\

Sampling and Aggregation Issues in Random Utility Model Estimation

Peter M. Feather

Measurement of nonmarket values often involves subjective judgments. Since these judgments may influence results, they should be carefully considered. I focus on an aspect of subjective choice relating to the estimation of random utility models. Such models require specification of each recreationalist's choice set. Whether an individual perceives his choice set as composed of all possible alternatives, a few popular alternatives, or collections of spatially aggregated alternatives is an important judgment affecting the conclusions.

Key words: aggregation, choice sets, importance sampling, random sampling, random utility models.

Accurate measurement of economic losses resulting from environmental degradation is important due to recent legislation requiring polluters to pay for these damages. For example, the Oil Pollution Act of 1990 stipulates that restoration costs as well as the diminished value of the affected resource can be recovered from guilty parties. These diminished values include both use and nonuse values. The inclusion of nonuse values created a controversy concerning the accuracy of these values measured by contingent valuation (CVM) methods. To settle the debate, the National Oceanographic and Atmospheric Administration (NOAA) commissioned a "blue ribbon panel" to determine if CVM is capable of providing reliable information. The panel (Arrow et al.) identified "a number of stringent guidelines for the conduct of CV(M) studies" (p. 41) that allow, when followed, "CV(M) studies (to) convey useful information" (p. 42).

Throughout the debate, travel cost methodologies (TCMs) used to determine lost use values went unchallenged. Although the NOAA report describes these damages as "somewhat more difficult to value than the more obvious out-of-pocket losses" (p. 2), their estimation with TCMs usually requires subjective decisions that affect welfare measures. Such decisions range from the type of model to apply (e.g., discrete choice models, pooled models, zonal models, hedonic models) and construction of the travel cost variable (e.g., what distance cost to use, how to value time) to defining the relevant choice set and affected population. While a few studies have investigated differences in welfare measures from competing models (Caulkins, Bishop, and Bouwes; Bockstael, Hanemann, and Kling; Kling), the issue of subjectivity in TCM deserves more attention to improve the reliability of the approach.

I focus here on an aspect of this issue—the selection of a choice set for the estimation of random utility models (RUMs). The estimation of RUMs can become burdensome due to large choice sets. This requires the investigator to subjectively decide how to reduce the choice set to facilitate estimation. Examples include Morey, Rowe, and Watson, who defined destinations as either "primary" sites, or the center of several sites to determine travel distance; and Kling, who selected six (out of seventeen) "popular" beaches for use in a simulation experiment.

Two common choice set reduction strategies are site aggregation and random draws from the choice set. The aggregation approach consists of combining similar alternatives into fewer aggregate alternatives. Random draw approaches

Peter M. Feather is an environmental economist with the Economic Research Service, USDA.

The views in this paper do not necessarily reflect those of the U.S. Department of Agriculture. The author wishes to thank Daniel Hellerstein, Mary Jo Kealy, Frank Lupi, Michael Needelman, and two anonymous reviewers for helpful comments. The content and conclusions are the sole responsibility of the author.

Review coordinated by Richard Adams.

Amer. J. Agr. Econ. 76 (November 1994): 772–780

use a randomly drawn subsample of destinations from the choice set. Applications of these approaches have been the subject of two recent articles. Parsons and Kealy examined the random draw approach and found that the procedure produced relatively stable coefficients even when small sample sizes were drawn. They suggested that the procedure be compared to an aggregation scheme. This was done by Parsons and Needelman, who found that aggregated models and random draw models produced widely different welfare estimates. On the basis of their experiments, Parsons and Needelman favor the random draw approach: "...if individual site data are available and the number of alternatives is large, estimate the Random Utility Model with random draws of alternatives" (p.431).

This conclusion was based on a comparison of welfare measures between the two methods. Another way to evaluate these approaches is to compare them with a "complete alternatives" model estimated using all individual sites as alternatives. If each individual's true choice set is made up of these alternatives, then the complete alternatives model can be treated as the true model providing a means of comparing the random draw and aggregation approaches.

In this paper, a complete alternatives RUM is estimated using a choice set of 286 lakes visited by a sample of fishermen. Next, two random draw models and an aggregated model are estimated from the same data set. The sampling strategies employ either equal (uniform) selection probabilities (as done in Parsons and Kealy), or importance sampling of alternatives. The former strategy implies that each individual perceives all 286 lakes in their choice sets since each lake has an equal probability of being drawn. The latter strategy implies that individuals have a more narrowly defined choice set where the majority of the lakes have very small choice probabilities while a few "important" lakes have much larger choice probabilities. These different definitions of the underlying choice set add another dimension of subjectivity which can be examined by comparing the methods.

Aggregation and sampling strategies are reviewed in the next two sections. The discussion follows that of Ben-Akiva and Lerman. Next, the data are described and the estimation results are presented. Finally, the welfare effects resulting from a change in quality are calculated using each type of model.

Aggregation of Alternatives in RUMs

The alternatives that make up an individual's choice set are termed "elemental alternatives" (Ben-Akiva and Lerman). For example, lakes may form the set of elemental alternatives for an individual participating in water-based recreation. The random utility of elemental alternative l to individual k (U_{lk}) is written as

$$(1) \quad U_{lk} = V_{lk} + \varepsilon_{lk}, \; l = 1, \dots, N \qquad k = 1, \dots, K$$

where V_{lk} is the deterministic portion of the utility function and ε_{lk} is an independently and identically distributed extreme value random variable with mode 0 and scale parameter μ. It is well known that the parameters of V_{lk} can be estimated using a multinomial logit model

$$(2) \quad P_k(l) = \exp(\mu V_{lk}) \,/\, \sum_{j=1}^{N} \exp(\mu V_{jk})$$

where $P_k(l)$ is the probability that individual k chooses elemental alternative l.

If the size of the choice set precludes estimation, it can be reduced by aggregating similar elemental alternatives into disjoint sets termed "aggregate alternatives." This approach may also be necessary in situations where data describing the elemental alternatives are only available on an aggregate basis (e.g., county or regional data). Another reason to aggregate may arise when little or no information is available for some alternatives, but these alternatives are assumed to be similar to others with known information.

In the aggregated model, disjoint sets (L_i) of aggregate alternatives are formed from the set of elemental alternatives. The random utility of aggregate alternative i to individual k (U_{ik}) is

$$(3) \quad U_{ik} = max(V_{lk} + \varepsilon_{lk}) \quad \forall\, l \in L_i.$$

It has been shown (see Parsons and Needelman) that (3) can be decomposed into

$$(4) \quad U_{ik} = \bar{V}_{ik} + (1/\mu)\ln(M_i) + (1/\mu)\ln(B_i) + \varepsilon_{ik}$$

where $\bar{V}_{ik}$ is the average of the ith aggregate alternative, M_i is the number of elemental alternatives in the ith aggregate alternative, and B_i is a measure of the variability of the utilities of the elemental alternatives in the ith aggregate alternative

$$(5) \quad B_i = \sum_{l \in Li} \exp\left[\mu(V_{lk} - \overline{V}_{ik})\right] / M_i.$$

Estimating an aggregated model using only $\overline{V}_{ik}$ results in bias unless either $1/\mu$ equals zero, or B_i and M_i are constant across aggregate alternatives. Neither of these conditions are likely. Because B_i is unknown, aggregate models will contain some bias even when $\ln(M_i)$ is included. Careful definition of aggregate alternatives in terms of similarity between elemental alternatives in each set and inclusion of $\ln(M_i)$ reduces, but does not eliminate this bias.

Sampling of Alternatives in RUMs

McFadden has shown that consistent RUM estimates can be obtained using a subset of randomly drawn alternatives. Let $\pi_k(\mathbf{D} \mid l)$ denote the conditional probability of drawing a subset [**D**] of alternatives for the kth individual given that the lth alternative is chosen. The joint probability of drawing both l and **D** can be written as

$$(6) \quad \pi_k(l, \mathbf{D}) = \pi_k(\mathbf{D} \mid l) P_k(l).$$

Using Bayes' theorem, the conditional probability of l being chosen given a subset of alternatives **D** is

$$(7) \quad \pi_k(l|\mathbf{D}) = \pi_k(\mathbf{D}|l)P_k(l) / \left\{\textstyle\sum_{j \in D} \pi_k(\mathbf{D}|j)P_k(j)\right\}$$
$$= \exp\left\{\mu V_l + \ln \pi_k(\mathbf{D}|l)\right\} / \textstyle\sum_{j \in D}$$
$$\left\{\exp[\mu V_j + \ln \pi_k(\mathbf{D}|l)]\right\}.$$

If each lake is assigned a uniform selection probability and J lakes are drawn without replacement from the $N - 1$ lakes not chosen (the chosen lake is added to the sample after the draw so that **D** contains $J + 1$ lakes), then $\pi_k(\mathbf{D} \mid l)$ is

$$(8) \quad \pi_k(\mathbf{D} \mid l) = J!(N - J - 1)!/(N - 1)! \quad \forall\, l \in \mathbf{D}.$$

This model is simple to estimate because (8) cancels out during estimation.

Ben-Akiva and Lerman suggest that random draws using uniform selection probabilities may not be an efficient sampling scheme. They argue that most alternatives have small choice probabilities while a few, more important alternatives have much larger choice probabilities. Sampling methods taking this into account have been termed "importance sampling" (Ben-Akiva and Lerman). One form of importance sampling requires assigning each elemental alternative a selection probability (q_{lk}). A sample of size J is then drawn with replacement from the choice set of all N alternatives. Once the sample is drawn, duplicate draws are deleted and the chosen alternative is added if it was not drawn. Under this approach, the size of **D** is bounded between 1 and $J + 1$. The conditional probability of **D** given l is

$$(9) \quad \pi_k(\mathbf{D} \mid l) = (q_{lk})^{-1} Q_k(\mathbf{D}) \qquad \forall\, l \in \mathbf{D}.$$

where $Q_k(\mathbf{D})$ is independent of the chosen alternative (see Ben-Akiva and Lerman, p. 266). The resulting model to estimate is

$$(10) \quad \pi_k(l \mid \mathbf{D}) = \exp[\mu V_{lk} - \ln(q_{lk})] / \sum_{j=l}^{N} \exp[\mu V_{jk} - \ln(q_{jk})].$$

Therefore, importance sampling requires the natural logarithm of the selection probability be included for each alternative in **D**, with the parameter constrained to equal minus one.

Data

A survey of Minnesota anglers and two supplementary water quality data sets are used in the analysis. The Survey of the Socioeconomic Aspects of Minnesota Sport Fishing was conducted in late 1989 and early 1990 by the Minnesota Center for Survey Research for the Minnesota Department of Natural Resources (MDNR). Survey participants were selected randomly from a population of individuals who purchased a fishing license in 1988. For survey details see Estenson. The lake water quality data set was constructed using information from the MDNR and the Minnesota Pollution Control Agency (MPCA). The combined MPCA and MDNR data sets contain information describing approximately 3,500 individual lakes. Both data sets contain numerous missing values and a large variance in observations per lake. Because many lakes in Minnesota are small and relatively inaccessible, the MPCA and MDNR concentrate their sampling efforts on lakes that are heavily used for water-based recreation. Acreage is known for almost all lakes, but other quality measures are known for only a small percentage of the lakes. The more numerous existing quality measures are lake depth, secci disk depth (a measure of water clarity), and the percentage of lake acreage considered littoral (near the shoreline) acres. Other quality measures such as pH, water temperature, and dissolved oxygen are known for fewer lakes.

The 286 lakes with known measures of total acreage, depth, percentage of littoral acres and secci depth measurements are used in the analysis. Observations collected during the "open water season" (June 24 to September 11) from 1985 through 1989 were averaged for each lake to describe the elemental alternatives. The "open water season" has been used by the MPCA in previous water quality analysis because: "Summer data are preferred for assessment purposes as they generally correspond to the maximum productivity of the lake, yield the best agreement between trophic variables, and reflect the period of maximum use of the resource" (Heiskary and Wilson, p. 5). The natural logarithm of lake acres is used in estimation to capture the diminishing effect of lake size described by Parsons and Kealy (i.e., marginal differences in lake size is perceived differently in small lakes than in large lakes). Aggregate alternatives were created by averaging quality observations over each county. The county lake size variable is the average natural logarithm of lake size.

The survey sample consists of 969 fishing trips taken to the 286 lakes in the water quality data set.[1] Lake visitation frequencies suggest that importance sampling may be a more reasonable method of estimating the choice set than random sampling. Three lakes in the sample account for 30% of all fishing trips. Another eight lakes account for an additional 11% of all fishing trips. In general, these lakes tend to be larger, murkier, shallower, and have less shore line acres than the average lake in the sample. A few exceptions occur in lakes located on the northern border of the state which are also large, but tend to be deeper, clearer, and have more shoreline acres than the average lake in the sample. This pattern indicates that a few lakes may indeed have high probabilities of being visited while most others have low visitation probabilities.

Travel cost is computed as the round trip travel distance times the American Automobile Association's estimated cost per mile based on the average costs of driving three selected mid-size cars 15,000 miles per year in 1989 ($0.306 per mile). Two distance matrices were constructed using the computer program ZIPFIP (Hellerstein et al.) to measure travel distance. Straight line distances between the centroid of each individual's zip code zone and the centroid of each lake (county) are used in the sampling (aggregated) models.

Estimation Results

The random sampling procedure is estimated using draws without replacement of size 5, 11, and 23 lakes from the choice set of $N - 1$ lakes not visited by the individual. On each draw, each lake has an equal chance of being selected. The lake actually visited is then added to the sample choice set. The importance sampling procedure is estimated using draws of size 6, 12, and 24 lakes with replacement from the entire choice set of N lakes. For each individual, and on each draw, the jth lake has a selection probability (q_j) equal to t_j/T where t_j is the total number of trips taken to the jth lake and T is the total number of trips. Duplicate draws are deleted and the lake actually visited is added if it was not drawn. The size of the draw is varied to investigate the effect of sample size on the estimates. Parsons and Kealy followed such an approach when they used the random sampling procedure described above and found that reliable estimates could be obtained with as few as six randomly drawn lakes. To lessen the possibility of drawing an unrepresentative sample, three replications are performed for each sample size.

RUM estimates using the sampling strategies to represent the choice set appear in tables 1 and 2. The models are estimated under the assumption that $\mu = 1$. For purposes of comparison, each table also includes the complete alternatives model estimated using all 286 lakes. The parameter estimates of the complete alternatives model indicate that anglers prefer lakes that are cheaper to visit, clearer, and larger with less (percentage-wise) shoreline

[1] The entire survey consisted of 1,099 trips taken to 360 individual lakes. All 286 lakes in the water quality data set are elements of this set of 360 lakes. The remaining 74 lakes are missing one or more quality measures. The 130 trips taken to these 74 lakes were discarded leaving 969 trips for estimation purposes.

Table 1. RUM Parameters Estimated Using Random Draws

	Variables[a]				
Sample size[b]	Cost	Depth	Secci	Plitt	ln(Area)
6	−0.0163	0.0055	−0.0658	−0.4053	0.4929
	(−14.0)	(2.80)	(−1.55)	(−1.87)	(21.2)
6	−0.0159	0.0014	0.0270	−0.3983	0.5228
	(−13.7)	(0.69)	(0.65)	(−1.84)	(22.2)
6	−0.0157	0.0032	0.0225	−0.3149	0.5280
	(−13.4)	(1.58)	(0.54)	(−1.47)	(22.4)
12	−0.0165	−0.0039	0.0724	−0.4677	0.5998
	(−15.3)	(−2.23)	(1.87)	(−2.34)	(27.4)
12	−0.0170	0.0024	0.0193	−0.1734	0.5923
	(−15.6)	(1.32)	(0.49)	(−0.83)	(27.0)
12	−0.0161	−0.0013	0.0649	−0.6041	0.5711
	(−15.2)	(−0.74)	(1.69)	(−3.06)	(26.9)
24	−0.0162	−0.0021	0.0964	−0.3711	0.6199
	(−15.8)	(−1.30)	(2.70)	(−1.96)	(31.3)
24	−0.0162	−0.0018	0.0923	−0.3488	0.6328
	(−15.8)	(−1.11)	(2.56)	(−1.84)	(31.6)
24	−0.0165	0.0026	0.0808	−0.1674	0.6172
	(−16.1)	(1.58)	(1.19)	(−0.87)	(30.9)
Complete	−0.0179	−0.0002	0.0931	−0.2757	0.6895
	(−18.9)	(−0.17)	(2.67)	(−1.54)	(38.1)

Note: Models based on random draws using uniform selection probabilities from the choice set of lakes not visited. The lake actually visited was then added to the sample. Draws are without replacement.

[a] Cost is travel cost. Depth is maximum lake depth in meters. Secci is the secci disk depth in meters. Plitt is the percentage littoral acres in each lake. ln(Area) is the natural logarithm of lake area in acres. t-statistics for the null hypothesis that the parameter equals zero appear in parenthesis.

[b] Sample size is the number of alternatives used to estimate the model ("sample size" minus one draws were made without replacement from the set of lakes not visited, then the visited lake was added). Number of trips is 969.

acreage. The depth variable is insignificant, but is retained because it may prove to be significant under different assumptions regarding the true underlying choice set.

The random draw models appear in table 1. Generally, as the sample size increases, the parameter standard errors decrease and the parameter estimates approach those of the complete alternatives model. This is especially true for the travel cost and log area variables which are significant at the 1% level in the complete alternatives model. Conversely, the depth variable often changes signs and is frequently insignificant in many of the models. These results agree with observations by Parsons and Kealy, who found travel cost and log area parameters to be relatively stable, but observed variation in other parameters. Examining estimates within a given sample size reveals that the parameter values often differ by significant amounts. As the sample size increases, the effect lessens, but does not disappear. Evidently, the estimates are sensitive to the randomly chosen choice set, especially when few alternatives are drawn.

Estimates using importance sampling appear in table 2. In general, the parameters of these models are more stable across draws within a given sample size than the parameters estimated using random draws. The parameter values remain relatively stable across sample size with the parameter standard errors decreasing as sample size increases. Exceptions are the percentage of littoral acres and log area parameters which increase (in absolute value) with sample size. The importance sampling results differ somewhat from the complete alternatives model. The depth parameter is negative and similar in magnitude to the value in the complete alternatives model, but much more significant (at the 1% level). The secci disk parameter is significant, but negative for all repetitions. The remaining parameters [Cost, ln(Area) and Plitt] agree in sign, magnitude, and significance with those of the complete alternatives model.

Differences in these models are the result of the sampling strategies used. The random draw procedure produces better estimates of the complete alternatives model. As the number of draws used approaches the size of the choice set, the randomly drawn sample will converge to the underlying choice set. This will not be true for the importance sampling strategy. Be-

Table 2. RUM Parameters Estimated Using Importance Sampling

Sample size[b]	Cost	Depth	Secci	Plitt	ln(Area)
	Variables[a]				
6	–0.0206	–0.0087	–0.1614	–0.4398	0.7309
	(–17.9)	(–4.38)	(–3.50)	(–1.79)	(32.2)
6	–0.0197	–0.0098	–0.1782	–0.4496	0.7159
	(–17.2)	(–5.00)	(–3.91)	(–1.83)	(31.8)
6	–0.0204	–0.0083	–0.1301	–0.4466	0.7180
	(–17.6)	(–4.21)	(–2.82)	(–1.83)	(32.3)
12	–0.0215	–0.0107	–0.1429	–0.5770	0.7763
	(–19.4)	(–5.96)	(–3.39)	(–2.55)	(36.3)
12	–0.0207	–0.0108	–0.1509	–0.6106	0.7824
	(–18.8)	(–6.03)	(–3.61)	(–2.70)	(36.7)
12	–0.0211	–0.0103	–0.1495	–0.6937	0.7829
	(–19.1)	(–5.78)	(–3.57)	(–3.08)	(37.0)
24	–0.0228	–0.0127	–0.1099	–0.7350	0.8423
	(–20.8)	(–7.49)	(–2.78)	(–3.37)	(39.9)
24	–0.0231	–0.0111	–0.1278	–0.7035	0.8503
	(–20.9)	(–6.51)	(–3.22)	(–3.18)	(40.3)
24	–0.0192	–0.0097	–0.3915	1.0606	0.8597
	(–18.9)	(–5.83)	(–7.06)	(4.24)	(43.3)
Complete	–0.0179	–0.0002	0.0931	–0.2757	0.6895
	(–18.9)	(–0.17)	(2.67)	(–1.54)	(38.1)

Note: Models based on random draws with replacement from the entire choice set with the probability of being selected on each draw equa to the proportion of total trips that were taken to each lake. Duplicate draws were deleted, and the lake actually visited was added to the sample if it was not drawn. Models include the natural logarithm of selection probabilities with the parameter constrained to equal minus one.

[a] Cost is travel cost. Depth is maximum lake depth in meters. Secci is the secci disk depth in meters. Plitt is the percentage littoral acres in each lake. ln(Area) is the natural logarithm of lake area in acres. t-statistics for the null hypothesis that the parameter equals zero appear in parenthesis.

[b] Sample size is the number of draws made from the choice set ("sample size" draws were made from the entire choice set with replacement, duplicates were deleted and the chosen alternative added if it was not drawn). Complete refers to the complete alternatives model. Number of trips is 969.

cause selection probabilities are unequal, and sampling is performed with replacement, lakes with higher selection probabilities will invariably appear in each individual's sample choice set much more frequently than when random sampling is used. The differences in the results of the two methods reflects a fundamental difference in defining the "true" choice set faced by an individual. If each individual perceives the choice set to be all 286 lakes used in the analysis, then the complete alternatives model can be treated as the truth and the random sample models can be treated as reasonable estimates of it. On the other hand, if individuals predominantly favor a few lakes, then the importance sampling models are reasonable estimates of the "true" unknown model.[2]

[2] The signs of the secci disk parameters emphasize this difference in estimating the individual's choice set through the sampling procedures. Clearer water appears to be more desirable when the random sampling approach is used, but less desirable when importance sampling is used. However, because the relationship between angler perceptions and water clarity is not known, the differences resulting from the two approaches cannot be used to validate one and refute the other. It is unlikely that a high quality angling location has water clarity at either extreme.

The models estimated with aggregated alternatives appear in table 3. Two models are estimated: the "Adjusted" model is adjusted for aggregation bias,[3] the "Simple" model is not. These models also represent differences in defining the "true" choice set. If the perceived choice set is the 286 lakes in the lake data set, then the adjusted model is appropriate because it is an estimate of the true complete alternatives model. If individuals base their perceptions on the average quality over spatial boundaries, then the simple model may be a better estimator of the "true" unknown model.

The adjusted model provides better estimates of the complete alternatives model's travel cost, log area, and depth parameters than does the simple model. In both of these models, the secci disk parameter is significant at the 1% level and has a positive sign, but is several times larger than in the parameter in the alternatives model. Neither model produces reason-

[3] The correction for aggregation in the adjusted model is the natural logarithm of the number of lakes used as a county descriptor and constraining the parameter to equal one.

Table 3. RUM Parameters Estimated Using Aggregated Data

Model[a]	Cost	ln(Area)	Depth	Secci	Plitt
Simple	−0.0121	0.3014	0.0069	0.5556	−1.7074
	(−14.2)	(12.2)	(1.72)	(14.8)	(−5.48)
Adjusted	−0.0192	0.7626	−0.0013	0.4910	0.2460
	(−19.6)	(24.3)	(−0.29)	(8.91)	(0.59)
Complete	−0.0179	0.6895	−0.0002	0.0931	−0.2757
	(−18.9)	(38.1)	(−0.17)	(2.67)	(−1.54)

Note: Results from aggregating quality data from 286 individual lakes into county aggregates. Number of trips is 969. Cost is travel cost. Depth is maximum lake depth in meters. Secci is the secci disk depth in meters. Plitt is the percentage littoral acres in each lake. ln(Area) is the average natural logarithm of lake area in acres. t-statistics for the null hypothesis that the parameter equals zero appear in parenthesis.
[a] The "Simple" model is the aggregated model excluding the natural logarithm of the number of lakes [ln(M)] in each aggregated alternative; the "Adjusted" model includes ln(M) and constrains its parameter to equal one. Complete refers to the complete alternatives model estimated with all 286 lakes.

able estimates of the percentage of littoral acres parameter. If the complete alternatives model is assumed to be the true model, then differences in the parameter estimates between the two aggregated models can be attributed to the variance term (B_i) in (4), which is unknown. However, these differences could also be attributed to alternative definitions of the underlying "true" choice set.

Welfare Assessments

The models discussed above are used to calculate the compensating variation (CV) resulting from a 10% increase in the logarithm of acreage at all lakes. The lake acreage variable is chosen because it is consistent in sign and significant at the 1% level across all models. Following Parsons and Kealy, all individual lake data are used to estimate the CV measures in the random draw and importance sampling models. Both individual and aggregated data are used to compute welfare measures in the aggregated models since one is adjusted for aggregation bias while the other is not. The kth individual's per choice occasion CV for a change in quality (see Parsons and Needelman) is

$$(11)\quad CV_k = \left[\ln \sum_{i=1}^{N} \exp(V_{ik}'') - \ln \sum_{i=1}^{N} \exp(V_{ik}') \right] / \beta \qquad \forall\, k = 1, \ldots, K$$

where β is the marginal utility of income, V_i'' (V_i') is the indirect utility of visiting the ith lake evaluated at the final (initial) quality level and N is the number of alternatives in the choice set.

Table 4 shows the mean CV computed using the sampling models divided by the mean CV computed using the complete alternatives model.[4] The accuracy of the random sampling approach increases with sample size. Compared to the complete alternatives model, the travel cost parameters in these models are relatively similar, but the log area parameters are (decreasingly) smaller over sample sizes. This re-

Table 4. Ratio of Average Welfare Measures - Sampling Models

	Sampling procedure	
Sample size[c]	Random[a]	Importance[b]
6	0.711	0.928
6	0.786	0.946
6	0.807	0.919
12	0.906	0.966
12	0.860	1.017
12	0.873	1.000
24	0.966	1.020
24	0.992	1.017
24	0.935	1.188

Note: Ratio of mean "estimated" welfare to mean "complete alternatives" welfare for a 10% increase in the natural logarithm of lake size.
[a] Random sampling procedure refers to assigning uniform selection probabilities to each lake and then sampling from the choice set of lakes not visited without replacement. The lake actually visited was then added to the sample.
[b] Importance sampling procedure refers to assigning actual visitation probabilities to each lake and then sampling with replacement from the entire choice set. Duplicates were deleted, and the chosen alternative was added if it was not drawn
[c] For the random sampling procedure, sample size is the number of alternatives used to estimate the model ("sample size" minus one draws were made without replacement from the set of lakes not visited, then the visited lake was added). For the importance sampling procedure, sample size is the number of draws made from the choice set ("sample size" draws were made from the entire choice set with replacement, duplicates were deleted and the chosen alternative was added if it was not drawn).

[4] The CV for the quality change using the complete alternatives model is $34.56.

Table 5. Ratio of Average Welfare Measures - Aggregated Models

Model[c]	Aggregated ratio[a]	Lake ratio[b]
Simple	0.588	0.538
Adjusted	1.196	1.072

[a] Estimated average compensating variation divided by average compensating variation from the complete alternatives model resulting from a 10% increase in the natural logarithm of lake acres. The aggregate welfare measure is estimated using aggregate data. The natural logarithm of lake acreage is increased 10% for each individual lake and then averaged for each county.
[b] Estimated average compensating variation divided by average compensating variation from the complete alternatives model resulting from a 10% increase in the natural logarithm of lake acres. The aggregate welfare measure is estimated using individual lake data.
[c] The "Simple" model is the aggregated model excluding the natural logarithm of the number of lakes [$\ln(M)$] in each aggregated alternative; the "Adjusted" model includes $\ln(M)$ and constrains its parameter to equal one.

sults in underestimates of the CV.[5] Even though the importance sampling procedure implies a different underlying "true" choice set than the random sampling procedure, the CV estimates using this sampling strategy are remarkably similar to the complete alternatives model. Because the parameters are more stable across and within sample size, the welfare estimates are more stable compared to the random sampling approach. Except for what appears to be an "outlier" sample choice set on one of the twenty-four draw samples, the CV measures calculated with importance sampling are generally invariant with respect to sample size. Apparently, large sample sizes and the need to perform replications appear to be more important when using the random sampling approach.

Similar CV ratios for the aggregated models appear in table 5. Both individual and aggregated data are used because one of these models is adjusted for aggregation bias while the other is not. Using either aggregated or individual data, the simple aggregated model produces much lower CV measures than the complete alternatives model. These differences can either be attributed to aggregation bias, or to differences in the definition of the underlying choice set. The simple model is an estimate of the "true" model only if individuals perceive their choice set as aggregate areas having similar quality levels. If the underlying choice set is actually individual lakes, then the adjusted model is the more appropriate aggregated model because it removes some of the aggregation bias. It produces CV measures that are more similar to the complete alternatives model, especially when individual lake data are used.

[5] This may be because frequently visited lakes are larger than average. Since the chosen alternative is always included in each individual's choice set, large lakes will make up a large proportion of lakes in the sample choice set when the number of draws is small. This effect diminishes as the number of draws increases.

Conclusions

I consider an area of subjectivity in travel cost models: the selection of a choice set for the estimation of random utility models. Estimating such models often involves large choice sets that must be reduced either by aggregation or by some form of sampling. Whether to aggregate or utilize a sampling strategy is a subjective decision regarding the "true" underlying choice set. If the true choice set is composed of all sites a recreationalist could potentially visit, the complete alternatives model can be treated as the truth, and the random sampling and adjusted aggregated models can be treated as estimates of that model. If individual perceptions are based upon limited information, the importance sampling models may be a better representation of the true unknown model. Finally, if perceptions are based upon spatial averages, the simple aggregated model may be closer to the truth.

These decisions affect the estimated parameters and welfare measures. Both the sampling and adjusted aggregation models yield slightly different results from the same (complete alternatives) model. The random sampling approach does much better at approximating the complete alternatives model when large numbers of draws are used. Because of the parameter variation observed within sample sizes, the investigator should perform multiple draws to avoid reliance on a single, possibly unrepresentative choice set. This is less true for the importance sampling procedure. The latter yields fairly stable parameter and welfare estimates across sample size and may be the better procedure to use—especially if only a few draws can be made from the choice set. Comparison of aggregated models shows that, in terms of parameter and welfare estimates, the adjusted approach provides a much better estimate of the complete alternatives model than does the simple approach. If the complete alternatives model is assumed to be true, the results indicate that aggregate models should be adjusted for aggregation bias. Welfare estimates from an adjusted model should be based on individual site-specific data rather than on aggregated data.

None of the three alternative views of the true underlying choice set, or of the appropriate model, are necessarily advocated in this paper. Little is known about how recreationalists perceive their choice sets when deciding upon which destination to visit. The pertinent point is that this perception may significantly affect welfare measures and should be carefully considered. Future work towards refining perceptions of "true" underlying choice sets may include Monte Carlo simulations and better survey information.

[Received May 1993; final revision received April 1994.]

References

Arrow, K., R. Solow, P. Portney, E. Leamer, R. Radner, and H. Schuman. "Report of the National Oceanographic and Atmospheric Administration Panel on Contingent Valuation." *Federal Register* 58(January 1993):4601–14

Ben-Akiva, M., and S. Lerman. *Discrete Choice Analysis: Theory and Application to Travel Demand.* Cambridge: The MIT Press, 1985.

Bockstael, N., W. Hanemann, and C. Kling. "Estimating the Value of Water Quality Improvements in a Recreational Demand Framework." *Water Resour. Res.* 23(May 1987):951–60.

Caulkins, P., R. Bishop, and N. Bouwes. "The Travel Cost Model for Lake Recreation: A Comparison of Two Methods for Incorporating Site Quality and Substitution Effects." *Amer. J. Agr. Econ.* 68(May 1986):291–97.

Estenson, J. *Survey of the Socioeconomic Aspects of Minnesota Sportfishing.* Minnesota Department of Natural Resources, Center for Survey Research Tech. Rep. #90-9, 1990.

Heiskary, S.A., and C.B. Wilson. *Minnesota Lake Water Quality Assessment Report.* Minnesota Pollution Control Agency, Program Development Section, Division of Water Quality, 1990.

Hellerstein, D., D. Woo, D. McCollum, and D. Donnelly. *ZIPFIP: A Zip and FIPS Database.* Washington DC: U.S. Department of Agriculture, ERS-RTD, 1993.

Kling, C. "Comparing Welfare Estimates of Environmental Quality Changes from Recreational Demand Models." *J. Environ. Econ. and Manage.* 15(November 1988):331–40.

McFadden, D. "Modeling the Choice of Residential Location." *Spatial Interaction Theory and Residential Location.* A. Karlquist et al., eds. pp.75–96. Amersterdam: North Holland Publishing, 1978.

Morey, E., D. Rowe, and M. Watson. "A Repeated Nested-Logit Model of Atlantic Salmon Fishing." *Amer. J. Agr. Econ.* 75(August 1993):578–92.

Parsons, R., and M. Kealy. "Randomly Drawn Opportunity Sets in a Random Utility Model of Lake Recreation." *Land Econ.* 68(February 1992):93–106.

Parsons, R., and M. Needelman. "Site Aggregation in a Random Utility Model of Recreation." *Land Econ.* 68(November 1992):418–33.