Familiar and Favorite Sites in a Random Utility Model of Beach Recreation

GEORGE R. PARSONS
University of Delaware

D. MATTHEW MASSEY
West Chester University

TED TOMASI
Entrix, Inc

Abstract *We estimate a random utility model of recreation demand accounting for choice set familiarity and favorite sites. Our approach differs from existing approaches by retaining all sites in estimating the parameters of site utility. Familiar and unfamiliar sites are specified with different utility functions. Favored sites are assumed to have higher utility than nonfavored sites in estimation.*

Key words Beach, random utility, recreation.

Introduction

In conventional applications of the random utility model to recreation demand a researcher specifies a broad universe of sites and uses it as the choice set in estimation.[1] This approach has been questioned for at least two reasons. One, individuals may be unfamiliar with many of the sites. And two, individuals may only seriously consider a much narrower set of favorite sites. To improve matters it is suggested that individual-based (endogenous) choice sets be used in place of researcher-based (exogenous) choice sets. The individual-based choice sets hinge on gathering data from individuals on their familiar and favorite sites.

There is a small literature on incorporating individual-based choice sets into random utility models of recreation demand. In one of the earliest RUM applications, Caulkins, Bishop, and Bouwes (1986) limited the choice set for each individual to sites actually visited. The typical choice set size for a person was about 3 or 4 sites from a universe of over a thousand sites. Individuals visiting only one site were eliminated from the data set. The final choice set certainly used familiar and favorite sites. Whether this motivated the definition of choice set is unclear.

Peters, Adamowicz, and Boxall (1995) estimated a RUM of fishing in Alberta.

[1] For some recent applications of the random utility model see Kaoru, Smith, and Liu (1995), McConnell and Strand (1994), Montgomery and Needelman (1997), Morey, Rowe, and Watson (1993), Needelman and Kealy (1995), Shaw and Jakus (1996), and Train (1998).

In their in-person survey each respondent was shown a list of sites and asked to indicate 'which of these sites they have visited in the past or would consider when choosing a site to go fishing.' They estimated a conventional model using the full set of 67 sites for each individual and compared it to a model that included only past visited and considered sites. The average choice set size for the latter model was 33 sites. They report large differences in the parameter and welfare estimates between the two models.

Hicks and Strand (2000) follow a similar approach in an application to beach use on the Chesapeake Bay. Individuals were asked in an in-person survey to identify beaches they were familiar with either through previous experience or having heard about it through another source. A conventional model with the full set of 10 sites was compared to a model using only familiar sites. The familiar choice set averaged 2.5 sites. Again, they found large differences in parameter and welfare estimates. They also compare the familiar based choice sets with some distance based choice sets and found differences here as well.

Adamowicz, Swait, Boxall, Louviere, and Williams (1997) estimate a model of moose hunting in Alberta. In an in-person survey hunters were asked to provide their perceptions of several attributes (*e.g.*, moose population and hunter congestion) at each of 14 sites. If the individual was unsure of an attribute's quality or simply unfamiliar with the site, he or she was free to answer 'I don't know' for that attribute. If an individual said 'I don't know' for all attributes (except for travel distance and road quality), it was assumed that the site was not in the choice set. This was Adamowicz *et al.*'s perception model which was compared to a conventional model with objective measures and to several stated-preference models. It was also combined with the stated preference models. Again, wide variation is reported in the parameter and welfare estimates, but a number things beyond choice set configuration were being changed.

All four papers follow basically the same approach. The researcher finds the set of sites that the person 'really knows and/or thinks about when taking a trip' and only includes these in the choice set. They differ in their criteria for narrowing the choice set: currently visited sites only, past and considered sites, familiar sites, and perceived sites.

Another approach is to model the choice set as endogenous. In a two stage model one first predicts the probability that a site is in the narrow set of sites the person 'really thinks about' and then predicts which among these sites is actually visited. Both stages use site attributes as explanatory variables. Ben-Akiva and Boccara (1995) layout the technical approach and present an application without welfare analysis and in the context of transportation demand. Horowitz and Louviere (1995) present a similar application in the marketing context and again without welfare analysis. The closest application to recreation demand along these lines is by Haab and Hicks (1997). Their approach is different, however, in that they have no familiar and favorite site data and design a modelling approach that incorporates all possible choice sets in the first stage. Comparison of the endogenous choice set models with conventional approaches again yield large differences in parameter and welfare estimates.[2]

We offer a different approach for incorporating familiar and favorite data into the standard model. In our approach, sites are never deleted from the choice set. We

───────

[2] A number of other related articles dealing with choice set definition issues such as aggregation and number of sites to include the choice set are Parsons and Needelman (1992), Kaoru, Smith, and Liu (1995), Feather (1994), Parsons and Hauber (1998), Jones and Lupi (1999), Lupi and Feather (1998), and Parsons, Plantinga, and Boyle (2000). None of these articles, however, deals with individual-based choice sets.

believe that there is vital preference information in observing and thus accounting econometrically for which sites are familiar and favorite. For this reason (and we will elaborate in the following sections) we choose to keep the full set intact in estimation.

We handle familiar site data by specifying a different utility function for familiar and unfamiliar sites. Site characteristics, no doubt, play a different role in site choice for these two groups. We handle favorite site data following Horowitz and Louviere (1995). In their approach, favorite (considered) sites are not deleted from the model in estimation or used to model endogenous choice set formation. Instead, favorite sites are simply treated as sites that have higher utility than sites not identified as favorite by the respondent. In effect, the information is used to sharpen the parameter estimates in the basic random utility model.

We estimate a basic site choice model and then compare it to models that uses the familiar and favorite sites information as described above. We also estimate and compare our approach to a Peters *et al.* (1995) model and a Hicks and Stand (2000) model. We begin with a description of the basic model.

Basic Model

We estimate our random utility model as a multinomial logit. Our application is to recreational beach use. In our model an individual is selecting one site from a universe of C possible beaches. We assume that each beach has the following site utility

$$U_i = X_i\beta + \varepsilon_i \tag{1}$$

X_i is a vector of the characteristics at site i ($i = 1, .., C$), including travel and time cost, beach amenities and so on. β is an unknown parameter vector, and ε_i is a random disturbance. We assume an individual selects the beach giving the highest utility.

We also assume that the C random disturbance terms ε_i are independent and identically distributed type 1 Weibull random variables. This yields the well known logit probability for visiting any given site k in the choice set

$$\Pr(visit_k) = \frac{\exp(X_k\beta)}{\sum_{i=1}^{C} \exp(X_i\beta)}. \tag{2}$$

The numerator is the exponential of the visited site utility, and denominator is the sum of the exponential over all site utilities in the universe of sites. Given observational data on people having visited one or more of the sites in the set of C and characteristics of the sites X_i, the logit probabilities are used in a standard likelihood function to estimate the parameters β. If multiple visits are made by a person over the season, current practice is to enter each chosen site probability separately in the likelihood function. In this sense our observations are 'trips'.

An individual's expected utility of visiting a site on any given choice occasion in the logit model is

$$\ln \sum_{i=1}^{C} \exp(X_i\beta). \tag{3}$$

This term is just the natural log of the denominator from the logit probability, again it sums over the C sites in the choice set. In our application we are interested in in-

vestigating the value of lost beach use (site closure) and reduced beach width. The lost consumer surplus associated with the loss of site 1 in the set of *C* sites is just the change in expected utility divided by the individual's marginal utility of income

$$W^1 = \frac{\ln \sum_{i=1}^{C} \exp(X_i\beta) - \ln \sum_{i=2}^{C} \exp(X_i\beta)}{\beta_{price}}. \tag{4}$$

β_{price} is the marginal utility of income and is the parameter estimated on the travel plus time cost variable in the logit. Notice that the first term in the numerator sums over all *C* sites, while the second term drops site 1. The closure of more than one site is measured by dropping more sites from the second term.

The lost consumer surplus associated with reduced beach width is

$$W^2 = \frac{\ln \sum_{i=1}^{C} \exp(X_i\beta) - \ln \sum_{i=1}^{C} \exp(X_i^*\beta)}{\beta_{price}} \tag{5}$$

where X_i is the vector of site attributes at site i in the baseline case, and X_i^* is the set of attributes in the case with narrower beaches due to erosion.

In our application, the universe of *C* sites is a set of 62 beaches in the Mid-Atlantic region. In the Basic Model all 62 sites enter the choice set and the model is estimated using the choice probabilities formulated as above. This raises a rather thorny issue for any empirical study of choice set formation. The so called 'full set' or universe of sites *C* must be determined by the researcher. But, this is a choice set definition issue in and of itself. In our case, we use all ocean beaches within a days drive of Delaware residents. But, we excluded bay beaches, inland rivers and lakes, as well as other forms of recreation and leisure. It is an arbitrary, but necessary, starting point.

The data used to estimate the model are from a random mail survey of 1,000 Delaware residents in the fall of 1997. We have 565 completed surveys from that sample. Individuals were asked to record their number of day, short overnight, long overnight, and extended stays at the 62 beaches in the Mid-Atlantic region for the year. From the sample of 565 individuals, 400 took day trips. Our analysis is confined to this set of 400.

The 62 beaches run from Sandy Hook (NJ) to Assateague Island (VA) (see figure 1). We also gathered data on location of hometown and demographics. The beach characteristic data used in our current model are shown and documented in table 1. The distances and travel times to beaches from each respondent's home was computed using PC-Miler. For travel cost we assume 35 cents per mile. For hourly time cost we use annual income divided by 2,040 (the typical number of hours a person works in a year).

Familiar Sites Model

Modeling Strategy

Hicks and Strand (2000) and Peters, Adamowicz, and Boxall (1995) identify sites unfamiliar to respondents in their sample, and then, in estimation, eliminate these sites from the choice set. The reasoning is that these alternatives play no role in site selection and hence do not belong in a model that purports to measure tradeoffs

among site characteristics made by respondents. So long as an individual is entirely unaware of a site we agree.

Our concern is that this approach may confuse 'low utility' sites with unfamiliar sites. Consider our Delaware data for the moment. Most Delaware residents visit and report being familiar with some Delaware beaches. New Jersey beaches get occasional visits and mentions, but the attention is mostly on the beaches within the state. At the same time, it is safe to say that most if not all Delaware residents are aware that there are ocean beaches in New Jersey. They may not be able to name most of these beaches or report detailed characteristics at most of these beaches, but they know that they exist. Understandably, they are also likely to report being unfamiliar with most of these beaches.
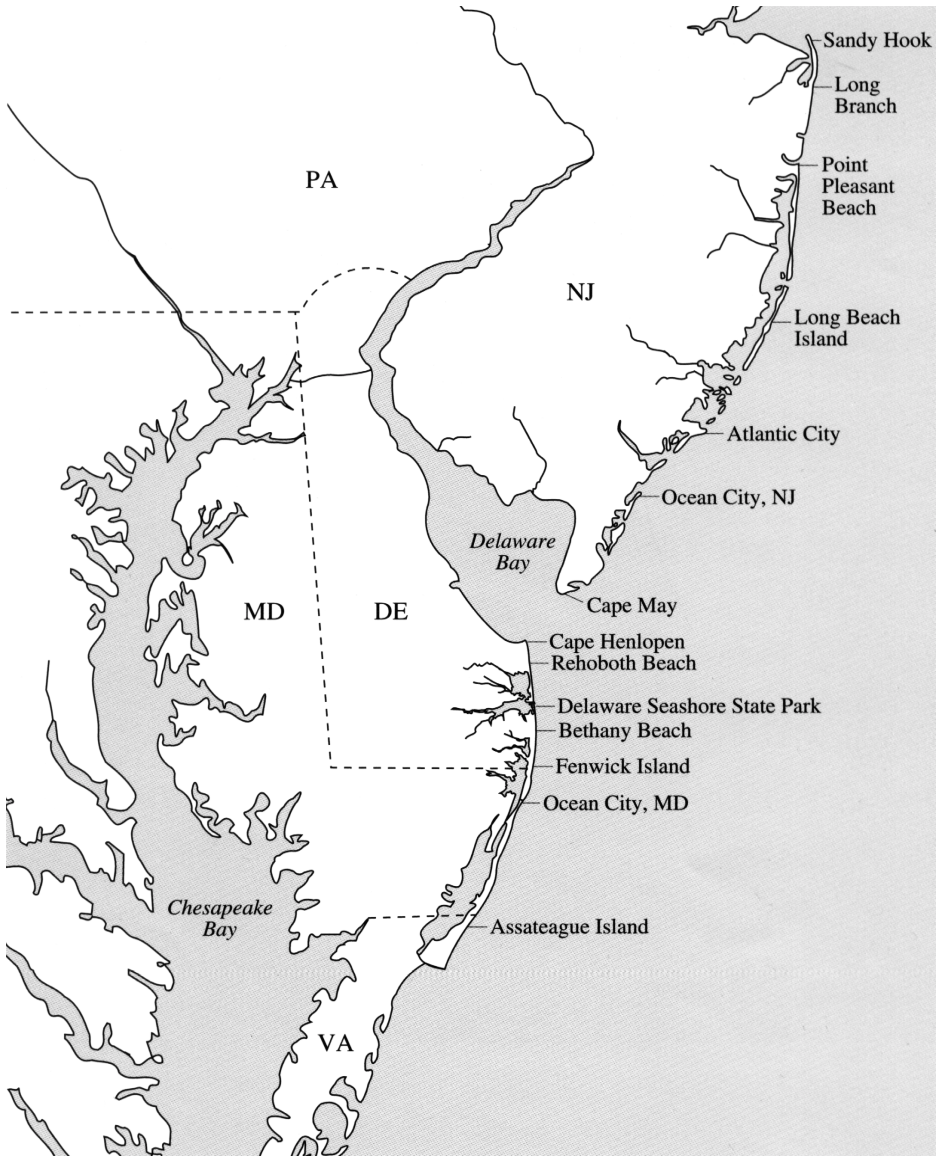

Figure 1. Mid-Atlantic Region

Table 1
Site Characteristics

Price	Trip cost (includes tolls, beach fees, transit cost and parking fees) + time cost
Length	Length of beach in miles
Boardwalk	Boardwalk with shops and attractions present (0,1)
Amusements	Amusement park, rides, or games nearby the beach (0,1)
Private	Private or limited access (0,1)
Park	State park, federal park, or wildlife refuge (0,1)
Wide beach	Width from dune toe to berm greater than 200 feet (0,1)
Narrow beach	Width from dune toe to berm less than 75 feet (0,1)
Atlantic City	Atlantic City (0,1)
Surfing	Recognized as a good location for surfing (0,1)
High rises	Highly developed (0,1)
Park inside	Part of the beach is a park area (0,1)

Note: We thank Tony Pratt and Michael Powell of the Delaware Department of Natural Resources and Environmental Control and Steven Hafner of the Coastal Research Center at Richard Stockton College for their help in creating the data set. Additional variable information was collected from state tourism guides, websites, commercial travel books, local newspapers, and regulatory agency information guides.

Our respondents are revealing that they prefer Delaware beaches to New Jersey beaches as they understand (perceive) them, albeit imperfectly. Most individuals probably have some approximate idea of distance to New Jersey beaches, some sense of the general character of New Jersey (versus Delaware) beaches, and perhaps an approximate idea of the length of or number of beaches at the New Jersey shore. In our judgement, the choice set should incorporate the New Jersey beaches as they are perceived by each individual. These are 'low utility' sites for most Delaware residents.

The pitfall of ignoring 'low utility' sites should be evident. Suppose that we estimate a model incorporating a narrow set of sites reported as being the set familiar to an individual. Suppose that these sites are all nearby the respondent's home. The problem in using this narrow choice set is that we fail to appreciate the importance of distance (time and travel cost) in the site choice. The individual has decisively rejected all distant sites but we do not account for this in the econometrics. Indeed, if the set of familiar sites are all nearby with approximately the same distance, distance would be measured as having little influence on site choice when in reality it is a central characteristic in site selection.

The prudent specification would include sites as they are perceived by individuals. Some site utilities would be specified in great detail, others in much less detail. If an individual is merely aware of some grouping of sites (*e.g.*, the New Jersey beaches) and otherwise aware of little detail about the sites perhaps some aggregated form is in order. Only in the event of complete unawareness of a site is it dropped from the choice set.

Given the impracticality of obtaining site-specific perception data in a mail survey, we resort to an approximation that divides the sites by familiar and unfamiliar beaches. Then, under the hypothesis that site characteristics will play a different role in the site selection at familiar versus unfamiliar beaches, we specify different site utilities for these beaches. This approach retains all the sites in the choice set addressing our concern about deleting 'low utility' sites. At the same time, it allows sites to play a lesser role in site selection through the parameter estimates. If site characteristics at unfamiliar sites play little or no role in site choice, we would expect this to show up in the parameter estimates.

For each observation then beaches are divided into two sets, $j \in C_{familiar}$ and $i \in C_{unfamiliar}$. The random utilities at the sites in the set $C_{familiar}$ are specified as

$$U_j = X_j\beta + \varepsilon_j. \tag{6}$$

The random utilities at the sites in the set $C_{unfamiliar}$ are specified as

$$U_i = X_i\delta + \varepsilon_i. \tag{7}$$

Both site utilities include the full set of beach characteristics X but have different parameters. If the characteristics at the unfamiliar sites play little or no role in site selection or a different role in site selection, it stands to reason that separating the sites should improve estimation.

In estimation we specify the Familiar Site model as a nested logit model. The unfamiliar sites are grouped together in a nest, and the familiar sites are left unnested. The probability of visiting site f, a familiar site, is

$$pr(f) = \frac{\exp(X_f\beta)}{\left\{\sum_{i \in C_{unfamiliar}} \exp[(X_i\delta)/\lambda]\right\}^\lambda + \sum_{j \in C_{familiar}} \exp(X_j\beta)}. \tag{8}$$

The probability of visiting site k, an unfamiliar site, may be decomposed as follows

$$pr(k) = pr(k|unfamiliar) \cdot pr(unfamiliar) \tag{9}$$

where $pr(k|unfamiliar)$ is the probability of selecting site k given that a trip is taken to an unfamiliar site, and $pr(unfamiliar)$ is the probability of visiting an unfamiliar site. Under our nesting assumptions, these probabilities have the form

$$pr(k|unfamiliar) = \frac{\exp[(X_k\delta)/\lambda]}{\sum_{i \in C_{unfamiliar}} \exp[(X_i\delta)/\lambda]} \tag{10}$$

and

$$pr(unfamiliar) = \frac{\left\{\sum_{i \in C_{unfamiliar}} \exp[(X_i\delta)/\lambda]\right\}^\lambda}{\left\{\sum_{i \in C_{unfamiliar}} \exp[(X_i\delta)/\lambda]\right\}^\lambda + \sum_{j \in C_{familiar}} \exp(X_j\beta)} \tag{11}$$

The term λ, the inclusive value, measures the degree of correlation among the unfamiliar sites. In theory we expect $0 < \lambda < 1$. The closer λ is to 0, the greater the correlation within the nest.

Defining Familiar Beaches

Defining familiar and unfamiliar beaches by mail survey was a more difficult task than we had anticipated. 'Familiar' to us meant a basic awareness of the characteristics of a beach. Unfortunately, in our questionnaire, using terms like 'beaches you are familiar with' or 'beaches you could comfortably describe to a friend considering a trip to the beach' or 'beaches you are aware of' evoked different meanings to people in our pretests. For example, the person that was perhaps most informed about beaches in the Mid-Atlantic region in all of our pretests generated a list of familiar sites smaller than average list size. He was reluctant to say he was familiar with a beach unless he could discuss it in great detail, its surfing conditions, the variations of a beach from its northern to its southern most stretches, its nearby restaurants, and so on. We also experimented with perceptions (and knowledge) of specific characteristics of beaches, but this was too time consuming and tedious for a mail survey.

After our pretests, we felt most confident using past trip information as a proxy for familiarity. We asked individuals to complete a table in which they designated beaches they had visited at least once during their adult life prior to 1997. For our analysis then, we assume that individuals are more familiar with sites they have visited in the past than sites they have not visited. With this definition the meaning of familiar was consistent across respondents. It was easy to administer in a mail survey. And, in our pretest, it appeared to be a good proxy for what we had in mind for being familiar with a beach.

As shown in table 2, the median number of familiar beaches for respondents in our sample was 10. About 22% of the population had five or fewer familiar beaches, and about 26% had sixteen or more.

Favorite Sites Model

Modeling Strategy

In studying consideration sets, Horowitz and Louviere (1995) hypothesize that:

> An individual's preferences among the entire set of available alternatives can be described by a utility function that determines both the consideration set and the choice. The utility of each alternative in the individual's consideration set is greater than the utility of every other alternative in the set.

Their work is a response to the modeling of consideration sets in the marketing literature, wherein consideration sets are regularly identified in stated preference surveys and then used in a separate choice-modeling stage distinct form the final choice model. In this approach the first stage model plays no role in determining the parameters of the utility function. Horowitz and Louviere (1995) depart from this standard practice arguing that *the structure of the utility function alone is sufficient to determine site choice and consideration set formation.*

Following this line of reasoning, treating the consideration set stage as distinct from the final choice stage purges vital preference information from making its way into the parameter estimates of the utility function. While this may be of little interest in the marketing context, where prediction is the primary goal, it is critical in the recreation context where welfare analysis is the goal. It is useful to note that one of the general findings from the marketing literature is that consideration sets can be

Table 2
Familiar, Favorite, and Visited Site Frequencies

# of Sites	Sample Reporting This Many Sites as Familiar (%)	Sample Reporting This Many Sites as Favorite (%)	Sample Reporting This Many Sites as Visited (%)
0	2.3	11.3	0.0
1–5	20.0	30.4	76.4
6–10	29.1	30.4	16.4
11–15	22.8	14.6	4.6
16–20	14.9	5.6	2.3
21–25	7.4	1.8	0.6
>25	3.5	5.9	0.3
Mean	11.5	9.4	4.1
Median	10	6	3

predicted with a great deal of accuracy but final choices cannot. This implies that perhaps a great deal of preference information is being purged or not making its way into the parameter estimates for the utility function.

In our Basic Model, the log-likelihood function used in estimation has the following form

$$\sum_{n=1}^{400} \sum_{s=1}^{62} t_{sn} * \ln\{\Pr(visit_{sn})\} \qquad (12)$$

where n denotes a person, s denotes a site, t_{sn} is the number of trips taken by person n to site s during the season, and $\Pr(visit_{sn})$ is the probability of person n visiting site s. This model relies solely on the observation of selected sites in the sample.

In the Favorite Sites Model, the log-likelihood function is amended as follows

$$\sum_{n=1}^{400} \sum_{s=1}^{62} t_{sn} * \ln\{\Pr(visit_{sn})\} + f_{sn} * \ln\{\Pr(favorite_{sn})\} \qquad (13)$$

where $f_{sn} = 1$ if site s is a favorite site, $f_{sn} = 0$ if s is not a favorite site, and $\Pr(favorite_{sn})$ is the probability that site s has greater utility than all the unfavorable sites. The latter probability for site r is just

$$\Pr(favorite_{rn}) = \frac{\exp(X_{rn}\beta)}{\exp(X_{rn}\beta) + \sum_{i \in C_{unfavored}} \exp(X_{in}\beta)}. \qquad (14)$$

Our Favorite Sites Model is really combining stated and revealed preference data, where the favorite site list is a form of stated preference data and the actual trip information is the revealed preference data.

Defining Favorite Beaches

We tried a variety of different questions for obtaining our set of favorite or considered beaches before settling on a final definition. The term 'consideration', in particular, was problematic in our pretests. It evoked vastly different interpretations from people and we eventually dropped it altogether. Some individuals insisted on a broad interpretation of the term, including virtually all beach sites in their consideration set. Others insisted on a narrow interpretation, only including sites that received serious (almost visited) consideration. As much as we tried to sharpen the definition, the question remained open to wide interpretation and had little consistent meaning across respondents.

Having lost confidence in using the terminology of consideration sets and other 'conversational' terms for defining these sets, we resorted to trying some definitions closer to actual behavior. We asked respondents to complete a table in which they specified the beaches they were somewhat likely or most likely to visit over the next three years. We assume that this set of beaches constitutes a person's set of favorite sites, the set of beaches the person is likely to think about when taking a trip. People seemed to be most comfortable answering the somewhat or most likely future trip scenarios. It got them to think in terms of actually taking trips, its interpretation was straightforward and consistent across respondents, and it was a rather easy to get in a mail survey.

As shown in table 2 the median number of favorite sites across our sample is 6. Only 13% of the sample had 16 or more sites as favorites.

Results

Parameter Estimates

The estimation results are presented in table 3. The Basic Model works much as expected. The coefficient estimate on price is negative and highly significant. Given its importance in welfare analysis this is encouraging. Other variables reducing the probability of a visit include: private, high rises, wide beach, and narrow beach. All are statistically significant. There are no surprises here. Private beaches with limited access have fewer day trips. There is some preference for less developed beaches. This is showing favor toward the Delaware beaches over the more developed Jersey shore among our Delaware residents. The negative coefficients on both wide beach and narrow beach show that beaches can be either too narrow or too wide. Narrow beaches (< 75 feet wide) have reduced space for use and are aesthetically less pleasing. Wide beaches (> 200 feet wide) make access to and from the water more difficult.

Variables increasing the probability of a trip include: length, boardwalk, amusements, park, Atlantic City, surfing, and park inside. Again, all are statistically significant. The larger, perhaps better known, beaches have a greater probability of visit, all else constant. Beaches with side attractions such as a boardwalk or amusements also have a greater draw. This was evident in a few attitudinal questions asked early on in our survey, and indeed is supported in the behavioral model. The park and park inside variables are intended to pick up beaches having a more natural character. We included the park inside variable to capture those beaches with some natural or public character that were not federal or state park lands. Surfing beaches have a higher visit probability due to the water itself as well as the special crowd appeal these beaches tend to engender. The concentration of casinos and gambling in Atlantic City give it a special character unlike any of the other beaches in the re-

Table 3
Regression Results

	Basic Model	Familiar Sites Model		Favorite Sites Model
		Familiar	Unfamiliar	
Price	−0.035 (−59)	−0.023 (−36)	−0.035 (−20)	−0.021 (−85)
Length	0.093 (3.2)	−0.052 (−1.6)	0.593 (6.4)	0.107 (5.7)
Boardwalk	0.862 (20)	0.830 (19)	−0.118 (−0.81)	0.648 (19)
Amusements	1.27 (26)	1.05 (20)	0.147 (0.82)	1.22 (34)
Private	−0.386 (−7.6)	0.079 (1.4)	−2.03 (−9.9)	−0.483 (−13)
Park	0.511 (9.4)	0.523 (8.8)	−1.95 (−7.8)	0.482 (12)
Wide beach	−0.754 (−16)	−0.625 (−12)	0.419 (2.2)	−0.605 (−17)
Narrow beach	−0.353 (−5.0)	−0.425 (−5.3)	0.554 (2.0)	−0.175 (−3.4)
Atlantic City	0.548 (5.3)	0.312 (2.9)	0.457 (1.2)	0.323 (4.4)
Surfing	0.907 (21)	0.930 (18)	−0.920 (−6.2)	0.592 (20)
High rises	−0.481 (−8.4)	−0.616 (−9.3)	−0.469 (−2.4)	−0.438 (−11)
Park inside	0.789 (9.8)	0.733 (8.2)	−0.287 (−1.6)	0.291 (5.8)
IV			0.699 (15)	
Day trips	9,330	9,330		9,330
# of people	400	400		400
Log likelihood	−23,390.3	−20,859.4		−36,424.2

Note: t-statistics are reported in parentheses beside the parameter estimates.

gion. For this reason we have an alternative specific constant to set that beach apart from the others. The coefficient is positive and significant reflecting the popularity of this beach and its casinos.

The Familiar Sites Model has separate coefficient estimates for the familiar and unfamiliar sites. Also, recall that we grouped the unfamiliar sites together in a nested logit model. The IV coefficient in table 3 is the parameter estimate for the inclusive value λ in the unfamiliar site nest. It stands to reason that many of the characteristics of sites unfamiliar to individuals would play a small role in determining site selection. That appears to be the case in our model. The parameter estimates on the characteristics of the unfamiliar sites tended to have lower t-statistics, but somewhat surprisingly many still seem to play an important role in site choice.

It is interesting to note that price, length, and private are among the set of significant site coefficients for the unfamiliar beaches. Even if individuals know little about the specifics of unfamiliar beaches, it stands to reason that they are less likely to try small, private, and distant beaches in this set. Individuals are apt to rule out distant beaches without having significant information on them, so it is no surprise that price plays and important role. It is also likely that among the set of unfamiliar beaches, individuals are more likely to have heard about or be aware of the larger beaches. For that reason, it is no surprise that length plays a role. Finally, the negative and significant coefficient on private makes sense. These are simply beaches that few individuals in the sample are familiar with and rarely visit because of their limited access. Two variables, park and surfing, unexpectedly give negative and significant coefficients. It may be that respondents are familiar with the better park beaches in the choice set. If so, that leaves the poorer park beaches in the unfamiliar set and when compared with other beaches do not fair so well or are simply not known to be parks.

Turning to the variables on the familiar sites, most are significant and have the expected sign. This result follows the arguments advanced by Peters, Adamowicz, and Boxall (1995) and Hicks and Strand (2000) that the characteristics of sites familiar to individuals should play the key role in site selection. These, after all, are sites the individuals know about and can make careful tradeoffs among the characteristics. A few results are worth some comment. The variable private has a positive coefficient and length has a negative coefficient. If an individual is familiar with a beach that is private, their likelihood of visiting it increases. It may be that they may own property there or know someone who does. Also, if an individual is familiar with a beach its size has little role in determining site utility. Size, it seems, plays a more important role in ranking unfamiliar sites than ranking familiar sites.

Notice that price is more significant statistically for the familiar than the unfamiliar sites but it is smaller in absolute value. The higher absolute value among the unfamiliar sites signals that travel cost is of greater relative importance among the characteristic set for the unfamiliar sites than among the familiar sites. This supports the notion that individuals may learn little about distant sites and rank their utility as low based primarily on high travel cost. (If this key characteristic alone rules out the site, why learn more?) The implicit prices for all the characteristics are reported in table 4. Each coefficient is divided by the coefficient on price to get its implicit price. The prices for the characteristics of the familiar sites tended to rise relative to the Basic Model. Among the unfamiliar sites the implicit prices of length and private increase rather dramatically relative to the Basic Model. All other implicit prices, but one, decline.

The Favorite Site Model, like the Basic Model, has a single set of coefficients. All are significant and have the same sign as in the Basic Model. Indeed, as expected, the Favorite Sites Model appears largely to reinforce the results of the Basic Model. Many of the coefficients become more significant when the favorite data are incorporated and none change sign. Had the set of favorite sites been quite different in character from the visited sites, the model would have been altered more significantly. The most significant change we detect is a smaller coefficient on the travel cost variable, which in turn increases the implicit prices on all but two of the beach

Table 4
Implicit Prices

	Basic	Familiar Sites		Favorite Sites
		Familiar	Unfamiliar	
Price	1.0	1.0	1.0	1.0
Length	2.66	−2.26	16.94	5.10
Boardwalk	24.63	36.09	−3.37	30.86
Amusements	36.29	45.65	4.20	58.10
Private	−11.03	3.43	−58.00	−23.00
Park	14.60	22.74	−55.71	22.95
Wide beach	−21.54	−27.17	11.97	−28.81
Narrow beach	−10.09	−18.48	15.83	−8.33
Atlantic City	15.66	13.57	13.06	15.38
Surfing	25.91	40.43	−26.29	28.19
High rises	−13.74	−26.78	−13.40	−20.86
Park inside	22.54	31.87	−8.20	13.86

Note: Prices are in U.S. dollars.

characteristics in the Favorite over the Basic Model. Again, the implicit prices are in table 4. It appears as though travel cost is less important once we consider favorite data along with visit data. This, as we will see in a moment, has implication in our welfare analysis.

We also estimated versions of the model suggested by Peters, Adamowicz, and Boxall (2000) and Hicks and Strand (2000). Peters, Adamowicz, and Boxall estimate a model including only familiar and favorite sites in the choice set. Hicks and Strand include only familiar sites. Our versions of these models are shown in table 5. The median number of familiar sites for our respondents is 10. The median number of familiar and favorite sites combined is 11, most of the favorite sites are familiar. Not surprisingly, these two models are quite similar in estimation. With the exception of length and Atlantic City, all variables are significant and have expected signs. The models depart most dramatically from the Basic Model in the coefficient estimate on travel cost, once again declining relative to the other parameters. Travel cost appears to be less important among the set of familiar and favorite sites. Table 6 shows the implicit prices for both models. Most rise relative to the Basic Model. In Peters, Adamowicz, and Boxall results with fishing sites in Canada the coefficient on price also fell relative to the other characteristics, while in Hicks and Strand results for beach use on Chesapeake Bay the coefficient increased.

Welfare Analysis

Table 7 presents the results of two welfare scenarios. The first is for beach closures at the 3 most visited and the 3 least visited sites in the choice set. The most visited sites are Rehoboth (DE), Ocean City (MD), and Cape Henlopen (DE). The least visited sites are Ortley (NJ), Chadwick (NJ), and Normandy (NJ). All results are losses per trip in dollars. For the Familiar Site Model we present the welfare loss using the coefficient on price from both the familiar sites and the unfamiliar sites. For

Table 5
Regression Results

	Peters, Adamowicz, and Boxall Model	Hicks and Strand Model
Price	−0.026 (−44)	−0.023 (−37)
Length	0.019 (0.62)	0.016 (0.51)
Boardwalk	0.643 (15)	0.645 (15)
Amusements	0.891 (18)	0.863 (17)
Private	−0.232 (−4.3)	−0.202 (−3.6)
Park	0.326 (5.8)	0.315 (5.7)
Wide beach	−0.586 (−12)	−0.603 (−12)
Narrow beach	−0.254 (−3.3)	−0.245 (−3.1)
Atlantic City	0.231 (2.2)	0.160 (1.5)
Surfing	0.653 (14)	0.641 (13)
High rises	−0.451 (−7.4)	−0.470 (−7.5)
Park inside	0.587 (7.0)	0.528 (6.2)
Day trips	9,330	9,330
# of people	400	400
Log likelihood	−20,387.2	−20,036.9

Rehoboth, Ocean City, and Cape Henlopen the Familiar and Favorite Models give larger losses per trip for a closure than the Basic Model. The Familiar Site Model using the familiar site price coefficient gives estimates that are 86–109% higher than the Basic Model. Using the unfamiliar site coefficient the estimates are only 21–36% higher. The Favorite Site estimates that are about 31–67% higher. The Peters *et al.* and Hicks and Strand Models also gave larger welfare losses. The Peters, Adamowicz, and Boxall estimates are 55–83% higher and the Hicks and Strand estimates are 78–114% higher than the Basic Model. These generally larger welfare estimates are driven largely by the relative drop in the coefficient on price in all the models that account for familiar and favorite sites.

For the least visited sites the Familiar Site, Peters, Adamowicz, and Boxall, and Hicks and Strand models give lower welfare estimates than the Basic Model. The

Table 6
Implicit Prices

	Peters, Adamowicz, and Boxall Model	Hicks and Strand Model
Price	1.0	1.0
Length	0.73	0.70
Boardwalk	24.73	28.04
Amusements	34.27	37.52
Private	−8.92	−8.78
Park	12.54	13.70
Wide beach	−22.54	−26.22
Narrow beach	−9.77	−10.65
Atlantic city	8.88	6.96
Surfing	25.12	27.87
High rises	−17.35	−20.43
Park inside	22.58	22.96

Table 7
Mean Welfare Estimates Per Trip Per Person

Welfare Scenario	Basic Model	Familiar Sites Model		Favorite Sites Model	Peters, Adamowicz, and Boxall Model	Hicks and Strand Model
		Familiar*	Unfamiliar**			
Beach Closure						
Rehoboth, DE	$7.97	$16.63	$10.85	$10.44	$14.60	$16.85
Ocean City, MD	2.55	4.83	3.15	4.27	4.06	4.68
Cape Henlopen, DE	3.45	6.42	4.19	4.51	5.35	6.13
Ortley, NJ	0.05	0.02	0.01	0.14	0.01	0.00
Chadwick, NJ	0.02	0.00	0.00	0.07	0.00	0.00
Normandy, NJ	0.01	0.00	0.00	0.06	0.00	0.00
Lost Beach Width	7.25	10.94	7.14	5.78	7.01	7.67

Notes: * Uses price coefficient from familiar sites. ** Uses price coefficient from unfamiliar sites.

Favorite Model gives higher estimates. All models give estimates in the pennies per trip per person, with several rounding to zero. The three models that account for familiar sites all assign unfamiliar sites lower site utilities. The Peters, Adamowicz, and Boxall and Hicks and Strand Models assign these sites no utility by dropping them altogether. The Familiar Sites Model sets unfamiliar sites apart from familiar sites in a model that effectively assigns them lower utility in estimation. Since the three sites being evaluated are rarely mentioned as being familiar sites, it is no surprise that the welfare loss for closing these beaches declines in these three models versus the Basic Model. The Favorite Site Model, on the other hand, does not assign lower values to an unfamiliar site, and the lower coefficient on price appears to drive up the welfare loss here in the same way it did for the three most visited sites.

Our second welfare scenario assumes all Delaware beaches decline in width to less than 75 feet. If the state were to discontinue its current beach nourishment policy while at the same time making no plans for a retreat from the beach, this might become a reality. Here the welfare loss in the Familiar Model is about the same as the Basic Model if the unfamiliar price coefficient is used and about 51% higher if the familiar price coefficient is used. The Favorite Model, on the other hand, reports a lower welfare loss for lost width, a decline of 20%. In the Favorite Model the narrow beach coefficient was lower than in the Basic Model, signaling that beach width may be less important in site utility once one considers the broader set of favorite sites. The Peters, Adamowicz, and Boxall and Hicks and Strand numbers are nearly the same as the Basic Model. In these models there is a commensurate drop in the price and narrow beach coefficients, enough so that the value of beach width stays fairly stable across the models.

Our welfare analyses assume that each individual's designation for a site as either familiar or unfamiliar is unchanged by the policy being considered. This may be unreasonable. If beach A, presently a familiar beach, is to be closed, other unfamiliar beaches may very well see a change in designation. It is reasonable that individuals seeking out other beaches to substitute for the loss of A might become familiar with beaches once designated as unfamiliar. A longer run welfare analysis might incorporate these adjustments by allowing for the change in designation in the welfare computation. The difficulty here is know which beaches to adjust. One might consider presenting two extremes as bounds: no adjustment as a short run change in welfare and complete adjustment as a long run change.

Conclusions

Most studies of recreation demand can realistically form universal site choice sets that are quite large. Yet, it stains credibility to include thousands, hundreds, or even dozens of sites in most individuals' choice sets since they are not likely to know or think seriously about more than a few sites. For this reason, researchers have attempted to estimate models using much narrower respondent-based choice sets. We are somewhat critical of these initial approaches because we believe that there is important preference information in understanding which sites people know about (familiar sites) and really consider (favorite sites). This information is certainly missing in the approaches that merely delete sites (Peters, Adamowicz, and Boxall 1995; and Hicks and Strand 2000). It is also lost in studies the predict choice sets (Ben-Akiva and Boccara 1995) because the predictive first stage of these models provide no information about the parameters of site utility.

Our approach differs from the existing approaches by retaining all sites in the choice set for estimation. Familiar sites and unfamiliar sites are specified with different site utility functions recognizing that the role site characteristics play in site

selection is likely to be quite different for a familiar versus an unfamiliar site. Favorite sites are simply incorporated in the likelihood function as being preferred to unfavored sites.

When familiar and favorite sites are accounted for in our analysis we find, among other things, that travel and time cost are a less important determinant of site choice. This tends to raise welfare estimates because the marginal utility of income in the model is lower. If it is indeed true that conventional models tend to overstate the importance of trip costs, this is an important finding because it implies the conventional RUMs that take no account of familiarity and favorites are tending to understate recreational values. This sweeping conclusion, of course, is not warranted at this stage. Replication with other data sets is needed to see if this is indeed a pattern. Hick and Stand's (2000) results, for example, already contradict the lower coefficient on price when familiar sites are taken into account. Also, if specific site characteristics are under study, it may very well be the case that preference differences are revealed for these characteristics that counteract the marginal utility of income change. We saw this effect in the welfare analysis for our nourishment scenarios.

We also find that unfamiliar sites tend to have a lower site utility when we account for their being unfamiliar in the model. That result persists even as the declining coefficient on travel cost works against it. This makes sense and has the overall effect of assigning more importance, economic value, to sites and attributes of the sites that are familiar to respondents.

It is interesting to note that our Familiar Sites Model gives results that are not widely different from Peters, Adamowicz, and Boxall and Hicks and Strand. Relative to the Basic Model, all the welfare estimates move in the same direction and the parameter estimates seem to be affected similarly. This lessens our concern somewhat about the loss of preference information in deleting sites. Again, replication is need before general conclusions can be drawn.

Finally, it worth noting the difficulty we had in getting consistent definitions for familiar and favorite sites. Our 'conversational' attempts in the questionnaire at identifying such sites were rather problematic. By conversational we mean using terms like 'familiar with' or 'could described to a friend' or 'know about' for isolating familiar and unfamiliar sites or terms like 'consider' or 'think about' for isolating the consideration set (what we came to call the favorite sites). The widely different interpretations given to these terms, most notably 'consider', rendered them of little practical analytical use in our judgement. In the end we opted for more behaviorially based information on past visits to help define familiar sites and future intentions to identify favorite sites. In person surveys have a distinct advantage here. More detailed definitions may be provided to respondents, including examples of what is meant by familiar and favorite (or consider), and one may even attempt to gather data on perceptions of sites and site characteristics.

References

Adamowicz, W., J. Swait, P. 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.

Ben-Akiva, M., and B. Boccarra. 1995. Discrete Choice Models With Latent Choice Sets. *International Journal of Research in Marketing* 12:9–24.

Caulkins, P., R.C. Bishop, and N.W. Bouwes. 1986. The Travel Cost Model for Lake Recreation: A Comparison of Two Methods for Incorporating Site Quality and Substitution Effects. *American Journal of Agricultural Economics* 68(2):291–97.

Feather, P.M. 1994. Sampling and Aggregation Issues in Random Utility Model Estimation. *American Journal of Agricultural Economics* 76(4):772–80.

Haab, T.C., and R.L. Hicks. 1997. Accounting for Choice Set Endogeneity in Random Utility Models of Recreation Demand. *Journal of Environmental Economics and Management* 34 (2):127–47.

Hicks, R.L., and I. Strand. 2000. The Extent of Information: Its Relevance for Random Utility Models. *Land Economics* In press.

Horowitz, J.L., and J.L. Louviere. 1995. What is the Role of Consideration Sets in Choice Modeling? *International Journal of Research in Marketing* 12:39–54.

Jones, C.A., and F. Lupi. 1999. The Effect of Modeling Substitute Activities on Recreational Benefit Estimates. *Marine Resource Economics* 14(4):355–72.

Kaoru, Y., V.K. Smith, and J.L. Liu. 1995. Using Random Utility Models to Estimate the Recreational Value of Estuarine Resources. *American Journal of Agricultural Economics* 77(1):141–51.

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

McConnell, K.E., and I.E. Strand. 1994. The Economic Value of Mid and South Atlantic Sportfishing, Volume 2. Report on EPA Cooperative Agreement, University of Maryland.

Montgomery, M., and M. Needelman. 1997. The Welfare Effects of Toxic Contamination in Freshwater Fish. *Land Economics* 73(2):211–23.

Morey, E.R., R.D. Rowe, and M. Watson. 1993. A Repeated Nested-logit Model of Atlantic Salmon Fishing. *American Journal of Agricultural Economics* 75:578–92.

Needelman, M.S., and M.J. Kealy. 1995. Recreational Swimming Benefits of New Hampshire Like Water Quality Policies: An Application of a Repeated Discrete Choice Model. *Agricultural and Resource Economics Review* 24(1):78–87.

Parsons, G.R., and A.B. Hauber. 1998. Spatial Boundaries and Choice Set Definition in a Random Utility Model of Recreation Demand. *Land Economics* 74 (1):32–48.

Parsons, G.R., and M.S. Needelman. 1992. Site Aggregation in a Random Utility Model of Recreation Demand. *Land Economics* 68(4):418–33.

Parsons, G.R., A.J. Plantinga, and K.J. Boyle. 2000. Narrow Choice Sets in Random Utility Models of Recreation Demand. *Land Economics* 76(1):86–99.

Peters, T., W. Adamowicz, and P. Boxall. 1995. The Influence of Choice Set Considerations in Modelling the Benefits of Improved Water Quality. *Water Resources Research* 613:1781–87.

Shaw, W.D., and P.M. Jakus. 1996. Travel Cost Models of the Demand for Rock Climbing. *Agricultural and Resource Economics Review* 25:133–42.

Train, K. 1998. Recreation Demand Models With Taste Differences Over People. *Land Economics* 74 (2):230–39.