Title

Heterogeneous Preferences for Water Quality: A Finite Mixture Model of Beach Recreation in Southern California

Authors
Hilger, James
Hanemann, Michael

Publication Date
2006-02-01

Heterogeneous Preferences for Water Quality: A Finite Mixture Model of Beach Recreation in Southern California*

James Hilger and Michael Hanemann†
University of California, Berkeley
hilger@are.berkeley.edu

February 1, 2006

Abstract

This paper uses a finite mixture logit (FML) model to investigate the heterogeneity of preferences of beach users for water quality at beaches in Southern California. The results are compared with conventional approaches based conditional logit (CL) and random parameters logit (RPL). The FML approach captures variation in preferences by modeling individual recreator preferences as a mixture of several distinct preference groups, where group membership is a function of individual characteristic and seasonal variables. The FML parameter estimates are used to calculate welfare measures for improvements in beach quality through a reduction of water pollution. These bound the traditional CL and RPL mean welfare estimates, and have the advantage of highlighting the distribution of the population sample's preferences. The data indicate the existence of four representative preference groups. As a result, willingness to pay measures for improvements in water quality can be weighted across individuals to calculate the distribution of individual welfare measures.

One of the groups is people who go to the beach with small children. An interesting finding is that these people have a lower mean WTP for improving water quality than people who go without a small child. This may well be an example of cognitive dissonance: parents find they go to the beach more often than others who don't have small children, since that keeps the children occupied and happy, and they adapt their perception of the water quality to be consistent with their behavior.

Previous environmental and resource economic applications of the FML have been limited to applications with small choice sets (6) and group membership variables (4). This paper extends the FML model through the estimation of a large (51) choice set with 9 membership variables. This application is the first to incorporate seasonal variables into the group membership function to capture seasonal heterogeneity.

*This reasearch was made possible by funding from the National Marine Fisheries Service and California Sea Grant. This paper is posted at the eScholarship Repository, University of California.
http://repositories.cdlib.org

†Ph.D. Candidate and Chancellor's Professor, respectively; Department of Agricultural and Resource Economics, University of California, Berkeley, California 94720. Address correspondences to hilger@are.berkeley.edu

1 INTRODUCTION

Environmental resource protection and management requires the ability to assign values to non-market goods. While the literature has generally focused on the *average* valuation and preferences for these goods, the importance of the *distribution* of preferences for environmental amenities for populations with diverse preferences has often been neglected.

With over 150 million visits a year, the beaches of Southern California provide environmental and recreational amenities to a diverse user population (USFS , 2000). However this coastal resource is strained by pollution that has led to increases in beach closures and pollution advisories in Southern California for the past nine years, with 1,469 postings in Los Angeles County in 2004 alone. The closure of beaches and the poor water quality even when beaches are open impair the public's use and enjoyment of the beaches of Southern California; and create loss of business and a significant reduction in the public welfare. However the costs associated with the widespread water management plan is estimated to be in the billions (Gordon, 2002). The magnitude of the cost estimates raises serious policy questions about the economic benefits of storm water pollution control.[1]

Not only is there a loss of welfare due to impaired use and enjoyment of the public beaches , but there can also be health impacts from swimming in polluted ocean waters, including upper respiratory infection and other illnesses.[2] Public concern regarding this degradation has prompted the approval of several Legislative and Assembly bills,[3] promoted cleanup and monitoring efforts, and increased the need for careful estimation of the welfare impact of water pollution at Southern California beaches for cost benefit analyses used in litigation, remediation, and general management. This paper investigates the willingness to pay for a reduction in beach water pollution and illustrates how these values vary by recreator characteristics and season.

Varying preferences of recreational users and the multiple use nature of beach sites complicate the estimation of willingness to pay measures for improvements in water quality and

[1] Another arena where the valuation of beach recreation has arisen is the prevention of oil spills and the measurement of damages caused by oil spills and benefits from oil spill prevention. In 1969 there was a major oil spill near Santa Barbara which attracted serious attention from economists interested in measuring the economic value of the damages caused by the spill (Sorensen, 1975). In February 1990 there was an oil spill off the coast at Huntington Beach which triggered a law suit by the State of California against the vessel owner that actually went to trial in the fall of 1997 and led to the award of damages amounting to $12.75 million as the value of the public's lost use and enjoyment due to the closure of beaches following the spill (Chapman and Hanemann 2001).

[2] A large epidemiological study, The Santa Monica Bay Restoration Project study, found an increase in the risk of contracting an illness when swimming near storm drains. Recreators that swam near storm drains were 57% more likely to suffer symptoms of a fever than other swimmers (Haile et. al., 1996). For a recent review of health risks associated with beach water pollution see NRDC (2005).

[3] California Assembly Bill 411, the "Right to Know" bill, was passed in 1999.

other beach attributes. Systematic preference heterogeneity can lead to bias in parameter estimates if left unaccounted for. This paper addresses systematic preference heterogeneity by utilizing a finite mixture logit (FML) random utility model. A panel trip diary data set documenting 4,462 Southern California winter and summer beach trips for 595 recreators from December 2000 to November 2001 is analyzed using the FML approach. Estimation results indicate that beach recreators can be characterized by one of four distinct representative groups by beach recreational decisions, and individual and seasonal attributes. This information is then used in calculating welfare estimates for each individual in the sample and the weighted average measure for the population. I find that the welfare estimates associated with an environmental improvement vary significantly both within the population and across seasons. One interesting result is that the presence of children on a beach trip that involves water contact is associated with lower mean WTP estimates for improvements in water quality.

1.1 CONTRIBUTION OF THIS PAPER

While FML models have been estimated previously in the environmental and resource economics literature this paper makes three main contributions. First, this paper is the first FML application focused on modeling the welfare and behavioral impacts of an environmental good associated with health outcomes. Water pollution is a widespread problem in many coastal and fresh water areas, and recreational swimming is the second most popular recreation activity in the United States with over 90 million participants (NRDC, 2004). Unfortunately, countless recreators swim in water that does not meet the EPA health standards (NRDC, 2004). This drives an increasing interest in determining what draws recreators to specific beaches (Hanemann et. al., 2004, and Lew and Larson, 2005) and what influences where they choose to recreate once at a particular beach (Pendleton, 2001). This paper furthers the understanding of the impact that water pollution has on beach recreation through the estimation of preferences coefficients for a diverse group of beach recreators. Preference estimates can be used to forecast and explain beach choice behavior conditional on beach attributes, such as water pollution. The ability to investigate the composition of these preference groups conditional of individual and season data is a useful tool for managers and policy makers in both the resource and public health arenas.

Second, this application contributes to the modeling of heterogeneity with the FML through the incorporation of a seasonal variable in the beach choice occasion preference membership function. This enables the analyst to capture seasonal variation in preferences for beach attributes. Other studies have not utilized the model to account for seasonal changes in site attribute preferences.

Lastly, this paper represents a substantial step forward in the technique's empirical application. Previous applications have been applied to fairly restrictive choice sets, primarily modeling binary participation choice or multinomial choice for up to 6 options (Boxall and Adamowicz, 2002). In contrast, this application models recreator decisions among a choice set of 51 beaches using a revealed choice data set. This model specification utilizes 9 individual trip membership function variables consisting of seasonal, activity participation, and demographic variables. This marks a substantial increase in the number of parameters estimated relative to other applications in the literature (Provencher et. al. use three (2002), Boxall and Adamowicz use 6 (2002), and Shonkwiler and Shaw use 3 (2003)).

I generate statistical estimates of the welfare impact to beach users due to changes in water quality for beaches in Southern California. Accounting and controlling for preference heterogeneity is the key objective of this research. To this end, I utilize the finite mixture logit random utility model which allows for variation in preferences across individuals and seasons.

Comparison of the welfare estimation results from the competing models indicates that the FML model provides an important insight into the heterogeneity of individual's willingness to pay (WTP) for improvements in water quality. The FML model's average estimate of mean WTP is roughly 4.64 times that of the standard logit model's mean WTP, while the RPL's WTP point estimate of the mean WTP is roughly 0.13 of the standard logit model's mean WTP. However, the estimated mean WTP for individual beach recreators ranges from zero or negative to 14 times the logit's mean WTP, depending on the type of recreators and the type of quality change. For example, trips involving water-contact recreation during the winter by male college graduates who are working full-time are associated with a high value for an improvement in water quality. Trips taken during the summer by male college graduates who are not working full time and who do not enter the water are associated with a low value for an improvement in water quality.

The remainder of this paper is organized as follows. Sections 2 and 3 further motivate the application and the model, respectively. In section 4, I describe the base model framework. In section 5, I describe the problem of heterogeneity in a discrete choice setting. I will then review the standard conditional logit (CL) framework and describe several econometric techniques that have been developed in order to account for heterogeneity. In section 6, the finite mixture logit model and estimation strategies are discussed. Section 7 describes the data and the model specification is outlined in section 8. Section 9 reports on the estimation results for the CL, RPL, and FML models and is followed by the paper's conclusion in section 10.

2 APPLICATION BACKGROUND

Coastal and marine health play an important role not only in the prosperity of the fisheries industry but also in the welfare of the communities which border the California coast and rely on the coastal environment for recreation and tourism.[4] Beach trips serve as a primary recreational activity for some and as a source of income for others.

However, coastal environmental resources are increasingly strained and affected by pollution and overuse. In California there were 6,568 beach closures and advisories in 2001. This was a 14% statewide increase from 2000 and marked the fifth consecutive year that beach closures and advisories have increased (NRDC, 2002). The public awareness of poor water quality is so widespread in the Los Angeles area that, in a focus group a few years ago, eight out of ten participants said that they do not go into the water when they go to local beaches (Hanemann, 2005).

The main cause of the beach closures in the Los Angeles area is storm water runoff. Although rainfall events are infrequent, when they do occur they generate a large volume of runoff from the streets, parking lots and other paved surfaces containing high pollution loads that bypasses sewage treatment plants. This runoff is discharged directly to the ocean through storm drains at or the near the local beaches. Storm water pollution is now coming under increasing regulatory pressure, but is extremely costly to manage. An engineering study conducted for the California Department of Transportation in 1998 estimated that to divert and treat flows from about 90% of the annual expected storm events in Los Angeles County would cost almost $54 billion, and the Los Angeles County Sanitation District revised this cost estimate to $65 billon. A 2002 report by engineers at the University of Southern California estimated the cost at $156 billion to cover 97% of the expected storm events (Gordon et. al., 2002). Other pollution sources include discharges from point sources, and sewage spills. The resulting degradation of the coastal environment damages marine and coastal flora and fauna, and adversely affects the welfare and possible health of beach recreators.

The risk of becoming ill while swimming at the beach reduces the welfare of those who venture in the water and contract an illness,[5] and diminishes the welfare for those who

[4] Lew and Larson estimate the mean value of a recreational beach day to be $28.28 (2004).

[5] Rabinovici et. al. (2004) review the valuation of health status literature and report that Mauskopf and French (1991) estimate the WTP for government programs to aid in the avoidance of gastrointestinal

forgo swimming because of the risk.[6] Public concern regarding this coastal degradation has prompted several studies focusing on the adverse health effects of coastal pollution and has generated the approval of several Legislative and Assembly bills (NRDC, 2002).

The differences in the values placed on beach recreation by different user groups can have important practical implications for beach management. For example, shoreline anglers care about different aspects of the beach recreation experience than surfers or mothers taking young children to the beach. What is considered an amenity to one may be an unwanted nuisance to others. Moreover, the resource manager may be forced to make trade-offs in meeting the needs of different groups. The information developed in this research can lead to improved management of coastal and beach resources.

3 MODELING BACKGROUND

A rich diversity of preferences among decision makers creates difficulties in terms of accurately modeling recreational site choice and estimating the economic value associated with a change in resource attributes. Diverse user groups often value different attributes of recreation sites and demand different services from them. If preference heterogeneity can be easily controlled by segmenting the sample population by a variable known to the analyst, a standard logit random utility model (RUM) can be used to estimate coefficients and welfare measures for each group separately. For example, beach recreators who swim in the ocean are likely to have different preferences for water quality and other beach attributes than those lying on the sand. However it is often unclear where to draw the line in defining sub-samples of the population.[7] This may lead to bias in welfare measures for changes in site attributes and hinder proper aggregation of welfare measures across individuals or time periods and adversely affect policy and management decisions.

The logit model handles variation in preferences by averaging over the individuals. In cases where the population is fairly homogenous in their preferences this may not cause a major problem; however, if the population is characterized by considerable systematic preference heterogeneity, the model's results may be misleading due to an averaging out effect over preferences from distinct groups. Additionally, the distribution of preferences over individuals or time are commonly lost due to the restrictive single point or even modal distribution

symptoms at $280 for a 2-4 day case and $1,125 for a 5-7 day case.

[6] Walsh et. al. (1992) report a mean value per visitor day of recreational swimming at $35.60 ($2001)

[7] If the analyst differentiated between beach and water users there would still be heterogeneity within users. For example, among water users, surfers may care about different aspects of the beach recreation experience than mothers taking young children to the beach to swim. Although both of these groups likely view clean water as desirable, they may differ in the level of importance they place on water quality.

which the model imposes on the data. The preservation of the preference distribution may aid analysts focused on the welfare changes between user groups, or to a specific type of user, due to changes in the attributes of the choice set. Similarly, an understanding of the temporal fluctuations in preferences could have important policy implications.

Suppose there are two different groups of beach users: novice and expert surfers, who prefer small and large waves respectively. Membership in either group is unobservable to the analyst, but may be statically correlated with observable demographic and seasonal data. Imagine further that there are several means of undertaking a coastal project which can have the secondary effect of increasing or decreasing the size of waves. Estimation with the standard logit model causes the opposing preferences for waves of the two user groups to be averaged out, resulting in model estimates that call for the medium sized waves, which are not preferred by either group. In contrast, a model that could statistically distinguish the two different types of users and estimate their separate preferences could lead to a policy whereby a variety of waves are maintained at specific beaches, resulting in a welfare improvements for both groups.[8]

The finite mixture logit (FML) model used in this paper accounts for systematic heterogeneity by sorting the agents into separate behavioral groups or latent segments, with different attribute preferences.[9] Within each latent segment, individuals are assumed to have homogeneous preferences. The segments are termed "latent" since individual membership in a particular segment, as well as the segments themselves, are not observable. The FML model simultaneously assigns an agent a probability of membership to each latent segment, and estimates the discrete choice probabilities for the random utility model. This approach captures the variation in preferences across the population through a discrete distribution with multiple probability masses. The model is distinctive in that it not only accounts for heterogeneity, but is able to explain the sources of that heterogeneity.

4 BASIC FRAMEWORK: RANDOM UTILITY MODEL

Random utility models have a long history as a powerful tool for resource managers. The random utility model is the standard statistical framework used to estimate the value of the change in consumer welfare due to an incremental change in the level of resource attributes in a setting characterized by consumer choice between several alternative recreation sites

[8]Note, the emphasis on the word statistically. In practice, preference groups are often not clearly defined into easily identifiable groups.

[9]This model was first proposed by McFadden (1986) and implemented by Swait (1994).

with varying attributes.[10]

Consider the utility maximization problem that an individual solves in relation to a recreation choice occasion between a set of J alternatives ($j = 1, ..., J$):

$$\underset{j}{Max} : u_i = v_i(M_i - C_{ij}, Q_j, Z_i) + \epsilon_{ij}. \tag{1}$$

Where $u(.)$ is a function of individual income, M_i, the travel cost of individual i visiting site j, C_{ij}, the quality and attribute mix of the chosen site, j, in the recreational choice set, Q_j, and individual socioeconomic characteristics, Z_i. The unobservable portion of utility is denoted by ϵ_{ij} and is assumed to be a random variable. The decision to recreate at a particular area is viewed as the decision to consume, or incorporate into one's utility function, the specific attributes that uniquely identify the chosen recreational site from others in the choice set. When individual i chooses to consume bundle j out of her choice set J, $j \subset J$, it is assumed that u_{ij} is the maximum of the J possible utilities in the choice set. The conditional probabilities of individual i choosing choice j can be derived as

$$\Pr_{ij} = \Pr(u_{ij} > u_{ik}) \; \forall \; k \neq j. \tag{2}$$

Maximum likelihood estimation can then be used to estimate the parameters of the indirect utility function (McFadden, 1973; and Bockstael, Hanemann, and Strand, 1986).

The outcome of an individual choice occasion, designated by Y_i, is a random variable. If and only if the disturbances associated with j, $\forall j \subset J$, are independently and identically distributed with the generalized extreme value distribution, the choice probabilities are given by

$$\Pr_{ij} = \Pr(Y_i = j) = \frac{e^{\beta' \mathbf{\Gamma}_{ij}}}{\sum_{j=1}^{J} e^{\beta' \mathbf{\Gamma}_{ij}}} \tag{3}$$

where $\mathbf{\Gamma}_{ij}$ is a vector of individual and alternative specific variables (McFadden, 1973). This model is known as the conditional logit model (CL). This formulation of the CL model causes individual variables, $\mathbf{Z}_i$, that do not vary over the choice set to drop out of the choice probability. The choice probability is then determined by choice specific variables.[11]

[10] The Conditional Logit Random Utility Model (CL RUM) is a widely used research tool. An early application of this model to recreational choice application is Hanemann (1978). For technical discussions refer to Greene (2000) and Wooldridge (2002). For a discussion of the application of RUMs to environmental economics refer to Haab and McConnell (2003).

[11] However, through construction of interaction variables that vary over both individuals and choice attributes, individual specific information can be retained as an argument in the choice probabilities.

Several econometric and modeling issues commonly arise with the Random Utility Model. Econometric consideration should be given to the independence of irrelevant alternatives property[12] and to identification issues surrounding the scaling parameter. In terms of modeling, the construction of the travel cost variable[13] and the formation of the choice set are major issues that have been the focus of considerable research.[14]

5 ECONOMETRIC ACCOUNTING OF HETEROGENEITY

Heterogeneous preferences are difficult to account for in behavioral choice models due to the formulation of the conditional logit (CL) model which has historically been the tool of random utility models. Within demand system models, the analyst can directly incorporate demographic, temporal, or other individual characteristic data directly into the individual's utility function to address preference heterogeneity. However under the specification of the CL, these characteristics drop out of the probability of an individual selecting a specific choice, thus preventing the direct identification of these characteristics in the model.

If heterogeneity is not accounted for, RUM estimates are characterized by bias and lead to inaccurate forecasts pertaining to changes in resource attribute levels and management policies (Chamberlain, 1978, 1980; and Jones and Landwehr, 1988). This bias adversely affects welfare estimates for simulated changes in resource attributes and/or management decisions.

[12] The analyst must take note that in the standard multinomial or conditional logit models the odds ratios for a specific pair of choices, $\Pr_j / \Pr_k$, is independent of the remaining alternatives. This property is known as independence of irrelevant alternatives (IIA). This property is fairly restrictive because in implies that the relative probability of choosing between alternatives remains constant after the introduction of a perfect substitute of one of the alternatives to the choice set. Several models such as the nested logit and random parameters logit models have been developed, in part, as a solution to IIA (Haab and McConnell (2003)).

[13] The assumption that travel cost prices are exogenously determined deserves comment, as the endogeneity in prices assumption is one of the primary issues critiqued in the discrete choice literature (Berry, Levinsohn, and Pakes, 1995; and Nevo, 2000). However as discussed in Train (2003), this issue is not of great importance outside of market-level demand models. Within customer-level demand models it is assumed that individual demand does not affect price. Moreover, within the recreational demand literature the price associated with choosing a specific good is determined by the cost of travel to that location. One alternative is that the consumption of the good is of large enough proportion in the individual's utility function that the individual incorporates the location of the recreational site as an important argument in the residential location decision making process. Secondly, site characteristics to some degree all relate to visitation. For most site attributes individual trips do not affect the attribute level. However, some attributes, such as solitude, offered by the site are highly sensitive to small changes in the number of trips taken to the site. Assuming that individual residence location and travel infrastructure is determined exogenously, the travel cost price is exogenous. See Parsons (1991) for a discussion on housing location.

[14] For a thorough review on the optimal size of the choice set see Kurisawa (2003).

To address heterogeneity, researchers have primarily focused on structural approaches requiring the *a priori* selection of typically demographic or choice variables. In "cluster models" individuals are segmented into demographically homogenous/similar groups. An alternative method incorporates into the indirect utility function and interaction variable composed of individual demographic variables, such as income, race, and family composition and various choice attributes (Adamowicz et al., 1997). These methods are limited by the assumption that preference groups can be accurately determined *a priori* by demographic variables, and theoretical issues pertaining to the choice of an interaction variable (Boxall and Adamowicz, 2003). 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 heterogeneity structure is complex and the sample consists of a large number of decision makers.

An additional structural method, the Generalize Extreme Value (GEV) Logit (or nested logit) disaggregates the decision between alternatives into subsets of similar alternatives, relaxing the IIA restriction (McFadden, 1978). In the context of beach recreation, the GEV framework has been used to model recreational beach choice conditional on the type of activity engaged in during the beach visit (Hanemann et. al., 2004). The primary benefits of this approach are that the model may be useful in highlighting the differences in choice behavior and welfare estimates for different user groups, and that it is not restricted by the IIA property. However, the approach requires that the "nesting" rules are defined *a priori*.

Another approach, the random parameter model, controls for heterogeneity across preferences by allowing estimated coefficients to randomly vary across individuals according to a continuous probability distribution, typically the normal or log-normal. By allowing for variation in coefficients over people, the unobserved portion of the respondent's utility is correlated over sites and time (Train, 1997). To set up the single choice occasion RPL model, we begin with the standard logit choice probability for individual i and relax the standard assumption that preferences for all individuals are equal. Assuming that individual tastes vary in the population we can write the probability density function of the preference parameters as

$$f(\beta|\Theta^*),$$

where Θ^* are the true parameters of the distribution of the preference parameters. The actual probability that an individual i chooses choice j is the integral of the standard logit probability for all possible values of the preference parameter weighted by the density of the preference parameters:

$$\Pr_{ij} = \int \left(\frac{e^{v_i(Cji,Q_j)}}{\sum_{j=1}^{J} e^{v_i(Cji,Q_j)}} \right) f(\beta|\Theta^*)\mathrm{d}\beta.$$

The parameters are estimated by using simulation to evaluate the integral in the choice probabilities. The RPL can easily be extended to a multiple choice occasion panel data setting (Train, 1997). Additionally, the RPL model is not restricted by the IIA property due to interactions within the choice probabilities of the attributes of all elements in the choice set (Train, 2003).

The RPL approach has two weaknesses. First, it assumes that preferences vary continuously across economic agents. Second, it does not offer a behavioral explanation for the source of the heterogeneity across people. Although the continuous distribution assumption is likely to be valid in many applications, for example the spiciness that one likes their food, there are many situations where actual preferences may be more accurately captured by multiple discrete probability masses. Moreover, from a management perspective a coarse grouping of preferences may sometimes be more useful. For instance, the presence of motorized watercraft likely enters either positively or negatively into the majority of individual beach recreator's utility functions. Resource managers are often concerned with obtaining the best possible estimates for specific individuals or user groups relevant to policy and equity concerns.

6 FINITE MIXTURE LOGIT APPROACH

An alternative solution is the finite mixture logit (FML), or latent segmentation (LS) approach. This approach was suggested in a RUM setting by McFadden (1986), and was implemented by Swait (1994). There has been a recent increase in the application of this approach, including several recreational choice models applications (Provencher et. al (2002), Boxall and Adamowicz (2002), and Shonkwiler and Shaw (2003). The FML approach is based on two important assumptions. First, individual preferences are neither homogeneous nor continuously distributed, but can be represented as discretely distributed with multiple probability mass points. Second, individual preferences are not purely a function of demographic variables, but can also be formed by perceptions, attitudes, behavior, past experiences, and unobserved variables.

Each "latent segment" is composed of like-minded individuals with homogeneous preferences. The segments are termed latent because individual membership to a particular segment is not observable, nor are the segments themselves. The FML model simultaneously assigns the economic agent the probability of membership to each latent segment and estimates the discrete choice probability for the random utility model. This approach captures the variation in preferences across the population through a discrete distribution with multiple probability masses. The model is unique in that it not only accounts for heterogeneity, but is able to explain the sources of that heterogeneity. This is of particular importance in regards to management decisions where user groups may either be demographically homogenous or where there is little correlation between user group preferences and the standard demographic variables. The FML model can estimate the coefficients associated with the choice occasion for each latent segment's utility function. The FML model additionally estimates the composition of the latent segments and can be used to help researchers and managers understand the processes involved in the formation of behavioral groups. The ability to segment the sample population through the estimation of the latent segment type may aid resource managers with welfare analysis and management policy.

The FML RUM is an extension of the CL model, and follows the assumption that individual i's indirect utility is maximized on a choice occasion by selecting alternative $j \subset J$. The probability that alternative j is chosen is the probability that the utility gained from choice j is greater than or equal to the utility forgone by not picking one of the other alternatives in the choice set, J.

Under the assumption that there exists some degree of heterogeneity in preferences across the sample, let S be the number of segments that the population is to be grouped into.[15] Individuals are assumed to belong to a segment $s(s = 1, .., S)$ within the sample population. Individuals within a segment are assumed to be characterized by homogeneous preferences. Additionally, in all but the trivial case, $S = 1$, the probability ratio between any two alternatives includes arguments from all other alternatives in the complete choice set, J. It has been shown that in these cases the FML model is not constrained by the IIA property.(Shonkwiler and Shaw, 1997).

6.1 Single Choice Occasion

In a cross sectional data setting, the optimal solution to the recreational choice decision for individuals within a given segment s, is to maximize

$$u_{i|s} = v(\beta_s X_{ij}), \tag{4}$$

[15] The optimal choice of S is discussed below.

where the β_s vector is the coefficients representing individual preferences conditional on individual i's membership in segment s.

The parameter coefficients for a specific segment of the population are estimated using the following probabilities.

$$\Pr_{ij|s} = \frac{e^{v_{i|s}(-Cji,Q_j)}}{\sum_{j=1}^{J} e^{v_{i|s}(-Cji,Q_j)}}. \tag{5}$$

Consider a latent membership likelihood function M^* that assigns individuals to segment $s \subset S$ (Swait, 1994). Arguments to M^* can include variables associated with the unobservable tastes, attitudes, and preferences of the members of the group as well as socioeconomic variables represented by the vector $Z_{i.}$ Segments can be identified using standard demographic variables, behavioral and preference data. Assume the following equation:

$$M_{is}^* = \gamma_s' \mathbf{Z}_i + \zeta_{is},\ s = 1, ..S, \tag{6}$$

where γ_s is a vector of segment specific parameters and ζ_{is} represents the error terms.

The membership likelihood function, M^*, is a random variable. To use the function in an econometric model requires assumptions about the distribution of its error terms. Following Kamakura and Russell (1989), Swait (1994) and Boxall and Adamowicz (2003) the error terms are assumed to be independently distributed across individuals with Type I extreme value distribution. The probability of individual i belonging to segment s can then be calculated as

$$\pi_{is} = \Pr(M_i = s) = \frac{e^{\gamma_s' \mathbf{Z}_i}}{\sum_{s=1}^{S} e^{\gamma_s' \mathbf{Z}_i}}. \tag{7}$$

This probability is modeled as multinomial logit framework where the independent variables in this function vary over individuals, unlike the conditional logit where the variation is in the choice specific variables. Addressing an indeterminacy in the model caused by the lack of normalization the following restriction must be imposed:

$$\pi_{is} = \frac{e^{\gamma_s' \mathbf{z}_i}}{1 + \sum_{s=2}^{S} e^{\gamma_s' \mathbf{z}_i}} \text{ for } s = 2, ...S, \tag{8}$$

$$\pi_{i1} = \frac{1}{1 + \sum_{s=2}^{S} e^{\gamma_s' \mathbf{z}_i}}, \text{ and} \tag{9}$$

$$0 \leqq \pi_{is} \leqq 1, \text{ such that } \sum_{s=1}^{S} \pi_{is} = 1.$$

To model choice behavior under the assumption that the sample population can be grouped into finite segments, the researcher estimates individual i's utility maximizing choice between J alternatives conditional on membership to a specific segment, s. The joint probability $\Pr_{ins}$ that an individual i is a member of segment s, and chooses alternative j is defined as

$$\Pr_{ijs} = \pi_{is} \Pr_{ij|s}. \tag{10}$$

It follows that for a single choice occasion the probability of individual i choosing alternative j unconditional on segment membership can be written as

$$\Pr_{ij} = \sum_{s=1}^{S} \pi_{is} \Pr_{ij|s}. \tag{11}$$

Defining d_{ij} as and indicator variable that takes the value of 1 if an individual i chooses site j and 0 if not, allows the writing of the individual likelihood function as

$$L = \sum_{s=1}^{S} \left[\pi_{is} \left(\prod_{j=1}^{J} \Pr_{ij|s} \right)^{d_{ij}} \right]. \tag{12}$$

The individual likelihood function can be rewritten as

$$L = \prod_{j=1}^{J} \left[\sum_{s=1}^{S} \left(\pi_{is} \Pr_{ij|s} \right) \right]^{d_{ij}}, \tag{13}$$

and the cross section likelihood function as

$$L = \prod_{i=1}^{I} \prod_{j=1}^{J} \left[\sum_{s=1}^{S} \left(\pi_{is} \Pr_{ij|s} \right) \right]^{d_{ij}}, \tag{14}$$

which yields the log likelihood function

$$\ln L = \sum_{i=1}^{I} \sum_{j=1}^{J} d_{ij} \ln \left[\sum_{s=1}^{S} \pi_{is} \Pr_{ij|s} \right]. \tag{15}$$

6.2 Panel Data: Constant and Variation over Time Model Specifications

The extension of the single choice occasion likelihood function to incorporate a time dimension utilizing panel data introduces a few complications in terms of the assumptions of segment membership independence across choice occasions. One assumption is that preferences are constant over time, although there is preference heterogeneity across individuals (Figure 1). This assumption is most appropriate when the set of choice occasions are temporally close (such as multiple decision choice occasions), or when preferences and choice attributes are stable over time. This specification has been applied in both the marketing and transportation literature (Ramaswamy et al., 1999, Greene and Hensher, 2003).

An alternative modeling specification, implemented in this paper, assumes that preferences can be allowed to vary both over individuals and time. This can be useful as preferences often tend to vary with seasonal tastes as the underlying choice decision changes. Seasonal variation in unobserved or unmeasured attributes necessitates the need to allow for seasonal variation in an individual's segment membership. Allowing for variation over time in preference membership relaxes the correlation between individual segment membership (Figure 2).

Write the probability of individual i choosing alternative j at time t as

$$\Pr_{ijt} = \sum_{s=1}^{S} \pi_{is} \Pr_{ijt|s} . \tag{16}$$

This leads to the likelihood function

$$L = \prod_{i=1}^{I} \prod_{t=1}^{T} \left[\sum_{s=1}^{S} \pi_{ist} \left(\prod_{j=1}^{J} \left(\Pr_{ijt|s} \right)^{d_{ijt}} \right) \right] \tag{17}$$

which simplifies as

$$L = \prod_{i=1}^{I} \prod_{t=1}^{T} \prod_{j=1}^{J} \left[\sum_{s=1}^{S} \pi_{ist} \Pr_{ijt|s} \right]^{d_{ijt}} , \tag{18}$$

and leads to the log likelihood function[16]

$$\ln L = \sum_{i=1}^{I} \sum_{t=1}^{T} \sum_{j=1}^{J} d_{ijt} \ln \left[\sum_{s=1}^{S} \left(\frac{e^{\alpha \gamma_s' \mathbf{Z}_{it}}}{\sum_{s=1}^{S} e^{\alpha \gamma_s' \mathbf{Z}_{it}}} \right) \left(\frac{e^{\beta_s' \mathbf{X}_{ijt}}}{\sum_{j=1}^{J} e^{\beta_s' \mathbf{X}_{ijt}}} \right) \right] . \tag{19}$$

[16] Note the individual demographic variables, $\mathbf{Z}_{it}$, have a time index.

The above likelihood function has been utilized in both the marketing (Swait, 1994) and recreation (Boxall and Adamowicz, 2004) literature. Both applications utilized stated preference data where each respondent made a series of sequential choices from a structured choice experiment where all choice decisions are made at the same time, weakening the basis for the preference variation over time assumption. The basis of the FML is that decisions made by different members of the same preference segment will be more correlated than decisions made by members of different segments.[17] This holds true unless there is a mechanism for an individual's segment membership to change between choice decisions (Morey, 2003).

The correct time specification choice is dependent on the goals of the analysis and what data is used. As a general rule, the constant over time specification is appropriate for models over short time durations which do not utilize membership covariates that vary over time and where preferences are assumed to be constant. The varying over time specification better suits applications that seek to model FM membership as a function of seasonality, the effect of previous choices, or individual characteristics that vary over time (the decision to get into the water on a specific beach trip). This paper utilizes the varying over time specification, as individual preferences are expected to vary over time due to both seasonal effects and variety seeking throughout the survey year.

6.3 Additional Econometric Issues

6.3.1 Scale Parameter

In addition to attribute preference parameters, the variance of the disturbance terms may also differ across segments of the population. In the standard CL framework the analyst assumes that the unobserved factors have constant variance, hence utility is of the same scale across respondents. However, this restriction is not implicitly held in the FML specification. Therefore FML model parameter estimates cannot be compared across segments directly. Researchers that do not take the differences in scaling parameters into account may incorrectly infer that the members of the segment with a larger coefficient estimate care about the attribute more than those individuals in the other segment. To properly interpret parameter results across segments analysts can compare the signs or ratios of parameter point estimates.[18]

[17] This assumes that the information set and individual characteristics are constant across choice decisions.

[18] Alternatively, the scaling parameter can be normalized for one segment so that the variance of the disturbance term is the same across both segments. This leads to the identification of the scaling parameter (Train, 2003).

6.3.2 Determining the Number of Segments

The appropriate number of segments is not identifiable in the FM class of models and is treated as exogenous. However, one can statistically test for improvements in the appropriate number of segments by estimating a series of models that iteratively increase the size of S. Improvements in model specification in terms of the number of latent segments in the population can be tested for through the use of McFadden's ρ^2, Bayesian Information Criterion, and Akaike Information Criterion test statistics. The use of traditional Likelihood Ratio tests in determining the number of segments should be used with caution as the regularity conditions are violated (Ben-Akiva and Swait, 1986, Jedidi, 1997, and Boxall and Adamowicz, 2003). In addition to the statistical tests, the analyst's judgment in regards to which model specification in terms of the number segments best describes the respondent population and addresses the relative policy questions should be applied.

Upon inspection of the FML model it is clear that through the selection of the appropriate number of segments the above model can mimic both the traditional CL and the RP models.[19] For instance, when $\gamma_s = 0, \beta_s = \beta, u_s = u, \forall s$, the FML reduces to the CL.

6.4 Welfare Estimation Theory

The generation of welfare measures associated with a change in the attributes of the choice alternatives is a primary use of the RUM. The economic marginal value associated with a change in water quality grades and other beach attributes can be calculated for each classification of user groups using model parameter estimates. Note that an individual's income does not vary over the alternatives in the choice set, so this term drops out of the probability. However the relevant measure of income in regards to the choice occasion is the individual's total income less the cost of the utility maximizing choice.

The FML model allows the calculation of more accurate welfare measures. The probability of membership into a latent segment is a function of individual demographic variables. Welfare measures can be calculated for each individual by properly weighting the welfare measures of the representative consumer of each latent segment by the membership probability to each latent segment.

$$\Pr_{j^*}(i) = \frac{e^{v_i(-Cji, Q_j, Z_i)}}{\sum_{j=1}^{J} e^{v_i(-Cji, Q_j, Z_i)}} \tag{20}$$

[19] In the present form FML is theoretically similar to the RPL where each respondent undertakes one choice occasion.

Changes in welfare due to a marginal change in a given attribute can be calculated using the marginal mean willingness to pay measure (WTP). This measure is defined as the maximum amount of income a person will pay in exchange for an improvement in the level of a given attribute provided and can be calculated as:

$$WTP_i^* = \frac{\beta}{\gamma} \tag{21}$$

where β is the parameter on the attribute of interest and γ is the travel cost parameter. Both parameters measure the marginal utility of the object in question. This result can easily be applied using FML parameter estimates:

$$WTP_{i|s}^* = \frac{\beta_s}{\gamma_s} \tag{22}$$

Because the degree of heterogeneity in preferences is assumed to be considerable in many recreational choice optimization problems, the ability to segment the changes in welfare over latent user types is important. However if the resource managers are interested in aggregate welfare measures over the sample, these can be calculated by adding up the welfare measures weighted by the latent segment probability (Boxall and Adamowicz, 2003).

$$WTP_{ji}^* = \sum_{s=1}^{S} \pi_s \left[\frac{\beta_s}{\gamma_s}\right] \tag{23}$$

This welfare measurement is an improvement over the traditional welfare calculation using coefficient estimates from the standard CL model due to the proper weighting of each segments' marginal willingness to pay.

7 DATA

The empirical choice model application utilizes an extensive recreational beach choice panel data set. The data come from a survey of households in Southern California. Respondents were asked to keep a diary of all their trips to beaches in Southern California from December 2000 through November 2001. The data consists of observation for 4,642 beach recreation choice occasions of 595 beach recreators throughout the 12 month period. Recreators include fishers, boaters, divers, surfers, sunbathers, runners, cyclists and other beach users. Beach recreator data contains demographic and behavioral data. An attribute data set contains individual beach attributes including water quality data and the travel times and distances between each beach and respondent residence.[20] The CL and RPL models are estimated using the same choice probability specification as the FML model.

Each respondent is assigned a unique numeric identifier in order to link survey responses from all segments of the project and thus create a large panel data set. The screener and recruitment surveys collect standard socioeconomic household data, as well as beach and non-beach recreation data. Respondents were asked to keep a record of every Southern California beach trip in a bi-monthly diary throughout the survey period.[21] For each trip, respondents were asked a series of trip details including the date of the trip, the specific beach they went to, the number of minors in their group, and information about up to four beach activities. Beach recreational activities are expected to be affected by seasonal variables. To control for this effect the data set is split into three time periods: winter (December and January), summer (June through September) and the remaining shoulder season months. Summary statistics on the seasonal distribution of trips, the probability of the average beach recreator's immersion rate, the percentage of trips that involves water immersion, and the proportion of recreators that enter the water are listed in Table 1.[22] Summary statistics on beach site trip counts are displayed in Table 2.

The implicit price of visiting each beach used in modeling is the travel cost construct. This construct is a function of the respondents reported income, and the estimated vehicle operational cost ($0.145/mile), travel time and the distance between the respondent's residence and each beach in the choice set.[23] One way travel distance and travel time between a respondent's address and the beach address are calculated using the computer program PC-Miler. The time and distance data is transformed into the round trip travel cost of each trip, and is one of the model's primary explanatory variables.[24]

[20] The complete data set consists of a screener and recruitment survey, 6 bi-monthly diary surveys, and 7 supplementary modules that focus on a variety of topics. The original data set comes from a random telephone sample of 1,848 respondents. Of these, 824 respondents were classified as non-beach users and 202 declined to take part in the survey The remaining 822 respondents agreed to be included in a large panel data set. Analysis shows that the demographics of the final sample is similar to those who declined to participate and therefore it is assumed that there is not a substantial amount of systematic self-selection bias. For a thorough discussion of the data see Hanemann et. al. (2003).

[21] Individual recreators frequented several beaches. 73% of all beach trips were to the recreator's most frequently visited beach.

[22] Due to multiple site trips or inconsistencies among the screener, recruitment, and diary surveys 14.2% of the trip observations have been dropped from the dataset. Multiple site trips make up 3.9% of the dataset and have been excluded from this analysis due to complications in capturing the percentage cost of travel from one beach to another for the price matrix, and the proper weighting of beach attributes. Multiple site trips are commonly handled in the literature by assuming that they are independent trips.

[23] This cost is calculated as

$$Cost_{ij} = 2 * [one\ way\ travel\ dist * 0.145 + (one\ way\ travel\ time) * (0.5\ *\ hourly\ wage)] \tag{24}$$

[24] For a discussion on the percentage choice of wage rate in a travel cost model in a beach recreation

Beach water pollution data is obtained from Heal the Bay, a Southern California non-profit group. This data contains weekly ratings on a scale of A+ to F for beach water quality for dry days at many monitoring stations throughout Southern California between June 1998 and April 2001. Three water quality variables are constructed utilizing this data: yearly average grade, bimonthly average for all years, and the bimonthly worst grade reported during the survey year (Mohn et. al., 2003). In addition to these three measures a set of discrete water quality variables, indicating an F or D grade, were constructed. Table 3 reports summary statistics on the bi-monthly occurrence of water quality grade ratings, the bimonthly within beach variance for water grades, and the number of trips taken by water quality grade and variance category.

8 RECREATIONAL BEACH CHOICE MODEL

Following the literature, recreational site choice decision occasions are modeled using the discrete choice RUM as a function of site attributes, individual characteristics, and seasonal data holding the number of trips taken as exogenously determined. The CL, RPL, and FML variants of the RUM are estimated using an identical specification for the site choice probability. The FML model uses additional variables as arguments to the group membership function.

To capture the seasonal variation in preferences, a seasonal dummy is included into the segment membership function. Previous recreational modeling studies which have focused on trip temporal characteristics, such as season or part of the week, have operationalized the temporal data as an interaction variable or used it to segment the data set *a priori*. The use of the time variable in the FML enables the analyst to capture the probabilistic nature of seasonal influences on beach recreation in Southern California where there are often unseasonably warm and cold days during the winter and summer respectively.

Explanatory variables used in the RUM specifications can be categorized into beach choice and group membership variables. Modeling individual site choices for beach recreation requires explanatory variables in terms of how the beaches in the choice set differ from one another. Binary composite variables for development, very developed, wild, and ugly beaches serve to collapse twenty component attributes into four composite indicator variables (Table 4).[25]

application see Lew and Larson (2004).

[25] The data set includes a large number of beach attribute variables (42) relative to the number of beaches in the choice set (51). Therefore, a composite choice variable strategy for the appropriate right hand side variables was developed in part to handle correlation within the beach attribute data set (Mohn et. al., 2003). The variables that are used to construct the composites are 0/1 indicator variables for the absence/presence of the relevant attributes. For a detailed discussion on the formation of the composite choice set, see Hanemann

Beach choice variables incorporated into the CL, RPL, and FML model specifications include beach travel cost, water quality, the length of the beach, and a set of binary composite variables for capturing the developed, very developed, wild, or ugly nature of the beaches. Beach attribute summary statistics and correlation matrices are displayed in Table 5 and Table 6 respectively.

Group membership dummy variables used in the FML specifications indicate whether the trip occurred during winter, the recreator got in the water, the recreator is male, kids are present on the trip, the recreator is a student, the recreator works full time, and the recreator is a college graduate (Table 7 and Table 8).

The model specification reported upon in this paper is a preliminary specification designed to illustrate the level of heterogeneity which characterizes preferences for attributes that describe beach recreation site choices. The objective of this paper is to illustrate the importance of handling systematic preference heterogeneity in a discrete choice setting characterized by diverse user groups. Estimation results indicate that the FML model is a useful tool in analyzing Southern California beach choice recreational decisions. The choice model specification reported in this paper focuses on broad composite beach attribute variables and excludes several activity specific variables. Inclusion of these omitted variables is expected to impact the parameter and welfare estimates reported in this paper. Additionally, inclusion is expected to strengthen the preference group separation of the FML model due to an increase in the dimensionality of preference space.[26]

9 CHOICE MODEL ESTIMATION RESULTS

Estimation of the CL, RPL, and FML models is implemented using numerical solutions with the GAUSS programming language and the Maxlik maximum likelihood software.[27] The CL and RPL model estimation is performed using the Newton-Raphson (NR) search algorithm and the FML is estimated using the Broyden-Fletcher-Goldfarb-Shanno method (BFGS) followed by the NR method. The model specification for beach choice variables is the base model specification from the preliminary report by the Southern California Beach Valuation Project (Hanemann et. al., 2004). White's standard errors are calculated for all regressions to correct for violations of independence between observations from a respondent.

Three specifications of the model, with respect to the water quality variable, are esti-

(2004).

[26] Whereas use of composite categorical data variables as a data reduction tool leads to a loss of information in the pattern of data over the attributes and respondents; as it is the pattern of data which allows the identification of latent segments (Ramaswamy, 1999).

[27] Gauss code for the RPL is available on-line from Kenneth Train (2001).

mated. Although the results of all estimated models are qualitatively robust, the continuous yearly grade water quality variable specification is chosen over the competing specifications based on improved measure of fit statistics, improved coefficient robust standard errors, and ease of convergence. This result indicates that beach recreators may base their recreational decisions based on impressions about water quality that are formed over many years as opposed to current information. Hanemann et. al. report a similar finding regarding GEV beach choice model estimation (2004).

The CL model parameter estimates are of the expected and plausible sign, except for one variable. Parameter estimates for travel cost, and very developed are negative. Parameter estimates for water quality rating, beach length, and developed beach dummy variables are positive. Counter intuitively the ugly beach dummy variable coefficient is positive.[28] The wild beach dummy coefficient is negative and not statistically different than zero. CL model parameter estimates are presented in Table 9.

The RPL model parameter estimates are of the same sign as those of the CL model. This result is expected. However the coefficient estimate for water quality is negative and not statistically significant, and the wild beach dummy's coefficient estimate is negative and statistically significant. As expected, the RPL has greater explanatory power than the CL model indicated by high pseudo R^2 and other test statistics (Table 10). RPL model parameter estimates are presented in Table 9.

9.1 Finite Mixture Logit Segment Testing and Results

Model estimation using the FML specification allows for an increased focus regarding the heterogeneous nature of the sample population's preferences. The FML is estimated iteratively with an increasing number of preference segment groups per specification. For specification of the FML model, a complete set of beach attribute coefficients is estimated for each latent segment. Additionally, a set of probabilities for each segment is estimated assigning segment membership as a function of the individual characteristics incorporated into the model.[29] The FML model is estimated for specifications with 2, 3, and 4 segments. Following the segment testing methodology from the literature, the 4 segment model is chosen as having the greatest explanatory power. The 4 segment model (FML-4) has the highest R^2 compared to the CL, RPL, and 2 and 3 segment FML specifications. The 4 segment model also shows statistical significant improvements over the 3 segment model for several other test statistics: AIC, AIC-3, and BIC (Table 10). A 5 segment model is programmed in

[28] This is likely due to an ommited variable.

[29] The number of segments minus one set(s) of segment membership function coefficients are estimated in order to account for the indeterminacy in the model.

Gauss, but did not converge despite using a variety of parameter starting values and search techniques.[30] The lack of convergence with the 5 segment model signals that 5 segments is too many, as parameter estimates are known to tend towards negative and positive infinity when an $N+1$ segment FM model is implemented on data which actually has N preference segments (Beard et. at., 1991). Additionally the literature cautions against absolute reliance on statistical tests to determine the number of segments in a finite mixture and suggests the use of common sense (Beard et. al., 1991, McKachlan and Barford, 1988, and Boxall and Adamowicz, 2003). It is suggested that in most cases no more than 5 segments are needed in the FM framework (Heckman and Singer, 1984). The maximum number of feasible segments for a 7 dimensional preference space is 8 segments. Taking all of these factors into account, I conclude that a 4 segment FML model is the best model.

The 4 segment FML model estimates the probability that an individual is a member of each preference group conditional on the season of trip and individual recreator characteristics. Each individual choice occasion in the sample thus has a probability of being in each segment.[31] For some choice occasions the probability is high (up to 70%), while for others it approaches zero (Table 9). Segments 4, 3, and 2 are the most likely preference groups to characterize the largest number of beach choice occasions at 34.8%, 33.8%, and 25.1% of the total number of trip. Segment 1 is least likely preference group to characterize a choice occasion (6.4%) with the lowest mean percentage of group membership, 10.6%. However it has a 64.2% probability of characterizing some choice occasions.

9.2 Marginal WTP Estimates

The average beach recreator in the sample has an estimated marginal willingness to pay (WTP) of $5.71 for a water pollution rating increase of one letter grade when estimated using the 4 segment FML specification. This FML estimate is 4.64 times greater relative to the CL specification estimate of $1.23 (the WTP measure for the RPL is $0.16[32]). However, this valuation estimate ranges from negative to $17.66 for individual beach recreators (roughly 14.35 times the CL WTP measure).[33]

[30] FML model for 1 to 6 segments are programmed and estimated with simulated data consisting of 1 to 6 preference segments.

[31] Choice occasions are the individual recreator, water use, season triples that characterize each trip.

[32] Note the parameter estimate on water quality is not statistically different than zero for the RPL model.

[33] Theoretically I expect that WTP is greater or equal to 0. However, a non-negativity constraint is not imposed during the process of estimation. In the case of RPL, although the RPL mWTP is positive, a portion of the distribution of the mWTP takes on negative values. In the case of FML, I believe the negative estimates of mWTP for Group 1 and 4 are likely due to an omitted variables bias, because the model fitted here does not include certain activity-specific beach characteristic variables that are expected to impact the

Latent groups 3 and 1, respectively, have the highest and lowest mean WTP estimates for a one letter grade increase in water quality. With a mean WTP point estimate of roughly 20 times the CL WTP estimate, Group 3 membership is particularly likely for winter trips taken by male college graduates that work full time and do not have children accompanying them to the beach. Individuals with Group 3 preferences are likely to choose beaches that have long beach length, development, but are not very developed, wild, or ugly (Table 9 and Table 10).

On the bottom half of the WTP distribution, Group 1 has a mean WTP point estimate of roughly negative 10 times the CL WTP point estimate. Trips that occur during the winter, where the respondent went into the water by recreators that work full time are less likely to be characterized by Group 1 preferences. Additionally, recreators that are male, students, do not work full time, and are not college graduates are more likely to be characterized by Group 1 than Group 3. Those with Group 1 preferences are likely to choose beaches that are very developed, ugly, and have poor water quality.

The ability to construct the distribution of welfare estimates for the sample population is one of the primary benefits of the FML model. In the beach choice application each trip occasion is characterized by a constant and 8 individual and trip specific binary variables. This simple characterization of each trip by agent and seasonal characteristics results in 256 different probability assignments which are used to assign beach choice preference group membership to each choice occasion triple. The existence of multiple preference groups allows the construction of a multi-modal welfare distribution. A major strength of the FML approach is that the location within the distribution of specific welfare measures is recoverable conditional on individual and trip specific characteristics. The WTP distribution for an improvement in water quality of one letter grade illustrates the heterogeneity in preferences for coastal water quality (Figure 3). Trips that occur during the winter, involve getting in the water, and are taken by male college graduates are associated with the representative groups that have a high valuation for an improvement in water quality. Conversely, trips taken during by students are strongly associated with representative groups with low WTP for water quality.

To analyze the relationship between the estimated WTP for individual trip occasions and the group membership variables, I regress the weighted estimated WTP for each beach trip on the individual and seasonal characteristics of the trip with ordinary least squares (OLS), and both cross-section and panel specifications of generalized least squares (GLS)(Table 11). Coefficient estimates for the winter trip, in the water, and college graduate variables are positive for all three estimators. The coefficient estimate for the student variable is negative

parameter estimates.

for all three estimators.

The coefficient estimate on the children present on trip variable is of particular interest. The OLS coefficient estimate for this variable is negative, whereas both GLS models produce coefficient estimates that are not significantly different than zero. However, the introduction of an interaction term for trips characterized by both the presence of children and getting into water produces negative and significant coefficient estimates for all three estimators. One would expect that beach trips that are taken with children and involve water recreation would have a higher probability of occurring at beaches with higher levels of water quality and would be associated with higher WTP for water quality. One explanation for this result may be that the polluted beaches are characterized by features that are perceived by parents to provide safer environments for their children to swim, such as a lack of surf, but at the same time perpetuate water pollution. This result suggests further research. As illustrated by the paradoxical above result above the ability to construct the distributions of the relative importance which site attributes have on site choice is an important tool for resource and health officials charged with the management of resources used by diverse user groups.

10 CONCLUSION

Coastal water quality impacts recreation and tourism. Southern California beach recreators cite pollution as a primary reason for abstaining from swimming, a belief supported by studies linking swimming in polluted water with illness. While there is interest in understanding the impact of water quality on beach recreation to improve resource and public health management, this task is complicated by the diversity of user preferences and the multiuse nature of the beach.

This paper implements the FML RUM to highlight the importance of capturing preference heterogeneity. Exploiting an extensive beach recreational panel data set, this paper furthers the literature by applying the FML approach to model preference heterogeneity regarding the impact of an environmental variable related to health and seasonality on recreational choice. The application also increases the number of choice alternatives and the number of variables included in the segment membership function estimated with the FML model in the literature.

Application of the FML model to the Southern California beach recreation data set recovers 4 preference groups, highlighting the variation in the importance of water pollution on beach choice for a diverse sample of beach users. For these groups, the impact that water quality has on recreational site choice, as measured by the mean WTP, ranges from negative to \$17.66, with an average of \$5.71, for an improvement in water quality of one letter grade.

The WTP estimate calculated with the CL model is $1.23. The RPL coefficient estimate on water quality is not significantly different from 0, and yields a WTP estimate of $0.16.

The FML approach facilitates the estimation of the distribution of water quality preferences and welfare measures across a diverse user-base, and enables researchers to identify user preference groups characterized by several variables. This increases the ability of resource managers to forecast the impact that changes in site characteristics will have on the beach choice and welfare across segments of society.

Estimation results indicate that recreators who enter the water have a higher estimated WTP for water quality. Gender, employment, education, and seasonal variables are also important in estimating one's preferences. One troubling result of the model is the finding that the presence of children on beach trips which include water activity is **not** associated with a higher WTP for improvements in water quality. This result highlights the model's ability to identify groups that resource managers and public health officials may desire to concentrate their educational outreach efforts.

The FML approach is likely to become increasingly important as diversity continues to grow, and the identification of user groups by a small number of variables becomes less feasible. The application of the model to a unique beach recreation data set is of major significance from the environmental management perspective. The powerful combination of being able to specify a model which simultaneously estimates the marginal benefits associated with different attributes for different groups and assigns group membership makes FML a particularly attractive model for policy analysis.

Figure 1:
Constant over Time LS Membership

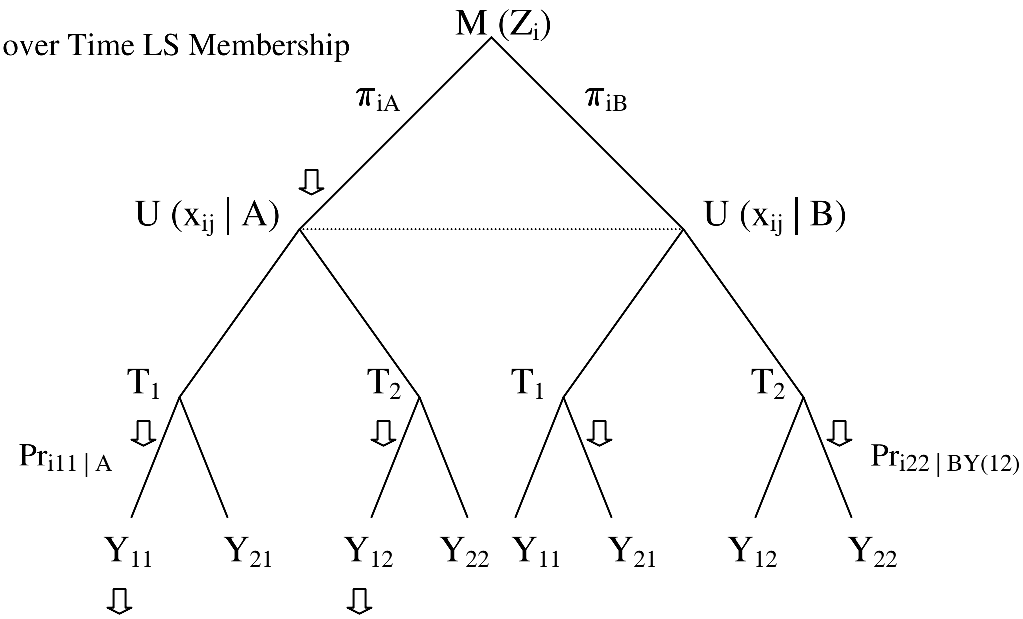
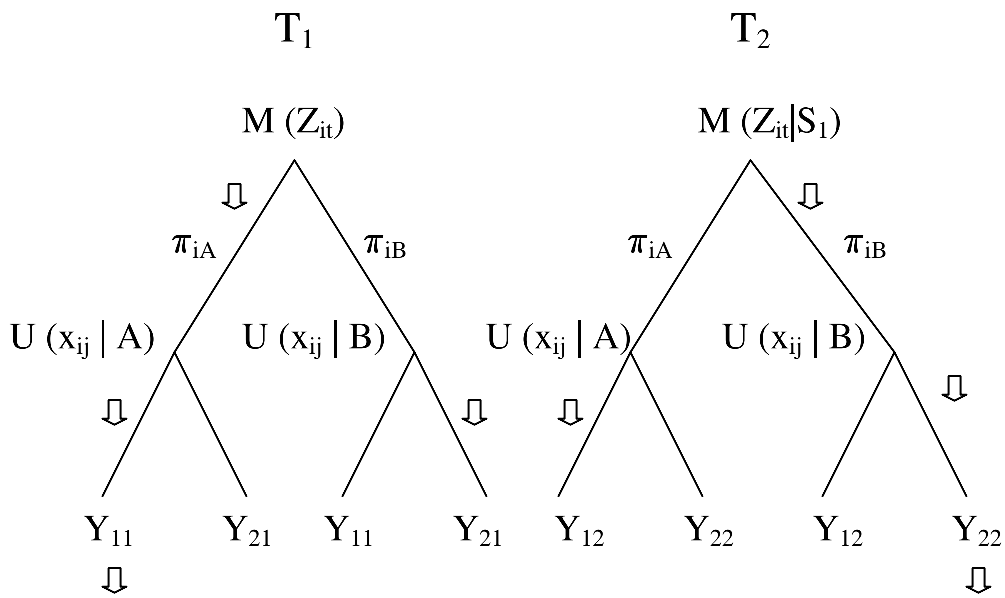


Figure 2:
Variation over Time LS Membership



Note:
Indexes: i= individual, j=site choice, t=time
Segment membership types are denoted by S={A,B} where A & B are segment types
$Pr_{i11 | A}$ indicates the probability of individual *i* choosing site 1 in time period 1, conditional on membership in segment A.
Choice outcomes, Y_{jt}, are indexed by choice then by time.

Table 1: Probablility of Water Recreation

	Trips	Recreators	Average individual %	Trip %	Recreator %
Total	4642	595	22%	27%	23%
Winter	987	222	9%	14%	6%
Summer	1749	378	30%	38%	62%
Shoulder Season	1906	377	22%	23%	58%

Table 2: Beach Site Trip Counts

Trips Count per Beach	Total Trips	Winter	Winter In Water	Shoulder Season	Shoulder Season In Water	Summer	Summer In Water
Minimum	0	0	0	0	0	0	0
Average	*91.02*	*16.67*	*2.69*	*28.90*	*8.47*	*21.27*	*13.02*
Maximum	*659*	*214*	*32*	*154*	*66*	*124*	*110*
Standard Deviation	*357.11*	*119.03*	*17.75*	*81.85*	*35.91*	*66.31*	*60.10*
Total	*4642*	*850*	*137*	*1474*	*432*	*1085*	*664*

Table 3: Grade and Grade Variance

GRADE	F	D	C	B	A		
Occurrence	7	8	22	62	207		
% Occurrence	2.5%	2.5%	7.0%	17.0%	58.0%		
Trips	53	329	483	1542	2232		
% Trips	1.2%	7.0%	10.0%	33.0%	48.0%		
GRADE VARIANCE	0	0.25	0.5	0.75	1	1.25	1.5
Occurrence	34	6	2	0	2	0	4
% Occurrence	71.0%	13.0%	4.0%	0.0%	4.0%	0.0%	8.0%
Trips	3883	369	5	0	289	0	96
% Trips	84.0%	8.0%	0.0%	0.0%	6.0%	0.0%	2.0%

Table 4: Composite Variables and Their Components

COMPOSITE VARIABLES		COMPONENT VARIABLES
Developed Beach	3 or more	Access_Street
Very Developed Beach	8 or more	Public Transit
		Restaurants
		Stores
		Concessions
		Rentals
		Beach Clubs
		Houses
		Condos/Hotels
		Pier
		Concerts
		Volley Ball Tournaments
Wild Beach	1 or more	Pedestrian Access Only
		Rocky
		Tide pools
		Dogs Allowed
Ugly Beach	1 or more	Oilpumps
		Oilrigs
		PowerSewer
		Stormdrains

Table 5: Choice Variable Summary Statistics

Choice Variables	Mean	Min	Max	Std Dev.
Cost	5.546	3.183	10.027	2.375
Water Quality	3.602	1.373	4.333	0.757
Beach Length (ln)	0.352	-2.207	2.088	0.940
Developed	0.549	0.000	1.000	0.503
Very Developed	0.196	0.000	1.000	0.401
Wild	0.314	0.000	1.000	0.469
Ugly	0.275	0.000	1.000	0.451

Table 6: Correlation of Choice Variables

	Cost	Water Quality	Beach Length (ln)	Developed	Very Developed	Wild	Ugly
Cost	1	0.034	-0.347	-0.168	-0.215	-0.013	-0.329
Water Quality	0.034	1	-0.099	-0.293	-0.009	0.094	0.038
Beach Length (ln)	-0.347	-0.099	1	0.302	0.385	-0.137	0.108
Developed	-0.168	-0.293	0.302	1	0.448	-0.236	0.116
Very Developed	-0.215	-0.009	0.385	0.448	1	-0.227	-0.082
Wild	-0.013	0.094	-0.137	-0.236	-0.227	1	-0.132
Ugly	-0.329	0.038	0.108	0.116	-0.082	-0.132	1

Table 7: Membership Variable Summary Statistics

Membership Variables	Mean	Min	Max	Std Dev.
Constant	1.000	0.000	1.000	0.000
Winter	0.213	0.000	1.000	0.409
Summer	0.377	0.000	1.000	0.485
In Water	0.266	0.000	1.000	0.442
Male	0.561	0.000	1.000	0.496
Kids	0.266	0.000	1.000	0.442
Student	0.175	0.000	1.000	0.380
Work Fulltime	0.649	0.000	1.000	0.477
College Grad	0.534	0.000	1.000	0.499

Table 8: Correlation of Membership Variables[a]

	Winter	Summer	In Water	Male	Kids	Student	Work Fulltime	College Grad
Winter	1	-0.404	-0.149	0.061	-0.079	-0.032	0.068	0.089
Summer	-0.404	1	0.201	-0.056	0.098	-0.017	-0.064	-0.011
In Water	-0.149	0.201	1	0.082	0.064	0.010	0.003	-0.010
Male	0.061	-0.056	0.082	1	-0.210	-0.075	0.238	0.037
Kids	-0.079	0.098	0.064	-0.210	1	-0.014	-0.058	-0.125
Student	-0.032	-0.017	0.010	-0.075	-0.014	1	-0.148	-0.116
Work Fulltime	0.068	-0.064	0.003	0.238	-0.058	-0.148	1	0.134
College Grad	0.089	-0.011	-0.010	0.037	-0.125	-0.116	0.134	1

[a] Constant Excluded

Table 9: Parameters (t-statistic) on Choice and Membership Variables

Variable	Logit	Random Parameters Model		4 Segment Model			
Choice Variables							
		Mean	SD	Segment 1	Segment 2	Segment 3	Segment 4
Cost	-0.085	-0.182	0.109	-0.653	-0.021	-0.408	-0.366
	(-50.887)	(-34.016)	(23.921)	(-6.074)	(-11.919)	(-15.980)	(-10.415)
Water Quality	0.105	0.028[a]	-0.007[a]	-7.950	0.047[a]	10.382	-0.637
	(4.316)	(1.008)	(-0.055)	(-4.395)	(0.852)	(10.673)	(-5.606)
Beach Length (ln)	0.470	0.567	-0.006[a]	-0.871	0.259	2.160	0.814
	(18.627)	(19.184)	(-0.114)	(-2.320)	(5.166)	(9.73)	(7.508)
Developed	0.789	1.192	-1.885	1.422	0.527	1.998	-0.448
	(17.456)	(5.770)	(-4.317)	(3.541)	(5.200)	(11.693)	(-2.226)
Very Developed	-0.097	-2.271	9.546	8.857	0.637	-6.347	1.836
	(-2.458)	(-2.728)	(3.252)	(4.482)	(5.746)	(-14.289)	(8.261)
Wild	-0.008[a]	-0.662	2.200	-2.291	0.206	-6.253	1.706
	(-0.192)	(-4.040)	(7.537)	(-3.995)	(2.271)	(-7.288)	(8.754)
Ugly	0.073	0.100	-0.364[a]	10.343	0.537	-8.461	0.748
	(2.122)	(2.186)	(-0.889)	(4.156)	(6.220)	(-13.663)	(6.369)
Membership Variables							
				Segment 1	Segment 2	Segment 3	Segment 4
Constant				-2.674	-0.788	-1.322	0
				(-8.781)	(-4.943)	(-8.871)	
Winter				0.299[a]	-0.204[a]	0.704	0
				(1.392)	(-1.295)	(5.162)	
Summer				1.195	0.282	0.516	0
				(6.006)	(2.058)	(3.994)	
In Water				-6.814	0.294	0.028[a]	0
				(-2.119)	(1.962)	(0.205)	
Male				1.907	-0.129[a]	0.855	0
				(7.321)	(-0.929)	(6.334)	
Kids				0.204[a]	0.446	0.120[a]	0
				(0.977)	(3.219)	(0.888)	
Student				0.083[a]	0.313[a]	-0.786	0
				(0.287)	(1.775)	(-3.823)	
Work Fulltime				-0.980	0.480	-0.309	0
				(-4.972)	(3.757)	(-2.597)	
College Grad				1.235	0.209[a]	1.006	0
				(5.764)	(1.616)	(8.206)	
Segment Membership Probability				Segment 1	Segment 2	Segment 3	Segment 4
Minimum				0.0%	5.7%	3.2%	10.2%
Mean				10.6%	29.8%	27.2%	32.4%
Maximum				64.2%	70.4%	69.8%	55.9%
Percent Sgement Membership by Max. Probability				Segment 1	Segment 2	Segment 3	Segment 4
				6.4%	25.1%	33.8%	34.9%

[a]Indicates that the parameter is **not** significantly different than 0 at the 5% level.
T-statistics are calculated using White's standard errors.

Table 10: Model Selection Statistics and Welfare Estimates
Finite Mixture Logit, Conditional Logit, and Random Parameters Logit Model Estimation Results[a]

Model	Logit	FML[b]			RPL[c]
Number of Segments	1	2	3	4	
Log Likelihood at Convergence (LL)	-14014.08	-12863.50	-12317.03	-12066.10	-13380.74
Log Likelihood Evaluated at 0 (LL0)	-18251.55	-18251.55	-18251.55	-18251.55	-18251.55
Number of Parameters (P)	7	23	39	45	14
AIC[d]	28042.16	25773.01	24712.06	24222.20	26789.48
AIC-3[e]	42063.25	38659.51	37068.08	36333.31	40184.23
BIC[f]	14043.63	12960.60	12481.66	12256.07	13439.84
ρ^2[g]	0.2322	0.2952	0.3252	0.3389	0.2669
Willingness to Pay by Segment		-\$6.87	-\$7.66	-\$12.18	
		\$18.40	\$7.37	\$2.19	
			\$21.03	\$25.46	
				-\$1.74	
Average Willingness to Pay[h]	\$1.23	\$5.64	\$5.89	\$5.71	\$0.16

[a]Sample size is 4642 choices from 595 individuals (N).
[b]FML represents the finite mixed logit model.
[c]RPL represents the random parameters logit model.
[d]AIC (Akaike Information Criterion) is calculated using {-2(LL-P)}.
[e]AIC-3 (Akaike Information Criterion-3) is calculated using {-3(LL-P)}.
[f]BIC (Bayesian Information Criterion) is calculated using {-LL+[(P/2)*ln(N)]}.
[g]ρ^2 is calculated as 1-(LL)/LL(0).
[h]Average Willingness to Pay is a weighted average of the willingness to pay by segment, using estimates of segment membership. Weighted WTP ranges from -$4.06 to $17.66

Table 11: Results for Estimated mWTP for Improvements in Water Quality

Regressors from Group Membership Function	Coefficients[a]					
	OLS		GLS		GLS panel	
	restricted		restricted		restricted	
Intercept	**2.451**	**2.366**	**2.680**	**2.691**	**2.294**	**2.234**
	(0.058)	(0.058)	(0.041)	(0.039)	(0.149)	(0.148)
Winter	**3.803**	**3.780**	**3.640**	**3.746**	**4.120**	**4.043**
	(0.058)	(0.057)	(0.400)	(0.035)	(0.213)	(0.211)
Summer	**0.138**	**0.129**	**0.464**	**0.324**	0.247	0.228
	(0.049)	(0.048)	(0.040)	(0.038)	(0.182)	(0.180)
Water	**3.626**	**4.016**	**2.241**	**3.023**	**3.941**	**4.273**
	(0.050)	(0.058)	(0.049)	(0.063)	(0.141)	(0.160)
Male	**1.085**	**1.057**	**1.478**	**1.365**	**1.050**	**1.028**
	(0.046)	(0.045)	(0.032)	(0.031)	(0.100)	(0.098)
Kids Present	**-0.145**	**0.237**	0.033	0.023	-0.165	0.218
	(0.050)	(0.058)	(0.038)	(0.029)	(0.110)	(0.142)
Student	**-3.656**	**-3.660**	**-3.355**	**-3.516**	**-3.654**	**-3.657**
	(0.057)	(0.056)	(0.045)	(0.045)	(0.125)	(0.123)
Work Full-time	**-0.272**	**-0.260**	**-0.369**	**-0.263**	**-0.273**	**-0.261**
	(0.047)	(0.046)	(0.039)	(0.036)	(0.102)	(0.101)
College Graduate	**3.154**	**3.165**	**2.769**	**2.671**	**3.134**	**3.149**
	(0.044)	(0.043)	(0.033)	(0.031)	(0.096)	(0.095)
Kids Present on Water Trip		**-1.318**		**-1.119**		**-1.312**
		(0.105)		(0.088)		(0.312)
Regression Statistics						
R Squared	0.828	0.83318111	0.8552	0.8736		
R Squared - overall					0.8261	0.8322
Adjusted R Squared	0.827	0.83285698	0.8549	0.8733		
Observations	4642	4642	4177	4264	4642	4642

[a] All coefficient estimates are significant at the 1% level.

[b] Standard Errors are in parenthesis.

[c] **bold** indicates signicantly different from 0, at the 1% level.

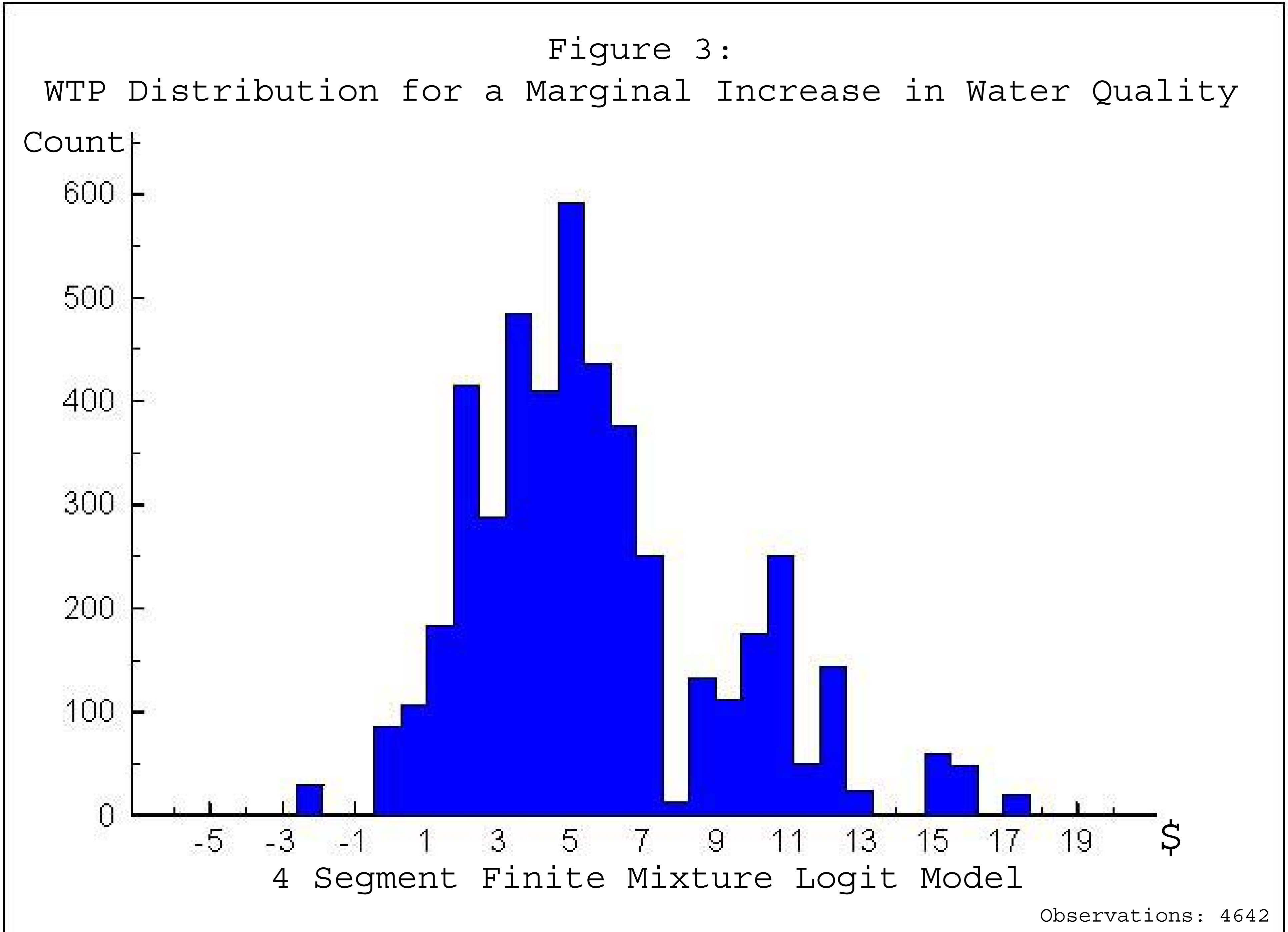


Figure 3:
WTP Distribution for a Marginal Increase in Water Quality
Count
600
500
400
300
200
100
0
-5
-3
-1
1
3
5
7
9
11
13
15
17
19
$
4 Segment Finite Mixture Logit Model
Observations: 4642

References

[1] 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.

[2] Andrews, R. and I. Currim. 2003. A Comparison of Segment Retention Criteria for Finite Mixture Logit Models. *Journal of Marketing Research*. Vol. 40. pp. 235-243.

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

[4] Ben-Akiva, M. and J. Swait. 1986. The Akaike Likelihood Ratio Index. *Transportation Science*. 20: 133-136.

[5] Bockstael, N.E., W.M. Hanemann, and C.L. Kling. 1987. Estimating the Value of Water Quality Improvements in a Recreational Demand Framework. *Water Resources Research*. Vol. 23, no. 5, pp. 951-960.

[6] Bockstael, N.E., W.M. Hanemann, and I.E. Strand, Jr. 1986. Measuring the benefits of water quality improvements in a recreational demand framework. *Water Resources Research*. Vol. 23, no. 5, pp. 951-960

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

[8] Chapman, D.J. and W. Michael Hanemann. "Environmental Damages In Court: The American Trader Case," In Anthony Heyes' (ed) The Law and Economics of the Environment, pp. 319-367, 2001.

[9] Deacon, R. T., C.D. Kolstad. 2000. Valuing Beach Recreation Lost in Environmental Accidents. *Journal of Water Resources Planning and Management*, Vol. 126, No. 6, November/December 2000, pp. 374-381

[10] Dempster, A.P., N. M. Laird, and D. B. Rubin. 1977. Maximum Likelihood from Incomplete Data via the EM Algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)*, Vol. 39, No. 1. (1977), pp. 1-38.

[11] Ebert, U. 1998. Evaluation of Nonmarket Goods: Recovering Unconditional Preferences. *American Journal of Agricultural Economics*. Vol 80. pp. 241-254.

[12] Genius, M. and E Strazzera. December, 2000. Evaluation of Likelihood Based Tests for Non-Nest Dichotomous Choice Contingent Valuation Models. Working Paper. Centro Ricerche Economiche Crenos. http://www.crenos.it/working/working.html.

[13] Gordon, Peter, John Kuprenas, Jiin-Jen Lee, James E. Moore, Harry W. Richardson, Christopher Williamson. An Economic Impact Evaluation of Proposed Storm Water Treatment for Los Angeles County. Los Angeles: School of Engineering and School of Policy, Planning, and Development, University of Southern California. November 2002.

[14] Greene, W.H. 2000. Econometric Analysis. Fourth Edition. Prentice-Hall, New Jersey.

[15] Greene, W, and D. Hensher. 2003. A Latent Class Model for Discrete Choice Analysis:Contrasts with Mixed Logit. *Transportation Research Part B*. Vol 37. pp. 681-698.

[16] Haab, T,C,, and K.E. McConnell. 2003. Valuing Environmental and Natural Resources: The Econometrics of Nonmarket Valuation. Edward Elgar Publishing, Massachusetts.

[17] Haaijer, R., and M. Wedel. 2000. Habit Persistence in Time Series Models of Discrete Choice. *Marketing Letters*. Vol. 12, pp 25-35.

[18] Haile RW, Alamillo J, Barrett K, et al. An Epidemiological Study of Possible Adverse Health Effects of Swimming in Santa Monica Bay: Final Report. Santa Monica, Calif: Santa Monica Bay Restoration Project; 1996.

[19] Hanemann, W.M. 1982. Applied welfare analysis with qualitative response models, Working Paper. No. 241. University of California, Berkeley, 26 pp.

[20] Hanemann, W.M., J.R. Hilger, C. Mohn, and L. Pendleton. Novermber, 2003. Data Reduction for Los Angeles Area Beach Site Attributes. Southern California Beach Project.

[21] Hanemann, W.M., L. Pendleton, C. Mohn, J. Hilger, K. Kurisawa, D. Layton, C. Busch, and F. Vasquez. August, 2004. Using revealed prefernce models to estimate the effect of coastal water quality on beach choiec in Southern California. Southern California Beach Project.

[22] Hanemann, W.M. Personal communication. November, 2005. Berkeley, California.

[23] Hausman J., D. McFadden. 1984. Specification Tests for the Multinomial Logit Model. *Econometrica*. Vol. 52. pp. 1219-1240.

[24] Heckman, J., and B. Singer. 1984. A Method for Minimizing the Impact of Distibutional Assumptions in Econometric Models for Duration Data. *Econometrica*, Vol. 52, No. 2., pp 271-320.

[25] Helmbold, D, R. Schapire, Y. Singer, and M. Warmuth. 1995. A comparison of new and old algorithims for a mixture estimation problem. *Annual Workshop on Computational Learning Theory*. Santa Cruz, California, United States. pp. 69 - 78

[26] Jedidi, K., H.S. Jagpal, and W.S. DeSarbo. 1997. Finite_mixture Structural Equation Models for Response-Based Segmentation and Unobserved Heterogeneity. *Marketing Science*. Vol. 16, pp 39-59.

[27] Kamakura, W., and G. Russell. 1989. A probabiliitstic Choice Model for Market Segmentation and Elasticity Structure. *Journal of Marketing Research*. Vo. 26, pp 379-390.

[28] Kamakura, W. and M. Wedel. 2004. An Empirical Bayes Procedure for Improving Individual-Level Estimates and Predictions from Finite Mixtures of Multinomial Logit Models. *Journal of Business and Economic Statistics*. Vol. 22. pp. 121-125.

[29] Lew, D.K., Larson, D.M. 2005. Valuing Recreation and Amenities at San Diego County Beaches. *Coastal Management*, 33:71-86.

[30] McFadden, D. 1973. Conditional Logit Analysis of Qualitative Choice Behavior. In P. Zarembka, ed., *Frontier in Econometrics*. New York: Academic Press.

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

[32] Mohn, C., M.H. Hanemann, L. Pendleton, and D. Layton. 2003. Shifting Sands: A Preliminary Exploration of Seasonal Variation in Preferences for Beaches. Working Paper.

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

[34] Morey, E.R. 2003. 'Latent Segmentation Likelihood Functions.' Personal Communication.

[35] Naik, P.A, M.R. Hagerty, and C. Tsai. 2000. A New Dimension Reduction Approach for Data_Rich Marketing Environments: Sliced Inverse Regression. *Journal of Marketing Research*. Vol 37, pp 88-101.

[36] Natural Resources Defense Council. 2002. Testing the waters: A guide to water quality at vacation beaches, 2002. Volume 12, pp 24-57.

[37] Natural Resources Defense Council. 2005. Testing the Waters: A Guide to Water Quality at Vacation Beaches.

[38] Office of Environmental Health Hazard Assessment. June, 2001. California: Sport Fish Consumption Advisories. www.oehaa.ca.gov

[39] Parsons, G. and D. Massey. 2003. "A Random Utility Model of Beach Recreation" in The New Economics of Outdoor Recreation, N. D. Hanley, ed. Edward Elgar. pp 241267.

[40] Pendleton, L. 2001. Managing Beach Amenities to Reduce Exposure to Coastal Hazards: Storm Water Pollution. *Coastal Management*, 29:239-252.

[41] 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*. Vol 84, pp 1066-1075.

[42] Provencher, B. and R. Bishop. 2004. Does accounting for preference heterogeneity improve the forecasting of a random utility model? A case study. Journal of Environmentatl Economics and Management. Vol. 48. pp. 793-810.

[43] Rabinovici, S.J.M., Bernknopf, R.L., Wein, A.M., Coursey, D.L., and Whitman, R.L. 2004. Economic and Health Risk Trade-Offs of Swim Closures at a Lake Michigan Beach. *Environmental Science and Technology*. Vol. 38, No. 10, pp 2737-2745.

[44] Ramaswamy, V., R. Chatterjee, and S. Cohen. 1999. Reply to "A Note on Ramaswamy, Chatterjee, and Cohen's Latent Joint Segmentation Models." *Journal of Marketing Research*. Vol 36, pp 115-119.

[45] Richardson, S. and P. Green. 1997. On Bayesian Analysis of Mixtures with an Unknown Number of Components. *Journal of the Royal Statistical Society. Series B*. Vol 59. pp. 731-792.

[46] Ruud, P.A. Extension of estimation methods using the EM algorithm. *Journal of Econometrics*. Vol 49, pp 305-341.

[47] Scarpa, R., K. Willsi, ad M. Acutt. 2004. Comparing Individual-Specific Benefit Estimates for Public Goods: Finite Versus Continous Mixing in Logit Models. Social Science Research Network Electronic Paper Collection. http://feem.it/Feem/Pub/Publications/WPapers/default.htm

[48] Shonkwiler, J.S. and W.D. Shaw. 2003 "A Finite Mixture Approach to Analyzing Income Effects in Random Utility Models" in The New Economics of Outdoor Recreation, N. D. Hanley, ed. Edward Elgar, pp. 268-278.

[49] Sorensen, B. 1975. Energy and resources. *Science*. Vol. 189 : 255-60

[50] Swait, J. and J. Louviere. 1993. The role of the scale parameter in the estimation and use of multinormal logit models. *Journal of Marketing Research*. Vol. 30. pp 305-314.

[51] Swait, J.R. (1994) A structural equation model of latent segmentation and product choice for cross-sectional revealed preference choice data. *Journal of Retailing and Consumer Services* Vol. 1, no 2, pp. 77-89.

[52] Train, K. (1997) Recreation Demand Models with Taste Differences Over People. *Land Economics*. Vol. 74, No. 2, May 1998.

[53] Train, K. 2001. Discrete Choice with Simulation Gauss Code. http://elsa.berkeley.edu/Software/abstracts/train0401.html

[54] Train, K. 2003. Discrete Choice Methods with Simulation, Cambridge. Cambridge University Press.

[55] United States Forest Service. National Survey on Recreation and the Environment: American's Participation in Outdoor Recreation. Southern Research Station.

[56] Vermunt, J.K. 2004. An EM algorithm for the estimation of parametric and nonparametric hierarchical nonlinear models. *Statistica Neerlandica.* Vol. 58, nr. 2, pp. 220-233.

[57] Wedel, M., W.S. DeSarbo, J.R. Bult, and V. Ramaswamy. 1993. A Latent Class Poisson Regression Model for Heterogeneous Count Data. *Journal of Applied Econometrics*. Vol. 8, No. 4, pp 397-411.

[58] Wooldridge, J.M. 2002. Econometric Analysis of Cross Section and Panel Data. MIT Press, Massachusetts.

[59] Yatchew, A. and Griliches, Z. 1984. Specification error in probit models. *Review of Economics and Statistics*. Vol. 66, pp. 134-139.