Recreation Demand Models with Taste Differences Over People

by

Kenneth E. Train

Department of Economics

University of California, Berkeley

July 1997

Forthcoming, LAND ECONOMICS, Vol. 74, No. 2, May 1998.

ABSTRACT: We estimate random-parameter logit models of anglers' choice of fishing site. The models generalize logit by allowing coefficients to vary randomly over anglers rather than being fixed. The models do not exhibit the restrictive "independence from irrelevant alternatives property" of logit and can represent any substitution pattern. Estimation explicitly accounts for the fact that the variation in coefficients over anglers induces correlation in unobserved utility over trips by the same angler. Willingness-to-pay for improved fish stock and the value to anglers of specific sites are calculated from the models and compared with the estimates obtained from a standard logit model.

I. Introduction

Recreation demand models are used to forecast demand for recreational activities as well as to determine the value that recreators place on the various factors that affect their choices. A prominent example is fishing models,[1] which describe anglers' choices of whether to take a fishing trip during a given period (e.g., week), which species of fish to target, and/or where to go fishing (i.e., the site). The models relate an angler's choices to the characteristics of the available options (such as the time and cost of traveling to each site, the availability of fish at each site, the availability of campgrounds, and so on) and to the characteristics of the angler (such as age, gender, and income.) The models provide estimates of anglers' willingness to pay for changes in site attributes, such as increased fish stocks or reduced contaminants at specified sites, as well as the value of individual sites themselves.[2]

Recreational demand models have usually been specified as logit or nested logit models.[3] These specifications have several advantages, including simplicity of estimation. However, they impose several well-known restrictions (McFadden 1973, 1978; Train 1986.) First, the coefficients of variables that enter the model are assumed to be the same for all people. This assumption implies that different people with the same observed characteristics have the same value (i.e., "tastes") for each factor entering the model. Second, logit and nested logit exhibit the "independence from irrelevant alternatives" (iia) property. (Logit exhibits this property for all alternatives, and nested logit exhibits it within each nest.) Because of this property, the models necessarily predict that a change in the attributes of one alternative changes the probabilities of the other alternatives proportionately. This substitution pattern can be unrealistic in many settings. Third, in situations with repeated choices over time, logit and nested logit assume that unobserved factors are independent over time for each decision-maker. In reality, however, one would expect unobserved factors that affect a decision-maker to persist, at least somewhat, over time.

In the current paper we estimate recreation demand models with a specification that is a generalization of logit and avoids these limitations. In particular, we estimate a random-parameters logit (RPL) model[4] of fishing site choice. RPL generalizes logit by allowing that the coefficients of observed variables to vary randomly over people rather than being fixed. With this generalization, the model does not exhibit the iia property with its restrictive substitution patterns. In fact, any pattern of substitution can be represented arbitrarily closely by an RPL (McFadden and Train 1997). The variation in coefficients over people implies that the unobserved utility associated with any alternative is necessarily correlated over time for each decision-maker. This correlation is incorporated into the estimation when there are observations on more than one choice situation for each person. The specification and estimation of RPL are described in the following section. The application to fishing site choice is described in section III.

II. Random-Parameters Logit

RPL models have taken different forms in different applications; their commonality arises in the integration of the logit formula over the distribution of unobserved random parameters. The early applications (Boyd and Mellman 1980, and Cardell and Dunbar 1980) were restricted to situations in which explanatory variables do not vary over decisionmakers, such that the integration, which is computationally intensive, is required for only one "decisionmaker" using aggregate share data rather than for each decisionmaker in a sample. Advances in computer speed, as well as greater understanding of simulation methods to approximate integration, have allowed estimation of models with explanatory variables varying over decisionmakers. Examples include Erdem (1995), Ben-Akiva and Bolduc (1996), Bhat (1996a,b), Brownstone and Train (1996), Mehndiratta (1996), and Revelt and Train (1996). The form of the RPL that we utilize in our investigation is described as follows.

An angler chooses among J possible sites each time he/she takes a fishing trip. The utility that angler n would obtain from site j in trip t is $U_{njt} = \beta_n' x_{njt} + \epsilon_{njt}$ where x_{njt} is a vector of observed variables, β_n is a vector of coefficients that is unobserved for each n and varies randomly over anglers representing each angler's tastes, and ϵ_{njt} is an unobserved random term that is distributed iid extreme value, independent of β_n and x_{njt}.

This specification is the same as for logit, except that now the coefficients β_n vary in the population rather than being fixed. The variance in β_n induces correlation in utility over sites and trips. In particular, the coefficient vector for each angler, β_n, can be expressed as the sum of the population mean, b, and individual deviation, η_n, which represents the angler's tastes relative to the average tastes in the population of all anglers. Utility is $U_{njt} = b'x_{njt} + \eta_n'x_{njt}+\epsilon_{njt}$. The researcher estimates b (see below) but does not observe η_n for each angler. The unobserved portion of utility is therefore $\eta_n'x_{njt}+\epsilon_{njt}$. This term is correlated over sites and trips due to the common influence of η_n. That is, the same tastes are used by the angler to evaluate each site; since the researcher does not observe these tastes completely (i.e., does not observe η_n), the portion of utility that the researcher does not observe is correlated over sites. Similarly, the same tastes are used by the angler for each trip; since the researcher does not observe these tastes completely, the portion of utility that is not observed by the researcher is correlated over the trips made by a given angler.

Because the unobserved portion of utility is correlated over sites, RPL does not exhibit the independence from irrelevant alternatives property of standard logit. Very general patterns of correlation over sites, and hence very general substitution patterns, can be obtained through appropriate specification of variables and parameters. As stated above, it has been shown (McFadden and Train 1997) that any random utility model, representing any subsitution patterns, can be approximated arbitrarily closely by an RPL.[5]

Because the unobserved portion of utility is correlated over trips for a given angler, RPL differs from standard logit which assumes that unobserved utility is uncorrelated over trips. However, the correlation over trips is of a specific type, which might not be appropriate in all situations. In particular, we do not subscript β_n by t; that is, tastes vary over anglers but not over time for each angler. This specification is consistent with the notion that an angler has particular tastes that stay with the angler. In general, however, an angler's tastes may change over time, and in particular may change in response to previous trip experiences (e.g., the angler's value of trip cost may rise after the angler has taken a number of expensive trips.) Our specification can be generalized to allow for these possibilities; this is an important direction for future work.[6]

We now derive the choice probabilities. If we knew the angler's individual tastes, that is, if we knew the value of β_n, then we could easily calculate the angler's probability of choosing a given site. Since ϵ_{nit} is iid extreme value, as in a standard logit model, the probabilties are logit given the value of β_n. In particular, if β_n were known to take the value β, the probability that angler n chooses site i for trip t would be standard logit:

$$L_{nit} = \frac{e^{\beta' x_{nit}}}{\sum_j e^{\beta' x_{njt}}} \qquad [1]$$

We do not, however, know the angler's individual tastes. Tastes vary in the population with density denoted $f(\beta|\theta^*)$, where θ^* are the parameters of this distribution (representing, for example, the mean and standard deviation of tastes in the population of anglers.) Since the researcher does not observe the angler's actual tastes, the probability that the researcher ascribes for the angler is the integral of equation [1] over all possible values of β weighted by the density of β. That is, the actual probability for the angler's choice of site is:

$$Q_{nit}(\theta*) = \int L_{nit}(\beta)\ f(\beta|\theta*)\ d\beta$$

For maximum likelihood estimation we need the probability of each sampled angler's sequence of choices (unless we only observe one trip for each angler.[7]) We obtain this probability in a manner similar to that above for one trip. If we knew β_n, then the probability of the angler's choices for several trips would be the product of logit formulas. In particular, let i(n,t) denote the site that angler n chose in trip t. If $\beta_n=\beta$, then, since ϵ_{nit} is iid over trips[8], the probability of angler n's observed sequence of choices is:

$$S_n(\beta) = \prod_t L_{ni(n,t)t}(\beta) \qquad [2]$$

Since we do not know β_n, the actual probability is the integral of [2] over all values of β:

$$P_n(\theta*) = \int S_n(\beta)\ f(\beta|\theta*)\ d\beta \qquad [3]$$

Note that there are two concepts of parameters in this description. The coefficient vector β_n is the parameters associated with angler n, representing that angler's tastes. These tastes vary over anglers; the density of this distribution has parameters θ^* representing, for example, the mean and covariance of β in the population of all anglers. The goal is to estimate θ^*, that is, the population parameters that describe the distribution of individual parameters.

The log-likelihood function is $LL(\theta)=\sum_n \ln P_n(\theta)$. Exact maximum likelihood estimation is not possible since the integral in [3] cannot be calculated analytically. Instead, we approximate the probability through simulation and maximize the simulated log-likelihood function (see e.g., Hajivassiliou 1993, and Hajivassiliou and Ruud 1994, for a discussion of simulation methods in estimation.) In particular, $P_n(\theta)$ is approximated by a summation over randomly chosen values of β. For a given value of the parameters θ, a value of β is drawn from its distribution. Using this draw of β, $S_n(\beta)$ -- the product of standard logits -- is calculated. This process is repeated for many draws, and the average of the resulting $S_n(\beta)$'s is taken as the approximate choice probability:

$$SP_n(\theta) = (1/R) \sum_{r=1,\dots,R} S_n(\beta^{r|\theta})$$

where R is the number of repetitions (i.e., draws of β), $\beta^{r|\theta}$ is the r-th draw from $f(\beta|\theta)$, and $SP_n(\theta)$ is the simulated probability of angler n's sequence of choices. By construction $SP_n(\theta)$ is an unbiased estimator of $P_n(\theta)$ whose variance decreases as R increases. It is smooth (i.e., twice-differentiable) which helps in the numerical search for the maximum of the simulated log-likelihood function. It is strictly positive for any realization of the finite R draws, such that the log of the simulated probability is always defined.[9]

The simulated log-likelihood function is constructed as $SLL(\theta) = \sum_n \ln(SP_n(\theta))$, and the estimated parameters are those that maximize SLL. Lee (1992) and Hajivassiliou and Ruud (1994) derive the asymptotic distribution of the maximum simulated likelihood estimator based on smooth probability simulators with the number of repetitions increasing with sample size. Under regularity conditions, the estimator is consistent and asymptotically normal. Furthermore, when the number of repetitions rises faster than the square root of the number of observations, the estimator is asymptotically equivalent to the maximum likelihood estimator. We use a thousand repetitions in our application below.[10]

We estimate models with normal and log-normal distributions for elements of β; other distributions are of course possible. The distributional assumptions for the model are described in the next section after the data are discussed.

III. Application

We estimate a model of anglers' choice among river fishing sites. This model is a component of the more complete angler-behavior model developed by Desvousges, Waters, and Train (1996) which describes anglers' choices of whether to take a fishing trip in each given week, and, for each trip, whether the trip is to a lake or river, the lake site for lake trips, the river site for river trips, and the duration of the trip. We concentrate on the river site component to illustrate the concepts of RPL; this component contains the parameters that are most central to estimation of anglers' willingness to pay for improved fish stocks in rivers.

A random sample of Montana anglers was obtained through random telephone solicitation and screening. For each sampled angler, records were obtained for each fishing trip taken from July 1992 through August 1993, including the location at which fishing occurred. Data were obtained on a total of 962 river fishing trips taken by 258 anglers. A total of 59 possible river fishing sites were defined based on geographical characteristics and other relevant factors. Each site contains one or more of the stream segments used in the Montana River Information System (MRIS). Variables that relate to the sites were obtained from MRIS and other sources. In particular, the following variables enter the model for each site:

1. Fish stock, measured in fish per 1000 feet of river (from the 1994 MRIS data, calculated as the weighted average over the MRIS stream segments within each site, with weights proportional to segment length.) In estimation, we rescale this variable to represent 100 fish per 1000 feet of stream.

2. Aesthetics rating, measured on a scale of 0 to 3, with 3 being highest (from MRIS with category 4 combined with category 3 for a rating of 3, averaged over stream segments.[11])

3. Log of size of each site, in US Geological Survey (USGS) blocks (from USGS maps.) This variable captures the fact that the angler has the option of many locations within the site, and the number of locations increases with the size of the site. The variable enters in log form as indicated by a nested logit structure with location choice below site choice (McFadden 1978; Train 1986).[12]

4. The number of campgrounds per USGS block in the site.

5. The number of State Recreation Access areas per USGS block in the site. These areas are indicated by a sign on the road designating access to fishing sites.

6. Indicator that the Angler's Guide to Montana lists the site as a major fishing site.

7. Number of restricted species at the site, such that catching and keeping these species is illegal during certain times of the year (from Montana fishing regulations, 1992-3.)

8. Travel cost. For each angler, the cost of travelling to each site was calculated, including both the variable cost of driving (i.e., gas, maintenance, tires, and oil) and the value of the time spent driving (with the value of time taken to be one-third of the angler's wage.)

Extensive information on the sample, site definitions, and variable construction are provided by Desvousges et al. (1996) and Desvousges and Waters (1995).

Table 1 presents a standard logit model estimated on these data. The coefficients of fish stock, trip cost, aesthetics, log(size), and restricted species enter with the expected signs. Number of campgrounds and number of access areas could logically take either sign: some anglers might prefer to have campgrounds and access areas that they could use, while other anglers might prefer the privacy that comes from there not being campgrounds and access areas nearby. Similarly for the major-site indicator: being listed as a major site in the Angler's Guide could deter some anglers who prefer privacy, while other anglers find the Guide's assessment as an inducement to go there[13].

It is doubtful that all anglers place the same value on each of these site attributes. It is also doubtful that the iia property of logit models holds across the 59 sites. We specify an RPL model to account for these issues. In particular, we estimate a model in which the coefficients of fish stock, trip cost, and aesthetics have coefficients that are distributed log-normal in the population; the coefficients of campgrounds, access areas, major-site, and restricted species are normally distributed; and the coefficient of log(size) is fixed. The log-normal distribution assures that each angler in the population has a positive coefficient for the variable, whereas with the normal distribution, some anglers in the population necessarily have negative signs while others have positive signs (the share of the population with each sign is determined by the mean and standard deviation of the distribution, which are estimated.) The coefficients of fish stock and aesthetics are expected to be positive for all anglers, such that they are given log-normal distributions. The trip cost coefficient is expected to be negative for each angler; in estimation, the negative of trip cost is entered such that its log-normally distributed coefficient is negative for all anglers.[14] Campgrounds, access areas, and being listed as a major site in the Angler's Guide could be negative or positive factors for different anglers; giving their coefficients a normal distribution allows the estimates to indicate the share with each sign. The presence of restricted species would be expected to have a negative impact during the periods of restriction; however, during unrestricted fishing periods, the impact could be positive. This variable is therefore given a normal distribution. The coefficient of log(size) is a measure of the correlation in unobserved utility across locations within each site (see footnote 12 above); under this interpretation, it is constant, rather than varying over anglers.

Let m denote an element of β that has a normal distribution. This coefficient is calculated as $\beta_m = b_m + s_m\mu_m$, where μ_m is an independent standard normal deviate. The parameters b_m and s_m, which represent the mean and standard deviation of β_m, are estimated. Each element of β that has a log-normal distribution is expressed $\beta_k = \exp(b_k + s_k\mu_k)$, where the parameters b_k and s_k, which represent the mean and standard deviation of $\ln(\beta_k)$, are estimated. The median, mean, and standard deviation of β_k are: $\exp(b_k)$, $\exp(b_k+(s_k^2/2))$, and $\exp(b_k+(s_k^2/2))*\sqrt{[\exp(s_k^2)-1]}$, respectively. Note that in this specification, the coefficients are independent; correlated coefficients are considered later.

Table 2 gives the estimated parameters of this RPL model. The estimated standard deviations of coefficients are highly significant, indicating that parameters do indeed vary in the population.[15] Also, the likelihood ratio index rises substantially compared to the logit model, indicating that the explanatory power of the RPL model is considerably greater than with standard logit. All of the parameters are significantly different from zero at 99% confidence, except for the mean of the campground coefficient. The standard deviation of the campground coefficient is, however, significant and fairly large. Taken together, the results for the campground coefficient imply that campgrounds do indeed affect anglers' choices, with some anglers preferring sites with campgrounds and other anglers preferring sites without campgrounds; the mean is not significantly different from zero because the different tastes regarding campgrounds tend to balance out in the population. The point estimates for the coefficient of access areas imply that about thirty percent of the population prefer having numerous access areas while the other seventy percent prefer having the privacy associated with fewer access areas. The Angler's Guide listing is estimated to constitute a positive inducement for about two-thirds of the population and a negative inducement for the other third.

The point estimates for the parameters for the log-normal distributions imply the following:

	Median	Mean	Std. dev.
Fish stock	0.0563	0.0944	0.1270
Aesthetics	0.4519	0.6482	0.6665
Trip cost	0.0906	0.1249	0.1185

The model in Table 2 specifies the coefficients to be independently distributed while they could in reality be correlated. For example, customers who are especially concerned about fish stock might also be highly concerned about aesthetics. For the estimation of willingness to pay for improved fish stock, we are particularly concerned with the possibility that the coefficients of fish stock and trip cost are correlated. To investigate these possibilities, we specify the log-normally distributed coefficients to be correlated. For computational convenience (i.e., to prevent the introduction of numerous covariance parameters), we maintain independence for the normally distributed coefficients. In particular, letting β represent the vector of coefficients for fish stock, aesthetics and trip cost, we specify $\log(\beta) \sim N(b,\Omega)$ for general Ω. The coefficient vector is expressed $\beta=\exp(b+L\mu)$ where L is a lower-triangular Choleski factor of Ω, such that $LL'=\Omega$, and μ is a vector of independent standard normal deviates. The top part of Table 3 gives the estimates of b and L, and the bottom part gives statistics for the coefficients implied by the point estimates of b and L. The elements of L are all significant at the 95% confidence level, indicating that there is indeed correlation among the three log-normally distributed coefficients. The correlations are positive, indicating that anglers who place a higher-than-average value on fish stock also tend to place higher-than-average values on aesthetics and trip cost. Essentially, the positive correlations among the three factors imply that these factors tend to be valued as a group relative to the other factors.

We now calculate willingness-to-pay estimates from the RPL models and compare them with those from standard logit. Given β, the change in consumer surplus (or, more precisely, the compensating variation) that is associated with a change in site attributes is calculated the same as for standard logit (see, e.g., Parsons and Kealy 1995):

$$C_{nt}(\beta) = \{ \ln\sum_j \exp(\beta' x^*_{njt}) - \ln\sum_j \exp(\beta' x^{**}_{njt}) \} \ / \ \beta^c \qquad [4]$$

where x^*_{njt} represents the original attributes, x^{**}_{njt} the new attributes, and β^c is the cost coefficient. The compensating variation for person n and trip t is therefore $CV_{nt} = \int C_{nt}(\beta)\, f(\beta|\theta^*)\, d\beta$, and the average compensating variation is the average of CV_{nt} over all trips by all sampled anglers.

We first calculate the compensating variation associated with increasing the fish stock at each site by 100 fish per 1000 feet of river. The results are given in the first row of Table 4. The logit model gives a compensating variation of $1.40 per trip. This, of course, is simply the ratio of the fish stock coefficient to the trip cost coefficient in Table 1 (since the fish stock coefficient represents the change in utility associated with a 1 unit change in the fish stock variable, which is scaled in units of 100 fish per 1000 feet of river.) The RPL with uncorrelated coefficients (the model of Table 2) gives an estimated compensating variation of $1.44, which is practically the same as from the standard logit. The model that allows correlations among coefficients (Table 3) provides a lower estimate of 0.93.

We next estimate the compensating variation associated with a doubling of the fish stock at each site, given in the second row of Table 4. The estimated compensated variation is higher for a doubling of the fish stock at all sites than for adding 100 fish per 1000 feet of river at each site. There are two reasons. First, the average fish stock in the base situation is 180. The average fish stock therefore rises more when the stock is doubled than when it is increased by 100. Second, doubling stocks increases the range of fish stocks while adding a fixed amount to each site does not. Greater variety allows for greater matching of sites with anglers' tastes. In comparing across models, the RPL without correlation gives a higher estimate than the standard logit, and the RPL with correlation gives a lower estimate. [16]

As a third scenario, we calculate the change in consumer surplus that arises from eliminating the Madison River sites from anglers' choice sets. This value is calculated by not including the sites in the second summation in equation [4] while holding $x^{**}_{njt} = x^{*}_{njt}$ for all other sites. The third row of Table 4 gives the average over all trips of the compensating variation associated with the elimination of the Madison River sites. Twenty-four of the 962 sampled trips were to the Madison River. Since only these trips are affected by the elimination of the Madison River sites, the compensating variation can also be expressed in terms of loss per trip to the Madison. These figures are given in the fourth row, calculated simply as the estimate in the third row (which is the average over all trips) times 962 sampled trips, divided by 24 trips to the Madison. The figures, which range from \$22 for the RPL with correlations to \$30 for the RPL without correlations, represent estimates of the amount that anglers must be compensated for each trip that they would take to the Madison for the lower utility that they obtain from taking the trip to another site instead. The RPL without correlations gives a higher estimate than the logit model, and the RPL with correlations gives an estimate that is nearly the same as the logit model.

In all three scenarios, the compensating variation from the logit model is between those from the two RPLs. Bhat (1996a) found that the estimated willingness to pay for travel attributes were somewhat but not greatly different in an RPL than a logit. These results might suggest that the logit model is fairly robust with respect to estimating compensating variations. However, these results are undoubtedly situation-specific. Bhat (1996b), in a different situation, found large differences between a logit and an RPL in estimated willingness to pay for travel mode attributes. Revelt and Train (1996) found that estimated willingness to pay for appliance attributes differed between a logit and RPL for some attributes and were similar for others. There is probably no general answer to whether logit obtains reliable estimates of compensating variations; to answer the question for any specific situation, estimation of an RPL is needed for comparison.

References

Ben-Akiva, M., and D. Bolduc. 1996. "Multinomial Probit with a Logit Kernel and a General Parametric Specification of the Covariance Structure." Department d'economique, Working Paper, Universite laval, Quebec, Canada.

Bhat, C. 1996a. "Accommodating Variations in Responsiveness to Level-of-Service Measures in Travel Mode Choice Modeling." Department of Civil Engineering, Working Paper, University of Massachusetts at Amherst.

Bhat, C. 1996b. "Incorporating Observed and Unobserved Heterogeneity in Urban Work Travel Choice Modeling." Department of Civil Engineering, Working Paper, University of Massachusetts at Amherst.

Bockstael, N., W. Hanemann, and C. Kling. 1987. "Estimating the Value of Water Quality Improvements in a Recreational Demand Framework." Water Resources Research 23: 951-960.

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

Bockstael, N., K. McConnell, and L. Strand. 1991. "Recreation." In Measuring the Demand for Environmental Quality Vol. 8, eds. J. Braden and C. Kolstad. Amsterdam: North-Holland Press.

Boyd, J. and R. Mellman. 1980. "The Effect of Fuel Economy Standards on the U.S. Automotive Market: An Hedonic Demand Analysis." Transportation Research 14A (5-6): 367-378.

Brownstone, D., and K. Train. 1996. "Forecasting New Product Penetration with Flexible Substitution Patterns." Paper presented at American Economic Association meetings, New Orleans.

Cardell, N. and F. Dunbar. 1980. "Measuring the Societal Impacts of Automobile Downsizing." Transportation Research 14A (5-6): 423-434.

Caulkins, P., R. Bishop, and N. 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: 291-297.

Desvousges, W., and S. Waters. 1995. "Report on Potential Economic Losses Associated with Recreational Services in the Upper Clark Fork River Basin." Vol. III and Appendices, Triangle Economic Research, 1000 Park Forty Plaza, Suite 200, Durham, NC 27713.

Desvousges, W., S. Waters, and K. Train. 1996. "Supplemental Report on Potential Economic Losses Associated with Recreational Services in the Upper Clark Fork River Basin." Triangle Economic Research, 1000 Park Forty Plaza, Suite 200, Durham, NC 27713.

Erdem, T. 1995. "A Dynamic Analysis of Market Structure Based on Panel Data." Haas School of Business, Working Paper, University of California, Berkeley.

Hajivassiliou, V. 1993. "Simulation Estimation Methods for Limited Dependent Variable Models." In Handbook of Statistics Vol. 11, Ch. 19, pp. 519-543, eds. G. Maddala, C. Rao, and H. Vinod. Amsterdam: Elsevier Science Publishers B.V.

Hajivassiliou, V. and P. Ruud. 1994. "Classical Estimation Methods for LDV Models using Simulation." In Handbook of Econometrics, Vol. IV, eds. R. Engle and D. McFadden. Amsterdam: Elsevier Science Publishers B.V.

Hausmann, J., G. Leonard, and D. McFadden. 1995. "A Utility-Consistent, Combined Discrete Choice and Count Data Model: Assessing Recreational Use Losses Due to Natural Resource Damage." Journal of Public Economics 56: 1-30.

Hausman, J., and D. Wise. 1978. "A Conditional Probit Model for Qualitative Choice: Discrete Decisions Recognizing Interdependence and Heterogenous Preferences." Econometrica 46 (2): 403-426.

Lee, L. 1992. "On Efficiency of Methods of Simulated Moments and Maximum Simulated Maximum Estimation of Discrete Response Models." Econometrica 8: 518-552.

McFadden, D. 1973. "Conditional Logit Analysis of Qualitative Choice Models." In Frontiers of Econometrics, ed. P. Zarembka. New York: Academic Press.

McFadden, D. 1978. "Modelling the Choice of Residential Location." In Spatial Interaction Theory and Planning Models, eds. A. Karquist et al. Amsterdam: North-Holland Press.

McFadden, D., and K. Train. 1997. "Mixed MNL Models for Discrete Response." Department of Economics, Working Paper, University of California, Berkeley.

Mehndiratta, S. 1996. "Time-of-Day Effects in Inter-City Business Travel." Ph.D. thesis, Department of Civil Engineering, University of California, Berkeley.

Morey, E., R. Rowe, D. Waldman. 1995. "Assessment of Damages to Anglers and Other Recreators from Injuries to the Upper Clark Fork River Basin." Revised report and rebuttal, state of Montana natural resource damage program, Hagler Bailly Consulting, Inc., P.O. Drawer O, Boulder CO 80306-1906.

Morey, E., R. Rowe, and M. Watson. 1993. "A Repeated Nested-Logit Model of Atlantic Salmon Fishing." American Journal of Agricultural Economics 75: 578-592.

Morey, E., W. Shaw, and R. Rowe. 1991. "A Discrete-Choice Model of Recreational Participation, Site Choice, and Activity Valuation When Complete Trip Data are Not Available." Journal of

Environmental Economics and Management 20: 181-201.

Parsons, G., and M. Kealy. 1992. "Randomly Drawn Opportunity Sets in a Random Utility Model of Lake Recreation." Land Economics 68: 93-106.

Parsons, G., and Needleman. 1992. "Site Aggregation in a Random Utility Model of Recreation." Land Economics 68: 418-433.

Revelt, D., and K. Train. 1996. "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level." Forthcoming, Review of Economics and Statistics.

Train, K. 1986, Qualitative Choice Analysis. Cambridge: MIT Press.

Train, K. 1997. "Mixed Logit Models for Recreation Demand." In Valuing the Environment Using Recreation Demand Models, eds. C. Kling and J. Herriges. Forthcoming, Elgar Press.

Wegge, T., R. Carson, and W. Hanemann. 1991. "Site Quality and the Demand for Sport Fishing in Southcentral Alaska." Report, Jones & Stokes Associates, Inc., 1725 23rd Street, Suite 100, Sacramento, CA 95816.

Footnotes

1. E.g., Bockstael et al. 1989; Wegge et al. 1991; Parsons and Needelman 1992; Parsons and Kealy 1992; Morey et al. 1991, 1993, 1995; Hausman et al. 1995; Desvousges et al. 1996).

2. See Bockstael, McConnel, and Strand 1991, for an overview of the use of recreation demand models for assessment of attribute values. The cites in note 1 provide specific examples.

3. E.g., all the fishing models cited above; the general lake recreation model of Caulkins et al 1986; the swimming choice model of Bockstael et al 1987; the boating, swimming and lake viewing models of Parsons and Kealy 1992; and the boating, hunting, and hiking/viewing models of Hausman et al 1995.

4. These models are also called "mixed logit" (Bhat 1996a; McFadden and Train 1997; Train 1997), "random-coefficients logit" (Bhat 1996b), and "error-components logit" (Brownstone and Train 1996).

5. An analog to nested logit is obtained by defining a dummy variable for each nest, with the dummy taking the value of one for all alternatives in the nest and zero for alternatives outside the nest. Allowing the coefficient of each nest-specific variable to vary randomly induces correlation in unobserved utility among alternatives within each nest, while not inducing correlation across nests. As such, RPL can represent a nested logit. RPL can also allow other coefficients to vary, such that the correlation pattern over alternatives is potentially richer than that provided by nested logit. In some situations, however, nested logit is easier to estimate than RPL and can be preferred on those grounds, especially if the richer correlation patterns that RPL allows are not needed to adequately represent the choice process.

6. As suggested by a reviewer, the change in tastes over time for each angler could be specified as a Markov chain with a complete connection between the angler's current tastes and the other sites that the angler has previously visited and in what order (or, for a more tractable specification, to the number of previous trips to other sites without referencing the sequence.)

7. It is important to note that estimation of RPL does not require observations on more than one trip for each angler. In fact estimation is faster with only one trip per angler. However, for most recreational demand data (including the data we use), repeated choices are observed for each sampled recreator.

8. Even though ϵ_{nit} is independent over trips, the unobserved portion of utility is correlated over trips, as discussed above. The correlation over trips is induced by β_n, such that if β_n were known, the remaining error term would be uncorrelated over trips.

9. The simulated probabilities for a sequence of choices sum to one over all possible sequences. Similarly, simulated choice probabilities for each trip (that is, simulated versions of $Q_{nit}(\theta)$) sum to one over alternatives, which is useful in forecasting.

10. Software to estimate RPLs, written in GAUSS with users manual and sample runs, can be downloaded (free of charge) from the author's home page at http://elsa.berkeley.edu/~train.

11. A rating of 4 is given to a "stream of national renown" while the rating of 3 is given to a "stream of outstanding natural beauty in a pristine setting." We combined the two ratings to avoid the possibility that the difference between a 3 and a 4 was based on the quality of fishing at the stream, rather than aesthetic quality per se.

12. To be precise: if the choice of site and location within site is a nested logit, and there are no observed differences in utility across locations within each site (though there are unobserved differences in utility that vary randomly), then the expected utility associated with the locations within each site is the log of the number of locations within the site.

13. Since the listing as a major site usually indicates high quality fishing, there is the possibility that its estimated coefficient captures some of the effect of fish stock. When the model is estimated without the major site dummy, the coefficient of fish stock rises by 15%.

14. When the model is estimated with normal distributions for the fish stock, aesthetics, and cost coefficients, the estimated mean for each coefficient is between the median and mean that are estimated with the log-normal distribution.

15. Part of this variation in tastes could perhaps be captured by characteristics of the anglers, which are not included in the model. In a RPL model of appliance choice, Revelt and Train (1996) found considerable variation remaining after including demographic variables, indicating that tastes vary considerably more than can be explained by observed characteristics of people.

16. To interpret the estimates for these scenarios it is important to distinguish increases in fish stock from increases in the number of fish that angler catch. Morey et al. (1995), for example, find sharply decreasing returns of catch rate to fish stock.

TABLE 1

Standard Logit Model

	Coefficient	Standard error
Fish stock	0.1061	0.0264
Aesthetics	0.5654	0.0628
Trip cost	-0.0756	0.0022
Guide lists as major	0.3718	0.1339
Campgrounds	-0.1380	0.2230
Access areas	0.4592	0.1661
Restricted species	-0.3084	0.0542
Log(Size)	0.5847	0.0764
Likelihood ratio index	.4324	
Log-Likelihood at convergence	-2201.2965	

TABLE 2

RPL Model

		Parameter	Standard error
Fish stock	Mean of ln(coefficient)	-2.876	0.6066
	Std. dev. of ln(coefficient)	1.016	0.2469
Aesthetics	Mean of ln(coefficient)	-0.7942	0.2287
	Std. dev. of ln(coefficient)	0.8493	0.1382
Trip cost (neg.)	Mean of ln(coefficient)	-2.402	0.0631
	Std. dev. of ln(coefficient)	0.8012	0.0781
Guide lists major	Mean of coefficient	1.018	0.2887
	Std. dev. of coefficient	2.195	0.3518
Campgrounds	Mean coefficient	0.1158	0.3233
	Std. dev. of coefficient	1.655	0.4350
Access areas	Mean coefficient	-0.9499	0.3610
	Std. dev. of coefficient	1.888	0.3511
Restricted species	Mean coefficient	-0.4989	0.1310
	Std. dev. of coefficient	0.8989	0.1640
Log(size)	Coefficient	0.9835	0.1077
Likelihood ratio index		0.5018	
Log-likelihood at convergence		-1932.3348	

TABLE 3

Model with Correlations Among Log-normal Coefficients

I. Estimated means and Choleski factor of covariance matrix of Ln(coefficients)
Standard errors in parentheses.

	Fish stock	Aesthetics	Trip Cost
Mean	-3.1641	-0.7128	-2.3743
	(0.6401)	(0.2171)	(0.0682)

Choleski Factor

1.5157		
(0.2539)		
0.3715	0.8489	
(0.1785)	(0.1799)	
0.5315	0.2074	0.6501
(0.0788)	(0.1051)	(0.0627)

II. Statistics for coefficients

	Fish Stock	Aesthetics	Trip Cost
Median	0.043	0.493	0.093
Mean	0.131	0.773	0.137
Standard deviation	0.321	0.924	0.152

Correlation Matrix

1.		
0.236	1.	
0.404	0.365	1.

TABLE 4

Compensating Variation in Dollars per Trip

	Logit	RPL with uncorrelated coefficients	RPL with correlated coefficients
Increase fish stock at all sites by 100 fish per 1000 feet of river	1.40	1.44	0.93
Double fish stock at all sites	3.18	4.25	2.70
Eliminate Madison Rives sites:			
per trip to any site	-0.58	-0.74	-0.54
per trip to Madison	23.	30.	22.