile cost is multiplied by the round trip distance to arrive at trip cost. Tolls, if any, are added as well.

Since travel costs may be shared by several people, efforts are sometimes made to apportion the costs. For example, one might ask respondents to report the number of people sharing travel cost on a last trip and divide the cost equally. Or, one might ask directly for an individual's share of the cost. In either case, this is an example of a last trip question mentioned in Step 6. Since it difficult to pose these questions for each trip and since there is no logic way to handle separate trips within a season differently in a single site model, the analyst typically relies on last trip data.

If the site or any of its substitutes have an access fee, that fee is included in the trip cost. Sometimes sites will have annual or weekly passes, senior discounts, or admit children for free. Making adjustments for seniors and children is easy, but accounting for discounts is more difficult and usually ignored. Typically the daily fee is used.

Equipment costs vary by type of recreation. In fishing one needs bait, tackle, a rod, and sometimes use of a boat. For beach use there may be chairs, umbrellas, surf boards and so on. For bird watching, there are binoculars and film. For something like bait the cost is simply the market price of bait for a day of fishing. For durable goods an imputed rent is needed. If one rents or charters a boat for fishing the cost is simply the fee. If one owns a boat the rent or cost of its service flow needs to be imputed. One approach for imputing such costs is to use the rental fee for comparable services. This is, no doubt, an overstatement of the cost, but is usually easy to obtain. Often times equipment cost is excluded from the trip cost estimate. It is difficult to estimate and, in some cases, is a negligible portion of trip cost. If incorporated, equipment cost is usually obtained using a last trip question.

An alternative strategy for estimating trip cost, access fees, and equipment cost is simply to ask individuals to report their expenses on the last trip to the site. This is usually done by expense category. The estimates are then summed to arrive at trip cost. The advantage of this approach is that it uses perceived cost information and the researcher need not construct the cost estimates. Since individuals base trip decisions on perceptions of cost, which may diverge from actual costs, the respondent reported estimate is compelling. However, objective estimates based on researcher computation as described above is most common. The data are cleaner in the sense that they are uniform across individuals and there are no missing or otherwise peculiar numbers.

The most difficult issue in computing trip cost, and certainly the one that has received the most attention in the literature, is estimating the <u>time cost</u> of the trip. The time lost traveling to and from the site as well as the time spent on the site is time that could have been devoted to other endeavors. The value of those lost opportunities is the time cost of the trip. Time cost often accounts for a sizable portion of the total trip cost and deserves careful attention.

In most applications the estimate of time cost is related to a person's wage in some way. This has a theoretical basis so long as the individual has a flexible working arrangement and can substitute work time for leisure time at the margin. Under such conditions, in theory, an individual increases the number of hours worked until the wage at the margin is equal to the value of an hour in leisure. Multiplying the hourly wage times travel and on-site time, in this case, is a fair estimate of time cost. Unfortunately, this simple model breaks down for many individuals. The simple leisure/work trade off does not apply to individuals working a fixed 40 hour a week job for a salary. These individuals do not have the flexibility to shift time in and out of work in exchange for leisure. The tradeoff is also implausible for retired folks, homemakers, students, and unemployed persons.

Despite the difficulty of extrapolating the simple flexible leisure/work model to many individuals in a recreation data set, the most commonly used approach to value time is still wage-based. For people with fixed work schedules most studies impute an hourly wage using annual income. Reported annual income is divided by the number of hours worked in a year -- a number in the range 2000 to 2080. Another approach is to impute an individual's wage using a simple wage regression over the subset of individuals in the sample earning an hourly wage (Smith, et. al. (1983)). In this case, wage is regressed on income and a vector of individual characteristics such as age, gender, and education. The fitted regression is then simulated over non-wage earners to impute a wage. For a recent application see McConnell and Strand (1994, p.100).

In wage-based applications, it is also common to see some fraction of the imputed wage used to value time, anywhere from 1/3 of the wage to the full wage, as the value of time. According to Feather and Shaw (1999), this practice stems from early transportation literature wherein analysts had imputed the time cost in empirical travel studies in this range. The recreation literature has more or less accepted 1/3 as the lower bound and the full wage as the upper bound, but neither is really on firm footing. For example, Feather and Shaw (1999) show that in theory for those on a fixed work week it is possible for the value of time to be greater than the wage. Finally, there are approaches for inferring values of time from market data in the recreation context. See McConnell and Strand (1981), Bockstael et. al. (1988), and Feather and Shaw (1999).

Time traveling to the site as well as time spent on-site should be included in any calculation of time cost. While the time of getting to and from the site is more or less fixed, time at the site is chosen by each individual and may vary across the sample. Nevertheless, on-site time is typically assumed to be constant across individuals and valued the same as travel time. Sometimes analysts use the sample average length of stay on the last trip as an estimate of the fixed on-site time. Others allow the on-site time to vary across the sample using last trip data as each person's on-site time estimate for each site.

It should be evident that measuring trip cost calls for considerable researcher judgement. Moreover, many of the trip cost components are endogenous or chosen by individuals. For example, the purchase of equipment, number of people a person fishes with, mode and route of travel, time on-site, choice of residence (starting point for trip), and so forth are all chosen. Yet, the model assumes that individuals take price as given. This creates the potential for bias parameter estimates. There are a few attempts in the literature to deal with this endogenity. McConnell (1992) suggests an approach that allows on-site time to be endogenous, and Parsons (1991) suggests an approach for purging trip cost of its endogeneity due to choice of residence. See Randall (1994) for a criticism of the trip cost model fueled chiefly by the issue of trip cost endogeneity.

2.2.8 *Step 8:* Estimate the Model

The next step is to estimate the model specified in Step 4. In most modern single site applications, the model is estimated as a count data model. The dependent variable (number of trips) is a nonnegative integer, and the frequency of zero and small numbers of trips typically make up a sizable fraction of the data set. Count models are well suited for these data. See Hellerstein (1999), Creel and Loomis (1990), and Greene (1997, p. 931-946) for details.

The basic count data travel cost model is a Poisson regression. The number of trips taken by a person to a site in a given season is assumed to be generated by a Poisson process. The probability of observing an individual take r trips in a season is

$$(5) \quad \Pr(r) = \frac{\exp(-\lambda) \cdot \lambda^r}{r!}.$$

The parameter λ is the expected number of trips and is assumed to be a function of the variables specified in the demand model. To ensure nonnegative probabilities, λ usually takes a log-linear form

$$(6) \quad \ln(\lambda) = \beta_{tc_r} tc_r + \beta_{tc_s} tc_s + \beta_y y + \beta_d z .$$

Substituting equation (6) into (5) gives an expression for the probability of observing an individual take r trips as a function of trip cost, income, and individual characteristics. Equation (6) is the Poisson form of the recreation demand specified in equation (2).

The parameters in equation (6) are estimated by maximum likelihood. For each person in the sample the analyst knows r, tc_r, tc_s, y, and z. Using these data and equations (5) and (6) the probability of observing the number of trips actually taken is constructed for each person in the sample. The likelihood of observing the actual pattern of visits then is the product of these probabilities

$$(7) \quad L = \prod_{n=1}^{N} \frac{\exp(-\lambda_n) \cdot \lambda_n^{r_n}}{r_n!}$$

An individual is denoted by $n = 1, \ldots, N$, so r_n is the number of trips taken by person n. In estimation, the parameters β, on which λ depends according to equation (6), are chosen to maximum L. There are many software packages available for estimating by maximum likelihood. GAUSS and LIMDEP are popular for Poisson forms.

Consumer surplus, or access value, for each person in the sample (area A in Figure 1) has an explicit form in the Poisson model. For individual n the surplus is

$$(8) \quad S_n = \lambda_n / -\beta_{tc_r} .$$

where λ_n is the expected number of trips from equation (6). Once the parameters of the model are estimated, equation (8) is used to calculate the surplus value for each individual in the sample and then aggregated over the population of users to arrive at a total access value. This is discussed in detail in step 9.

The actual form of the probability used in estimation varies somewhat from equation (5) depending on whether one uses on-site or off-site sampling. Recall from step 3 that <u>on-site</u> random samples are truncated at one trip and over sample more frequent users. Either of these complications will bias parameter estimates unless corrected statistically.

The corrected probably for an on-site sample is a slight variation on the basic Poisson probability in equation (5). It takes the form

$$(9) \quad pr(r_n \mid r_n > 0) = \frac{\exp(-\lambda_n)\lambda_n^{r_n - 1}}{(r_n - 1)!} .$$

This corrects for truncation at one trip and endogenous stratification and differs from the basic Poisson regression only by r_n-1 replacing r_n. See Haab and McConnell (2002, p. 174-81) for the derivation. With on-site sampling then equation (9), instead of (5), enters the likelihood function for each individual. Consumer surplus is still measured as shown in equation (8). See Shaw (1988) and Greene (1997, p. 936-7) for more on truncated regressions in the Poisson model.

Off-site random samples avoid the problem of truncation and endogenous stratification and have the advantage of including non-participants. This allows the researcher to model the decision to participate in recreation at the site or not. For example, in the simple Poisson model of equations (5) and (6) the probability of not participating *(r = 0)* can be modeled along with the decision of how many trips to take *(r = 1,2,3 ,....)*. The same Poisson process is assumed to generate the outcome for any count of trips, including zero. For individuals taking zero trips, $\Pr(r_n=0) = \exp(-\lambda_n)$ enters the likelihood function. Again, modeling participation is important because it helps pin-down the choke price or intercept on the demand function by using observed choices to estimate $\Pr(r_n=0)$.

A more complicated version of the model is the hurdled Poisson which assumes that the decision of whether or not to take a trip and the decision of how many trips to take are generated from different Poisson processes. The probability of taking zero trips is assumed to be $\exp(-\theta_n)$, where θ_n is some function of individual characteristics that governs whether or not a person takes trips at all. This may include some of the same variables used in the trip frequency portion of the model as well as some new variables. The individual's probability of taking one or more trips then is just $(1\text{-}\exp(-\theta_n))$. The model becomes

$$(10)\qquad \Pr(r_n = 0) = \exp(-\theta_n)$$
$$\Pr(r_n \mid r_n > 0) = (1-\exp(-\theta_n))\frac{\exp(-\lambda_n)\lambda_n^{r_n}}{r_n!(1-\exp(-\lambda_n^{r_n}))}$$

The term $1\text{-}\exp(-\lambda_n)$ scales the all the non-zero probabilities so that they sum to one. The likelihood function is then constructed using the probabilities in equation (10).

In some cases the hurdled model is estimated with a double hurdle. For example, the population of nonparticipants in a given season may be divide by those who never take trips to the site of interest and those who do but for one reason or another have not in this particular season. For example, in a study of fishing on a river, there may be two types of nonparticipants: those who never fish and those who fish but not at this site in this season. The analyst may wish to treat these groups differently -- the latter group may suffer from a site loss while that former will not (at least in terms of use values). See Shonkweiler and

Shaw (1996) and Haab and McConnell (1996) for more on the double hurdled model.

Finally, it is common to see a variation of the Poisson Model know as the Negative Binomial Model to estimate travel cost models. In the Poisson Model the mean and variance of r_n are constrained to be equal. To the extent that this constraint is unreasonable, the model is in error. The Negative Binomial Model is an approach for relaxing this constraint. The basic reasoning and structure above still applies, but the estimation is somewhat more complex. See Greene (1997, pp. 939-40) and Haab and McConnell (2002).

2.2.9 ***Step 9*:** Calculate Access Values

In the final step, access value for the site is computed using the estimated model. Access value may be reported as

- a mean seasonal value per person,
- a total seasonal value for the population,
- a per trip value per person, and/or
- a total discounted present value of the site.

The seasonal per person estimate for any observation in the sample is the area A in Figure 1. In the Poisson Model this is S_n in equation (8) . The estimated seasonal access value for the n^{th} individual in the sample then is

$$(11)\quad \hat{S}_n = \frac{\hat{\lambda}_n}{-\hat{\beta}_{tc_r}} = \frac{\exp(\hat{\beta}_{tc_r} tc_{rn} + \hat{\beta}_{tc_s} tc_{sn} + \hat{\beta}_y y_n + \hat{\beta}_d z_n)}{-\hat{\beta}_{tc_r}} \ ,$$

where ˆ denotes an estimated value using the results of the Poisson regression, and the subscript *n* on an explanatory variable denotes the value of that variable for individual *n*.

If one has estimated a Poisson Model using a randomly drawn <u>off-site sample</u>, the sample mean access value is

$$(12)\quad \bar{S}_{off} = \frac{\sum_{n=1}^{N} \hat{S}_n}{N}$$

where $\bar{S}_{off}$ is an unbiased estimate of the mean access value over the population of participants and nonparticipants in the market area sampled. A reasonable estimate of aggregate seasonal access value is

$$(13)\quad AS = \bar{S}_{off} \cdot POP_{off}$$

where POP_{off} is the total number of people in the relevant geographic market. POP_{off} is typically from census data. For example, if one sampled the entire population over 16 years old within a days drive, that population is the relevant value for POP_{off}. If one is using a targeted population such as all individuals holding a fishing license, POP_{off} is the total number of people holding a license in the relevant time period.

If one is estimating a model with an <u>on-site sample</u>, the sample mean access value is a biased estimate of the population mean because is over samples more frequent visitors to the site. A 'corrected' sample mean is

$$(14)\quad \bar{S}_{on}^{c} = \frac{1}{N^*}\sum_{n=1}^{N}\frac{\hat{S}_n}{r_n}\text{, where } N^* = \sum_{j=1}^{R}\frac{n_j}{j}.$$

where dividing the individual surpluses by r_n and the summation by N^* are corrective weights. The term n_j is the number of persons in the sample taking j trips, and R is the largest number of trips taken by a person in the sample. A reasonable estimate of aggregate seasonal surplus in this case is

$$(15)\quad \mathrm{A}S = \bar{S}_{on}^{c} \cdot POP_{on}$$

where POP_{on} is the total number of participants at the site over the season. POP_{on} may be gathered from an independent data source on participation rates at the site or may be estimated from survey data. In the latter case one must again adjust for over sampling frequent visitors.

In principle, an off-site study using equation (13) and an on-site study using equation (15) estimate the same aggregate value *AS*. $\overline{S}_{on}^{c}$ is larger than $\overline{S}_{off}$ because it excludes nonparticipants. But, POP_{on} is smaller than POP_{off} for the same reason. One way of thinking about $\overline{S}_{on} \cdot POP_{on}$ is that it implicitly assumes that anyone who has not taken a trip has zero surplus and is excluded in the calculation of *AS*. $\overline{S}_{off} \cdot POP_{off}$, on the other hand, includes nonparticipants and each will have a $\hat{S}_n > 0$ since everyone has some positive probability of taking one or more trips in a Poisson model. And, the nonparticipants contribute many low surplus values to the overall mean giving the off-site model its lower seasonal mean.

An alternative method of estimating aggregate surplus is to compute an average per trip per person value and then multiple this by an estimate of the total number of trips taken to the site. Since average per trip values are the per person seasonal value of the site divided by the number of trips taken, the average per trip value in a Poisson Model is

$$(16) \quad \hat{t} = \frac{(\hat{\lambda}_n / -\hat{\beta}_{tc_r})}{\hat{\lambda}_n} = \frac{1}{-\hat{\beta}_{tc_r}} \ .$$

This applies to on-site and off-site models alike. To arrive at an aggregate value for the site multiply the average per trip value by the number of trips taken to the site during the relevant season. This gives

$$(17) \quad AS = \hat{t} \cdot TRIPS$$

where *TRIPS* is the total number of day trips to the site over the relevant season. Many parks and major recreation sites collect data or at least estimate *TRIPS* which makes this a popular method of measuring aggregate surplus.

Finally, it is common to see the discounted present value of a site computed using the seasonal aggregate estimate. The easiest approach is to assume no change in the use of the site over time, no change in the change in the character of the site, and a constant rate of discount. Then, using the conventional formula for a the value of a perpetuity, the discounted present value of the site is

(18) $\mathrm{PV} = AS / i.$

where i is the real rate of discount, usually a number in the range of .01 to .05. See Krutilla and Smith ( ) for a discussion of examples of measuring PV when use and values change over time.

In reporting site values, be it S, AS, t, or PV, it is important to be clear what is and is not included in the value. For example, it may be a value of day trips for rocking climbing in a park. If so, it should be noted that this excludes overnight and side trips, other types of recreation taking place in the park, and nonuse values.

2.3 A Single Site Application

Sohngen's (2000) model of beach recreation on Lake Erie in 1997 is good example of a modern application of the single site model. He estimated two models -- one for Maumee Bay State Park and the other for Headlands State Park. Muamee Bay is in the western Ohio. Headlands is in eastern Ohio. Maumee Bay offers opportunities for recreation beyond beach use including golfing, camping and so forth. Headlands is more natural.

Shongen defines the sites as the beaches located withing these two parks (step 1). The recreation users are people visiting a beach within either park for recreation or pleasure (step 2). All forms of beach recreation are aggregated into a single type. The data were gathered on-site (step 3).

Both models were specified with own price (tc_r), income (y), one or two substitute prices (tc_{s1} and tc_{s2}), and a variety of explanatory variables (z) (step 4). The Maumee Bay model used one substitute site; the Headlands model used two. In both cases the substitutes were nearby beaches similar in character to the beaches under study. The Headlands substitutes were located on either site of Headlands Park. Income was measured as annual household income divided by 10,000. The demand shifters were the scaled responses to five attitudinal questions. "These questions asked individuals to rank (from 1 to 5) how important certain issues were in the choice of making a visit. The issues included water quality, maintenance, cleanliness, congestion, and facilities.

Higher rankings indicate the issue is more important to the individual visitor." Finally, the model included a dummy variable for whether or not the primary purpose of the last trip was for beach use.

Multiple purpose trips were included in the demand model along with single purpose trips (step 5). Since all trips were single day visits by people living within 150 miles of the site, other purposes are likely to be incidental side trips of little consequence. As noted above, they included a demand shifter in the model for trips with the sole purpose of visiting the beach. *Sole = 1* if the sole purpose was to visit the beach and *Sole = 0* otherwise.

The survey was conducted on-site. Random beach users were handed a survey and asked to return it by mail (step 6). Respondents reported their number of day as well as overnight trips to the beach over the entire season. Over 90% of the trips to Headlands were day trips; over 66% to Maumee Bay were day trips. Only day trips were considered in the analysis. Their hometown (and zip code) was given when respondents were asked for their home address. They also gathered data on other activities while on a typical beach trip and the attitudinal data on factors that affected a decision to visit the beach. The latter were gathered to constructed the demand shifters specified in step 4. The response rate was 52% for Headlands and 62% for Maumee Bay.

Trip cost was measured as the sum of travel expenses and time cost (step 7). Distances to the site were measured as the linear distance from the center of an individual's home zip code to the beach using latitude and longitude coordinates. That distance was doubled (round trip) and then multiplied by 33 cents per mile giving total transit cost. The average distance traveled to Headlands was 26 miles and to Maumee Bay was 35 miles. Time cost was measured as an imputed wage times travel time. Travel time was calculated using the estimated round trip distance and assuming people traveled at 40 miles per hour. The imputed wage was 30% of annual household income divided by 2040. On-site was ignored.

Four different functional forms were estimated (step 8). Two continuous (linear and log-linear) and two count (Poisson and Negative Binomial). Truncation at zero trips due to on site sampling was accounted for in each model. Endogenous stratification (over sampling of more frequent users in on site samples) was not.[2] The parameter estimates for the Poisson models for Maumee Bay and Headlands are shown in Table 3.

The coefficient on own trip cost (tc_r) is negative and significant in both regressions, so the demand function is downward sloping. The coefficient on income is positive in both regressions but significant in only the Headlands regression. The effect of substitute sites on trips works as expected in the Maumee Bay model, the higher the cost of reaching a substitute all else constant, the more trips taken to the site. The results is a bit weaker in the Headlands model. One of the substitute prices has a negative but insignificant coefficient and the other is positive and significant.

The dummy variable for whether or not the primary purpose of the trip is for beach use works in opposite directions in the two models. The attitudinal variables also produced variable results. Recall that these variable are characteristics of the respondents, not sites. One would expected the coefficients on these variables, however, to be correlate with the characteristics of the site. For example, if a site is congested, one would expected a negative and significant coefficient on congestion. This would indicate that people who consider congestion important in making a trip decision are less likely to make a trip to a congested site.

Table 3: Single Site Poisson Models for Headlands and Maumee Bay

	Maumee Bay	***Headlands***
Variable	*Parameter Estimate*	*Parameter Estimate*
tc_r	-.040***	-.026***
Income	.018	.040***
Sole	-.016	.292***
tc_{s1}	.004***	.005
tc_{s2}	--	-.004
Water Quality	-.053	-.139***
Maintenance	-.270***	.033
Cleanliness	.176**	.028
Congestion	-.065*	-.066***
Facilities	.098**	-.004
Constant	2.648***	2.433***
R^2	.38	.29
Sample Size	230	345

Notes: Three asterisks denotes that a coefficient that is statistically significant at the 99% level of confidence; two denotes significance at 95%; and one significance at 90%. *Water quality, maintenance, cleanliness, congestion,* and *facilities* are individual, not site characteristics, based on respondents ranking of the importance of each beach characteristic in making a trip decision. The ranking ran from 1 to 5, where 1 = strongly disagree it is important and 5 = strongly agree it is important. Source: Sohngen (2000).

The results are then used to estimate access values (step 9). A per trip, an annual aggregate, and a discounted present value estimate was computed for each park. The per person per trip values using the model above are $25 (= 1/.04) for Maumee Bay and $38 (= 1/.026) for Headlands. The range of per trip values from all models (not reported here) was $14 to $33 for Maumee Bay with a midpoint of $23.50, and $11 to $39 for Headlands with a midpoint of $25.

The Ohio Department of Natural Resources reported the total number of trips taken to the beaches during the 1997 season at 224,000 for Maumee Bay and 238,000 for Headlands. The total access value of the beaches for day trips was estimated by multiplying total trips to each beach by its per trip value. This is my equation (18). Sohongen reported these using the midpoint of the per trip values across the models, so I report the same. These were \$5.6 million (= 238,000·\$23.50) for Maumee Bay and \$5.6 million (= 224,000·\$25) for Headlands. Last, assuming a real discount rate of 3% and no future change in the use of the beaches, the total discounted present value of each beach is \$187 million (= \$5.6/.03). This accounts for day trip beach use only, excluding all other uses as well as non-use value.

2.4 Variations

While the single site travel cost model using individual data to estimate the access value of a site is the most defensible and widely used application, there are some variations on this theme. I will briefly mention a few of these here and direct the reader to the latest literature.

First, there are single site zonal travel cost models. These are estimated using aggregate visitation rate data and average trip costs from predefined zones near a recreation site. The zonal model has fallen out of favor due to its lack of consistency with basic theory. Nevertheless, when data are limited, the zonal model can provide a useful approximation. See Hackett (2000) for a recent application and Loomis and Walsh (1997) for more detail.

Second, there have some efforts at valuing quality changes with single site models. These all seek to estimate a shift in a single site demand function due a change in quality and to then use the area between the functions as an estimate of the value of the quality change. There are two versions of these models: pooled and the varying parameter. These can be estimated with cross section data on many sites (see Smith and Desvouges (1985) and Loomis (1988)), using time series data on a single site (see Brown et. al. (1983)), or using data that combines actual trips to a site with hypothetical visits to a site (see Layman et. al. (1996) or McConnell (1986)).

Third, there are systems of single site demand equations. These are stacked single site models for a group of substitute sites. See Bockstael, McConnell, and Strand (1991, pp.254-6) for a discussion and Burt and Brewer (1971) for the first application to a system of demands. Morey (1981) estimated a system of share equations and Ozuna and Gomez (1994) and Shonkwiler (1999) have estimated systems of count data models. These models account for substitutes and allow one to estimate access value for more than one site simultaneously. However, as Bockstael, McConnell, and Strand (1991) explain, it is impossible to value quality changes without placing rather stringent restrictions on the model. Furthermore, it is difficult to estimate such models when the number of sites rise above more than a half dozen or so. The RUM model, on the other hand, can hundreds or even thousands of substitute sites and works well in valuing quality changes.

3. THE RANDOM UTILITY MODEL

This section is divided into four parts like the previous section: basic model, steps in estimation, an application, and variations. I will discuss the use of the RUM model for valuing site access and changes in site quality. As a general rule the RUM model is preferred to a single site model. It does a better job of capturing site substitutes and it is better for valuing quality changes.

3.1 Basic Model

The RUM model considers a person's choice of a site for a recreation trip. Instead of a 'quantity demanded' as in the single site model, there is a site chosen. In choosing a site a person is assumed to consider its 'price' and its characteristics. The 'price' is the trip cost. The characteristics are things about the site that matter to people such as ease of access and environmental quality. While the time frame for a single site model is a season, the time frame for the RUM model is a choice occasion. When analyzing day trips a choice occasion is simply a day.

The theory works as follows. On a given choice occasion, a person considers visiting one of S sites denoted as $i = 1,2,....,S$. Each site is assumed to give the person a site utility v_i. The utilities are assumed to be a function of trip cost and site characteristics. The utility for site i assuming a linear form is

$$(19) \quad v_i = \beta_{tc} tc_i + \beta_q q_i + e_i$$

where tc_i is the trip cost of reaching site i, q_i is a vector of site characteristics, e_i is a random error term, and the βs are parameters. One expects site utility to decline with trip cost ($\beta_{tc} < 0$), to increase with desirable characteristics such a easy access and good environmental quality, and to decrease with undesirable characteristics. The random error term accounts for unobserved factors.

In theory, a person chooses the site with highest utility. Site *k*, for example, is chosen if

(20) $\beta_{tc}tc_k + \beta_q q_k + e_k \geq \beta_{tc}tc_i + \beta_q q_i + e_i$ for all $i \in S$.

A useful way of expressing this result is in terms of trip utility. An individual's trip utility is

(21) $u = \max(v_1, v_2, ..., v_S)$.

Trip utility is the maximum attainable site utility on a given choice occasion assuming a person visits a site. If site *k* gives the highest utility, the person visits site *k* and attains trip utility $u = v_k$.

Since people may choose not to take a trip on a given choice occasion, it is common to see a no-trip utility included the choice set. Let a person's no-trip utility be

(22) $v_0 = \alpha_0 + e_0$.

No-trip utility is the highest utility a person can attain in any activity other than visiting one of the S sites. A person now chooses from S + 1 alternatives -- S sites and no-trip. Now, it is useful to define a choice occasion utility as

(23) $u^* = \max\{v_0, v_1, ..., v_S\}$.

Choice occasion utility is the maximum attainable utility on a given choice occasion. If no-trip gives the highest utility, choice occasion utility is $u^* = v_0$ If visiting site *k* gives the highest utility, choice occasion utility is $u^* = v_k$. And, so forth.

Individual characteristics similar to the vector *z* in the single site model (age, family size, years of experience in a recreation activity, and so on) may also enter the model and do so in one of two ways. First, they may be used to capture differences in participation in recreation across the sample. Some people like to fish, and some do not. Some people like to go to the beach, and some do not. And, so forth. To capture differences in participation characteristics are entered as shifters in the no-trip utility function

$$(24) \quad v_0 = \alpha_0 + \alpha_1 z + e_i$$

where z is a vector of characteristics believed to influence a person's propensity for recreation. In this way, no-trip utility varies across the population. Consider a model of recreational fishing. If gender is a dummy variable in the vector z ($z_g = 1$ if female) and men like to fish more than women, a positive coefficient on z_g allows women to have a higher no-trip utility than men and captures this effect. Or, consider rock climbing. If interest in climbing decreases with age, a positive coefficient on age in the vector z would capture this effect.

Individual characteristics may also be used to capture differences in preferences across the population for different sites. Some people like natural beaches, while others like more developed beaches. Some people trailer a boat and prefer a site with a boat ramp, while others look to rent a boat or moor a boat at some preselected site. To captures differences in preferences for different sites, individual characteristics are interacted with the relevant site characteristics.

Assume there are m site characteristics and let $q_i = (q_{1i}, q_{2i}, \ldots q_{mi})$. Suppose in an analysis of beach use that there are people for whom surf fishing at a beach is an important part of their recreation experience and for others of whom it is not. Suppose I have data on whether or not a person owns a surf fishing license. I might specify site utility as

$$(25)\quad v_i = \beta_{tc}tc_i + \beta_{q_1}(q_{1i} \cdot z_1) + \beta_{q_2}q_{2i} + ... + \beta_{q_m}q_{mi} + e_i$$

where q_{1i} is a measure of the quality of surf fishing and $z_1 = 1$ if the person owns a surf fishing license and $= 0$ if not. In this way $\beta_{q1}q_{1i}$ affects site utility only for individuals who own a surf fishing license. Sites with good surf fishing give higher utility for the fishing population but not for the general population. It is possible to concoct numerous interactions of this sort.

Now, lets turn to how the RUM theory can be used to value site access at one of the S sites. Suppose site 1 is closed due to an oil spill. Using RUM theory, I can express a person's choice occasion utility with and without a spill. The utility without a spill, the baseline, is

$$(26)\quad u^*(baseline) = \max\{v_0, v_1, v_2 ..., v_S\}\ .$$

The utility with a spill is

$$(27)\quad u^*(spill) = \max\{v_0, v_2, v_3 ..., v_S\}\ .$$

The utility $u^*(spill)$ excludes site 1 since it is closed. The decline in utility or welfare loss due to the spill is $u^*(spill) - u^*(baseline)$. This is the difference in the utility a person can attain when the site is and is not available. The change in welfare is

$$(28)\quad \Delta u^* = \max\{v_0, v_2 ..., v_S\} - \max\{v_0, v_1, v_2 ..., v_S\}\ .$$

Using equation (28), suppose a person visits site 1 if there is no closure. It must be true for that person that $v_1 > v_i$ for all i. If site 1 where to close, the person would choose the alternative with the second highest utility. Suppose that is site k. If so, the person visits k instead of 1 and utility declines from v_1 to v_k. By the same reasoning, if no-trip has the next highest utility, the person no longer takes a trip and utility declines from v_1 to v_0. The closer the second highest utility is to site *1* utility, the smaller the decline in welfare.

If a person visits a site other than site *1* if there is no closure, then there is no loss in utility. For example, if a person visits site *k* when site *1* is open, then *u*(spill)* = *u*(baseline)* = v_k. The same alternative that maximized utility without closure maximizes it with the closure. Note that there is no accounting here for congestion at site *k* that may follow from the closure of site *1*.

To convert the decline in utility in equation (28) into monetary terms, divide by an individual's marginal utility of income. In the RUM model the negative of the coefficient on trip cost, $-\beta_{tc}$, is a measure of the marginal utility of income. It tells us how much an individual's site utility would increase if trip cost were to decline or, what is the same, if income were to rise for that trip.[3] The welfare loss in monetary terms then is

$$(29) \quad \Delta w = \left[\max\{v_0, v_2 ..., v_S\} - \max\{v_0, v_1, v_2 ..., v_S\}\right] / -\beta_{tc}$$

Δw is a *per choice occasion* value. In a day trip model it gives the welfare loss for a day due to the oil spill. To convert it to a seasonal value comparable to the single site model, multiply by the number of choice occasions in the season. The seasonal value for the loss of site *1* is

$$(30) \quad \Delta W = T \cdot \Delta w$$

where T is the number of choice occasions in a season.

The value of access to more than one site is computed by dropping many sites from equation (27). For example, if sites 1 through 5 are closed due the oil spill, sites 1 through 5 are dropped in equation (27). It also possible to value new sites by adding them to the choice set. In this case, the trip cost and site characteristics for the new site or sites must be specified as well.

The model is used in a similar way to value changes in site characteristics. Instead of dropping a site from the choice set, a site characteristic is altered which, in turn, changes site utility. For example, if water quality is a site characteristic in the vector q_i, a decline in water quality at site *1* is captured by writing q_1 as $q_1^{\#}$, where the element in q_i measuring water quality is smaller in $q_1^{\#}$ than in q_1. The utility at site *1* then declines from $v_1 = \beta_{tc}tc_1 + \beta_q q_1 + e_1$ to $v_1^{\#} = \beta_{tc}tc_1 + \beta_q q_1^{\#} + e_1$.

In this case an individual's choice occasion utility without the decline in water quality is *u*(baseline)* in equation (26). The choice occasion utility with the decline in water quality is

$$(31) \quad u^*(dirty) = \max\{v_0, v_1^{\#}, v_2 ..., v_S\},$$

where $v_1^{\#}$ is the utility at site *1* with the now poorer water quality. Following the same reasoning as a site closure, the welfare loss is

$$(32) \quad \Delta w = \left[\max\{v_0, v_1^{\#}, v_2 ..., v_S\} - \max\{v_0, v_1, v_2 ..., v_S\}\right] / -\beta_{tc}.$$

Equation (32), like equation (29), captures an array of possible behavioral responses. If an individual does not take a trip or visits a site other than site *1* without the decline in water quality, there is no change in utility. The same alternative maximizes utility with and without the change in water quality, so choice occasion utility is unchanged. If an individual visits site *1* without the change, he or she may either continue to visit site *1*, visit another site, or take no trip with the change. In each case, the choice occasion utility declines.

If the individual continues to goes to site *1*, utility drops from v_1 to $v_1^{\#}$. The individual maximizes utility by making the same trip but the recreation experience is diminished due to the decline in the water quality. If the individual goes to another site (say *k*) or no longer takes a trip, then utility declines from v_1 to v_k or from v_1 to v_0. All of these manifestations are captured in equation (32). And again, this gives a *per choice occasion* value. A seasonal value is computed as before using equation (30). And again, as with site access value, quality changes can be evaluated at many sites simultaneously by altering the relevant attribute for each affected site in equation (31).

So far, I have treated the RUM model as deterministic -- assuming all parameters and the error terms are known. In application, however, the parameters are estimated and the error terms are unknown. The error terms are assumed to come from some know random distribution. The choice occasion utility in equation (23) then is also random. This is easy to see by substituting the random site and no-trip utilities into equation (23) giving

$$(33)\quad u^*(baseline) = \max\{\alpha_0 + e_0, \beta_{tc}tc_1 + \beta_q q_1 + e_1, ..., \beta_{tc}tc_S + \beta_q q_S + e_S\}$$

If the error terms in equation (33) are random, choice occasion utility u^* is random. For this reason, in application, one uses the expected value of choice occasion utility, instead of a deterministic value. Expected choice occasion utility, the applied counterpart to equation (23), is

$$(34)\ eu^*(w/o) = E[\max\{\alpha_0 + e_0, \beta_{tc}tc_1 + \beta_q q_1 + e_1, ..., \beta_{tc}tc_S + \beta_q q_S + e_S\}]$$

Equation (29) for access value and (32) for a quality change now take the forms

$$(35)\quad \Delta w = \left[E(\max\{v_0, v_2, ...v_S\}) - E(\max\{v_0, v_1, v_2, ...v_S\})\right]/-\beta_{tc}$$

$$(36)\quad \Delta w = \left[E(\max\{v_0, v_1^{\#}, v_2, ...v_S\}) - E(\max\{v_0, v_1, v_2, ...v_S\})\right]/-\beta_{tc}$$

In estimation, the form of the distribution for the error terms determines the form of the expected value of the choice occasion utility. Each of the behavioral responses to the a site closure and a quality change described above still apply. However, now, each response occurs with some probability. Annual values, once again, are calculated by multiplying the per choice occasion values in equations (35) and (36) by the number choice occasions in the season.

In estimation then, one seeks to estimate the parameters of the site and no-trip utilities in equations (19) and (24) using visitation data. Table 3 shows the form of a typical data set -- a count of trips to each site in the choice set, trip cost to each site, detailed site characteristics for each site, and demographic data across a sample of individuals. For simplicity the table pertains a simple model with three sites and one site characteristic (water quality). In most applications the number of sites and site characteristics is much larger. Once the data are assembled, the parameters of site and no-trip utility are estimated using some form of a discrete choice multinomial logit model (more on this in the following sections). The estimated parameters are then used to value site access or quality changes at specified sites for each individual in the sample using equation (36) and (37). Seasonal measures and means across the sample are computed. And finally, the estimates are extrapolated to the population. The next section covers these steps in detail.

Table 3: Typical Data Set for Estimating a RUM Model (3 site choice set)

ID	Number of Trips to Site			Trip Cost to Site $			Water Quality Index at Site *1-10*			Income *$000*	Age *yrs*
	Site 1 (r_1)	Site 2 (r_2)	Site 3 (r_3)	Site 1 (tc_1)	Site 2 (tc_2)	Site 3 (tc_3)	Site 1 (q_1)	Site 2 (q_2)	Site 3 (q_3)	(y)	(d_1)
1	2	0	17	45	158	15	10	2	1	67	43
2	1	0	3	111	201	35	8	7	5	22	28
3	0	3	0	29	33	345	2	8	9	109	39
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.	.	.	.
N	12	0	0	12	66	123	5	2	1	78	51

3.2 Steps in Estimation

The steps in estimating the basic RUM model are listed in Table 4 and are similar to the steps for a single site model. To avoid repetition I will focus on aspects of the steps in the RUM model that are different from the single site model. It will be useful to refer the single site counterpart for most steps.

I will assume that a RUM model is being estimated for the purpose of valuing site access and quality change at several sites, that the analyst is using day trip data only, and that the data are individual-based.

Again, I will follow this section with a brief discussion of a RUM application by Matt Massey and myself. We estimated a model of Mid-Atlantic beach use by Delaware residents to value beach closures and loss due to erosion. I will describe how we handled each step of the analysis.

Some other good applications to read if you are just beginning to learn about the use of trip cost RUM models are Montgomery and Needelman (1997) and Murray, Sohngen, and Pendelton (2001). I also recommend two publications which go into considerable depth in laying out the development of their model including basic theory, survey design, data collection, computer programs, and so forth. These are McConnell and Strand (1994) and Hoehn et. al. (1996).

Table 4: Steps in Estimating a RUM Model

Step 1	Identify the Impacts to be Valued
Step 2	Define the Population of Users to be Analyzed
Step 3	Define the Choice Set
Step 4	Develop a Sampling Strategy
Step 5	Specify the Model
Step 6	Gather Site Characteristic Data
Step 7	Decide on the Treatment of Multiple Purpose Trips
Step 8	Design and Implement the Survey
Step 9	Measure Trip Cost
Step 10	Estimate Model
Step 11	Calculate Access and/or Quality Change Values

3.2.1 ***Step 1***: Identify the Impacts to be Valued

The analysis begins by identifying the impacts to be valued. These will take the form of site closures, openings, or quality changes at one or more sites. Some examples of impacts are the closure of many lakes and rivers due to a fish consumption advisory, the opening of a new park or skiing area, the expansion of several hunting areas, a change in water quality across several urban beaches, and so forth. Identifying the impacts to be evaluated at the outset defines the research problem and gives direction for the succeeding steps.

The impacts may be hypothetical. For example, a site is currently open and you want to analyze the loss associated with its possible closure, or a site presently has no hiking trials and you want to consider the gain associated with adding trails. The impacts may also be actual events. For example, several lakes are closed due to contamination and you want to analyze the welfare loss.

With a site closure or opening one identifies the areas and, in turn, the sites that are affected. With a quality change one is identifying the affected sites and the changes that will occur there. In this case, it is useful to begin thinking early in the analysis about how quality will be measured for each site. For example, will objective measures of quality be used -- like levels of dissolved oxygen for water quality? Or, will perceived measures of quality be used -- like a rating based on the reporting from respondents in a survey ? Also, are there published measures or will you need to construct new measures as part of the analysis? And, how will these quality changes map into the policy being evaluated? It is also necessary to establish early on that there is sufficient variation in quality across sites to measure its effect on site choice. If quality is more or less uniform across the sites for the specific characteristic of interest, it is impossible to measure its effect on site choice. See Parsons and Kealy (1992) for an example using objective measures and Adamowicz et. al. (1997) for an analysis using objective and perceived measures.

In some cases the impact may be analyzed as either a change in access or a change in quality. For example, a fish consumption advisory might be analyzed as a quality change instead of a closure, especially if people continue to use the site. In this case one includes a dummy variable for fish consumption advisories in the characteristic vector q. For this approach to work some sites in the choice set must have advisories currently.

3.2.2 ***Step 2***: Define the Population of Users to be Analyzed

The next step is to define the population of users for whom values will be estimated. In principle, this includes all users and potential users of the affected sites -- persons who use the sites without the changes and who might use the site with the changes. A geographic area encompassing all users is the market for the changes under consideration.

One way to capture the market in a day trip analysis is to define the population as all individuals residing within a days drive of the affected sites. In practice one approximates this market using boundaries that are convenient for sampling. The most common market definition is the set of residents in one or more states within a days drive of the sites. In some cases analysts will focus on a particular state. This may be called for by the decision makers funding the study or simply due to limited resources. As with the single site model, one also needs to consider what types of recreation to include and whether or not to aggregate recreation uses.

Here are some examples of day trip market definitions: Parsons and Kealy (1992) define the market for water based recreation on Wisconsin lakes as all Wisconsin residents. Adamowicz et. al. (1997) define the market for hunting in Alberta as Alberta residences holding provincial moose licenses. Bockstael, Hanemann, and Kling (1987) define the market for Boston beaches as the Boston metropolitan area. See step 4 for a discussion of defining the market before sampling begins.

3.2.3 ***Step 3***: Define the Choice Set

The next step is to define the choice set *S* in equation (22). This includes defining and determining which sites belong. In principle, one wants to include all the sites with impacts identified in step 1 plus all other sites that may serve as substitutes for these sites for the population of users defined in step 2. In practice, one always approximates this set. Again, political boundaries play a role in constructing the choice set and defining the sites.

It is easiest to see how analysts go about defining sites and determining what to include in the choice set by giving some examples. In Parsons and Kealy (1992), we analyze lake recreation in Wisconsin. Individual lakes in the state over a certain size are defined as the sites. There are over 1000 such lakes. If a lake is within 150 miles of a persons home, it is included in his or her choice set. This is based on the fact that few people traveled over 150 miles to make a day trip in the data.

Andrews (1996) analyzes trout fishing in eastern Pennsylvania. He defines sites as the management units used by the state fish and wildlife service. These are stream segments and lakes known to have trout. Each is different enough in character to be managed separately and data are organized by these units by the state. There are over 2000 sites. The choice set includes any site within 185 miles of a person's home. The longest day trip in the data set is 183 miles.

Parsons, Jakus, and Tomasi (1999) study lake recreation at the reservoirs operated by the Tennessee Valley Authority in the southeastern United States. Each person's choice set is defined as 14 specific reservoirs in the TVA system. Shaw and Jakus (1996) study rock climbing in northeastern United States. Their sites are defined as the four major climbing areas in the region. Each person has all four sites in his or her choice set. One major climbing area is excluded on the grounds that it is distinctly different from these four. Morey, Watson, and Rowe (1993) study salmon fishing in Maine and Canada. Nine rivers in the region known for salmon fishing are included each person's choice set.

McConnell and Strand (1994) study marine recreational fishing on the east coast of the United States and define coastal counties as the sites. Again, if the site is within 150 miles of the person's home it is included in the choice set. Hausman, Leonard, and McFadden (1995) study recreational fishing in Alaska and defined sites as one of 17 large fishing regions in the state. These regions serve as each person's choice set.

As you see, choice sets vary in size from 3 or 4 sites to over a thousand sites. Site definitions can vary from highly aggregated regions or counties to rather narrowly defined units such as beaches or small segments of a river. In choosing the number of sites and the degree of aggregation of sites it is best to error on the side of too many sites and narrow site definitions. Modern econometric software packages can handle a large number of sites and the feasibility of randomly drawing sites in estimation to approximate larger choice sets is also an option (see Parsons and Kealy (1992)).

If you must aggregate sites into regions or counties, the general rule is to group similar sites together. The sites should be similar in all characteristics including trip cost. The less similar the sites, the more bias one is likely to encounter (see Parsons and Needelman (1992)). For approaches that mix aggregated and non-aggregated sites together see Lupi and Feather (1998) and Parsons, Plantinga, and Boyle (2000).

Finally, there has some concern about using researcher versus individual defined choice set in RUM models. The definitions I have described above are all researcher defined. Peters et. al. (1995) and Hicks and Strand (2000) have argued that people cannot perceive this many sites in making a choice and suggest an approach using choice sets determined by people in the survey. Individuals identify sites they consider in site choice and this makes up the choice set. For a counter argument to this approach see Parsons, Massey, and Tomasi (1999). Researcher defined choice sets still dominate the RUM literature.

3.2.4 *Step 4:* Develop a Sampling Strategy

Next, a sampling strategy is developed. Virtually all published studies use some form of off-site random sampling. The users identified in step 2 are contacted by phone or mail and asked to report trips taken to the sites defined in the choice set in step 3. See Champ's chapter 3 or Dilman (1999) for random sampling with mail and phone surveys. In some cases a targeted population of users, such as people with boating licenses, will be contacted randomly. In most applications participants and non-participants will be sampled. This enables the analyst to model the decision of whether or not to participate in the recreation activity.

Like the single site model, sampling for a RUM model runs into the issue of low participation rates from the general population. The fraction of people participating in most forms or outdoor recreation is low enough that random samples from the general population can yield a small sample of actual users. The conventional way of dealing with this problem is to stratify the sample -- sampling counties or communities nearer the sites more heavily. Participation rates are higher in the nearby communities. In some cases stratification is the only way to get nearby users in the sample if local communities are relatively sparsely populated. This not only increases the number of participants but also allows for a more even geographic distribution across the sample of users which can sharpen parameter estimates by ensuring variation of trip cost across the set of sites. As with the single site model, a target population of license holders circumvents the low participation rate problem.

Unlike the single site model, on-site sampling is usually not an option. On-site sampling or what is known as 'choice-based sampling' is covered extensively in the RUM literature. To estimate parameters without bias, site weights are needed. The weights are based on the actual distribution of day trips by the population of users being studied and are used to adjust the count of trips to the different sites to reflect their true occurrence. I know of no RUM applications of the trip cost model that estimated using a choice based sample. See Train (1986, pp. 48-9), Ben-Akiva and Lerman (1985), or Laitila (1999) for more on working with on-site samples.

3.2.5 ***Step 5***: Specify the Model

Before data collection begins one specifies the elements that will be included in the model in equations (19) and (24).

Site utility in equation (19) includes trip cost (tc_i) and a vector of site characteristics (q_i) for to each site. In some cases it will include interactions with individual characteristics. Trip cost is the same as in the single site model, the sum of travel and time cost. The vector q_i includes characteristics that matter to people in making a site choice. This will vary depending on the type of recreation being studied. Fishing will include catch rate of fish, rock climbing will include difficulty of the climb, boating will include presence of boat ramps and so forth. It is difficult to make a generic list, but here are some common attributes one sees in a RUM model

- amenities
- size
- access
- environmental quality
- park (yes/no)
- opportunities on-site
- sport success
- remote location (yes/no)
- character of nearby area
- special features

Amenities are measures such as acreage of tree coverage. Size might be total acres in the site, length of beach, or number of trails. Access might be a variable indicating the presence of a boat ramp or one indicating that the site is only accessible by 4 wheel vehicle. Environmental quality includes measures of water quality or a variable for the presence of a fish consumption advisory. A variable designating whether or not it is a park is a common site characteristic. By opportunities on-site, I mean things like camping available or a museum is present. Sport success is catch rate of fish, bag rate for hunting, sightings for birding watchers, and so forth. The character of the nearby area might be indicating if a beach is developed or natural. And finally, special features are unusual characteristics that may make difference in site choice such as a dam on a river. Finally, if a site or group of sites is distinguished from others sites in the choice set in some more general way, it is common to see an alternative specific constant assigned to that site. An alternative specific constant is a dummy variable the site utility for one site or more sites.

The no-trip utility in equation (24) is a vector of individual characteristics that govern how often and whether or not a person makes a trip. These play the same role as the vector *z* in the single site model. Some common characteristics are age, occupation, and urban/rural residence – essentially the same list shown in single site step 4.

In some cases individual characteristics enter the site choice model as interaction terms. As explained in earlier, this happens when one believes that a certain characteristic affects people differently. For example, the presence a boat ramp may matter to people who trailer a boat but may not to people who keep a boat at a marina. Or, the difficulty of a rock climb may matter to expert climbers but not to novice climbers or at least the degree of its importance may be different. An interaction is entered into the model as shown in equation (25).

Specification always depends on available data. For this reason steps 5 and 6, model specification and collection of site characteristic data, work in parallel. The analyst has a certain specification in mind but amends it to accommodate available data. For some good of examples of differences in specification by recreation types see Sidderelis (1995) for a boating, Shaw and Jakus (1996) for rock climbing, Karou (1995) for fishing, Morey (1981) for skiing, Parsons and Massey (1999) for beach recreation, and Adamowicz (1997) for hunting.

3.2.6 *Step 6*: Gather Site Characteristic Data

The next step is to gather site characteristic data. As noted, step 5 works in tandem with step 6. The typical sources for characteristic data are the state and federal agencies responsible for managing the resource. Environmental protection, natural resource management, and fish and wildlife agencies are often good sources. These agencies may have data on indicators of environmental quality and physical measures such as size and elevation Furthermore, the agencies often have ready made site definitions. A fish and game agency, for example, may have management units for which data are gathered for their own purposes. In most cases these agencies are a good starting point for data collection.

Other sources of data are tourist bureaus, clubs and associations, universities, scientists, and newspapers. In some cases fieldwork where the analyst constructs the primary data him or herself or interviews knowledgeable people is necessary. Even in cases where variables are not constructed by field observation, such visits can confirm or amend existing site and variable definitions.

In some circumstances the site characteristic data are gathered in the survey. For example, individuals may be asked to rate the quality of hunting at each site visited in the current season. Using these responses the analyst constructs a hunting quality index for each site. One could do the same with catch rate of fish, view amenities and so on.

There are several problems with this approach. First, the data for each site are usually confined to those who have visited the site in the current season.[4] This is likely to bias quality measures upward at each site. People who visit a site are those who find the site desirable. Second, popular sites have more data than less popular sites. Many sites may a single or no visitor. This causes an asymmetry in the quality of variable measurement across sites. Third, perceived variables are more difficult to map into actual policy changes.

An alternative method for working with survey data is to preform an auxiliary regression using the reported measure for the characteristic (such as number of fish caught) as a dependent variable and observable site characteristics such as lake size, depth, elevation, and presence of regulations as explanatory variables. The unit of observation is a site and the number of observations is the number of respondents times the number of different sites visited by each. The fitted regression is then used to predict catch at each site. For some examples see McConnell, Strand, and Blake-Hedges (1995).

3.2.7 ***<u>Step 7</u>*: Decide on the Treatment of Multiple Purpose Trips**

In a day trip RUM model multiple purpose trips are handled in much the same way as in the single site model. Either all trips are assumed to be single purpose or multiple purpose trips are identified in the survey and dropped from the analysis. Again, this is a larger issue when overnight trips are being analyzed.

Another way of dealing with multiple purpose trips is to define site characteristics in such a way that they account for the potential of other activities while visiting the site. For example, a beach characteristic in a RUM model might be a dummy variable for the presence of shopping nearby. This method at least recognizes that an experience may be broader than recreation at the site alone.

It also possible to imagine a variant of Mendelsohn, et. al.'s (1992) approach for the RUM model. Alternatives would be defined as portfolios of sites. Each site would be represented by a dummy variable in the RUM model (like a site characteristic) and trip costs would be measured as the cost visiting the entire portfolio. People would choose portfolios. Welfare losses could be computed for losses of one or more sites from the portfolios.

3.2.8 ***<u>Step 8</u>*: Design and Implement Survey**

Next is the design and implementation of the survey. It is organized in the same four parts as a single site survey

- Introductory material
- Trip count questions
- Last trip questions
- Demographic/Household Characteristic questions

The introductory material, last trip questions, and demographic questions are essentially the same as in a single site survey. Unlike the single site survey, one gathers data on trips to every site in the choice set. The responses are used to form a data set like that in Table 3. RUM surveys face the same recall and trips categorization issues faced with single site models. See single site step 6 for a discussion. Recall is perhaps a larger problem in a RUM survey, because we are concerned about many sites.

There are essentially three approaches for counting trips to the sites. One provides the respondent a preestablished list of sites. The respondent reviews the list and records the trips over the relevant time period. This approach is easiest for data handling. The sites are defined and respondents' count fits neatly into the definition. For sites off the list, there is usually a catch-all category such as 'some other site'. This approach is also somewhat less burdensome. The site names serve as a reminder to people and recalling exact site names is not necessary.

The second approach is open-ended. People are asked to list all the sites they have visited over the relevant time period along with a count of trips to each. Once gathered one must rework the data such that site definitions are consistent across respondents. This can be time consuming. Sites often have more then one name, people using nearby town names, and sites with similar names can lead to confusion. When the number of potential sites runs into the hundreds or thousands, this approach may be hard to avoid. Even in these cases an insert, separate from the survey, listing (and numbering) the sites is worth considering.

One useful innovation to use in conjunction with the open-ended approach is a map-and-sticker. Respondents are asked to place a sticker on a map for each site visited. The sticker is numbered to correspond to a table in which the respondent records the number of trips taken to the site. This eases the final assembly of the data and avoids any confusion about the name and location of the site. It also has the advantage of aiding in finer site definitions depending on individuals' use of sites. For example if the there is heavy visitation to a site which seems divided by trips to the northern and southern portion of the site, one might opt for dividing the site along these lines.

Mail surveys or mail-phone surveys favor the approach where a list of sites is provided to the respondents. Phone surveys call for an open-ended approach unless the number of sites is less than a dozen or so. However, one can approximate the preestablished list approach by being interactive. The caller has a list of sites. During the survey respondents are asked to count trips region by region. For example, “Did you visit any beaches on the Outer Banks?” If so, which beaches? When the respondent gives a beach, the caller verifies that the beach fits the list and proceeds. If the respondent has trouble recalling the name, the caller can ask about the general area and provides some clues till the correct beach is identified. If the respondent gives and unknown name, it is even possible to find the closest beach. This can be an extremely valuable way to construct the trip data file and can eliminate the post survey site definitions.

A third approach is to work with last trip data alone. In this case one gathers information only on the last site visited – what site was it, how many trips were taken to that site over the season, and how many were taken to all sites in total? A basic RUM model can be estimated with these data. The survey is less complex and recall is a smaller issue. This comes at the expense of far fewer trips and hence less data for the final analysis.

3.2.9 ***Step 9***: Measure Trip Cost

This step is essentially the same a single site step 7 except that trip cost is measured to every site in a persons choice set, not just the visited sites. This invariably calls for a software package to compute travel time and distances between respondents hometowns and the sites. PCMiler is a popular one. Respondents’ hometowns and the site names must defined so they map into the chosen software package. Small communities are sometimes excluded. Spellings must be exact. Some packages require state abbreviations as part of the title. Most software packages will compute time and distances between zip codes as well as an approximation.

For more discussion of the issues surrounding the measurement of trip cost such as the value of time and using individual versus household measures see single site step 7.

3.2.10 ***Step 10***: Estimate Model

Next, the analyst seeks to estimate the parameters α and β in equations (19) and (24) specified in step 5. This is done in a probabilistic framework. An expression for the probability of visiting a site is formed. The probability that a person visits site k is

$$(37) \quad pr(\beta_{tc}tc_k + \beta_q q_k + e_k \geq \beta_{tc}tc_i + \beta_q q_i + e_i \text{ for all } i \in S \text{ and } \geq \alpha_0 + \alpha_1 z + e_0) \quad .$$

Under different assumptions about the distribution of the error terms e_i, different forms for equation (37) are derived. The simplest is the multinomial logit (ML) for which the probability that an individual visits site k is

$$(38) \quad pr(k) = \frac{\exp(\beta_{tc}tc_k + \beta_q q_k)}{\exp(\alpha_0 + \alpha_1 z) + \sum_{i=1}^{S} \exp(\beta_{tc}tc_i + \beta_q q_i)}$$

Notice that the probability of visiting site k depends the characteristics of site k (in the numerator and denominator of equation (38)) and on the characteristics of other sites (in the denominator). Also notice that each person has a probability for each site and no trip. The no-trip probability has $\exp(\alpha_0 + \alpha_1 z)$ in its numerator, and its denominator is the same as the denominator in equation (38). This ML probability assumes that error terms in equation (19) and (24) are independent and identically distributed with Weibull distribution. See Greene (1997, p. 913).

The parameters are estimated by maximum likelihood using the probabilities in equation (38). If one has data on N persons each visiting one of the S sites or taking no trip for a given choice occasion, the likelihood of observing that pattern visits in the data using the ML probability model is

$$(39) \quad L = \prod_{n=1}^{N} \prod_{i=0}^{S} pr(i)^{r_{in}}$$

where $r_{in}= 1$ if individual n visited site i and $= 0$ otherwise. The probability $pr(i)$ is the logit form from equation (38). The parameters are estimated by choosing values of α and β to maximize L. These are the maximum likelihood estimates for the model. This is the form of the model that is most likely to generate the patterns of visits actually observed in the data.

In circumstances where individuals are observed taking multiple trips to sites over the season the same likelihood function is used with r_{in} now equal to the number of trips taken to site i by individual n.[5] Many software packages are available for estimating logit models. LIMDEP, GAUSS, and TSP are common.

This ML model is criticized for a restriction know as the independence of irrelevant alternatives (iia). This restriction implies that the relative odds of choosing between any two alternatives is independent of changes that may occur in other alternatives in the choice set. For example, the iia property implies the following. If an improvement in a characteristic at site k causes a 10% increase in the probability of visiting that site, then the percent change in the probability of visiting each of the remaining sites in the choice set must decrease by 10% (a proportional reduction in all other probabilities). If some sites are better substitutes for the site experiencing the improvement, this result is unrealistic. One expects good substitutes to have a larger percentage reduction in probability than poorer substitutes.

There are essentially two methods for relaxing iia: the nested logit (NL) and the mixed logit (MX). Both introduce correlation among the site and no-trip utility error terms which allows for more general patterns of substitution in the model. The NL model nests alternatives in groups believed to be close substitutes. For example, in an application of beach use, where sites include ocean and bay beaches, nesting the ocean and bay beaches into separate groups may make sense. This has the effect of allowing beaches within each nest to be better substitutes for one another. The iia assumption still holds for sites within a nest, but is relaxed for sites from different nests. In this way the model allows for a richer pattern of substitution. The error terms within each nest are viewed as having shared unobserved characteristics which leads the correlation of the error terms within a nest. (eg. all ocean beaches may have large waves and similar view amenities). A common application of the NL model is to nest sites and no-trip in separate nests. The reasoning here is that sites are more likely to serve as a better substitutes for one another than no-trip. Or, put differently, sites are likely to share unobserved characteristics. See Morey (1999) for more on NL models in recreation demand, Greene (1997, p.121) for an introductory discussion, and Morey, Rowe and Watson (1993) or Parsons and Hauber (1998) for applications.

The second method for relaxing the iia assumption is a mixed (MX) or random parameters logit. MX models are estimated used simulated probability techniques and are fast becoming the standard for estimating trip cost RUM models. Like the nested model, the mixed model is a generalization of a basic multinomial logit model that allows the parameters α and β to be random. The variation in each parameter is interpreted as a component of the error term in the site utilities which leads to correlation among the utilities and a more general pattern of substitution. See Train (1999) for an example. NL and MX models may be estimated using GAUSS or LIMDEP.

3.2.11 ***Step 11***: Calculate Access and Quality Change Values

In the final step, access or quality change values are computed. Value may be reported as

- a mean per choice occasion value per person,
- a mean seasonal value per person,
- a total seasonal value for the population,
- a per trip value per person, and/or
- a total discounted present value of the site.

The parameters estimated in the previous step are used to calculate welfare changes using equations (35) and (36). The form of the expected maximum utility depends on the assumed distribution for the error terms in the model. With the Weibull distribution in the ML model, the expect maximum utility of a choice occasion is

$$\text{(40)} \quad \text{eu} = \ln\left\{\exp(\hat{\alpha}_0 + \hat{\alpha}_1 z) + \sum_{i=1}^{S} \exp(\hat{\beta}_{tc} tc_i + \hat{\beta}_q q_i)\right\}$$

This is the log of the denominator in equation (38). It is a preference-weighted utility index for a choice occasion in the sense that all alternatives are included and the higher the alternative's utility the larger its roll in the expression. Again, its form follows directly from the assumed distribution for the error terms. With equation (40) then, one can value per choice occasion site access and quality changes for one or more sites using equations (35) and (36).

Per choice occasion site loss for person *n* is

$$\text{(41)} \quad \hat{\text{s}}_{\text{n}} = \ln\left\{\exp(\hat{\alpha}_0 + \hat{\alpha}_1 z_n) + \sum_{i=6}^{S} \exp(\hat{\beta}_{tc} tc_{in} + \hat{\beta}_q q_{in})\right\} - \ln\left\{\exp(\hat{\alpha}_0 + \hat{\alpha}_1 z_n) + \sum_{i=1}^{S} \exp(\hat{\beta}_{tc} tc_{in} + \hat{\beta}_q q_{in})\right\}$$

where the first 5 sites are lost. A ˆ denotes an estimated value using the estimation results and the subscript *n* on an explanatory variable that denotes the value of that variable for individual *n*.

For a quality change the per choice occasion value is

$$\text{(42)} \quad \hat{s}_n = \ln\left\{\exp(\hat{\alpha}_0 + \hat{\alpha}_1 z_n) + \sum_{i=1}^{S} \exp(\hat{\beta}_{tc} tc_{in} + \hat{\beta}_q q_{in}^*)\right\} - \ln\left\{\exp(\hat{\alpha}_0 + \hat{\alpha}_1 z_n) + \sum_{i=1}^{S} \exp(\hat{\beta}_{tc} tc_{in} + \hat{\beta}_q q_{in})\right\}$$

where q^*_i is a vector indicating a quality change at some or all of the *S* sites. Again, as the distribution of the error term changes, the form of the expected maximum utilities in equation (42) change. If the sample is randomly drawn a simple mean of $\hat{s}_n$ is presented for the per choice occasion value. If the sample is stratified, the sample mean is adjusted to represent the population.

The seasonal value for each individual is the total number of choice occasions times his or her per choice occasion value. So, the mean seasonal per person value is

$$(43) \quad \overline{S} = T \cdot \overline{s}$$

where $\overline{s}$ is the sample mean per choice occasion value (adjusted for stratification if necessary) and *T* is the total number of choice occasions in the seasons. In a day trip model, *T* is the number of days in the season. The aggregate seasonal value over the population is

$$(44) \quad AS = \overline{S} \cdot POP$$

where *POP* is population of users and potential users. This might be all residents within driving distance of the site, all people owning a fishing license, or whatever defined the population in sampling. *AS* is occasionally converted to a discounted present value assuming a constant following of recreation services from the site or sites. As shown with the single-site model in section 3.2.9 this is

$$(45) \quad \mathrm{PV} = AS / i$$

where *i* is rate of discount. Equtions (43) through (45) work for site access and quality changes alike.

Last, it is not unusual to see per trip access or quality change values presented is a RUM analysis. A per trip per person value is

$$(46) \quad \hat{t} = AS / trips$$

where *trips* is the total number of day trips by the relevant population. One can use an external estimate of the total number of day trips taken to the site over the relevant season or estimate the number of trips using the RUM model. In either case the per trip value applies to the same scenario consider for site access or quality change above.

An alternative approach for estimating per trip values arises when one has estimated a model that excludes no-trip form the choice set. In this case, the one estimates per trip, instead of per choice occasion values, in the basic RUM model and per trip values flow naturally from the results. See Parsons and Massey (2002) for an example. Although perhaps convenient, one must keep in mind that these results come from a restricted model that disallows no-trip as and alternative.

3.3 A RUM Application

Matt Massey and I have estimate several RUM models of beach use in the Mid-Atlantic region of the United States using a 1997 beach use survey. See Parsons and Massey (1999) or Parsons and Massey (2002). Here, I present another version of that model following the steps outlined in the previous section. The Mid-Atlantic beaches in our analysis include all of New Jersey, Delaware, and Maryland ocean beaches -- 62 beaches.

We identify two impacts for two analyses in our study (step 1). The first is the potential closure of beaches in the state of Delaware due to oil spills, water pollution, other environmental episode. These are analyzed as lost site access. The second is beach erosion or narrowing of beaches in the state. This is analyzed as a quality change. Our quality measure is beach width. Beach width data is provided by the states and there is sufficient variation across the beaches in this characteristic.

We define our market as all residents in the state of Delaware (step 2). We recognize that there were a large number of out of state users. However, budget limitations and our key interest in exploring some methodological issues in the model result in our using this narrow market definition. We also treat all uses of the beach as a single recreation type. We aggregated sunbathing, swimming, surf fishing and so on.

We define our choice set as all ocean beaches within a days drive of Delaware residents (step 3). This includes 62 beaches in four states. Beaches (our sites) were defined using the political boundaries of beach communities which is consistent with how people identify beaches in this area. For example, Ocean City MD, Rehoboth DE and Cape May NJ were all separate beaches in our analysis. Every person had all 62 sites in their choice set.

Delaware is a small state with 3 counties only. The most populated county is located in the north. The ocean beaches are along the southern most county. We randomly sampled an equal number of residents over the age of 16 from each of the three counties (step 4). We stratified in this way to avoid a population dominated by residents from the northern most county. We were not too concerned about low participation rates. Historic data convinced us that half or more of the population used the beaches in a typical year.

There are a number of things we thought would influence day trips to the beaches in this region (step 5). Among these were:

- trip cost
- natural vs developed beach
- private or limited access
- presence of boardwalk
- availability of parking
- availability of bathhouses and other facilities
- beach width

These, more or less, were our targets as we set out to gather the data in the next step. For individual characteristic data we thought occupation, education, family composition, and flexibility in work would be important.

The site characteristic data were gathered from various sources: state departments' of natural resources (including interviews with experts in those agencies), field trips, interviews with scientist working on a data on the physical characteristics of the New Jersey beaches, tourist guides, maps, newpapers, and locate web sites (step 6).

We analyzed day trips only and assumed all trips were single purpose or at least that side trips were incidental and easily ignored without introducing error (step 7).

We used a random mail survey of 1000 Delaware residents in the Fall of 1997 (step 8). An initial mailing of the survey was followed by a reminder postcard one week later and a second mailing of the survey after three weeks. Our response rate was 55%. Individuals were asked to report day, short overnight, long overnight, extended stay, and side trips separately. The survey was eight pages long. Respondents were asked to complete a 2-page table of trips to 62 beaches. A map insert was provided to help identify beaches.

Trip cost was measured as the sum of travel expense, time expense, and beach fees (step 9). Travel cost was 35 cents times round trip distance plus tolls and parking fees. Many trips to New Jersey beaches are via toll roads and on some routes a ferry is used to cross the mouth of the Delaware Bay. After the shortest route was determined using PC Miler, the toll routes were identified and their cost computed. Many New Jersey beaches have fees. We used a per day fee which was published by beach for the 1997 season in a local newspaper. Time costs were estimated a wage proxy times round trip travel time. The wage proxy was annual household income divided by 2080. PC Miler was used to compute round travel times.

We estimated a nested logit model. The results are in Table 4. The site characteristics are variables appearing in site utility. The individual characteristics are variables appearing in no-trip utility.

As shown, the site characteristics that increase a site's day trip utility are boardwalk, amusements, good surfing, having a park, and good parking. The characteristics that decrease a site's day trip utility are private, being too wide or too narrow, and having high rises nearby. None of these are surprising. Length of the beach and park (state or federal) were insignificant and facilities had the 'wrong' sign. The alternative specific constant for Atlantic City and for New Jersey were positive and significant.

The inclusive value coefficients are parameters which characterize the degree of substitutability of among alternatives within a nest. To be consistent utility maximization, these coefficients should fall between 0 and 1. The closer the coefficient is to 0 the greater the degree of substitutability. A coefficient equal 1 is the same as not nesting. See Morey (1999) for an excellent discussion of NL models and interpreting inclusive values. Out results indicate a higher degree of substitution among sites within the same state. The coefficient on IV beaches suggests that the no-trip/sites nest is a mis-specification.

The individual characteristics that decrease no-trip utility and hence increase the probability of taking a beach trip are number of kids under 10 years old in the household, flexible time, owning a beach cottage in Delaware or New Jersey, student, working part time, or volunteer. No-trip utility increases with age, retired, and working at home.

For site access values, we considered the closure of each of the 62 beaches separately and the closure of groups of beaches (step 11). For the groups of beaches we considered the 6 northern most Delaware beaches and 8 southern most beaches. For the beach erosion scenarios we considered a narrowling of all developed beaches in Delaware to less than 75 feet wide. We report seasonal per person, seasonal aggregate (for Delaware residents), and discounted present values (for Delaware residents) in each case.

The mean seasonal per person value, equation (43), for the loss a single beach ranged from about \$5 for the northern most beaches in New Jersey to about \$135 for the most popular beaches in Delaware and Maryland. These were Rehoboth (DE), Cape Henlopen (DE), and Ocean City (MD). Recall that our sample considers only Delaware residents and the scenario assumes all other beaches remain open. These relative sizes make sense. The northern New Jersey beaches are far away and have many nearby substitute. The higher valued beaches are close to population centers and have many of the desirable characteristic noted in the results above.

These values translate into mean seasonal aggregate losses, using equuation (43), that range from \$2.9 million for the least valued northern New Jersey beaches to \$77.8 million for the highest valued Delaware beaches. The relevant population is all residents of Delaware over the age of 16 in 1997. That is consistent with the sample frame. It is important to keep in mind that this value excludes overnight trips, people from other states, and non-use values.

We are particularly interested the loss of groups of beaches. For example, if an oil spill should occur, it is likely that more than a single a beach would be lost. The northern beaches in Delaware are particularly vulnerable. Losing these simultaneously gives a mean seasonal per person loss of \$698 which translates into an aggregate seasonal loss of \$402 million. Losing the southern beaches gives a mean seasonal per person loss of \$554, an aggregate loss of \$319 million. Again, overnight trip, out of state residents, and non use value is excluded.

Loss of beach width due to erosion is a major issue on beaches in the Mid-Atlantic. So, we considered a scenario where all developed beaches with a width of 75 feet or greater are narrowed to less than 75 feet. If a beach is not developed it not likely to erode and so was excluded. To calculate an individual's expected maximum utility with erosion, all developed beaches have site utilities computed with *wide = 0* and *narrow = 1*. The mean seasonal per person loss for a narrowing of Delaware's beaches is about \$76. The corresponding aggregate loss is \$44 million.

3.4 Variations

There are a number of variations on the basic RUM model and an active research agenda advancing the technique (see Herriges and Kling (1999)). I will mention a couple variations on the model I consider particularly important.

First, there is the application of Kuhn-Tucker models to recreation demand. These models have features of both RUM models and demand systems. They can be used to value quality changes and access and, it has been shown recently that these models may be used in settings with a large number of sites (see Phaneuf et. al. (2000) and Phaneuf et al. (2002)). Kuhn-Tucker models are noted for their utility theoretic link between participation and site choice.

Second, there is increasing use of revealed and stated preference data in combination in the context of RUM models. This enables an analyst to consider behavior beyond the range of observable data and hence a much wider range of policy options (see Adamowicz (1994)).

Third, a time dimension is being introduced into RUM models. This allows trips to have interdependence over time and allows for time periods (eg., weekend versus weekday) to be treated differently. See Adamowicz (1994).

Fourth, the mixed logit (MX) mentioned in section 2.2.10 has been widely adopted as a means of introducing complex patterns of substitution and unobserved heterogeneity into the model. This model is almost certain to see even wider use. See Train (1999).

Fifth, RUM models are often extended to include more than site choice. They may incorporate choice target species for fishing, choice of boat or shore in fishing, and even type of recreation. See Parsons and Hauber (1998) or McConnell and Strand (1994).

Finally, the basic model presented here may be extended to overnight trips but care must be taken to account for multiple purpose trips and to calculate lodging cost accurately. Neither is easy. See Shaw and Ozog ( ) and Hoehn et. al. (19 96).

Table 4: RUM Model for Mid-Atlantic Beaches

Site Characteristics:		
Variable	*Definition*	*Parameter Estimate (t-stat)*
tc	*travel plus time cost*	*-.04 (63.9)*
Length	*Length of beach in miles*	*.13 (8.3)*
Boardwalk	*Boardwalk present = 1*	*.41 (6.3)*
Amusement	*Amusements nearby = 1*	*.48 (15.4)*
Private	*Private or limited access beach = 1*	*-.17 (6.3)*
Park	*State or federal park*	*.04 (0.6)*
Wide	*Wide beach (= 1 if > 200 feet)*	*-.33 (12.8)*
Narrow	*Narrow beach (= 1 if < 75 feet)*	*-.20 (5.4)*
AC	*Atlantic City = 1*	*.42 (7.2)*
Surf	*Good surfing = 1*	*.40 (15.5)*
HighRise	*High rises present on beach = 1*	*-.30 (9.4)*
ParkWithin	*Park located within the beach = 1*	*.25 (5.1)*
Facilities	*Bathhouse, restroom facilities present = 1*	*-.05 (1.1)*
Parking	*Parking available at beach = 1*	*.13 (1.8)*
New Jersey	*New Jersey beach = 1*	*.51 (33.9)*
IV(NJ)	*Inclusive Value on New Jersey Beaches*	*.49 (36.9)*
IV(DE)	*Inclusive Value on Delaware Beaches*	*.99 (38.7)*
IV (Beaches)	*Inclusive Value on Beaches*	*2.06 (11.0)*
Individual Characteristics:		

Variable	Definition	Parameter Estimate (t-stat)
Variable	Definition	Parameter Estimate (t-stat)
Constant		.25 (5.3)
ln(age)	Log of age	.20 (7.0)
Kidsu10	Number of kids under 10 in household	-.26 (9.4)
Flextime	Flexible time available in work schedule	-.14 (3.4)
Cottage (DE)	Own beach property in Delaware = 1	-1.3 (25.5)
Cottage (NJ)	Own beach property in New Jersey = 1	-.80 (16.4)
Retired	Retired = 1	.53 (10.5)
Student	Student = 1	-.90 (19.5)
Parttime	Work part time =1	-.56 (13.3)
Workhome	Work at home = 1	.94 (12.0)
Volunteer	Work as a volunteer = 1	-.16 (2.6)
Sample Size	565	
Mean Log- Like	-94.05	

4. CONCLUSION

The traditional single site model and contemporary RUM model are the two most widely used travel cost models in recreation demand. The RUM model is the modern workhorse. It is able to account to a broad array of substitutes. It is possible to value changes in site access as well changes in quality at one or more sites. It tells a convincing and defensible story. And, software is readily available for estimation. Most advances in travel cost modeling are taking place in the context of RUM models. The single site model requires less data and is easier to apply. In circumstances where one is interested in access value at only one site and the number of substitute sites is not large, the single site model is often used and is defensible.

NOTES

1.This overlooks other multiple site approaches such as conventional demand systems and the hedonic travel cost model. Neither have been as popular as the RUM model. See Ward and Beal (2000) for a discussion of conventional demand systems and Brown and Mendelsohn (1984) for the hedonic travel cost. See Bockstael, McConnell, and Strand (1991) for a critique of the hedonic travel cost model.

2.For this reason, their probabilities corresponded to Greene's truncated model (1997, p. 937) instead of my equation (9).

3.Sometimes site utilities are written as $v_i = \beta_{tc}(y-tc_i) + \beta_q q_i$, and no-trip utility is written as $v_{0i} = \beta_{tc}y + \alpha$ where y is a measure of per trip income. In this way trip cost is seen quite explicitly as reducing income available for other uses and hence lowering welfare. The coefficient β_{tc} (now positive) is also readily interpreted as a marginal utility of income. In estimation y is constant across sites and no-trip utility, provides no explanatory power for choice of alternative, and hence drops out in estimation.

REFERENCES

Adamowicz, Wiktor. L. "Habit Formation and Variety Seeking in a Discrete Choice Model of Recreation Demand," Journal of Agricultural and Resource Economics 19(1): 19-31. 1994.

Adamowicz, W., J. Louviere, and M. Williams. "Combining Revealed and Stated Preference Methods for Valuing Environmental Amenities," Journal of Environmental Economics and Management 26:271-92. 1994.

Adamowicz, Wiktor L. Joffre Swait, Peter Boxall, and Michael Williams. "Perception Versus Objective Measures of Environmental Quality in Combined Revealed and Stated Preference Models of Environmental Evaluation," Journal of Environmental Economics and Management 32: 65-84. 1997.

Andrews, Tom. 1996.

Ben-Akiva, M. and S. Lerman. *Discrete Choice Analysis.* Cambridge MA: MIT Press, 1985.

Bockstael, Nancy E. "Travel Cost Models," Handbook of Environmental Economics. Blackwell Publishers: Cambridge, MA. 1995.

Bockstael, Nancy E. and Ivar E. Strand. "The Effect of Common Sources of Regression Error on Benefit Estimates," Land Economics 63(1): 11-20. 1987.

Bockstael, Nancy, W. Michael Hanemann, and Catherine L. Kling. "Estimating the Value of Water Quality Improvements in a Recreational Demand Framework," Water Resources Research 23(5): 951-960. 1987.

Bockstael, Nancy E., W. Michael Hanemann, and Ivar E. Strand, Jr. Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models. Report to the U.S. Environmental Protection Agency. College Park, MD: University of Maryland. 1984.

Bockstael, Nancy E., Kenneth E. McConnell, and Ivar E. Strand. Benefits from Improvements in Chesapeake Bay Water Quality. Report to the U.S. Environmental Protection Agency. College Park, MD: University of Maryland. 1988.

Bockstael, N.E., K.E. McConnell, and I.E. Strand. "A Random Utility Model for Sportfishing: Some Preliminary Results for Florida," Marine Resource Economics 6: 245-260. 1989.

Bockstael, Nancy E., Kenneth E. McConnell, and Ivar E. Strand, Jr. "Recreation" Measuring the Demand for Environmental Quality, North Holland: New York: 1991.

Bockstael, Nancy E., Ivar E. Strand, and W. Michael Hanemann. "Time and the Recreational Demand Model," American Journal of Agricultural Economics 293-302. May 1987.

Boyle, Kevin. Marine Resource Economics.

Brown, G. and R. Mendelsohn. "The Hedonic Travel Cost Method," Review of Economics and Statistics 66: 427-433. 1984.

Chen, H.Z. and S.R. Cosslett. "Environmental Quality Preference and Benefit Estimation in Multinomial Probit Models: A Simulation Approach," American Journal of Agricultural Economics 78(3). 1998.

Cicchetti, Charles J., Anthony C. Fisher, and V. Kerry Smith. "An Econometric Evaluation of a Generalized Consumer Surplus Measure: The Mineral King Controversy," Econometrica 44(6): 1259-1276. 1976.

Creel, Michael D. and John B. Loomis. "Theoretical and Empirical Advantages of Truncated Count Data Estimators for Analysis of Deer Hunting in California," Amer. J. Agr. Econ. 434-441. May 1990.

Feather, Peter and W. Douglas Shaw. "Estimating the Cost of Leisure Time for Recreation Demand Models," Journal of Environmental Economics and Management 38:49-65.

1999.

Freeman, A. Myrick III. The Measurement of Environmental and Resource Values: Theory and Methods. Resources for the Future: Washington, D.C. 1993.

Greene, William H. Econometric Analysis (3rd Ed.). Prentice Hall: New Jersey. 1997.

Haab, Timothy C. and Kenneth E. McConnell. "Count Data Models and the Problem of Zeros in Recreation Demand Analysis," American Journal of Agricultural Economics 78: 89-102. 1996.

Haab, Timothy C. and Kenneth E. McConnell. *Valuing Environmental and Natural Resources.* Edward Elgar: Cheltenham, UK 2002.

Hanemann, W. Michael. "Welfare Analysis with Discrete Choice Models." Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice, Edward Elgar: Cheltenham, UK. 1999.

Hauber, Albert H. IV and G.R. Parsons. "The Effect of nesting Structure on Welfare Estimation in Random Utility Models: An Application to the Demand for Recreational Fishing," American Journal of Agricultural Economics 82: 501-514. 2000.

Hausman, Jerry A., Gregory K. Leonard, and Daniel McFadden. "A Utility-Consistent, Combined Discrete Choice and Count Data Model: Assessing Recreational Use Losses Due to Natural Resource Damage," Journal of Public Economics 56:1-30. 1995.

Hellerstein, Daniel. "Can We Count on Count Models?" Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice, Edward Elgar: Cheltenham, UK. 1999.

Hellerstein, D. and R. Mendelsohn. "A Theoretical Foundation for Count Data Models," Amer. J. Agr. Econ. 75: 604-11. 1993.

Herriges, J.A. and C.L. Kling, (eds.). Valuing Recreation and the Environment: Revealed Preference Methods in Theory and Practice. Edward Elgar: Aldershot U.K. 1999.

Herriges, Joseph A., Catherine L. Kling, and Daniel J. Phaneuf. "Corner Solution Models of Recreation Demand: A Comparison of Competing Frameworks," Chapter 6 in Herriges and Kling (1999). Edward Elgar: Cheltenham, UK. 1999.

Hicks, R.L. and I. Strand. "The Extent of Information: Its Relevance for Random Utility Models," Land Economics 76 (2). 2000.

Hoehn, J.P., Theodore Tomasi, Frank Lupi, and Heng Z. Chen. "An Economic Model for Valuing Recreational Angling Resources in Michigan," Michigan State University, Report to the Michigan Department of Environmental Quality. 1996.

Kolstad, Charles D. Environmental Economics. Oxford University Press: New York, 2000.

Krutilla and Smith (1975)

Laitila, Thomas. "Estimation of Combined Site-Choice and Trip-Frequency Models of Recreational Demand Using Choice-Based and On site Samples," Economic Letters 64: 17-23. 1999.

Layman, R. Craig, John R. Boyce, and Keith R. Criddle. "Economic Valuation of the Chinook Salmon Sport Fishery of the Gulkana River, Alaska, under Current and Alternate Management Plans," Land Economics 72(1): 113-28. 1996.

Lupi, F. and P.M. Feather, "Using Partial Aggregation to Reduce Bias in Random Utility Travel Cost Models," Water Resources Research 34(12): 3595-603. 1998.

McConnell, K.E. "The Damages to Recreational Activities from PCB's in New Bedford Harbor." Prepared for the Ocean Assessment Division, National Oceanic and Atmospheric Administration, December 1986.

McConnell, K.E. "On site Time in the Demand for Recreation," American Journal of Agricultural Economics 74: 918-25. 1992.

McConnell, K E., I. Strand, and L. Blake-Hedges. "Random Utility Models of Recreational Fishing: Catching Fish Using a Poisson Process," Marine Resource Economics 10: 247-261. 1995.

McConnell, K.E. and I.E. Strand. "The Economic Value of Mid and South Atlantic Sportfishing." University of Maryland, Report to the USEPA and NOAA. 1994.

McFadden, D. "Economic Choices," American Economic Review 91(3): 351-378. 2001.

Mendelsohn, R., J. Hof, G. Peterson, and R. Johnson. "Measuring Recreation Values with Multiple Destination Trips," American Journal of Agricultural Economics 74: 926-33. 1992.

Morey, Edward. "The Demand for Site-Specific Recreational Activities: A Characteristics Approach," Journal of Environmental Economics and Management 8:345-371. 1981.

Morey, Edward. "Two Rums Uncloaked: Nested-Logit Models of Site Choice and Nested-Logit Models of Participation and Site Choice," Chapter 3 in Herriges and Kling (1999). Edward Elgar: Aldershot U.K. 1999.

Morey, Edward, Robert D. Rowe, and Michael Watson. "A Repeated Nested-Logit Model of Atlantic Salmon Fishing," American Journal of Agricultural Economics 75: 578-592. 1993.

Ozuna, T. and I.A. Gomez, "Estimating a System of Recreation Demand Function Using a Semmingly Unrelated Poisson Regression Approach," Review of Economics and Statistics 76: 356-60. 1994.

Parsons, George R. "A Note on Choice of Residential Location in Travel Cost Demand Models," Land Economics 67(3): 360-364. 1991.

Parsons, George R. and Mary Jo Kealy. "Randomly Drawn Opportunity Sets in a Random

Utility Model of Lake Recreation," Land Economics 68(1): 93-106. 1992.

Parsons, George R. and Michael Needelman. "Site Aggregation in a Random Utility Model of Recreation," Land Economics 68(4): 418-433. 1992.

Parsons, George R. and Aaron J. Wilson. "Incidental and Joint Consumption in Recreation Demand," Agricultural and Resource Economics Review. 1997.

Parsons, George R. and A. Brett Hauber. "Spatial Boundaries and Choice Set Definition in a Random Utility Model of Recreation Demand," Land Economics 74(1): 32-48. 1998.

Parsons, George R., Paul M. Jakus, and Ted Tomasi. "A Comparison of Welfare Estimates from Four Models for Linking Seasonal Recreational Trips to Multinomial Logit Models of Site Choice," Journal of Environmental Economics and Management 38: 143-157. 1999.

Parsons, George R., D. Matthew Massey, and Ted Tomasi. "Familiar and Favorite Sites in a Random Utility Model of Beach Recreation," Marine Resource Economics 14: 299-315. 1999.

Parsons, George R., Andrew J. Plantinga, and Kevin J. Boyle. "Narrow Choice Sets in a Random Utility Model of Recreation Demand," Land Economics 76(1): 86-99. 2000.

Parsons, George R. and D. Matthew Massey. "A RUM Model of Beach Recreation," in *The New Economics of Outdoor Recreation,* ed. Nick Hanely and Douglass Shaw, forthcoming 2002.

Peters, T., W.L. Adamowica, and Peter C. Boxall. "Influence of Choice Set Considerations in Modeling the Benefits from Improved Water Quality," Water Resources Research 31(7): 1781-1787. 1995.

Phaneuf, Daniel J., C. L. Kling, and J. Herriges . "Estimation and Welfare Calculation in a Generalized Corner Solution Model With an Application to Recreation Demand," Review of Economics and Statistics 82: 83-92. 2000.

Phaneuf, Daniel J. and V. Kerry Smith. "Recreation Demand Models" Manuscript, North Carolina State University, 2002.

Randall, Alan. "A Difficulty with the Travel Cost Method," Land Economics 70(1): 88-96. 1994.

Shaw, D. "'On site Samples' Regression Problems of Nonnegative Integers, Truncation, and Enodgenous Stratification," Journal of Econometrics 37: 211-223. 1988.

Shaw, W. Douglass and Paul Jakus. "" 1996.

Shaw and Ozog, 1999.

Shonkwiler, J.S. "Recreation Demand Systems for Multiple Site Count Data Travel Cost Models," Chapter 9 in Herriges & Kling (1999). Edward Elgar: Cheltenham, U.K.

1999.

Shonkwiler, J.S. and W. D. Shaw. "Hurdle Count-Data Models in recreation Demand Analysis," Journal of Agricultural and Resource Economics 21(2): 210-219. 1996.

Siderelis, Christos, Gene Brothers, and Phil Rea. "A Boating Choice Model for the Valuation of Lake Access," Journal of Leisure Research 27(3): 264-282. 1995.

Siderelis, Christos and Roger Moore. "Outdoor Recreation Net Benefits of Rail-Trails," Journal of Leisure Research 27 (4): 344-359. 1995.

Smith, V.K., W.H. Desvousges, and M.P. McGivney. "The Opportunity Cost of Travel Time in Recreation Demand Models," Land Econom. 59(3): 259-277. 1983.

Smith, V. Kerry and William H. Desvousges. "The Generalized Travel Cost Model and Water Quality Benefits: A Reconsideration," Southern Economics Journal 52(2): 371-381. 1985.

Smith, V. Kerry, William H. Desvousges, and Matthew P. McGivney. "Estimating Water Quality Benefits: An Econometric Analysis," Southern Economic Journal 50(2), 422-437. 1983.

Sohngen, Brent. "The Value of Day Trips to Lake Erie Beaches," Dept. of Agricultural, Environmental, and Development Economics, Ohio State University. 2000.

Train, Kenneth E. "Mixed Logit Models for Recreation Demand," Chapter 4 in Herriges and Kling (1999). Edward Elgar: Aldershot U.K. 1999.

Ward, Frank A and D. Beal. Valuing Nature with Travel Cost Models: A Manuel. Edward Elgar: Cheltenham, UK. 2000.