SOUTHERN CALIFORNIA BEACH VALUATION PROJECT

Using revealed preference models to estimate the effect of coastal water quality on beach choice in Southern California.

Prepared by

Michael Hanemann, Linwood Pendleton, Craig Mohn, James Hilger, Koichi Kurisawa, David Layton, Chris Busch, and Felipe Vasquez

Submitted to

U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration (NOAA),
U.S. Dept. of the Interior: Minerals Management Service,
CA Department of Fish and Game: Office of Spill Prevention and Response (OSPR)
CA State Water Resources Control Board, and
Santa Monica Bay Restoration Commission

September 17, 2004

Acknowledgments

The research team would like to acknowledge the Southern California Coastal Water Research Program and Heal the Bay of Santa Monica, California for their help in collecting and interpreting water quality data for Southern California. Additionally, we would like to thank Norman Meade, Bob Leeworthy, Guangyu Wang, David Chapman, Pierre Duvair, Adrian Griffith, James Lima, Eric English, Fred Piltz, Rodney Cluck, Gerald Horner, and Nancy Lee for their input into research design, analysis, and report preparation. Finally, numerous student research assistants have helped in the collection of beach attribute data including Sarah Kern, Miwa Tamanaha, D.G. Webster, Claudia Avendaño, and Stephanie Hada.

Table of Contents

INTRODUCTION

Wide sandy beaches, sunshine, volleyball nets, and myriad other factors help to attract as many as 238 million visits to California beaches each year (King and Symes 2003). Most visits are made by state residents who go to beaches along California's south and central coasts. The economic importance of these local beach trips is significant. A recent study by Pendleton (2003, 2004) estimates that local beach goers spend as much as $9.5 billion annually when they go to the beach. Furthermore, beaches contribute substantially to the economic well being of beach goers who enjoy the non-market benefits of outdoor recreation. In the same study, Pendleton estimates that the non-market values associated with beach going in California may be as high as $5.8 billion annually.

Despite the size of the beach economy, incidents of coastal marine pollution continue to diminish the economic potential of beach recreation in California. Chronic coastal water pollution due to bacteriological contamination at Huntington Beach during the summers of 1999 and 2000 led to steep declines in beach attendance, expenditures, which most likely also affected the non-market value of beach visits to Huntington Beach and the surrounding coast. Similarly, oil pollution has had spatially limited, but nevertheless dramatic effects on beach going in Southern California (see Chapman and Hanemann 2001).

The task of identifying and estimating the impacts of coastal water pollution is complicated by the variety and interconnectedness of factors that influence where and when people decide to go to the beach. Visitors differ by age, sex, physical ability, wealth, income, and outdoor interests. Visitors can participate in one or more of a variety of activities at the beach and the availability of these activities and the enjoyment derived from participating in different activities varies throughout the year. Seasons even influence the places that beach goers might otherwise visit if they did not go to the beach. Demographics, activity choices, and seasonality complicate the degree to which beach attributes, including coastal water quality, affect the decision to choose a given beach on a given day. In turn, the preferences people place on clean coastal water may vary tremendously across people or even over time and across different activities, even for the same person. Water quality is likely to be more important for a swimmer than a walker, even when the swimmer and walker are the same person on different occasions.

In this report, we use random utility models to estimate the influence that beach water quality has on people's choices of when and where to go to the beach. In doing so, we develop a model that will allow for the estimation of the economic value that people place on coastal water quality under a variety of scenarios including impairments and improvements in water quality. The goal is to determine the economic welfare impacts of water quality changes including beach closures and changes in beach water quality (as measured by Heal the Bay's Beach Water Quality Grading system – see http://www.healthebay.org/brc/gradingsystem.asp).

We explore the way that differences among users, activities, and seasons influence the value beach goers place on water quality and consequently the way in which these differences affect the economic and social distribution of impacts caused by changes in beach quality. We start with a basic, repeated logit model of beach choice in which user differentiation and seasonality are ignored. From this foundation, we build increasingly more complicated models of beach choice that allow us to model the spatial substitution possibilities that confront the beach goer in her choice of beach destinations. We end our analysis with a 3 tiered nested repeated logit model of beach choice that more accurately models the way in which preferences held by beach goers may vary over time and by activity. We draw conclusions about the methods and importance of activity choice and seasonality in models of recreational site choice. In subsequent reports, we will use this model to estimate the economic impact of a variety of scenarios including improvements and degradation in water quality and beach closures of varying duration.

BACKGROUND

As in the application of all discrete choice models, a number of issues must be addressed before the model can be developed and estimated. Issues include how to value time, how to estimate choices over time and space, and how to define choice sets. The literature is rich in its discussion of ways to handle the above topics. In this paper, we focus on two issues that are less well covered by the literature yet are especially important in estimating beach choice in a year round setting – seasonality and heterogeneity among users and their preferences.

Seasonality

Models of outdoor recreational site choice are complicated by the fact that choice behavior varies seasonally, presumably reflecting a seasonal variation in preferences, constraints, and/or substitutes. For example, beach goers may prefer wide sandy beaches in the summer and picturesque rocky coasts during the winter months. Summer months often offer times when schedules are more flexible than in winter months; days in the summer may be twice as long as in winter. Snowboarding may be a legitimate substitute for surfing, but only when there is snow. In extreme cases, beaches can disappear altogether in the winter, only to reappear in the summer as accreting currents deposit new sand on the coast.

Issues of seasonality are likely to be important determinants of outdoor recreational behavior whenever there are pronounced differences in seasonal climate. Despite the obvious importance of seasonality, models of recreational site choice rarely account for the influence of seasonal differentiation. In many cases, data used to estimate recreational choice models are either cross-sectional or collected over a very short period of time.[1] In many cases when data are collected over a period of time (whether as a cross-sectional time series or as a panel), the standard approach has been to treat observations as independent observations generated by a single data generating process (for alternative approaches see Provencher and Bishop 1997 and Swait et al. 2004). When a pooled model is estimated, preferences, choices, and constraints are assumed to be constant across the period modeled.

Mixed logit models including random parameter models (also known as random coefficients models, Train and McFadden 2000 and Train 1998, Breffle and Morey 2000, Morey and Rossman 2003), finite mixture models, and their variants (Boxall and Adamowicz 2002, Arcidiacono and Jones 2003) offer some help in handling preference heterogeneity and the more complicated error structures associated with panel data. Generally, these models allow preferences to vary over individuals. Further, to the

[1] Examples of recreational choice models estimated using cross-sectional data are too numerous to list here.

degree that individuals and choice occasions are differentiated in these models, preferences may also vary over time. Nevertheless, the standard application of these models has not accounted for the structural origins of preference heterogeneity over time. As a result, the application of these models to the prediction of seasonal welfare change is limited – site attribute changes of short duration cannot be valued if preferences for such attributes vary seasonally. Desvouges and Waters (1995) and Desvouges, Waters, and Train (1996) extend the general random parameters models to include seasonality by restricting the choice set of Montana anglers to account for seasonal closures; in Southern California, however, beaches are available and accessible year-round.

In this paper, we examine the econometric issues associated with seasonality by using a yearlong panel data of beach choice in Southern California. We use as our benchmark for comparison the standard pooled repeated choice model. From this baseline, we investigate how preferences change across seasons for beach visitors. We further decompose possible structural changes in seasonal preferences by differentiating among beach goers that engage in different types of activities that may vary from season to season. This decomposition is carried out in three stages. First we expand our basic model by adding variables that are used predominately or exclusively in particular activities. The inclusion of these activity specific variables serves to indicate whether what appears to be seasonal differences in preferences for water actually reflects differences in activity opportunities over seasons. In this model we can investigate how preferences across all attributes vary by wave.

Second, we take the activity choice of our respondent as given and focus on the person's choices conditional on the selection of a given activity. To represent preferences for site attributes as conditional on the choice of activities we create "activity variables" in which we interact a dummy variable indicating the beach goer participated in a specific type of beach activity (e.g. surfing or bicycling) with a dummy variable indicating that a beach has an attribute that may be more or less preferred by people participating in that activity (e.g. is a beach known to have a surf break or a bike/walking path). This approach allows us to model choice behavior for different activities in different seasons.

Finally, our third approach is to develop a three-nested multi-nomial logit model, in which we simultaneously model participation, activity choice, and seasonal beach choice. Because of the complicated nature of the three-nested model, we limit our examination of seasonality to its impact on preferences for water quality. By handling activity choice, beach choice, and seasonality simultaneously, we can investigate whether seasonal differences in preferences for water quality are driven by differences in activity choices at a beach or potentially by other factors that could vary seasonality (including non-beach substitutes).

Heterogeneous Preferences Among Users and Their Activities

Preferences for goods and their characteristics often vary across individuals. In particular, preferences over beaches are likely to be characterized by systematic heterogeneity. Which site a recreationist chooses to visit is a function of her preferences over the complete set of characteristics that describe the sites in her choice set. The presence of heterogeneity in preferences is of importance in the estimation of random utility models where it can result in bias (Train 2003). Biased attribute coefficient estimates lead to biased welfare measurements of changes in site attributes and hinder the proper aggregation of welfare measurements across individuals. These biases can adversely affect policy decisions and skew the welfare distribution of decisions regarding resource management. Additionally, resource managers may be interested in welfare changes between user groups or for a specific type of user due to changes in management policy.

Heterogeneous preferences are difficult to account for in behavioral choice models due to the formulation of the conditional logit (CL) model, historically the workhorse of random utility models. Within demand system models, the analyst can directly incorporate demographic or other individual characteristic data directly into the individual's utility function to address preference heterogeneity. However under the specification of the CL, individual characteristics drop out of the econometric choice model. The result is that individual characteristics are not directly identifiable in the choice model.

A simple solution to this problem is to interact specific individual variables, such as income, race, and family composition with various choice attributes (Adamowicz et al. 1997, Breffle and Morey 2000). This method is limited in practice due to the difficulty of knowing, *a priori,* what individual and choice variables should be used to construct a variable that accounts for heterogeneous preferences (Boxall and Adamowicz, 2002). Other related solutions to this problem include the fixed effects and random effects specification of the conditional logit model (McFadden, 1986). However, these methods are difficult to employ when the sample consists of a large number of different kinds of decision makers.

The current state of the art approach to address heterogeneity is the random parameter logit (RPL) model. This approach handles heterogeneity across preferences by allowing estimated coefficients to vary randomly across individuals according to a continuous probability distribution. Two possible shortcomings of this approach are that the RPL does not offer an explanation for the source of the heterogeneity and that it implicitly assumes that preferences vary continuously across economic agents. Breffle and Morey (2000) and Morey and Rossman (2003) begin to address these shortcomings by combining both classic preference heterogeneity and random parameter methods.

For many types of recreation, especially beach recreation, preferences for attributes are likely to be conditioned upon the choice of the recreational activity. In these cases, one way to model heterogeneity is to directly model the choice of activity in a second nest of the beach goer's decision model and then estimate separate preferences for attributes conditioned upon the activity chosen. If the nests are chosen properly, the nested model provides the analyst with information about possible sources of heterogeneity as opposed to solely being able to account for it. The ability to model the heterogeneity of the sample population may aid resource managers with welfare analysis and management policy.

THE DATA

The Southern California Beach Valuation Project panel dataset is unique in its scope and provides a wealth of information regarding beach and beach goer attributes. There are literally dozens of beach attribute variables that could enter the dataset. This abundance of data, which is often seen as a blessing, can also have its disadvantages. Whereas studies that lack this wealth of explanatory variables will often estimate models using all of the available variables, the variables used in the analysis for the Southern California Beach Projects must be carefully chosen from this set.

An Overview of the Data

We model beach-going behavior to fifty-three individual, mostly contiguous beaches in San Diego, Orange, Los Angeles, and Ventura County. We use several sources of data to model beach choice behavior: travel cost data, water quality data, beach attribute data, and geographic data. Because of the many issues surrounding the estimation of travel cost and travel time, we devote an entire section to this topic below.

Water Quality Data

Data regarding beach water quality are based on water quality information provided to the public by the not for profit organization, Heal the Bay. The Heal the Bay (HTB) water quality data consist of site-specific letter-grades (i.e. scores) for bacteriological water quality, measured at numerous data collection points in the study area. These collection points are mapped to the beach sites used in the economic model. Water quality grades were collected by Heal the Bay throughout the year, however the number of available observations varies both by beach and over time due to irregularities in sampling frequencies. Further, HTB data were collected for both wet periods (immediately after a rain) and dry periods; separate wet and dry HTB grades are made available to the public. In order to define a comprehensive and consistent measure of water quality, we calculate composite dry-weather grades for each beach in the study area based on annual averages across all corresponding HTB observations. There are not sufficient wet-weather grades to construct averages for all beaches in all waves.

Note that temporally varying attributes, like water quality, can be measured as point estimates, means over time, and even variance over time. In fact, the beach grades provided by Heal The Bay are themselves running geometric means of water quality measures over four week periods. In our analysis, we estimate models with daily grades and average weekly, monthly, and even annual grades. We find that average annual beach grades best explain beach goer behavior.

Geographic Data

We use geographic information to estimate the length of each beach. Because the size of a site may influence the probability of choosing the site, it is common to include the natural log of the site's size (in this case, length) as a regressor in the model.

The extreme northern and Southern beaches in our set are included in order to represent all beaches north and south of our study area.

Beach Attribute Data

Beach attribute data were collected to characterize the fifty-one primary beaches of the study area (herein referred to as the beach attribute data). Beach attribute data were not collected for the southern most and northern most beaches since these beaches capture all trips to beaches south and north of the choice set.

The beach attribute data consist largely of binary variables indicating the presence or absence of a specific non-seasonal beach characteristic; count variables measure the quantity or abundance of a resource present. We categorize these attributes into three groups of explanatory variables: policy variables, activity specific variables, and composite variables. Policy variables reflect attributes that can be directly managed by beach agencies. Activity specific variables reflect beach attributes that are necessary or important for certain kinds of beach activities (e.g. bike paths are an attribute that is important for bicycling.) Composite variables capture general suites of characteristics of beaches including the degree of development at beaches.

The primary policy variable of interest for the beach project is water quality. Additional variables with possible policy implications are included as candidates for explanatory variables. These secondary policy variables include the absence and presence or count of:

Firepits
Lifeguard Stations
Parking Lots
Public Facilities
Public Restrooms
Sandy Shoreline
Showers
Sidewalk
Street parking

Activity Specific Variables

Because modeling heterogeneous preferences is one of the goals of the project, consideration is given to how the beach site choice of individuals is related to their choice of activity and the presence of the appropriate amenities at that site.

Variables that are important to specific activities include the presence of:

Bikepaths
Camping
Diving (spots)
Fishing (spots)
Piers
Playgrounds
Rentals (concessions)
Surfing (breaks)
Volleyball Nets

In addition to these activity-specific variables for beach goers as a whole, we also explore the importance of these indicators to beach goers that participate in specific activities. Towards this end, we create interaction dummy variables for particular activities. For instance, the activity indicator variable is 1 if the individual participates directly or indirectly (e.g. watches an activity undertaken by others) in activity X. We then interact this activity indicator with the relevant activity-specific variable.

Geographic Variables

Many attributes of beaches reflect their geographic location. We examine three types of exogenously determined, "geographic" variables in our model. Each of these attributes indicates proximity to:

Harbors
Natural areas
Rivers

In addition, we use maps based on a geographic information system of our study beaches to estimate beach length. Due to the large variation in the size of the beaches in the study area, an approximate length variable for the useable portion of the beach is used.

In addition to beach length as an indicator of beach size (and thus an important explanatory variable in its own right), beach length also can be used to scale beach attributes for which we have continuous data. Although most of the beach attribute variables are binary data, several attributes are characterized by count or continuous variables. Some of the variables for which we have count data are: beach clubs, beachside restaurants, concession stands, fire pits, lifeguard stations, public restrooms, and volleyball nets. Model specifications were estimated which used beach variables scaled under the assumption that the attributes are uniformly distributed over the beach shoreline.

We apply several scaling strategies to the data to capture the way in which beach goers experience beach and water quality attributes. These different strategies include: 1) keeping the data in its raw form (i.e., a mix of binary presence/absence variables and continuous count and ordinal qualitative ratings); 2) normalizing the count variables by beach length while maintaining the raw data for presence /absence and qualitative variables; and 3) transforming the non-binary and non-policy variables into binary presence/absence variables for specific attribute levels.

In several cases the correct normalization strategy for variables was not clear *a priori.* For some variables the relevant question appeared to be whether or not the attribute was present or absent at the beach in question, whereas for other variables, the relevant question is the level of density of a specific attribute. As an example, lifeguard towers are approximately uniformly distributed over beaches while restaurants, concessions, and restrooms are typically clustered into specific areas.

The transformation of variables from count variables to binary variables also requires substantial judgment in determining the threshold levels of importance. For example how many restrooms, restaurants, or fire pits are enough in order for the attribute to be adequately measured by binary variables? To address this problem we asked two major questions: 1) intuitively how would the variable be interpreted and 2) what is the distribution of the count variables.

Composite Variables

A number of variables seemed to occur in groupings that could best be described by a single "composite attribute" rather than individual component attributes. Including all of these component attributes individually would likely reduce the degrees of freedom of the model and cause problems with multicollinearity while each attribute individually would only modestly improve the performance of the model. If these attributes were perfectly orthogonal to our other model variables, then the exclusion of these attributes would only influence the explanatory power of the model, but would leave the coefficient estimates unbiased. However, because of the large number of attributes included in this group, it is likely that combinations of these attributes could prove to be correlated with the other right hand side variables described above.

We consider several obvious candidates for composite variables including commercial activity, development, natural amenities, and scenic blight. To develop our composite variables, we combine the formal tool of cluster analysis with an informal, intuitive equivalent of discriminant analysis to isolate combinations of the attributes that characterize the beaches.

The cluster analysis approach to constructing the composite variables allows the similarities between the beaches to guide variable definition. This approach assumes that there are a finite number of "types" of beaches – families of beaches with key attributes that are sufficiently similar so that the characteristics can be considered approximately constant over the group, and thus each beach can be assigned to a "type" of cluster. We use cluster analysis to identify groupings of the beach attributes that are statistically "close" or similar to each other and "far" from the other groups. In this application, we use the simple, intuitive Euclidean distance between the multidimensional numerical descriptions of beaches as a collection of attributes to measure closeness.

The first step to creating composite variables is to examine beach groupings that we feel *a priori* might have similar characteristics. The key to constructing these variables is to identify attributes that capture the same, or very similar, information for the beach goer. One example of this “collapsing of variables” is to collapse Rocky and Sandy into a single dichotomous variable which characterizes the composition of the shoreline. Similarly, we create a simple composite variable termed Ugly (ugly view) to indicate that one or more aesthetically degrading conditions existed at a beach. We first present these simple composites below and then discuss in more detail the creation of more complicated composite variables.

Sandy: Shoreline Composition

The rocky and sandy variables can be interpreted by the following trichotomy:

Sandy = 1 and Rocky = 0 (== Not at all rocky, very sandy)
Sandy = 1 and Rocky = 1 (== somewhat rocky)
Sandy = 0 and Rocky = 1 (== very rocky)

This rockiness variable captures how rocky the beach is and can be considered to range from 0 to 2 (which implies a cardinal relationship between somewhat rocky and very rocky) or as a dichotomous variable that captures either no rocks or no sand, depending on definition.

Ugly View

An Ugly Beach variable is created to equal 1 if at least one aesthetically degrading condition existed at a beach and zero otherwise. Ugly is constructed using Oil pumps, Oil Rigs, Power/Sewer Plants, and Storm Drains. Four beaches have none, thirty four have one, twelve have two, two have three, and none have all four. It should be noted that there are no oil pumps on the ocean side of PCH, which probably mitigates the impact of the Oil Pumps.

Using Cluster Analysis to Create More Complicated Composite Variables

Cluster Analysis can be used to identify influential site attribute variables that account for the grouping of sites based on similarity of characteristics. Cluster analysis requires the analyst to specify some number of groups (which can be varied iteratively) and then to employ a multivariate distance metric to partition the full set of sites into the specified number of groups using the distance metric and the criterion of maximizing within-group similarity and between-group heterogeneity. Once an acceptable number of groups has been found, the analyst inspects the results and identifies the specific site attribute variables that can account for the partition, either informally or through the use of a technique such as multiple discriminant analysis. Alternatively, one might create new dummy variables reflecting group membership that act as surrogates for site attributes that are themselves associated with group membership.

In the cluster analysis, we include only those attribute variables that could be represented by a binary 1/0 designation. Most of the beach attributes already are measured as absent or present: we convert the rest into 0/1 indicator variables or sets of 0/1 indicators to indicate the rough level of the covariate if there is a wide range of values. We used the Euclidean distance as a measure of similarity.

The first step of the process is to use the cluster algorithm to decompose the beaches into similar groups. The “types” of beaches are characterized by estimating a multinomial logit model on cluster membership. The means of the excluded variables and the coefficients of the variables included in the multinomial-logit are examined to characterize groups. While each cluster contains groupings that are hard to characterize simply, two types of beaches stand out in most of the relevant clusterings. The first type could be described as a “developed beach”, characterized by having a high likelihood of having stores, volleyball tournaments, equipment rentals, access by public transit, houses, concerts, street access, concessions, beach clubs, a pier, restaurants, and/or condos and hotels. (For a complete list of the component attributes, see Table 1.) The second type could be described as a “wild beach,” characterized by having a high probability of being accessible by only pedestrian paths, tide pools, rocky shorelines, and allowing dogs.

The sandy, ugly, development, and wild variables serve to collapse twenty component attributes into four composite indicator variables. In the choice models estimated below, we include the composite “sandy” variable in the category of policy variables because beach nourishment is an important policy factor for beach managers in Southern California.

Table 1 summarizes the composite variables. The variables that are used to construct the composites are 0/1 indicator variables for the absence/presence of the relevant attributes. The “developed beach” composite variables are determined by the sum of the number of the attributes present, with develop 1 being used to indicate the presences of three or more of the underlying attributes, and develop2 being used to indicate eight or more of the underlying attributes. It is interesting to note that Nature, the variable that indicates that a beach abuts a natural area, is not included in the wild_beach composite variable. This is because many beaches lie across the highway or street from natural areas, but the actual beaches are developed and actively managed.

Table 1: Composite Variables and Their Components

Composite Variables	Component Variables
Sandy	Sandy Rocky
Ugly Beach (Ugly)	Oilpumps Oilrigs PowerSewer Stormdrains
Developed_Beach (Develop1) Very_Developed_Beach (Develop2)	Access_Street Public Transit Restaurants Stores Concessions Rentals Beach Clubs Houses Condos/Hotels Pier Concerts Volley Ball Tournaments
Wild_Beach	Pedestrian Access Only Rocky Tide pools Dogs Allowed

Final Explanatory Data Set

Table 2 summarizes the explanatory variables used in the choice model. Note that the table is split into sections. The first section captures water quality attributes. The second section captures the composite variables and length. The third group of variables represents attributes that can be managed through policy. The fourth group of variables includes beach features that are geographically exogenously determined (e.g. harbor). The final group consists of attributes which can primarily be thought of as relating to specific activities or demographic subgroups of the panel – a few of these attributes also are used in constructing the composite variable; this poses no modeling issues since they will be used in conjunction with demographic or activity variables only. Despite the large number of variables, pairwise collinearity among the right hand side variables is modest.

Table 2: Beach and Water Quality Attributes

Attribute Name	Range and Description	Mean	(standard deviation)
Water quality attributes			
HTB_yr	0-4.333, Average HTB dry grade for all months	3.597	(0.764)
Composite variables and beach length (a normalizing attribute)			
Length	0.11-8.07, Length of beach in miles	1.974	(1.498)
Develop1	0/1, beach has several characteristics of development	0.540	(0.503)
Develop2	0/1, beach has very many characteristics o development	0.180	(0.388)
Wild	0/1, beach has several characteristics o naturalness or lack of development	0.320	(0.471)
Ugly	0/1, beach has visible oilrigs, oilpumps, power/sewer facilities, or stormdrains	0.280	(0.454)
Policy Attributes			
Firepits	0-261, # of firepits	14.36	(45.20)
Lifeguards	0-24, # of lifeguard towers	6.200	(5.764)
Parking	1/0, presence of public parking	0.840	(0.370)
Pubfac	1/0 presence of public facilities	0.380	(0.490)
Restrooms	0-20,# of restrooms	0.840	(0.370)
Sandy	1/0, beach is sandy	0.860	(0.351)
Showers	1/0, presence	0.680	(0.471)
Sidewalk	1/0 presence of sidewalk adjacent to beach	0.520	(0.505)
Strparking	1/0, parking along street near beach	0.760	(0.431)
Geographically Determined Attributes			
Harbor	0/1, presence of harbor or marina	0.180	(0.388)
Nature	1/0, abuts natural area	0.420	(0.499)
Rivers	1/0, river or creeks flows through or abuts beach	0.080	(0.274)
Activity Relevant Attributes			
Bikepath	1/0, presence of bike path adjacent to beach	0.440	(0.501)
Camping	1/0, campgrounds or RV parking	0.160	(0.370)
Diving	1/0, diving allowed	0.340	(0.479)
Fishing	1/0, fishing allowed	0.960	(0.198)
Pier	1/0, presence	0.240	(0.431)
Playground	1/0, presence	0.360	(0.485)
Rentals	0/1, bike or skate rentals available	0.180	(0.388)
Surfing	1/0, surfing at beach	0.740	(0.443)
Volley	0-107, # of permanent volleyball nets	10.22	(19.50)
Interaction Attributes			
Surfer*beach	Respondent is a surfer and beach has a surf break (i.e. Surfing = 1)	0.020	(0.139)
Run*bikepath	Respondent is a runner and beach has a bikepath	0.093	(0.291)
Diver*diving	Respondents is a diver and beach allows diving	0.002	(0.046)
Fisher*pier	Respondent is a fisher and beach has a pier	0.013	(0.112)
Fisher*fishing	Respondent is a fisher and beach allows fishing	0.048	(0.214)
Boat*harbor	Respondent is a boater and beach is near a harbor or marina	0.011	(0.106)
Adjacent Beach Dummy Variables			
Oceanside	1/0 trip was to Oceanside Beach or south	n/a	
Point Mugu	1/0 trip was to Point Mugu or north	n/a	
Venice	1/0 trip was to Venice Beach	n/a	

Travel Time And Travel Cost

Overview of the Issues

Determining the cost to each individual of visiting each potential site is a critical step for modeling recreation behavior because this variable captures the crucial tradeoff between cost and preference for beach attributes. It is this tradeoff that allows the analyst to deduce the monetary value that the beach goer places on beach attributes and water quality.

Choice models require the analyst to have an estimate of the cost to visit every site for every person. Since most people visit only a few sites, the costs to visit the other sites cannot be based on direct observation, but must be imputed by the researcher. Moreover, since the costs must be comparable across sites and respondents, if one cost is constructed in a particular manner for one respondent, all costs must be constructed in the same manner for all respondents. This means that, even if the beach goer provides their own estimate of their costs for the sites they visits, these costs cannot necessarily be used by the researcher because they may not be consistent with how the researcher imputes costs to other sites.

In principle, there are three critical components to the cost of visiting a site: (i) the out-of-pocket costs of traveling to and from the site (e.g. gas, maintenance, and depreciation expenses as estimated by the American Automobile Club), (ii) the opportunity cost of the time used to travel to and from the site, and (iii) the opportunity cost of the time spent on the site.

Each of these components raises issues both of data availability and of economic model structure; whether or not some of these costs apply depends on how one conceptualizes the individual beach goer's choice. The conceptualization of beach choice is complicated and there is no "right way" to incorporate the concept of travel cost and time into the model; the more detailed are our attempts at accounting for all aspects of travel cost and time, the more complex our model becomes. Hence, there is a trade-off between what is realistic and what is tractable. We address each of these issues, starting with time on site and working backwards to the more fundamental issues of calculating travel costs and time.

Time On-Site

How on-site time is modeled depends, in part, on whether or not the researcher treats trip length as an endogenous decision on the part of the beach goer. If time on-site is endogenous it must be determined on the basis of some cost per unit time spent on-site; while this price per unit time on-site is exogenous to the beach goer; her actual on-site time expenditure is endogenous and reflects her decision how long to stay. Theoretical models with this structure have been considered by Smith, Desvouges and McGiveny (1983), McConnell (1992), Berman and Kim (1999) and others. The theoretical literature generally has focused on qualitative properties of the resulting demand functions -- the demand functions for the number of trips to each site, and the demand functions for the (average) length of a trip to each site. However, the literature has not provided examples of explicit functional specifications for these demand functions that are both tractable and consistent with utility maximization.

Discrete choice models can treat trip length along with site selection as a discrete choice. While ignoring the issue of trip frequency, these models make it possible to model trip length. To use the method, however, one needs data on the time cost per unit of on-site time; often, this is not readily available.

In this paper, we do not include time on site as a potential decision variable for several reasons. First, the respondents report very few multi-day trips and we do not include these trips in our estimation. Second, while one could still treat the precise length of a one-day trip as a decision variable, this adds a level of complexity that complicates our exploration of other important issues (e.g. seasonality and activity choice). Consequently, in this phase of the analysis we will treat the cost of on-site time as being zero, and as a result we may be undervaluing the cost of a trip to the beach. Nevertheless, we believe any error created by making this assumption will only slightly reduce our estimated welfare impacts and thus this approach is the most conservative available to us at this time.

Estimating Travel Cost and Travel Time

In our survey, we collected precise data (including street addresses) about respondents' origin location. We use PC Miler^{tm} to calculate how many miles each respondent would have to travel in order to drive from their residence to each beach in the region and how long this would take. Since there is usually a choice of routes, PC Miler makes an estimate of the shortest route and the time taken to drive this route under typical conditions for that type of road. Since individuals may be idiosyncratic in their choice of routes, there obviously is some possibility of measurement error when imputing the distances and times from PC Miler, but we believe this is likely to be quite small.

Given the PC Miler estimate of the distance and travel time from a person's residence to a beach, we have to make some further assumptions in order to convert this to a monetary travel cost. From the survey, we know the mode of transportation the respondent uses to get to the beach (e.g. automobile, bus, bike, or walk). If it is walk or bike, there is effectively no transportation cost. If the transportation mode were by bus (an occurrence rarely reported in our sample) we would use the cost of a bus ride. The overwhelming majority of respondents in the sample traveled to the beach via automobile. There are two issues to address in the calculation of travel cost by car. First, we calculate variable expenses based on the average figures for expenditure on gasoline and oil per mile in 2000 from the American Automobile Association; we use average values because we do not know what type of car the respondent used to get to the beach. Second, there is the issue of what other costs to include. McFadden (1997) argues that motorists should only pay attention to these variable costs when making their travel decisions. Whether that is what actually happens is an empirical question. In the literature, many researchers also include an estimate of vehicle maintenance and other operating costs. We use both variable fuel costs and maintenance costs provided by the American Automobile Association for Southern California in 2000.

While it is relatively straightforward to estimate distances and times from respondents' homes to each site, accounting for how they value that time is much less straightforward. In the transportation literature, Truong and Hensher (1985) as well as Bhat (1998) show that time is valued differently for different modes of transport and for different categories of activity during travel – waiting, walking, in-vehicle time, etc. Since virtually all of the trips made in this sample were made by car, we assume the same valuation of time applies regardless of the travel mode.

Assuming that travel time is valued in the same way by each individual, there is no general agreement in the recreation demand literature as to how to value this time spent traveling. The early recreation demand literature used a fraction of the individual's wage rate, usually one third or one half of the wage rate. We follow the standard approach in this study and value travel time at a fixed proportion of the wage rate. We allow this fraction to vary from zero to the full wage rate (0%, 33% and 100% of the wage rate).

<u>Measuring the Wage Rate</u>

To the extent that the value of time is imputed from wages – whether it is valued at the full wage rate or in some other proportion – the researcher needs an estimate of the respondents' wages. This is often somewhat problematic. In some cases, this is because the subject does not actually work (e.g., is unemployed, or retired, or a housewife, or otherwise outside the labor force). In other cases, while the respondent does report that they were employed, they are asked about salary, but not wage. Like many surveys, we asked respondents for their annual income, not their hourly wage. In this case, it is common to estimate the wage from the information on annual income by assuming the income is derived from some fixed number of hours worked per year (e.g., 2000 hours per year, corresponding to working 40 hours per week and 50 weeks per year) and dividing income by the fixed number of hours. However, even if the annual income is reported with perfect accuracy, the assumption of a specific number of hours worked per year inevitably introduces the possibility of some measurement error. Moreover, there is often likely to be some measurement error in the reporting of annual income.

In our study we use a sensitivity analysis to value travel time alternatively at 0%, 33%, 50% and 100% of the wage rate. The wage rate will be calculated from annual income by assuming 2000 hours worked per year. For individuals who did not report their income, we used the following imputation procedure.

Income is assumed to be lognormally distributed, with the mean determined by the following covariates: constant, male, kids, student, work fulltime, retire/disable, college graduate, high school graduate, black, Hispanic. Covariates are assumed to be zero if the respondent didn't provide them. Age is not used because it was missing for many respondents and could not be assigned a default value as easily as a 0/1 variable.

The respondents provided yes/no answers to a progressively narrow set of questions about income range. These answers yield intervals that bound the respondents' incomes. Some respondents only answered a few of the questions and their income had wide bounds. Other respondents answered all of the questions, and their income was determined to more precise bounds. If no lower bound was supplied, $1 is assumed. For the income distribution estimation an unspecified upper bound is assumed to be infinite, although later in calculating the expected income conditional on the individual's interval, it is assumed to be $3 million. (The smallest annual income generated by our procedure is $5177.68 per year. This is equivalent to about $2.60 per hour; so it is not bounded by the minimum wage. However, welfare and social security benefits are also not bounded by the minimum wage, and so we believe this lower bound is reasonable.)

Assuming that ln(income) was normally distributed around the mean (conditional on covariates), we used maximum likelihood to find the values for the impacts of the covariates on the expected income as well as the standard deviation about that conditional mean, given that the unknown income lay in the respondent-supplied interval. If a respondent gave no upper bound to the interval containing income, we assumed that it was an unbounded interval when fitting the model. Given the coefficients, we calculate the (lognormal)income distribution for each individual, conditional on their demographic covariates. We then numerically calculated the expected value of the individual's income, conditional on the covariates and coefficients and on the income lying in the interval given, using 100,000 randomly drawn uniformly distributed points in the interval for the evaluation.

Choice Set Determination

To estimate the beach choice model, we must identify the set of all feasible choices for each respondent. In the simplest case, the analyst includes in the choice set all sites that the respondent has a non-zero probability of visiting. In many cases, the choice set can be assumed to include all available recreation sites, especially when the number of sites is small and there are no restrictions on site access. In practice, the number of potential sites may be large and some sites may be unknown to the respondent. In other cases, certain sites may not be accessible to an individual due to physical limitations, limitations in skill (see for instance Grijalva et al. 2002) or a lack of one or more attributes that are specific to the activity that a respondent may undertake at a site.

Haab and Hicks (1999), Haab and McConnell (2002), and Parsons (2003) provide a review of choice set studies of recreation. The marketing research and transportation literature has recognized the importance of choice set formation and have developed various models to define the choice set. See Shocker et al. (1991) and Roberts and Lattin (1997) for a review of the marketing studies, and Thill (1992) for a review of the transportation literature.

Site aggregation is common in random utility models (RUMs) of recreation. Empirical studies of aggregated choice sets in recreation models include Parsons and Needelman (1992), Feather (1994), Kaoru, Smith, and Lui (1995), Lupi and Feather (1998), Jones and Lupi (1999), and Parsons, Plantinga, and Boyle (2000). Site aggregation is the process where a group of recreation sites is defined as a single choice alternative. For example, a site may be defined as a county or region made up of several lakes or beaches. When the number of potential sites is large, site aggregation often is used to reduce the choice set to a manageable size. If the characteristics of aggregated sites are homogeneous, the aggregation should be fairly straightforward. Otherwise, aggregation can result in a loss of information and thus a loss of estimation accuracy.

Ben-Akiva and Lerman (1985) provide an economic model of aggregated choice sets. Parsons and Needelman (1992) shows that the utility function in Ben-Akiva and Lerman (1985) decomposes into the average utility at sites in each aggregated choice set and a measure of the heterogeneity of sites in aggregated choice set. Parsons and Needleman also show that estimating a model using aggregated choice sets will bias coefficient estimates toward zero when utilities produced by aggregated sets are similar; increasing the heterogeneity within each aggregate group or increasing the number of beaches in the aggregated choice set tends to increase this aggregation bias. In other words, aggregating dissimilar sites will bias model estimates.

Kaoru, Smith, and Lui (1995) analyzed marine recreational fishing site choice in North Carolina and considered experiments that compared thirty-five disaggregated sites with smaller aggregated choice sets of eleven and twenty-three sites. The welfare benefits estimated by aggregated models differed considerably from disaggregated models. Parsons and Needelman (1992) used data on fishing trips to lakes in Wisconsin. They aggregated 1,133 sites into smaller choice sets of sixty-one and nine sites. Their results indicated that extreme aggregation could seriously impact parameter and thus welfare estimates.

Parsons, Plantinga, and Boyle (2000) and Jones and Lupi (1999) suggest the possibility of using partial aggregation to define choice sets. Parsons, Plantinga, and Boyle (2000) analyzed data on fishing trips to lakes in Maine. Their results showed that benefits estimated by the aggregated or narrow choice set models were lower than the baseline model.

Jones and Lupi (1999) considered an experiment similar to Parsons, Plantinga, and Boyle (2000). Jones and Lupi (1999) narrowed 83 counties of fishing site in Michigan along 6 lines of species and resource type using the factor analysis. Their empirical results showed that the benefits estimated by the narrow choice set models were relatively similar to those by the original choice set model. Models in which the respondents face a small set of alternatives in the choice set do not permit substitution away from sites when sites are lost; this tends to over-state marginal and total losses. At the same time, extremely narrow choice sets reduce the size of the population affected by the policy, thereby tending to under-state the value of total losses.

Our respondents visited more than 300 named beaches. Most of these beach names, however, were redundant names for the same beaches or were specific locations within larger, better known beaches. We reduced our unmanageably large initial choice set of beaches by mapping beaches named by respondents to fifty one primary public beaches listed in the California Beach Access Guide (Coastal Commission 1997). This aggregation groups contiguous and similar sub-sites of beaches together into larger beaches. Because of the relative homogeneity of attributes within these larger beaches and because many of these beaches have one primary access point, very little information is lost in the aggregation. Further, because the beach sites in our final choice set correspond to beach management jurisdictions, the results from the model will apply more directly to policy needs and decisions. Additionally, we represent all beaches to the South of our choice set by the indicator beach Oceanside and those beaches to the North by the indicator beach Point Mugu. Despite the aggregation of the choice set, we still maintain a substantially large choice set of beach sites (fifty one primary beaches and two composite beaches). A complete list of beaches is given in Table 3.

Table 3. Aggregated Beaches in the Choice Set

Code	beach name	code	beach name	code	beach name
1	Oceanside	21	Surfside	41	Las Tunas
2	San Onofre South	22	Seal	42	Malibu (Surfrider)
3	San Onofre North	23	Alamitos Bay	43	Dan Blocker (Corral)
4	San Clemente State	24	Belmont Shores	44	Point Dume
5	San Clemente City	25	Long Beach	45	Free Zuma
6	Poche	26	Cabrillo	46	Zuma
7	Capistrano	27	Point Fermin	47	El Matador
8	Doheny	28	Royal Palms	48	La Piedra
9	Salt Creek	29	Abalone Cove	49	El Pescador
10	Aliso Creek	30	Torrance	50	Nicholas Canyon
11	Laguna	31	Redondo	51	Leo Carrillo
12	Crystal Cove	32	Hermosa	52	County Line
13	Corona Del Mar	33	Manhattan	53	Point Mugu
14	Balboa	34	El Segundo		
15	Newport	35	Dockweiler		
16	Santa Ana River	36	Mother's		
17	Huntington State	37	Venice		
18	Huntington City	38	Santa Monica		
19	Bolsa Chica	39	Will Rogers		
20	Sunset	40	Topanga		

MODEL ESTIMATIONS

In the following section we estimate a series of increasingly sophisticated specifications of the standard logit random utility models including a three nested logit specification. For all specifications we assume that the ultimate choice is for single day trips to beaches in Southern California. We consider only day trips and we assume that each choice occasion is independent of all others.

From the start, we believed that seasonality and activity choice were important in explaining beach choice and thus in determining the value that people place on water quality. To demonstrate the importance of seasonality and activity choice, we begin by estimating three increasingly more fully specified multinomial, repeated choice models of beach choice using data pooled across all seasons (the results from all three models using pooled data are summarized in Table 4). We start with a primitive model (model 1) that includes only water quality, travel cost, and our composite variables that describe beach type. Then we add more explanatory information regarding activities. We also estimate these standard logit random utility models separately for individual seasons and in doing so demonstrate the way in which seasons affect the sign and magnitude of coefficient estimates (model 2). We finally arrive at what we believe is the most complete and accurate model of beach choice behavior; this is our final three tiered nested model of participation, activity choice and beach choice. The purpose of this exposition is to demonstrate the independent influences that activities and seasonality have on site choice and to show the reader that a failure to account for both activities and seasons can significantly alter the results of the model. If the reader is interested only in the final model, the next two sections can be skipped.

Applying the Standard Repeated Choice Random Utility Model

We model the choice of a beach on each occasion as being independent of choices on all other occasions. Specifically, we estimate a repeated random utility model that assumes the probability that an individual *i* chooses site *j* depends on the relative utility of site *j* compared to all other sites. As in most applications of the multinomial logit, we estimate a model of choice in which the beach goer chooses a beach destination to maximize an indirect utility function that consists of both a deterministic component and a random component. Specifically, we estimate a model in which the deterministic component of indirect utility is a function of attributes that remain constant over time, x_{ij}, and attributes that change over time W_{ijt}. In this analysis, temporally constant attributes include "land attributes" (e.g. restrooms); temporally varying attributes include the measures of water quality. Note that temporally varying attributes, like water quality, can be measured as point estimates, means over time, and even variance over time. In fact, the beach grades provided by Heal The Bay are themselves running geometric means of water quality measures over four week periods. In our analysis, we estimated models with daily grades and average weekly, monthly, and even annual grades. In every model we found that average annual grades provide more explanatory power than other measures (based on more significant coefficients estimates and greater likelihood measures). This may indicate that beach goers are using past experiences and general levels of water quality to inform their beach decisions. The deterministic component of indirect utility enjoyed by beach goer *i* from choosing site *j* at time *t* is given as

$$U_{ijt} = \mathbf{X}_{ij} * \boldsymbol{\beta} + \mathbf{W}_{ijt} \tilde{a}$$

The probability that individual *i* chooses beach *j* at time *t* is given by

$$\Pr_{it}(j) = \prod_{ijt} = \frac{e^{U_{ijt}}}{\sum_{j=1}^{J} e^{U_{ijt}}}$$

The vectors of coefficients, $\boldsymbol{\beta}$ and ã, reflect the preferences the beach goer places on the attributes in $\mathbf{X}_{ij}$ and $\mathbf{W}_{ijt}$ respectively. In addition, there exists some randomness in the choice of beaches by beach goers. This randomness is captured by the random component of the indirect utility function, ε_{ijt}. This random component may reflect true stochastic processes in the choice process and unobserved site attributes that cannot be modeled by the analyst. We follow the literature and assume that the random component of the indirect utility function is distributed as a Type I Extreme random variable.

In the first model specification (Model 1), we include only the most basic beach attributes. The point of this model is to demonstrate how failure to account for seasons and activities in a model of beach choice will produce a model that does not find the true value people place on water quality. In this basic model, we include travel cost, water quality, and our basic composite variables. In Table 4 we provide the results for the estimated models across the pool of all waves and observations. Without describing all of the results here, we note that the basic repeated choice model estimates that beach goers prefer beaches that are closer, longer, and more developed (but not too developed). The pooled model does not yield an estimated preference for water quality that is either significant or of the sign we predicted (we predict that people prefer clean water). This reflects the common observation that people do go to dirty beaches. The estimated coefficients on water quality are more in line with our intuition when we estimate the model for individual waves (see below and in Table 6). The models estimated for waves 1, 2, and 6 – the wet weather months when water quality varies the most and has the most number of impaired water quality days – all yield coefficients estimates for water quality that are positive and significant (see our discussion of seasonality below). Obviously, pooling across seasons masks the differing preferences for water quality that people hold in different seasons.

In the second model (Model 2), we estimate a more fully specified model of beach choice in which we include many beach attributes that are of general relevance (e.g. public facilities and parking) as well as some other attributes that are required for specific types of beach activities. For instance, we include whether a beach has a surf break, a bikepath, or volleyball nets. This more fully specified model has greater explanatory power than the basic model (higher psuedo R_2 and higher log likelihood). In this more fully specified model, the estimated preference for water quality is positive and significant. The model demonstrates the importance of having a more complete description of beach attributes. As described above, it is easy to observe beach goers choosing beaches that are decidedly dirty from a water quality perspective. On closer examination, however, it can be seen that many of the "dirtiest" beaches in Los Angeles and Orange Counties also provide the most man-made desirable attributes (including lifeguard towers and parking). Model 2 captures the opposing preferences for these different beach attributes.

Model 2 also demonstrates that many activity specific beach attributes (e.g. fishing and surfing opportunities) are important factors in explaining beach choice. First, activity specific opportunities are important in their own right. Second, in our discussion of seasonality (in a following section), we show that the dramatic differences across waves in the estimated coefficient on water quality become significantly less pronounced in the second model (the coefficient is positive in all waves and significant in all waves but Wave 3). The added attributes in the second model do not vary over wave and season, but the activities people participate in do vary and, thus, so does the relevance of activity specific opportunities. The results of the second model suggest that differential participation in activities may account for what appears to be changing seasonal preferences for water quality among beach goers.

In our third model (Model 3), we explore the importance of activity participation in beach choice further by interacting indicator variables for participation in specific activities with indicator variables for activity specific opportunities. For example, if a respondent reported that he fished then we give this respondent a 1 for fishing, while we give a zero for non-fishers. We then interact this fishing variable (fisher) with attributes that we think are important for fishing (e.g. the beach is a known fishing spot or there was a pier). We find that the estimated coefficients on the attributes shared by the second and third model remain stable and generally of the same level of statistical significance. The estimated coefficients of the new activity interaction variables tend to be generally significant. The two exceptions are running/bike path and boating/harbor. Both of those variables are somewhat distinctive: running opportunities are available at almost every site, while boating/harbor is the opposite – it applies at only a handful of sites in our sample.

Table 4: Coefficient Estimates for the Non-nested Repeated Choice Models Using Pooled Data

Variable	Estimated Coefficients Model 1	Model 2	Model 3
Cost	-0.1007 [a]	-0.0978[a]	-0.0986 [a]
HTB yr	-0.0180	0.2070 [a]	0.2170 [a]
ln(length)	0.5384 [a]	0.2996[c]	0.3198 [c]
Develop 1	0.5373 [a]	1.2383 [a]	1.2592 [a]
Develop2	0.2775 [a]	0.3039[c]	0.3086 [c]
Wild	0.1681	1.0565[a]	1.0164 [a]
Ugly	0.0561	-0.2333	-0.2611 [c]
Pubfac		0.3597[b]	0.3677 [b]
Restrooms		2.0531 [a]	1.9679 [a]
Sandy		0.9699 [c]	0.9945 [c]
Sidewalk		0.4242[b]	0.4654 [b]
Harbor		-0.0815	-0.0621
Nature		0.7522[a]	0.7112 [a]
Rivers		0.2407	0.3362
Camping		-0.2747	-0.3341
Diving		-0.0034	-0.0475
Fishing		-0.8996 [c]	-0.53 85
Pier		-0.0349	-0.1542
Playground		-1.1190 [a]	-1.1046 [a]
Rentals		0.6681 [a]	0.6745 [a]
Surfing		0.7278[a]	0.7001 [a]
Volley		0.0055	0.0039
Oceanside		5.5946[a]	5.7791 [a]
Surfer*beach			0.5593 [b]
Run*bikepath			-0. 1265
Diver*diving			2.4445 [a]
Fisher*pier			0.9113 [c]
Fisher*fishing			-3.0011 [a]
Boat*harbor			-0.6 100
Pseudo R2	0.25	0.30	0.31
Log-likelihood	-14114.51	-13154.77	-13066.36
Significance	a=<.001	.001<b<0.05	.05<c<.10

It is important to note that this third specification has the limitation that both activity choice (which is presumed to be exogenous) and beach choice may both be endogenously determined by the explanatory variables in the model (including the activity variables with which the activity indicators are interacted.) As a result, this internal endogeneity is likely to lead to bias in the estimated coefficients of the model. Nevertheless, we present these results here because the additional activity interaction variables improve the fit of the model slightly and they suggest the need for further investigation of the role of activities in beach choice. This is accomplished through the nested model to be presented below.

Accounting for Seasonality in the Beach Choice Model

To explore the effects of seasonality in beach choice, we estimate separately our basic beach choice models for Wave 1 (December and January), Wave 2 (February and March), Wave 3 (April and May), Wave 4 (June and July), Wave 5 (August and September), and Wave 6 (October and November). By estimating the models for each wave separately, we can investigate how seasonality influences the coefficients that we estimate for our beach choice models. While all of the estimated coefficients could potentially vary across waves, we focus our attention here on how the estimated preference for water quality varies in different waves.

Differences in coefficient estimates across the waves could be caused by a variety of factors with differing impacts on the results of our models. First, the significance of our coefficient estimates could vary substantially across waves. One possible cause of variation in estimated coefficient significance could be the fact that the frequency of visitation by our respondents varies considerably by wave. Table 5 summarizes the number of beach trips taken by the survey respondents in each wave.

Table 5: Summary of Trips by Wave

WAVE	sum	mean	max	min	Range	Std. Dev.
Wave1	1027	5.010	47	1	46	6.839899
Wave2	744	3.875	28	1	27	3.996726
Wave3	681	3.547	27	1	26	3.665017
Wave4	1501	4.289	30	1	29	3.460731
Wave5	938	4.043	30	1	29	4.415424
Wave6	527	3.847	30	1	29	4.712141

Second, it may also be the case that differences in estimated preferences for water quality reflect real seasonal differences in the strength and nature of preferences that beach goers hold for water quality. Preferences could vary seasonally for several reasons. First, different kinds of beach goers (e.g. swimmers vs. runners) may have different preferences for water quality. In some waves, certain types of beach goers may be more or less represented than others. We investigate user-differentiated preferences for water quality and beach choice in the next section. A second potential reason for differences in estimated preferences across waves is that individual recreational behavior may change over the seasons. For instance, an individual beach goer may be more likely to swim during the summer and run at the beach during the winter and so changes in activity choice alone may influence preferences. Third, it may be the case that offsite recreational possibilities change during the year (e.g. snowboarding is a substitute activity for beach going only during the winter months). While we expect variation in offsite recreation possibilities to primarily influence participation decisions, it is possible that changing offsite possibilities may influence onsite preferences for water quality. Finally, it is always possible that beach goers' preferences change in some systematic way across seasons that we have not yet determined.

To compare estimated preferences for water quality across the seasons, we examine wave-by-wave data for the most basic of our non-nested models, Model 1, and the most fully specified of the basic models, Model 3 (Table 6). (The coefficient estimates of models 2 and 3 do not vary significantly.) By looking at seasonal differences in the estimated preferences of the two models, we can determine how much of what appears to be seasonal differences in estimated preferences can be accounted for simply by more fully specifying the models to include activity relevant explanatory variables.

Table 6: Seasonal Differences in the Estimated Coefficients on Quality

Coefficient	wave 1 Dec-Jan	wave 2 Feb-Mar	wave 3 April-May	wave 4 June-July	wave 5 Aug-Sept	wave 6 Oct.-Nov
Model 1						
Cost	-0.107	-0.101	-0.091	-0.097	-0.101	-0.099
Standard Error	0.004	0.005	0.005	0.003	0.005	0.006
P Values	0.000	0.000	0.000	0.000	0.000	0.000
Beach Grade	0.317	-0.018	-0.028	-0.018	-0.059	0.119
Standard Error	0.049	0.048	0.057	0.042	0.049	0.062
P Values	0.000	0.706	0.620	0.670	0.228	0.055
Model 3						
	wave 1	wave 2	wave 3	wave 4	wave 5	wave 6
Coefficient						
Cost	-0.098	-0.099	-0.096	-0.096	-0.100	-0.101
Standard Error	0.004	0.005	0.005	0.004	0.005	0.006
P Values	0.000	0.000	0.000	0.000	0.000	0.000
Beach Grade	0.573	0.217	0.141	0.316	0.503	0.373
Standard Error	0.080	0.074	0.101	0.099	0.117	0.116
P Values	0.000	0.004	0.164	0.001	0.000	0.001

Notice first that the estimated coefficients on cost are robust and consistent across all specifications, while the estimated coefficients on water quality vary significantly. In the simplest specification, Model 1, the coefficient on water quality ranges from –0.059 in the late summer wave to 0.3 17 in the mid-winter waves. Of course, the significance of the estimated coefficients differ in the six waves. In Model 1, only the late Fall and mid-Winter waves yield estimates with significant coefficients; both of these estimates are of the expected (positive) sign. This greater significance in these two waves may reflect the fact that beach water quality during these months tends to be worse than during other, drier, times of the year, and so beach goers may have stronger preferences for water quality.

In the more fully specified Model 3, in which more activity-specific factors are included among the explanatory variables, the estimated coefficients on water quality are consistently of the expected sign; all but one coefficient estimate is significant at the 0.1% level or better. Clearly, the inclusion of activity-specific variables not only improves the fit of the model, it also improves our ability to differentiate seasonally varying preferences for water quality and attributes that support activities.

In many ways, the wave-by-wave estimation of Model 3 can be considered a state of the art estimation of the impact of water quality on recreational beach choice. Model 3 includes far more explanatory variables than most recreation site choice models. Further, Model 3 accounts for the influence that divergent activity choices may have on preferences for site quality. Model 3 even provides a seasonal resolution rare in recreation choice models. Despite the advances inherent in Model 3, the model yields an estimate of water quality that is averaged over all users, regardless of the activities they undertake during their beach trip. We believe, *a priori,* that beach visitors who recreate in the ocean will have stronger preferences for clean coastal water than other beach goers. Further, if the choice of activity is non-random, then welfare changes due to changes in water quality will be distributed across the population in non-random ways. Simply put, different people are likely to be harmed to different degrees by water quality impairment. A more complicated, nested model is required to more fully understand how the benefits and costs of water quality change are distributed in society.

Towards a More Comprehensive Model of Beach Choice: The Importance of Heterogeneity Among Users and Activities

The demographic and recreational diversity, or heterogeneity, of beach goers in Southern California complicates the accurate modeling of recreational beach site choice and the assessment of economic value associated with a change in beach characteristics. If heterogeneity is not accounted for, our random utility models could be biased and produce inaccurate information on the effects of changes in beach attributes (see Yatchew and Griliches, 1984). If present, this bias would adversely affect the model in terms of the distribution of welfare estimates due to changes in resource attributes and/or management decisions.

To address the effects of respondent heterogeneity in models of choice, researchers primarily have focused on three approaches. Two approaches involve the *a priori* selection of variables – most commonly demographic attributes. In the first approach, "the cluster models," the researcher places individuals into demographically similar groups or segments. The second approach incorporates individual demographic variables into the indirect utility function through the use of interaction variables.

A third approach to incorporating respondent heterogeneity in choice models is the random parameter logit (RPL) model. This method handles heterogeneity across preferences by allowing estimated coefficients to randomly vary across individuals according to a continuous probability distribution, typically the normal or lognormal distribution. One aspect of the RPL model is that it is not restricted by the independence of irrelevant alternatives (IIA) property. This is due to interactions within the choice probabilities of the attributes of all elements in the choice set (Train, 2003). The RPL model relaxes the restriction of the conditional logit (CL) that requires coefficients on observed variables to be fixed over all individuals. By allowing for variation in coefficients over individuals, the unobserved portion of the respondent's utility is correlated over sites and time (Train 1998). As a result, the RPL provides a better fit to the data.

The RPL method, however, has two important limitations. First, the RPL assumes that preferences vary continuously across economic agents (i.e. the respondents or beach goer). Although the continuous distribution assumption is likely to be valid in many applications, for example how spicy one likes their food, there are many situations where actual preferences may be more accurately captured by multiple discrete probability masses. For instance the presence of a fishing pier enters discretely into the typical beach goer’s utility function.

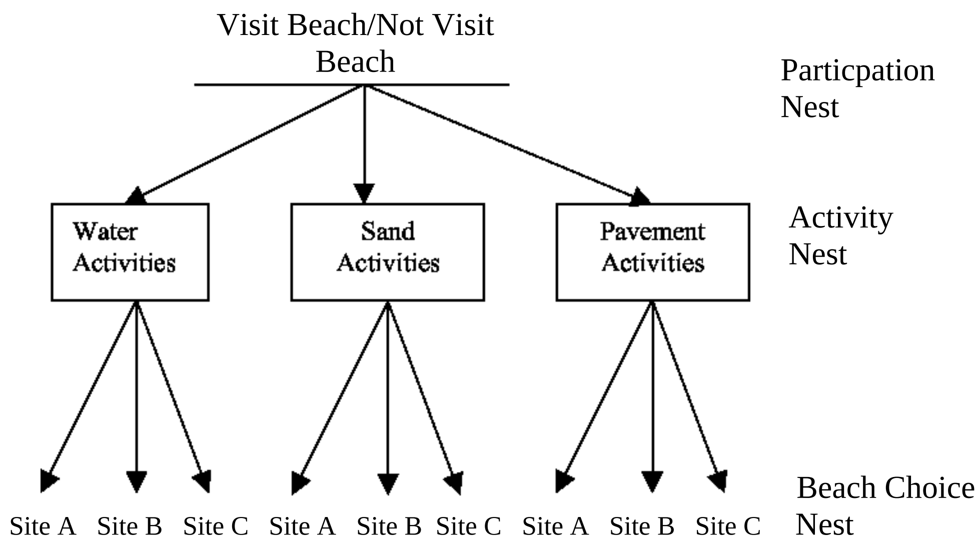
Second, random parameter models may not be the most appropriate for beach management decisions. Beach managers often are concerned with understanding how changes in water quality impact specific individuals or user groups. While the RPL advances the analyst's ability to estimate the most efficient model parameter coefficients (Boxall and Adamowicz, 2002), the standard application of the RPL does not provide a means for assessing the distributional impacts of changes in beach attributes, including water quality. The RPL can only provide information regarding a behavioral explanation for the source of the heterogeneity across people if the analyst also models the mean of the random parameter coefficients as a function of personal characteristics (see Breffle and Morey 2000).

Another approach to modeling preference heterogeneity is to use a nested repeated multinomial logit framework (applications include Bockstael, McConnell and Strand 1989, Kaoru 1995, Morey et al. 1993, Hauber and Parsons 2000). The nested approach is based on two important assumptions. First, individual preferences are neither homogeneous nor continuously distributed, but can be more accurately represented as being discretely distributed. Second, individual preferences are not purely a function of demographic variables, but can also be formed by expectations regarding the utility of site choice.

In our case, systematic heterogeneity could be accounted for by modeling the choice of activity and then by estimating the choice model conditioned upon the choice of activity. Within each activity group, preferences are assumed to be homogeneous; however preferences, and utility functions, can vary between groups. A primary benefit of the nested approach is that the model may help to explain variation in preferences across individuals conditional on the probability of membership to a group. The increased explanatory power provided by the nested model should be of benefit to beach managers in terms of welfare analysis and policy decisions. The results of the nested model also would allow beach managers to see how preferences and behavior vary for different kinds of beach goers. Further, beach managers could explore how welfare impacts differ among different user groups. The nested model can estimate the coefficients on explanatory variables associated with the recreational beach choice occasion for each activity type.

A Nested Model of Beach and Activity Choice

Nested models of recreational site choice and participation are now common in the literature (see for instance Morey et al. 1993, Kaoru 1995, McNair et al. 1999, and Morey 1999). A simple diagram of the three level nested logit approach we use to estimate models of beach choice, activity choice, and participation is given below. In the exposition that follows, we start at the bottom of our nest, understanding that the choice in any one nest is conditioned upon having made a decision in the previous nest. So for instance, the choice of an activity is made only after a respondent has decided to go to the beach and the beach chosen depends upon the activity undertaken. While the decisions run top to bottom, the modeling progresses from bottom to top – each choice is made given the expected utility of the nest below. So, the decision to visit the beach is made based on the expected utility of considering all possible activities and beach choices. All three nests are estimated simultaneously using Full Information Maximum Likelihood.



<u>The Beach Choice Nest</u>

Beach site choice is conditioned upon the choice of activity as assigned by the hierarchical method described above. The beach choice nest is a standard linear in attributes model of site choice in which the respondent is believed to chose the beach that provides the greatest (indirect) utility. The deterministic component of indirect utility enjoyed by beach goer *i* from choosing site *j* at time *t* is given as

$$U_{ijt} = \mathbf{X}_{ij}*\boldsymbol{\beta}$$

Where $\mathbf{X}_{ij}$ is the vector of all beach attributes, described in the data section above. The probability that individual *i* chooses beach *j* at time *t* is given by,

$$\Pr_{it}(j) = \prod_{ijt} = \frac{e^{U_{ijt}}}{\sum_{j=1}^{J} e^{U_{ijt}}}$$

The vector of coefficients, **β,** reflects the preferences the beach goer places on the beach attributes in $\mathbf{x_{ij}}$. In addition, there exists some randomness in the choice of beaches by beach goers. This randomness is captured by the random component of the indirect utility function, ε_{ijt}. This random component may reflect true stochastic factors in the choice process and unobserved site attributes that cannot be modeled by the analyst. We follow the literature and assume that the random component of the indirect utility function is distributed as a Type I Extreme random variable.

For each activity type, a separate utility function is estimated with separate coefficients, including separate coefficients on travel cost. In past studies, the cost coefficients of alternative submodels in a nested logit random utility model have generally been constrained to be the same, the implication being that the marginal utility of income does not vary between nests. Hensher and Green (2002), however, demonstrate that because scale parameters vary between submodels, constraining coefficients to be the same is not the equivalent of constraining marginal utilities to be the same. Because the magnitude of the stochastic term in the beach selection utility function almost certainly is not the same across the different activity types, we allow the scale of the coefficients to be determined by the choice data within each activity type. This implies that the coefficients on cost are not identical for the different activity types. Further we assume a constant inclusive value coefficient for the alternative activity types in the activity selection sub-model and so the marginal utilities of expenditures in the classes are not equal. The standard model, which

which imposes the same coefficient on cost in each choice sub-model, would distort the key coefficients in an unpredictable way, biasing welfare measures.

In the beach choice nest, two beaches north and south of the geographic choice set were also included. These beaches, Oceanside and Point Mugu, were characterized solely by binary indicator variables and travel costs because we did not have beach attribute data for them. Point Mugu is omitted from water activity and sand-activity choices, since there were no trips to that destination for those activities. We also use a binary indicator for Venice beach. Venice beach is an important tourist destination and offers many amenities and attractions that are not found at other beach sites (e.g. Muscle Beach, the graffiti pit, the skating pit, drum circles, etc.). This level of the nest is estimated for all trips that a) were to a single beach in the choice set, b) could be classified by activity, and c) were taken by a respondent who supplied income and address information for cost calculation.

<u>The Activity Choice Nest</u>

The activity choice nest, which models the probability that a respondent chooses an activity, depends upon the expected utility of participating in an activity (as estimated using the inclusive value from the first nest) and demographic characteristics of the respondent. Before we can proceed with the activity choice nest, we describe how we assign respondents activity choices to a limited number of activity choice categories.

Beach Activities

Before a nest of activity choice can be estimated, we first had to define and assign activity choices made by respondents. There were 42 distinct activities reported by the survey responses plus “other.” These 42 activities were grouped into the following categories: water contact activities (abbreviated to “WATER”); activities on the sand (abbreviated to “SAND”); and activities involving walking, running, bicycling, etc on the boardwalk or pavement (abbreviated to “PAVEMENT”) as well as activities involving shopping, dining, etc. Activities were assigned to one of these categories based on both the similarity of the distinct activities and also the relationship between the activity and the attributes of the sites. A detailed description of the categories is as follows:

Water

The “Water” category includes all activities that are characterized by direct contact with water such as splashing in the water, swimming, and SCUBA diving. In addition to these immersion activities “canoeing” and “kayaking” are included in the “Water” category. “Canoeing” and “kayaking” are similar to other “Water” activities in that participants have a relatively high probability of getting wet. Additionally many canoeists and kayakers participate in a version of surfing and seek beaches with similar characteristics.

Recreational site choice for participants of “Water” activities are expected to be sensitive to attribute levels for characteristics such as water quality, and the presence of life guards, storm drains, and rivers.

Sand

The “Sand” category includes activities that commonly are identified as “Beach” activities such as “playing in the sand”, “beach combing,” “sunbathing,” and “volleyball.” Additionally, activities such as “enjoying the view,” “reading,” and kite flying” are categorized as “Sand” because they are typically “open space” activities in which the participants are neither actively traveling (such as hiking or cycling), taking part in consumptive activities (such as dinning or shopping), nor are they taking part in “water” activities.

In addition to the above activities “fishing” is categorized as a “Sand” activity. In the dataset “fishing” is limited to “shore or pier fishing” and is therefore does not include fishing from a boat. Additionally, site attribute variables in the dataset may not be sufficient to estimate a “fishing” site choice model as the available attributes do not include water depth, fish species, or catch rate data. Nevertheless, shore and pier anglers do recreate directly on the beach and thus we believe that many of our beach attributes may help explain their site choices. In this light, the most appropriate categorization for “shore or pier fishing” is “Sand.”

Recreational site choice for participants of “Sand” activities are expected to be sensitive to attribute levels for beach amenities such as the rockiness and sandiness of the beach, the availability of facilities such as fire pits, volleyball courts, and piers, and the level of coastal development.

Pavement

The “Pavement” activity category includes both activities that are pedestrian in nature and those utilizing bike paths or sidewalks. Some of these activities such as cycling and roller-skating are limited to taking place on paved sidewalks or bike paths, while others such as “walking,” “hiking” and “jogging” can either take place on bike paths or on the sand. However a major similarity among the activities in the “Pavement” category is that they are not limited to being done at the beach –they can be carried out elsewhere.

We also include in the “Pavement” segment those activities that are consumption based, but not necessarily beach related, such as “shopping”, “dinning”, and “visiting amusement parks.” “Pavement” activities are different from “Water” or “Sand” based activities in that the actual choice set extends beyond the beach and it is expected that the beach goer taking part in a “Retail” activity will respond to the site attributes differently. Additionally it is noted that both the site attribute list and the choice set are incomplete for a complete “retail” choice model in Southern California. Through the separation of “Retail” based beach trips from “Water” and “Sand” better estimates of attribute coefficients are expected.

Other

A small number of respondents listed activities that were not easily classified into any of the above groups. These responses are assigned to the “Other” segment. For a list of what specific activities were assigned to each general category refer to Table 7.

Assigning An Activity To A Trip

The computer assisted telephone interview (CATI) software used to conduct the telephone surveys permits the interviewer to document up to four activities associated with each beach trip recorded. In fact, in roughly half of the trip observations the respondent only reported one activity -- out of 5411 beach trip observations, 2636 (49%) responses recorded one activity. However, 1687 (32%) responses recorded two activities, 763 (14%) recorded three activities, and 325 (6%) recorded four activities. When more than one activity is reported, it may be that some or all of these responses fall into the same broad activity category described above, in which case there is no problem in assigning an activity to the trip. In many cases, however, we still have multiple broad activity categories associated with a trip. To assign an activity for each trip, we use a hierarchical classification.

In the case of beach related activities, we created a hierarchical ordering of activities by ranking the broad activity categories according to their expected order of importance to the beach choice decision. A second consideration in assigning an activity to a trip is how well the site choice model is likely to explain the beach choice conditional on the activity chosen. For certain activities, our ability to forecast the choice of a site conditional on that activity is limited by fact that our data does not contain a full set of attributes relevant for that specific activity, nor do we have data on the full choice set for that activity. In the case of Pavement, for example, there are other potential locations besides the beaches to ride a bike, but we do not have attribute data about these sites. This inevitably constrains our ability to model site choice when retail is the target activity. In turn, this might also influence whether we should be to classify a trip as a Pavement trip when other activities are also conducted in the same trip. Similarly, with activities such as dining, running, and cycling, these activities can be conducted at other locations in the Los Angeles area besides the beaches covered by our data. We are working with an incomplete choice set compared to the situation that exists with sunbathing, say, or swimming in the ocean. We place more emphasis on sand- or water-related activities when assigning an activity to a trip. This logic leads to the following hierarchy when assigning an activity to a trip.

- Activities that involve some degree of getting into the "Water" are placed at the top of the activity hierarchy. This hierarchy is based largely on the hypothesis that in general beach goers participating in "water" based activities are more selective regarding their recreation site than others. Those who get in the water will have different preferences regarding water quality and other attributes than those beach goers who remain dry.
- Following the "Water" category, the available site attributes are best suited for explaining the recreational site choice for those participating in "Sand" activities.
- "Pavement" activities are ranked third, as the ability to take part in "Pavement" activities such as running are believed to be large draws for beach visitation.

TABLE 7: Activity Choices

	Hierarchical Categories	Count	Count	Count	Count
		Activity 1	Activity 2	Activity 3	Activity 4
Boating	W	27	2	0	0
Body boarding/body surfing/skimboarding	W	249	69	21	9
Canoeing	W	12	1	0	0
Jet boating/jet skiing/personal water craft	W	0	37	9	6
Kayaking	W	20	4	4	0
Sailing	W	5	0	0	0
Scuba diving	W	0	1	0	0
Snorkeling	W	2	14	0	0
Splashing in water	W	75	59	2	4
Surfing	W	418	64	2	0
Swimming	W	291	170	54	4
Wading	W	64	98	11	2
Water skiing	W	2	1	0	0
Windsurfing / boardsailing	W	1	1	0	0
Activities with children	S	111	114	40	21
Bar-b-q	S	19	21	20	6
Beachcombing	S	9	28	26	0
Enjoying the view	S	135	138	51	18
Fishing (shore or pier)	S	69	20	1	0
Frisbee	S	29	39	28	2
Kite flying	S	10	10	1	0
People watching	S	93	100	63	29
Picnicking	S	137	120	38	14
Played in the sand	S	14	26	9	6
Reading	S	28	45	32	0
Relaxing	S	1	1	2	0
Sand football/soccer	S	15	25	2	15
Sleeping	S	1	1	0	0
Sunbathing	S	305	156	88	12
Volleyball	S	91	160	31	1
Walking the dog	S	65	15	5	0
Watched fireworks	S	21	19	1	0
Bicycling	P	444	64	3	10
Hiking	P	7	0	9	5
Jogging	P	343	50	13	19
Rollerblading/roller skates	P	165	18	52	2
Walking	P	1478	450	155	12
Amusement park/ arcade	P	7	5	1	0
Eating/ drinking	P	31	53	11	4
Shopping/dining	P	207	278	109	17
Other	O	414	299	194	107
Total		5415	2776	1088	325

In the activity choice nest, we model the probability of choosing an activity as

$$\Pr(activity_{iat}) = \frac{e^{(s_{it}\alpha_a \mathbf{I1}_a \gamma)}}{\sum_{a=1}^{3} e^{(s_{it}\alpha_a \mathbf{I1}_a \gamma)}}$$

Where S_{it} is a vector consisting of demographic attributes of the respondent, some wave-specific constants, and the inclusive value from the first nest, I1. The coefficient $á_a$ is normalized to zero for the alternative of pavement-based activities, except that the coefficient on the inclusive value is the same for all three alternative activity types. The inclusive values are calculated from the results of the beach choice nest for the estimation of the activity nest of the model based on the usual formula

$$\text{inclusive value} = \ln\left(\sum_{j=1}^{J} e^{\mathbf{x}_j \beta}\right)$$

The inclusive values are calculated for each activity using the full choice set for that activity. These values are calculated for each individual and each wave, provided that cost, location and the explanatory variable data needed in the second level of the model were available. The inclusive value covariate is equivalent to the expected utility from the beach choices for each activity type. Note that we can calculate this expected utility even for respondents who did not report any trips.

The second level is estimated for all single-destination trips that could be classified as to activity type and where the user supplied income and location data to calculate costs, as well as all of the demographic variables needed to estimate this level and the top level of the model.

The Participation Nest

The final nest models the level of participation (in beach related activities) of the respondents. We include all observations where people a) reported trip counts for a month and b) supplied the income, location, and other covariates needed for all three levels of the model. Specifically, the participation model is a repeated logit model of participation for each month. We include all observations where people reported trip counts for a month and supplied the income, location, and other covariates needed for all three levels of the model. The probability of observing exactly k trips is given by:

$$\Pr(k) = \left(\frac{N!}{k!*(N-k)!}\right) * P^k * (1-P)^{(N-k)}$$

$$\Pr(trip) = P = \frac{e^{Z\delta}}{1+e^{Z\delta}}$$

where Z includes the inclusive value, I2 (expected utility) from the activity-choice nest, and other explanatory variables. P gives the probability of visiting a beach on a single choice occasion, and N is the number of days (choice occasions) in each month.

Inclusive values for the activity level are calculated using the same formula as before\

$$\mathbf{I2} = \ln(\sum_{a=1}^{3} e^{(S_{it}\alpha_a + \mathbf{I1}_a \gamma)})$$

The Results of the Nested Model

The results for the three-level nested repeated logit are given in Table 8. Note first that not all explanatory attributes enter into every submodel. Many attributes that are logical explanatory variables for beach choice conditioned upon a water-based activity are not logical explanatory variables for sand or pavement based activities. Second, note that our primary explanatory variables were collected with the intent of explaining beach choice by those who intended to undertake water based activities. For this reason, our water-based activity submodels (both the beach choice and activity choice models) are more fully specified and have greater explanatory power. In the discussion below, we focus primarily on the results of the model which pertain to water quality and beach choice by those that undertake water based activities.

The coefficients for travel cost and water quality are significant and have the expected sign and magnitude (see the welfare estimates in the following section). The coefficients of each submodel are designated by a capital letter indicating the activity type for which the coefficients correspond (e.g. W=water, S= sand, P=pavement). A number following the activity letter designation indicates that the coefficient is specific to a particular wave in the model (e.g. W1=water, wave 1). The coefficients on water quality (WHTB Yr) are positive for water-based activities, but not significantly different from zero for sand-based activities. A water quality variable did not enter into the right hand side of the “Pavement” sub-model because we had no reason to believe that water quality should matter to beach goers that did not go on the sand or get in the water. Note that combining water users with other users would have diminished our ability to detect the preference placed on clean water by swimmers and surfers. Within water-based activities, the preferences for water quality are slightly diminished for wave 2. While Development and Harbors proved to be utility degrading for water users, Development was a desirable attribute for sand users and pavement users preferred the presence of both Development and Harbors to their absence. The other coefficient estimates are generally in keeping with our intuition.

From a policy perspective, we would like to understand what factors determine the choice of water-based activities and how changes in water quality influence the choice to participate in water-based activities or substitute to other types of beach activities (especially when water quality becomes degraded). The model results indicate that seasonality is an important determinant of the decision to undertake water-based activities. All users are more likely to choose water based activities in the Spring, Summer, and early Fall. Race also plays a factor in the choice of beach activities. Black respondents were less likely overall to choose a water-based activity, while Hispanics were not significantly different from others (e.g. whites, Asians, and Native Americans) in their choice of water-based activities. Males were more likely to get in the water. Interestingly, age was not found to be an important factor in influencing the choice of a water-based activity, but families with children were more likely to participate in sand-based activities. Finally, all potential beach goers were more likely to go to the beach in the summer. Blacks, Hispanics, students, and households with children were less likely to go to the beach than others. Males and those with only part-time employment were more likely to visit area beaches.

Table 8: Three Nested Beach Choice, Activity Choice, and Participation Model
Time Valuation = 50% wage rate

BEACH CHOICE MODELS

Note, the first letters W,S, and P indicate coefficients that apply to the Water, Sand, and Pavement activity submodels. Numbers in the second position indicate coefficients that apply to the wave of that number.

	Mean log-likelihood		-2.73638	
	Observations		4545	

	Estimates	Standard Error	T-statistic	P-value
Cost, water	-0.0734	0.0050	-14.7110	0.0000
Cost, sand	-0.0982	0.0055	-17.8390	0.0000
Cost, pavement	-0.1164	0.0078	-14.9170	0.0000
	Water-based Activity Beach Choice Model			
	Water activities beach choice model, variables that affect all waves			
WHTB Yr	0.4158	0.0802	5.1810	0.0000
Wln(Length)	1.4381	0.0825	17.4230	0.0000
Wugly	-0.4670	0.0656	-7.1160	0.0000
WDevelop2	-0.4049	0.0763	-5.3040	0.0000
WWild	0.8046	0.1454	5.5320	0.0000
WLifeguard/length	0.2761	0.0283	9.7540	0.0000
Wsandy	2.0235	0.2249	8.9960	0.0000
Wsurfing	0.5366	0.1327	4.0460	0.0001
Wdiving	0.6735	0.1150	5.8550	0.0000
Wharbor	-1.1274	0.0990	-11.3900	0.0000
WOceanside dummy	6.7932	0.4580	14.8320	0.0000
Wvenice dummy	3.7928	0.3330	11.3910	0.0000
	Water activities beach choice model, variables for specific waves			
W2HTB Yr	-0.2494	0.0979	-2.5470	0.0109
W4HTB Yr	-0.0489	0.0953	-0.5130	0.6083
W4Lifeguard/length	0.1355	0.0389	3.4850	0.0005
W5HTB Yr	0.1079	0.1260	0.8560	0.3921

Table 8 (continued): Three Nested Beach Choice, Activity Choice, and Participation Model : Time Valuation = 1/2 wage rate

	Sand-Based Activities Beach Choice Model			
	Estimates	**Standard Error**	**T-statistic**	**P-value**
SHTB Yr	-0.0616	0.0587	-1.0490	0.2944
Sln(Length)	0.9255	0.0730	12.6810	0.0000
SUgly	-0.4312	0.0681	-6.3280	0.0000
SDevelop	0.5703	0.1262	4.5190	0.0000
SWild	1.0007	0.1086	9.2170	0.0000
SVolleyball/length	0.0028	0.0047	0.5980	0.5501
SLifeguard/length	0.4067	0.0276	14.7590	0.0000
SHarbor	-0.5332	0.0775	-6.8800	0.0000
SSandy	0.3518	0.2269	1.5510	0.1210
SPlayground	-0.1681	0.0819	-2.0510	0.0403
SRestroom	0.5587	0.2140	2.6110	0.0090
SFirepit/length	0.0000	0.0021	-0.0190	0.9850
SOceanside dummy	3.1894	0.5500	5.7990	0.0000
SVenice dummy	3.5158	0.2870	12.2510	0.0000
S4Volleyball/length	-0.0146	0.0080	-1.8260	0.0679
S5Volleyball/length	-0.0080	0.0091	-0.8790	0.3792
	Pavement-Based Activities Beach Choice Model			
Pln(Length)	1.7028	0.1178	14.4580	0.0000
PUgly	-0.6110	0.0791	-7.7230	0.0000
PDevelop2	-0.4309	0.1252	-3.4430	0.0006
PWild	0.4471	0.1405	3.1810	0.0015
PLifeguard/length	0.6228	0.0288	21.5930	0.0000
PParking	-0.6932	0.3054	-2.2700	0.0232
PPubfac	-0.5484	0.1516	-3.6170	0.0003
PSandy	0.5309	0.4908	1.0820	0.2793
PShowers	1.1947	0.2192	5.4500	0.0000
PStrparking	1.2397	0.2294	5.4040	0.0000
PHarbor	0.2207	0.0860	2.5650	0.0103
PNature	0.7228	0.1946	3.7130	0.0002
PRivers	0.8127	0.2632	3.0880	0.0020
PBikepath	0.1631	0.1414	1.1530	0.2489
PCamping	-2.2342	0.2942	-7.5950	0.0000
PPlayground	0.0727	0.0929	0.7830	0.4338
PRestrooms	0.2596	0.3984	0.6520	0.5146
PSidewalk	0.3584	0.1303	2.7510	0.0059
PRentals	-0.5841	0.1245	-4.6920	0.0000
PPointMugu dummy	5.8516	0.5881	9.9510	0.0000
POceanside dummy	5.9621	1.0794	5.5240	0.0000
PVenice dummy	6.4541	0.4756	13.5700	0.0000

Table 8 (continued): Three Nested Beach Choice, Activity Choice, and Participation Model : Time Valuation = 50% wage rate

ACTIVITY CHOICE MODEL				
	Mean log-likelihood		-1.01975	
	Observations		4837	
	Estimates	**Standard Error**	**T-statistic**	**P-value**
	Variables Affecting The Choice Of Water-Based Activities.			
I1	0.3072	0.0655	4.6920	0.0000
WConstant	-1.8684	0.2146	-8.7070	0.0000
WMale	0.4058	0.0709	5.7200	0.0000
WBlack	-1.0335	0.2467	-4.1890	0.0000
WHispanic	-0.1000	0.0929	-1.0760	0.2817
W3Constant	1.3466	0.1347	9.9940	0.0000
W4Constant	1.9405	0.1683	11.5280	0.0000
W5Constant	1.2422	0.1882	6.6000	0.0000
W6Constant	0.6229	0.1435	4.3400	0.0000
SConstant	-0.0028	0.1829	-0.0150	0.9877
SKids	0.2064	0.0730	2.8260	0.0047
S3Constant	0.0479	0.1273	0.3770	0.7065
S4Constant	1.1758	0.1136	10.3520	0.0000
PARTICIPATION MODEL				
	Mean log-likelihood		-3.89772	
	Observations		7686	
I2	0.3835	0.0849	4.5160	0.0000
Constant	-4.6445	0.1438	-32.3060	0.0000
Male	0.4350	0.0601	7.2350	0.0000
Kids	-0.1602	0.0578	-2.7700	0.0056
Student	-0.1823	0.0748	-2.4370	0.0148
Workparttime	0.1873	0.0879	2.1300	0.0331
Black	-0.6761	0.1338	-5.0520	0.0000
Hispanic	-0.5655	0.0746	-7.5830	0.0000
Summer	0.2089	0.0895	2.3350	0.0196

WELFARE ESTIMATES

Welfare estimates for changes in beach availability and/or beach attributes are being calculated separately and extrapolated to the entire population of beach goers in our four-county study area. The results of the welfare calculations will be presented in a separate report.

CONCLUSION

Despite anecdotal evidence to the contrary, coastal marine water quality is an important factor in determining when and where Southern Californians go to the beach. The decision about when and where to go the beach in California is complex and depends on many factors in addition to water quality, including the natural and managed attributes of beaches, the cost of getting to the beach, the activities one plans to undertake at the beach, the personal characteristics of the beach goer, and the season during which the choice takes place. Our analysis indicates that it is not possible to isolate the effect of water quality on the choice of a beach site unless this is simultaneously modeled along with other important components of beach choice such as the choice of beach activity and seasonal participation in beach recreation. We have shown the value of adopting a comprehensive approach to modeling beach behavior by first presenting the results of simpler model of beach choice and then contrasting this with a more complex, multi-nested model. The simpler models do not properly capture the impact of water quality on beach choice and thus on the economic welfare of beach goers. To address the complex nature of beach choice in Southern California, we use a three-level nested random utility model of beach choice to simultaneously model how the beach going public chooses beaches, the activities they undertake at the beach, and whether or not to go to a beach. This three level model allows us to estimate the ways in which changes in beach water quality, beach attributes, and beach closures impact beach goers; specifically we can estimate changes in attendance at the 51 principle beaches of Los Angeles and Orange Counties, changes in expenditures, and changes in the economic well-being of beach goers. This model can serve as the foundation for policy decisions regarding beach water quality management, oil spill prevention, and the determination of fines for events that result in the impairment of coastal water quality or the closure of beaches.

REFERENCES

Adamowicz, W.L., J. Swait, P.C. Boxall, J. Louviere, and M. Williams. 1997. Perceptions versus objective measures of environmental quality in combined revealed and stated preference models of environmental valuation. Journal of Environmental Economics and Management 32:65-84.

Arcidiacono, P. and J.B. Jones. 2003. Finite Mixture Distributions, Sequential Likehihood and the EM Algorithm. Econometrica. Vol 71, no. 3, pp 933-946 Ben-

Akiva, Moshe, and Lerman, Steven R. (1985) Discrete Choice Analysis: Theory and Application to Travel Demand, Cambridge: The MIT Press.

Berman, Matthew D, and Hong Jin Kim. 1999. "Endogenous On-Site Time in the Recreation Demand Model," Land Economics, 75 (4), November, 603-19.

Bhat, Chandra R. 1998. "Accomodating Flexible Substitution Patterns in Multi-Dimensional Choice Modeling: Formulation and Application to Travel Mode and Departure Tine Choice." *Transportation Research B,* 32(7):455-466

Bockstael, N.E., McConnell, K.E., and I.E. Strand, (1989). "A random utility model for sport fishing: some preliminary results for Florida: Marine Resource Economics. 6:245-260.

Boxall, P.C. and W.L. Adamowicz. 2002. Understanding heterogeneous preferences in random utility models: A latent class approach. Environmental and Resource Economics. Environmental and Resource Economics 23(4):421-446.

Breffle, W.S. and E.R. Morey. 2000. Investigating Preference Heterogeneity in a Repeated Discrete-Choice Recreation Demand Model of Atlantic Salmon Fishing. Marine Resource Economics, Vol 15, 1-20, 2000. (An earlier versions was presented at the 1999 W- 133 meetings in Tucson)

California Coastal Commision. 1997. California Beach Access Guide. E. Caughmann and J. Ginsberg, eds., University of California Press, Berkeley

Chapman, David and Michael Hanemann. 2001. Environmental damages in court: the American Trader case, in The Law and Economics of the Environment, Anthony Heyes, Editor, pp. 3 19-367

Desvousges, W.H., and S.M. Waters. 1995. Volume III: Report on Potential Economic Losses Associated with Recreation Services in the Upper Clark Fork River Basin. Submitted to United States District Court, District of Montana, Helena Division in the matter of State of Montana v. Atlantic Richfield Company. Case No. CV-83-31 7-HLN-PGH.

Desvousges, W.H., S.M. Waters, and K.E. Train. 1996. Supplemental Report on Potential Economic Losses Associated with Recreation Services in the Upper Clark Fork River Basin. Durham, NC: Triangle Economic Research.

Feather, Peter M. 1994. Sampling and Aggregation Issues in Random Utility Model Estimation, American Journal of Agricultural Economics, 76(4), (November), 772-780.

Grijalva, T., Berrens, R., Bohara, A., Jakus, P. and W.D. Shaw. 2002. Valuing the Loss of Rock Climbing Access. Land Economics. February. 78(1):103-120

Haab, T,C,, and K.E. McConnell. 2002. Valuing Environmental and Natural Resources: The Econometrics of Non-Market Valuation. Edward Elgar Publishing, Massachusetts.

Haab, Timothy C. and Hicks, Robert L. (1999) Choice Set Considerations in Models of Recreation Demand: History and Current State of the Art, Marine Resource Economics, 14(4), (Winter), 271-281.

Hauber, A. B. and G. R. Parsons, (2000). "The Effect of Nesting Structure Specification on Welfare Estimation in a Random Utility Model of Recreation Demand: An Application to the Demand for Recreational Fishing." American Journal of Agricultural Economics 82(3): 501 - 514.

Hensher, D. and W. Greene. 2002. Specification and estimation of the nested logit model: alternative normalizations. Transportation Research B - Methodological, vol.36:1, pp. 1-17.

Jones, Carol A. and Lupi, Frank (1999) The Effect of Modeling Substitute Activities on Recreational Benefit Estimates, Marine Resource Economics, 14(4), (Winter), 357-374.

Kaoru, Y. 1995. Measuring marine recreation benefits of water quality improvements by the nested random utility model. Resource and Energy Economics 17(2):1 19-136.

Kaoru, Yoshiaki; Smith, V. Kerry and Liu, Jin Long (1995) Using Random Utility Models to Estimate the Recreational Value of Estuarine Resources, American Journal of Agricultural Economics, 77(1), (Feb.), 141-151.

King, P. and Symes, D. 2003. The Potential Loss in Gross National Product and Gross State Product from a Failure to Maintain California’s Beaches. A report prepared for the California Department of Boating and Waterways. San Francisco State University

Layton, David. 2000. Random coefficients models of stated preference surveys. Journal of Environmental Economics and Management. 40:21-26.

Lupi, Frank and Feather, Peter M. (1998) Using Partial Site Aggregation to Reduce Bias in Random Utility Travel Cost Models, Water Resources Research, 34(12), (December), 3595-3603.

Morey, E.R. and K.G. Rossman. 2003. Using Stated-Preference Questions to Investigate Variations in Willingness to Pay for Preserving Marble Monuments: Classic Heterogeneity, Random Parameters, and Mixture Models. Journal of Cultural Economics, Vol. 27 (3/4), 2-5-229, November 2003.

McNair, Douglas J. and Cox, Sean D. (1999) A Heteroskedastic Nested RUM of Freshwater Fishing, Marine Resource Economics, 14(4), (Winter), 333-341. McConnell, K. E. 1992. On-Site Time in the Demand for Recreation. American Journal of Agricultural Economics, 1992, 918-925.

McFadden, D. 1986. The choice theory approach to market research. Marketing Science Vol. 5, pp. 275-297.

McFadden, Daniel. “Measuring Willingness-to-Pay for Transportation Improvements.” Dept. of Economics, UC Berkeley, July 1, 1997

Morey, E.R., R.D. Row, and M. Watson. 1993. A repeated nested-logit model of Atlantic salmon fishing. American Journal of Agricultural Economics 75:578-592.

Morey, E.R. 1999. TWO RUMS UNCLOAKED: A Nested Logit Model of Site Choice, and A Nested Logit Model of Participation and Site Choice. Chapter 4 in Valuing the Environment Using Recreation Demand Models, (C.L. Kling and H. Herriges,

Eds.), Edward Elgar Publishing Ltd. An earlier version was presentation at the 1994 W-133 meetings, Tucson, Arizona.
Parsons, George R. 2003. The Travel Cost Model, Chapter 9, in A Primer on Nonmarket Valuation, Champ, Patty, Boyle, Kevin and Brown, Tom (eds.) Kluwer.
Parsons, George R. and Needelman, Michael S. (1992) Site Aggregation in a Random Utility Model of Recreation, Land Economics, 68(4), (Nov.), 418-433.
Parsons, George R.; Plantinga, Andrew J. and Boyle, Kevin J. (2000) Narrow choice sets in a random utility model of recreation demand, Land Economics, 76(1), (Feb.), 86-99.
Pendleton, L. 2003. Estimating the Regional Economic Benefits of Improvements in the California Coastal Ocean Observing System. A report prepared for Ocean.us. Arlington, VA.
Pendleton, L. 2004. Harnessing Ocean Technology to Improve Beach Management. Prepared for Ocean.US. Washington, D.C.
Provencher, B. and R.C. Bishop. 1997. An Estimable Dynamic Model of Recreation Behavior with an Application to Great Lakes Angling. Journal-of-Environmental-Economic s-and-Management. June; 33(2): 107-27
Provencher, B., K.A. Baerenklau, and R.C. Bishop. 2002. A Finite Mixture Logit Model of Recreational Angling with Serially Correlated Random Utility. American Journal of Agricultural Economics 84:1066-1075.
Roberts, John H. and Lattin, James M. (1997) Consideration: Review of research and prospects for future insights, Journal of Marketing Research, 34(3), (Aug.), 406-410.
Shocker, Allan D.; Ben-Akiva, Moshe; Boccara, Bruno and Nedungadi, Prakash (1991) Consideration Set Influences on Consumer Decision-Making and Choice: Issues, Models, and Suggestions, Marketing Letters, 2(3), 181-1971.
Smith, V. Kerry, William H. Desvousges, and Matthew P. McGivney, 1983. The Opportunity Cost of Travel Time in Recreational demand Models. Land Economics, 59 (3), August 1983, 259-78.
Swait, J. , Adamowicz, V. and van Buren, M. (2004) Choice and Temporal Welfare Impacts: Incorporating History into Discrete Choice Models, Journal of Environmental Economics and Management, 47(1):94-1 16.
Thill, Jean-Claude (1992) Choice set formation for destination choice modeling, Progress in Human Geography, 16(3), 361-382 .
Train, K., 1998. Recreation Demand Models with Taste Differences over People. Land Economics, Vol. 74, No. 2, pp. 230-239
Train, K. 2003 Discrete Choice Methods with Simulation, Cambridge. Cambridge University Press.
Train, K. and D. McFadden. 2000. Mixed MNL Models for Discrete Response. : Journal-of-Applied-Econometrics. Sept.-Oct. 2000; 15(5): 447-70
Truong, T. and David A. Hensher. 1985. Measurement of Travel Time and Opportunity Cost from a Discrete-Choice Model. The Economic Journal, 95(378):438-451
Yatchew, A. and Griliches, Z. 1984. Specification error in probit models. Review of Economics and Statistics, Vol. 66, pp. 134-139.

Appendix I: Estimated Models with Alternative Values of Time (0%, 33% and 100 of the wage rate)

Table A1: Three Nested Beach Choice, Activity Choice, and Participation Model Time Valuation = Zero

BEACH CHOICE MODELS

Note, the first letters W,S, and P indicate coefficients that apply to the Water, Sand, and Pavement activity submodels. Numbers in the second position indicate coefficients that apply to the wave of that number.

	Mean log-likelihood		-2.69349	
	Observations		4545	

	Estimates	Standard Error	T-statistic	P-value
Cost, water	-0.3435	0.0416	-8.255	0.0000
Cost, sand	-0.3802	0.0180	-21.116	0.0000
Cost, pavement	-0.5048	0.0708	-7.130	0.0000
	Water-based Activity Beach Choice Model			
	Water activities beach choice model, variables that affect all waves			
WHTB Yr	0.3330	0.1281	2.600	0.0093
Wln(Length)	1.4317	0.0802	17.851	0.0000
Wugly	-0.5540	0.0711	-7.787	0.0000
WDevelop2	-0.5227	0.1024	-5.105	0.0000
WWild	0.8256	0.1525	5.414	0.0000
WLifeguard/length	0.2792	0.0288	9.679	0.0000
Wsandy	2.2092	0.2311	9.561	0.0000
Wsurfing	0.4984	0.1722	2.894	0.0038
Wdiving	0.8475	0.1534	5.525	0.0000
Wharbor	-1.2916	0.1153	-11.200	0.0000
WOceanside dummy	6.9491	0.5565	12.487	0.0000
Wvenice dummy	3.7657	0.3734	10.085	0.0000
	Water activities beach choice model, variables for specific waves			
W2HTB Yr	-0.2171	0.1650	-1.316	0.1882
W4HTB Yr	0.0364	0.2261	0.161	0.8722
W4Lifeguard/length	0.1443	0.0395	3.653	0.0003
W5HTB Yr	0.2125	0.1841	1.154	0.2485

Table A1 (continued): Three Nested Beach Choice, Activity Choice, and Participation Model : Time Valuation = Zero

	Estimates	Standard Error	T-statistic	P-value
	Sand-Based Activities Beach Choice Model			
SHTB Yr	-0.0416	0.0684	-0.609	0.5427
Sln(Length)	0.9066	0.2379	3.811	0.0001
SUgly	-0.4998	0.0807	-6.191	0.0000
SDevelop	0.6488	0.1371	4.732	0.0000
SWild	1.0998	0.5462	2.014	0.0441
SVolleyball/length	0.0011	0.0112	0.098	0.9220
SLifeguard/length	0.3904	0.0273	14.307	0.0000
SHarbor	-0.6169	0.0805	-7.660	0.0000
SSandy	0.4701	1.2146	0.387	0.6987
SPlayground	-0.2155	0.2958	-0.728	0.4663
SRestroom	0.4926	0.7341	0.671	0.5023
SFirepit/length	0.0019	0.0029	0.652	0.5145
SOceanside dummy	3.4020	2.7093	1.256	0.2092
SVenice dummy	3.3827	2.0498	1.650	0.0989
	Sand activities beach choice model, variables for specific waves			
S4Volleyball/length	-0.0137	0.0113	-1.214	0.2247
S5Volleyball/length	-0.0080	0.0151	-0.532	0.5948
	Pavement-Based Activities Beach Choice Model			
Pln(Length)	1.7612	0.2730	6.451	0.0000
PUgly	-0.6885	0.0999	-6.892	0.0000
PDevelop2	-0.5581	0.4839	-1.153	0.2487
PWild	0.5578	0.7549	0.739	0.4600
PLifeguard/length	0.6194	0.0685	9.041	0.0000
PParking	-0.7418	1.6210	-0.458	0.6472
PPubfac	-0.5985	0.3425	-1.747	0.0806
PSandy	0.7614	2.7042	0.282	0.7783
PShowers	1.2062	0.7196	1.676	0.0937
PStrparking	1.2679	0.5622	2.256	0.0241
PHarbor	0.2098	0.1210	1.734	0.0830
PNature	0.8203	0.7769	1.056	0.2910
PRivers	0.8504	1.2273	0.693	0.4884
PBikepath	0.0521	0.5446	0.096	0.9238
PCamping	-2.3191	0.5639	-4.113	0.0000
PPlayground	0.0857	0.1852	0.463	0.6434
PRestrooms	0.0933	0.4024	0.232	0.8167
PSidewalk	0.3298	0.1406	2.346	0.0190
PRentals	-0.4494	0.5822	-0.772	0.4401
PPointMugu dummy	5.1530	5.2537	0.981	0.3267
POceanside dummy	7.2516	4.4098	1.644	0.1001
PVenice dummy	6.2591	4.4703	1.400	0.1615

Table A1 (continued): Three Nested Beach Choice, Activity Choice, and Participation Model : Time Valuation = Zero

ACTIVITY CHOICE MODEL				
	Mean log-likelihood		-1.01975	
	Observations		4837	
	Estimates	**Standard Error**	**T-statistic**	**P-value**
	Variables Affecting The Choice Of Activities.			
I1	0.3712	0.2728	1.361	0.1736
WConstant	-1.7894	1.8450	-0.970	0.3321
WMale	0.4066	0.0813	5.001	0.0000
WBlack	-1.0503	0.2503	-4.196	0.0000
WHispanic	-0.1365	0.1000	-1.365	0.1723
W3Constant	1.3009	0.1775	7.331	0.0000
W4Constant	1.6919	0.3984	4.247	0.0000
W5Constant	1.0664	0.4689	2.274	0.0229
W6Constant	0.5409	0.1933	2.798	0.0051
SConstant	0.0116	3.3651	0.003	0.9973
Skids	0.1976	0.1152	1.716	0.0862
S3Constant	0.0152	0.9444	0.016	0.9872
S4Constant	1.1061	0.1430	7.735	0.0000
PARTICIPATION MODEL				
	Mean log-likelihood		-3.89772	
	Observations		7686	
I2	0.4940	0.3092	1.598	0.1101
Constant	-5.0187	0.8306	-6.042	0.0000
Male	0.3808	0.0661	5.763	0.0000
Kids	-0.1537	0.0642	-2.395	0.0166
Student	-0.1387	0.0746	-1.859	0.0630
Workparttime	0.2069	0.0873	2.370	0.0178
Black	-0.5940	0.1520	-3.908	0.0001
Hispanic	-0.4529	0.0739	-6.125	0.0000
Summer	0.1457	0.1785	0.816	0.4144

Table A2: Three Nested Beach Choice, Activity Choice, and Participation Model Time Valuation = 33.33% wage rate

BEACH CHOICE MODELS

Note, the first letters W,S, and P indicate coefficients that apply to the Water, Sand, and Pavement activity submodels. Numbers in the second position indicate coefficients that apply to the wave of that number.

	Mean log-likelihood		-2.72809	
	Observations		4545	

	Estimates	Standard Error	T-statistic	P-value
Cost, water	-0.1017	0.0066	-15.3940	0.0000
Cost, sand	-0.1323	0.0070	-18.9620	0.0000
Cost, pavement	-0.1589	0.0099	-16.0380	0.0000
	Water-based Activity Beach Choice Model			
	Water activities beach choice model, variables that affect all waves			
WHTB Yr	0.4057	0.0834	4.8650	0.0000
Wln(Length)	1.4358	0.0820	17.5080	0.0000
Wugly	-0.4787	0.0660	-7.2520	0.0000
WDevelop2	-0.4253	0.0764	-5.5670	0.0000
WWild	0.7975	0.1470	5.4250	0.0000
WLifeguard/length	0.2771	0.0286	9.7040	0.0000
Wsandy	2.0611	0.2253	9.1470	0.0000
Wsurfing	0.5232	0.1325	3.9500	0.0001
Wdiving	0.7234	0.1164	6.2130	0.0000
Wharbor	-1.1608	0.0993	-11.6850	0.0000
WOceanside dummy	6.8599	0.4581	14.9740	0.0000
Wvenice dummy	3.7893	0.3324	11.4010	0.0000
	Water activities beach choice model, variables for specific waves			
W2HTB Yr	-0.2377	0.0971	-2.4480	0.0144
W4HTB Yr	-0.0354	0.1076	-0.3290	0.7422
W4Lifeguard/length	0.1372	0.0392	3.5000	0.0005
W5HTB Yr	0.1181	0.1287	0.9180	0.3586

Table A2 (continued): Three Nested Beach Choice, Activity Choice, and Participation Model : Time Valuation = 33.33% wage rate

	Sand-Based Activities Beach Choice Model			
	Estimates	**Standard Error**	**T-statistic**	**P-value**
SHTB Yr	-0.0617	0.0580	-1.0640	0.2875
Sln(Length)	0.9207	0.0730	12.6160	0.0000
SUgly	-0.4450	0.0684	-6.5030	0.0000
SDevelop	0.5750	0.1275	4.5090	0.0000
SWild	1.0163	0.1106	9.1860	0.0000
SVolleyball/length	0.0022	0.0047	0.4660	0.6415
SLifeguard/length	0.4049	0.0275	14.7450	0.0000
SHarbor	-0.5519	0.0777	-7.1000	0.0000
SSandy	0.3743	0.2333	1.6050	0.1086
SPlayground	-0.1716	0.0823	-2.0860	0.0370
SRestroom	0.5520	0.2142	2.5770	0.0100
SFirepit/length	0.0001	0.0021	0.0370	0.9704
SOceanside dummy	3.2238	0.5692	5.6640	0.0000
SVenice dummy	3.4885	0.2913	11.9770	0.0000
S4Volleyball/length	-0.0146	0.0079	-1.8320	0.0670
S5Volleyball/length	-0.0080	0.0092	-0.8630	0.3884
	Pavement-Based Activities Beach Choice Model			
Pln(Length)	1.7078	0.1152	14.8240	0.0000
PUgly	-0.6324	0.0817	-7.7370	0.0000
PDevelop2	-0.4567	0.1314	-3.4750	0.0005
PWild	0.4271	0.1384	3.0870	0.0020
PLifeguard/length	0.6244	0.0287	21.7870	0.0000
PParking	-0.7205	0.3045	-2.3660	0.0180
PPubfac	-0.5271	0.1440	-3.6610	0.0003
PSandy	0.5058	0.4684	1.0800	0.2802
PShowers	1.1984	0.2233	5.3660	0.0000
PStrparking	1.2615	0.2322	5.4330	0.0000
PHarbor	0.2119	0.0865	2.4500	0.0143
PNature	0.7378	0.1896	3.8910	0.0001
PRivers	0.8530	0.2588	3.2950	0.0010
PBikepath	0.0945	0.1263	0.7480	0.4542
PCamping	-2.2419	0.3015	-7.4370	0.0000
PPlayground	0.0693	0.0902	0.7690	0.4419
PRestrooms	0.2099	0.4046	0.5190	0.6040
PSidewalk	0.3805	0.1280	2.9730	0.0029
PRentals	-0.5607	0.1162	-4.8260	0.0000
PPointMugu dummy	5.7086	0.5186	11.0090	0.0000
POceanside dummy	6.1926	1.0334	5.9930	0.0000
PVenice dummy	6.3319	0.3915	16.1750	0.0000

Table A2 (continued): Three Nested Beach Choice, Activity Choice, and Participation Model : Time Valuation = 33.33% wage rate

ACTIVITY CHOICE MODEL				
	Mean log-likelihood		-1.01975	
	Observations		4837	
	Estimates	**Standard Error**	**T-statistic**	**P-value**
	Variables Affecting The Choice Of Water-Based Activities.			
I1	0.3272	0.0688	4.7550	0.0000
WConstant	-1.9049	0.2136	-8.9160	0.0000
WMale	0.4068	0.0712	5.7100	0.0000
WBlack	-1.0352	0.2466	-4.1980	0.0000
WHispanic	-0.1021	0.0927	-1.1020	0.2705
W3Constant	1.3361	0.1349	9.9050	0.0000
W4Constant	1.8990	0.1810	10.4940	0.0000
W5Constant	1.2210	0.1996	6.1180	0.0000
W6Constant	0.6115	0.1442	4.2410	0.0000
SConstant	0.0048	0.1475	0.0330	0.9739
Skids	0.2038	0.0733	2.7810	0.0054
S3Constant	0.0418	0.1233	0.3390	0.7347
S4Constant	1.1671	0.1140	10.2350	0.0000
PARTICIPATION MODEL				
	Mean log-likelihood		-3.89772	
	Observations		7686	
I2	0.3995	0.0859	4.6520	0.0000
Constant	-4.7054	0.1429	-32.9300	0.0000
Male	0.4306	0.0600	7.1710	0.0000
Kids	-0.1539	0.0577	-2.6670	0.0077
Student	-0.1809	0.0746	-2.4250	0.0153
Workparttime	0.1882	0.0878	2.1430	0.0321
Black	-0.6758	0.1340	-5.0440	0.0000
Hispanic	-0.5641	0.0747	-7.5520	0.0000
Summer	0.1995	0.0902	2.2110	0.0270

Table A3: Three Nested Beach Choice, Activity Choice, and Participation Model
Time Valuation = 100% wage rate

BEACH CHOICE MODELS

Note, the first letters W,S, and P indicate coefficients that apply to the Water, Sand, and Pavement activity submodels. Numbers in the second position indicate coefficients that apply to the wave of that number.

Mean log-likelihood		-2.74775
Observations		4545

	Estimates	**Standard Error**	**T-statistic**	**P-value**
Cost, water	-0.0395	0.0028	-14.0130	0.0000
Cost, sand	-0.0548	0.0033	-16.4410	0.0000
Cost, pavement	-0.0639	0.0046	-13.8810	0.0000
	Water-based Activity Beach Choice Model			
	Water activities beach choice model, variables that affect all waves			
WHTB Yr	0.4286	0.0817	5.2440	0.0000
Wln(Length)	1.4412	0.0831	17.3410	0.0000
Wugly	-0.4523	0.0649	-6.9670	0.0000
WDevelop2	-0.3785	0.0762	-4.9680	0.0000
WWild	0.8160	0.1435	5.6850	0.0000
WLifeguard/length	0.2749	0.0278	9.9010	0.0000
Wsandy	1.9726	0.2239	8.8090	0.0000
Wsurfing	0.5559	0.1330	4.1790	0.0000
Wdiving	0.6036	0.1116	5.4110	0.0000
Wharbor	-1.0815	0.0981	-11.0300	0.0000
WOceanside dummy	6.6859	0.4575	14.6130	0.0000
Wvenice dummy	3.7987	0.3339	11.3750	0.0000
	Water activities beach choice model, variables for specific waves			
W2HTB Yr	-0.2661	0.1004	-2.6520	0.0080
W4HTB Yr	-0.0664	0.0989	-0.6710	0.5020
W4Lifeguard/length	0.1332	0.0386	3.4490	0.0006
W5HTB Yr	0.0959	0.1276	0.7520	0.4523

Table A3 (continued): Three Nested Beach Choice, Activity Choice, and Participation Model : Time Valuation = 100% wage rate

	Sand-Based Activities Beach Choice Model			
	Estimates	**Standard Error**	**T-statistic**	**P-value**
SHTB Yr	-0.0605	0.0576	-1.0500	0.2938
Sln(Length)	0.9323	0.0735	12.6830	0.0000
SUgly	-0.4100	0.0677	-6.0530	0.0000
SDevelop	0.5666	0.1254	4.5180	0.0000
SWild	0.9790	0.1098	8.9180	0.0000
SVolleyball/length	0.0039	0.0047	0.8300	0.4065
SLifeguard/length	0.4085	0.0272	15.0270	0.0000
SHarbor	-0.5048	0.0774	-6.5230	0.0000
SSandy	0.3187	0.2333	1.3660	0.1719
SPlayground	-0.1650	0.0823	-2.0050	0.0449
SRestroom	0.5641	0.2152	2.6210	0.0088
SFirepit/length	-0.0001	0.0021	-0.0680	0.9455
SOceanside dummy	3.1196	0.5437	5.7370	0.0000
SVenice dummy	3.5511	0.2917	12.1740	0.0000
S4Volleyball/length	-0.0147	0.0081	-1.8170	0.0692
S5Volleyball/length	-0.0081	0.0091	-0.8820	0.3779
	Pavement-Based Activities Beach Choice Model			
Pln(Length)	1.6977	0.1162	14.6120	0.0000
PUgly	-0.5784	0.0780	-7.4160	0.0000
PDevelop2	-0.3941	0.1284	-3.0700	0.0021
PWild	0.4849	0.1410	3.4380	0.0006
PLifeguard/length	0.6197	0.0290	21.3750	0.0000
PParking	-0.6520	0.3065	-2.1270	0.0334
PPubfac	-0.5851	0.1541	-3.7970	0.0001
PSandy	0.5822	0.4721	1.2330	0.2175
PShowers	1.1886	0.2161	5.5000	0.0000
PStrparking	1.2041	0.2246	5.3620	0.0000
PHarbor	0.2351	0.0867	2.7130	0.0067
PNature	0.7011	0.1868	3.7540	0.0002
PRivers	0.7490	0.2593	2.8890	0.0039
PBikepath	0.2726	0.1458	1.8700	0.0615
PCamping	-2.2273	0.2905	-7.6660	0.0000
PPlayground	0.0784	0.0969	0.8080	0.4189
PRestrooms	0.3375	0.4366	0.7730	0.4395
PSidewalk	0.3192	0.1287	2.4800	0.0131
PRentals	-0.6174	0.1227	-5.0310	0.0000
PPointMugu dummy	6.0381	0.5809	10.3950	0.0000
POceanside dummy	5.6402	1.1446	4.9280	0.0000
PVenice dummy	6.6473	0.4927	13.4920	0.0000

Table A3 (continued): Three Nested Beach Choice, Activity Choice, and Participation Model : Time Valuation = 100% wage rate

ACTIVITY CHOICE MODEL				
	Mean log-likelihood		-1.01975	
	Observations		4837	
	Estimates	**Standard Error**	**T-statistic**	**P-value**
	Variables Affecting The Choice Of Water-Based Activities.			
I1	0.2809	0.0590	4.7580	0.0000
WConstant	-1.8131	0.2086	-8.6900	0.0000
WMale	0.4045	0.0704	5.7460	0.0000
WBlack	-1.0302	0.2468	-4.1750	0.0000
WHispanic	-0.0979	0.0930	-1.0530	0.2924
W3Constant	1.3623	0.1333	10.2190	0.0000
W4Constant	1.9930	0.1645	12.1150	0.0000
W5Constant	1.2666	0.1794	7.0590	0.0000
W6Constant	0.6390	0.1419	4.5020	0.0000
SConstant	-0.0127	0.1750	-0.0720	0.9423
SKids	0.2091	0.0731	2.8620	0.0042
S3Constant	0.0569	0.1246	0.4570	0.6479
S4Constant	1.1878	0.1126	10.5520	0.0000
PARTICIPATION MODEL				
	Mean log-likelihood		-3.89772	
	Observations		7686	
I2	0.3609	0.0812	4.4440	0.0000
Constant	-4.5649	0.1416	-32.2380	0.0000
Male	0.4397	0.0602	7.3090	0.0000
Kids	-0.1690	0.0580	-2.9140	0.0036
Student	-0.1827	0.0751	-2.4340	0.0149
Workparttime	0.1870	0.0881	2.1230	0.0338
Black	-0.6743	0.1335	-5.0530	0.0000
Hispanic	-0.5631	0.0745	-7.5620	0.0000
Summer	0.2221	0.0878	2.5300	0.0114

Appendix II: DATA COLLECTION AND CLEANING

Work began on the project in January 1999 to design a panel survey of Southern California residents to track their usage of beaches in the region. Following extensive testing, the recruitment of a panel of residents commenced in November 1999, using a phone survey of a large random sample of area households. In August 2000, a second recruitment effort occurred to replenish the panel. The panel survey was conducted in waves of two-months duration. At the beginning of each wave, panel members were sent a map identifying the beaches of Orange and Los Angeles Counties and a calendar for the upcoming two months. At the end of each wave, the panel members were interviewed by phone. Each survey had a section designed to catalog every trip by the respondent to a beach in Southern California during that wave, plus an additional section with questions on a particular special topic. The waves and special topics are as follows:

Wave 1:	Dec 1999 – Jan 2000. Use of time.
Wave 2:	Feb – March, 2000. Health effects.
Wave 3:	April – May, 2000. Familiarity with beaches.
Wave 4:	June – July, 2000. Expenditures on beach recreation.
Wave 5:	August – September, 2000. Contingent behavior/contingent valuation.
Wave 6:	October – November, 2000. Attitudes regarding San Onofre power plant.

Phase I of the data analysis was conducted between June 2001 and January 2002. The results of this work were described in a series of reports:

Beach Recruitment Report (August, 2001)
Report on Choice Set Familiarity (August, 2001)
Report on Beach Trips by Wave (September, 2001)
Revised Report on Activities (November, 2001)
Report on Panel Participation and Attrition (December, 2001)
Beach Expenditures Report (January, 2002)
Report on Wave 4 Analysis (January 2002)
Report on Wave 4 Valuation (January, 2002)
Data Collection Production Report (January 2002)
Revised Contingent Behavior-Contingent Valuation Report (March, 2002) Report on Data Collection, Checking, Cleaning, and Archiving for Phase I of the Southern California Beach Project (March, 2002)
Report on Valuation Methodology (May, 2002)

The main focus of the Phase I work was data checking and summarization of the raw survey data. The data analysis was confined to the data from wave 4 and was intended as much as a vehicle for in-depth data checking, data cleaning and software development as for substantive data analysis. The main data analysis was intended to start once the data checking and cleaning had been completed.

The data supplied by the Chico Survey Research Center (SRC) had one record for each respondent if they took no trips to a beach in Southern California that wave, and one record for each distinct beach destination visited if the respondent did go to the beach during that wave.

In order to analyze the data, the sets of destinations had to be grouped into separate trips which reflect a single excursion from home. Although Chico SRC provided several variables which were supposed to convey this information, in many cases these variables were contradictory or were obviously incorrect. We combined information including starting and end dates and times, beach destinations, and the patterns of each panelist's other trips to correctly classify these trips and destinations. These corrections were done in close consultation with the Chico SRC, and in many cases involved going back to look at the original CASES datasets which had the CATI responses. This effort was complicated by the fact that the SRC used multiple releases of the CASES software over the year-long course of the survey, and there appeared to be differences in the way that values were filled in for multiple destinations on the same trip for the different versions.

We also checked that trip begin/end dates made sense, and corrected several errors. One diary dataset had a number of data-entry errors where the typist appeared to have shifted the digits on a subset of numeric entries by one digit to the right (e.g. 31 became 42). This showed up in dates as well as some categorical responses and was easily corrected in consultation with the Chico SRC. It is important to note that many trips did not contain an exact date because the respondent failed to supply one. The SRC made some effort to generate synthetic dates by looking at responses as to which week of the month and whether the trip was a weekend/weekday trip, and then randomly spreading the trips among the possible days. These "corrections" were removed, as they contain no information not in other parts of the dataset and are misleading. One trip contained a mis-keyed panelist ID, however the correction was obvious to us and was agreed to by SRC staff.

The initial contact Screener data is unavailable for 119 people in the Replenishment sample of respondents who supplied responses for waves 4, 5, or 6. 98 of those people reported taking trips. However, there is sufficient redundancy between the Screener data and the Demographics questions on the wave 4 diary for these people that the only data unavailable for them is data on the other household members (their number, ages and genders). These records appear to be missing at random, and the main variable which is unavailable for them is the presence of children in the household. That information can probably be reconstructed with a fair degree of accuracy by looking at whether they reported bringing minors along on any trips, information which is contained in the diary data.

Since we require accurate estimates of travel time and distance, we examined the home addresses of the respondents very carefully. We combined data from the initial responses for each panelist, and incorporated address corrections supplied with the responses to each wave. We cross-checked this information with the actual mailing addresses used by Chico, and then manually examined each address change to determine whether it was a correction or an actual move. The addresses were then examined using PC Miler 10.0, and addresses which were not precisely located were reviewed individually and corrected by employing information from PC-Miler, Yahoo Mapping website, and the US Postal Service zip-code look-up website. We tried as many variations on spelling of street names as we could, and had a high rate of successful corrections. A large part of the difficulty arose from the preponderance of Spanish-language names in Southern California and the unfamiliarity of the Chico survey takers with these place names. The task was further complicated by the fact that the mapping software could only locate addresses exactly using zip codes for most locations, but required the city name for others.

Some people who were contacted during the replenishment survey answered questions with information about their beach trips during June and July, but refused to join the panel for the purpose of reporting their future trips. If they provided useable address data, we will utilize the data on their trips during this wave. However, if they did not provide address data, their trips cannot be used in our analysis.[2] In addition, 72 of the people recruited in either the original recruitment survey or the replenishment survey elected to participate over the internet and therefore did not supply addresses.[3] [4] Some other respondents supplied only PO box addresses (which provide us a zip code only).

In total, out of total of 1308 individuals who supplied information to the survey[5], 1182 gave usable address information; of these, 1102 gave addresses that were located exactly, and 80 gave addresses that could be located only to within a zip code.

Of the 1182 people who supplied information to us and for whom we have usable address information (i.e., an exact address or a zip code), 359 did not make any trips to the beach in Southern California during the period they reported to us. The remaining 823 reported taking one or more trips to the beach in Southern California; in aggregate they provided information on a total of 6737 trips.[6]

In collecting trip information, the diary questionnaire distinguished between trips made to a single beach site in Southern California versus those made to multiple beach sites, and between trips lasting for one day or less versus those lasting for more than one day. The total of 6737 trips includes some of all four kinds of trips.

[2] 54 people who reported 179 trips fell into this category; these trips are excluded from the counts given below.

[3] 28 of these 72 individuals reported visiting the beach, and they took a total of 82 trips; these trips are excluded from the counts presented below.

[4] In any future implementations of these survey instruments we would make a point of asking internet participants for their mail address.

[5] We use this phrase rather than saying “1254 panel members” because, as noted above, some people contacted during the replenishment survey supplied information about their beach trips during June and July but refused to join the panel; hence these people were not panel members.

[6] This total excludes the 261 trips by 82 individuals who did not provide usable address information, as noted in footnotes 1 and 2.

Of the 823 individuals who reported on their trips and for whom we have usabele address information, 625 made *only* one-day trips to single-site destinations; these individuals made a total of 4096 trips. The other 198 individuals made some multi-day and/or some multi-site trips. Of these, 57 made *both* some multi-day trips *and* some multi-site trips; 99 made some multi-day trips but *no* multi-site trips; and 42 made some multi-site trips but *no* multi-day trips. The following table breaks down the 6737 trips by trip length and the number of beach site destinations:

	ONE DAY TRIPS	> 1 DAY TRIPS	
SINGLE-SITE TRIPS	6226 trips by 799 people.	241 trips by 129 people.	6467 trips by 817 people.
MULTI-SITE **TRIPS**	214 trips by 54 people.	56 trips by 49 people.	270 trips by 99 people.
	6440 trips by 803 people.	297 trips by 156 people.	6737 trips by 823 people.

Of the 823 individuals who reported on their trips and for whom we have usable address information, 704 supplied income information to us, but 119 did not. The 704 individuals for whom we have usable address and income information accounted for a total of 5689 trips.

Finally, the 6737 trips break down by wave as follows:

WAVE 1 (Dec-Jan)	1162 trips
WAVE 2 (Feb-Mar)	1005 trips
WAVE 3 (Apr-May)	708 trips
WAVE 4 (June-July)	2003 trips
WAVE 5 (Aug-Sept)	1179 trips
WAVE 6 (Oct-Nov)	680 trips