Population-Specific Recreation Demand Models and the Consequences of Pooling Sample Data

Rod F. Ziemer and Wesley N. Musser

This paper considers the implications of different geographical population definitions in analysis of demand for wildlife recreation. Demand functions for fishing, small game hunting, big game hunting, and wildlife enjoyment are estimated for individual Southeastern states and also for a pooled sample of all the states. Statistically significant differences between the state and regional estimates of the variable cost coefficient exist in 18 of the 40 cases. Consumer surplus values derived from state cost coefficients can differ greatly from values derived from pooled coefficients.

The demand for wildlife recreation increasingly is becoming the subject of economic analysis. This investigative interest most likely results from increased public demand for outdoor recreation in recent years, a demand that according to the Outdoor Recreation Resources Review Commission will increase three-fold by the end of the century. Cicchetti, *et. al.*, have classified past recreation demand studies into two types — site-specific and population-specific. The majority, site-specific, are concerned with participation rates at a particular recreation site. Population-specific models utilize data associated with observed households without regard to actual site visitation frequencies. Population specific models have been concerned with various population definitions: the U.S. Bureau of Outdoor Recreation and Kalter and Gosse used a national sample, Brown, *et. al.*, used a state sample, and Gum and Martin (1975) considered a set of regions within a state. While choice of population size is related both to the research problem and available data, results of empirical studies in general suggest that size of population could influence the demand parameters.

This paper considers the consequences of sample size on the parameters for population-specific recreation demand models. Demand models are estimated from separate samples for states in the Southeastern United States. Similar models are also estimated for a pooled sample of all the states within the Southeast. The coefficients for the average cost variables in the state models are then compared with the Southeastern regional coefficients resulting from the pooled sample. The impact of using the pooled coefficients for individual state estimates of consumer surplus for various recreation activities is then considered. In addition, regional estimates of consumer surplus derived with both the pooled and state coefficients are also considered.

A Regional Model of Wildlife Recreation Demand

Demand functions were estimated for warm water fishing, small game hunting, big game hunting, and general wildlife enjoyment which was defined as watching and photographing birds and animals. These de-

Rod F. Ziemer is Graduate Research Assistant and Wesley N. Musser is Associate Professor of Agricultural Economics, University of Georgia.

Carter Hill, Peter J. Barry, and anonymous reviewers provided very helpful comments at various stages in the preparation of this paper.

mand functions were based on the following general model:

$$(1) \quad \ln Q_{ih} = a_h + \sum_{k=1}^{n} b_{kh} X_{ik} + u_{ih}$$

where ln Q_{ih} is the natural log of quantity of occasions demanded by household 'i' for activity 'h', X_{ik} is the value of the independent variable 'k' for household 'i', a_h is the intercept term, b_{kh} is the coefficient for the independent variable 'k' for activity 'h' and u_{ih} is an error term. The semi-log mathematical form was chosen on the basis of equation fit and statistical significance in trial specifications.

States included in the empirical analysis were Arkansas, Georgia, Kentucky, Louisiana, Maryland, Mississippi, South Carolina, Tennessee, Virginia and West Virginia; data for three Southeastern states, Alabama, Florida, and North Carolina, were not included in the sample as these states did not participate in the survey. All data were for 1971 and obtained from a survey of wildlife recreation completed in 1974 by the Environmental Research Group at Georgia State University.

The independent variables, which are consistent with demand theory and past recreation studies, such as Brown, *et. al.*, Edwards, *et. al.*, and Gum and Martin (1975) were: (1) average cost per occasion, (2) average miles traveled per occasion, (3) total leisure time, (4) education, and (5) family income. The choice of average cost as a proxy for price is based on the approach of the standard Hotelling-Clawson model discussed and summarized by Edwards, *et. al.*, (pp. 4-7). Total leisure time was defined as total leisure hours both during the week and on weekends. The education variable was defined as either the total years of education for a single person, or as the average years of education for the husband and wife for a married household. Since respondents were asked to specify the income range corresponding to their annual income on the wildlife recreation survey, family income was represented as a set of dummy variables. Table 1 indicates the income ranges used in constructing the dummy variables; the first income variable was dropped in the regression analysis and an intercept term included.

Results for the ordinary least squares regression analysis as applied to the pooled sample are presented in Table 2. Standard t-scores appear in parentheses. Average cost and mileage per occasion were significant in all four activity equations. Total leisure time was significant in all but the wildlife enjoyment equation, while education was only significant in the small game hunting equation. Insignificant variables were deleted from the final specifications Family income, at least over some ranges, was significant in all activity equations.[1]

Comparability of Southeastern and State Cost Coefficients

The average cost regression coefficient obtained in outdoor recreation demand models is often used, as an indicator of participant's cost sensitivity, to value recreational resources in terms of consumer surplus (for example see Gum and Martin or Sawyer and

[1]Households which did not participate in any form of wildlife recreation were not considered in the analysis (for example, only households which actually fished were considered in the fishing equation). Non-participants are excluded because their average cost of zero is an inappropriate proxy for price. Thus, the price proxy for recreation is only observable for actual participants in a population specific study.

TABLE 1. Range of Family Income Dummies

Variable	Range
1	Under $3,000
2	$3,001 to $5,000
3	$5,001 to $7,000
4	$7,001 to $10,000
5	$10,001 to $15,000
6	$15,001 to $20,000
7	$20,001 to $25,000
8	Over $25,000

TABLE 2. Regression Results, Wildlife Recreation Demand in the Southeastern U.S., 1971

	Activity			
Independent Variable	Warm Water Fishing	Small Game Hunting	Big Game Hunting	Wildlife Enjoyment
Intercept	2.355813 (28.81)***	2.209727 (17.61)***	1.598179 (12.78)***	4.752367 (27.81)***
Average Variable Cost	−.010154 (−6.32)***	−.000649 (−2.00)**	−.000730 (−3.33)***	−.047437 (−4.19)***
Average Mileage per Occasion	−.000014 (−9.34)***	−.000009 (−11.54)***	−.000006 (−10.92)***	−.000016 (−7.14)***
Total Leisure Time	.012838 (4.57)***	.007487 (2.47)**	.012046 (3.39)***	-
Education	-	−.018970 (−1.85)*	-	-
Family Income 2	.235154 (2.55)**	.326397 (3.08)***	.201229 (1.45)	−.077916 (−.33)
Family Income 3	.348949 (3.89)***	.306494 (2.98)***	.119003 (.92)	−.263850 (−1.12)
Family Income 4	.442100 (5.30)***	.426091 (4.22)***	.090880 (.74)	−.538127 (−2.45)**
Family Income 5	.518186 (6.06)***	.449735 (4.23)***	.332651 (2.69)***	−.459308 (−2.14)**
Family Income 6	.467187 (4.40)***	.378893 (2.91)***	.235656 (1.62)	−.580888 (−2.41)**
Family Income 7	.457727 (3.38)***	.340303 (2.08)**	.257508 (1.34)	−.405037 (−1.37)
Family Income 8	.358317 (2.57)**	.504813 (3.05)***	.370108 (2.10)**	−.445145 (−1.52)
R^2	.074	.080	.122	.105
Number of Observations	2782	1983	1124	984

***significant at the .01 level
**significant at the .05 level
*significant at the .10 level

Shulstad). Because this coefficient is important in estimating consumer surplus, this analysis of the impact of population size on the demand parameters focuses strictly on the average cost coefficient.

The analysis of comparability was based on estimation of the pooled activity specifications (Table 2) for samples delineated by states. The coefficients for average cost for the various activities by states are presented in Table 3. The t-statistics for comparability of state and pooled coefficients appear in parentheses below the estimated coefficients. The comparability tests, which were based on a null hypothesis that the state coefficient was not statistically different from the pooled coefficient, were derived as follows:

$$t_{ij} = \frac{\hat{\beta}_{ij} - \hat{\beta}_j}{s_{ij}} \quad (2)$$

where t_{ij} is the t-statistic for the i'th state and the j'th activity, $\hat{\beta}_{ij}$ is the average cost coefficient for the i'th state and the i'th activity, $\hat{\beta}_j$ is the Southeastern (pooled) average cost coefficient for the j'th activity, and s_{ij} is the standard error of $\hat{\beta}_{ij}$.

The null hypothesis could be rejected at the five percent level for 18 of the 40 coefficients in Table 3. However, some of these 18 coefficients have such large standard errors that they are not themselves significantly different from zero (superscripted 'b' in Table 3). In such cases, the meaning of the comparability test is not clear. A reasonable criteria might be to only consider those coefficients which differ statistically from zero (indicated by asterisks in Table 3) in a comparability test. According to this criteria, the coefficients for Georgia and Louisiana differ significantly from the pooled coefficients for all activities except wildlife enjoyment, while those for West Virginia differ significantly for all activities except big game hunting. Only one of the activity average cost coefficients for Kentucky, Maryland, Mississippi, and Tennessee differs significantly. None of the states appear in general to have results that are similar to those of the region for all activities; Arkansas is the most similar with three coefficients not significantly different from the pooled estimates at the five percent significance level.

TABLE 3. Average Variable Cost Coefficients and Comparability t-statistics

	Activity			
Region Size	Warm Water Fishing	Small Game Hunting	Big Game Hunting	Wildlife Enjoyment
Pooled	−.010154***	−.000649**	−.000730***	−.047437***
Arkansas	−.019891*** (−1.92)	−.005242** (−1.85)	−.004810* (−1.46)	−.046098 (.02)
Georgia	−.038703*** (−3.96)[a]	−.025063*** (−2.67)[a]	−.010706** (−2.18)[a]	−.079596 (−.62)
Kentucky	−.008687* (.33)	−.000170 (.23)	−.013100*** (−2.88)[a]	.000718 (2.58)[b]
Louisiana	−.034978*** (−2.86)[a]	−.008993*** (−2.76)[a]	−.005736*** (−2.32)[a]	−.401724 (−1.24)
Maryland	−.001178 (3.01)[b]	−.002936 (−.63)	−.011122** (−2.17)[a]	−.014774 (1.45)
Mississippi	−.020653 (−1.80)	−.002045 (−.65)	−.008229*** (−2.54)[a]	−.238892 (−1.00)
South Carolina	−.024775** (−1.45)	−.000231 (1.09)	.000233 (2.48)[b]	−.203269 (−.62)
Tennessee	−.022164** (−1.11)	−.018109*** (−2.66)[a]	−.003383 (−1.08)	.141686 (1.89)
Virginia	−.000390 (2.74)[b]	−.002989 (−.79)	−.012911*** (−2.79)[a]	−.094939*** (−1.44)
West Virginia	−.065028 (−3.67)[a]	−.014962*** (−2.96)[a]	−.000547 (.67)	−.638632*** (−3.77)[a]

***significant at the .01 level

**significant at the .05 level

*significant at the .10 level

[a]The State coefficient is statistically different from the pooled coefficient at the .05 level and is itself statistically different from zero.

[b]The State coefficient is statistically different from the pooled coefficient at the .05 level but is itself statistically insignificant.

Consumer Surplus Results

To clearly illustrate the possible difference in consumer surpluses obtained with pooled and disaggregated data, consumer surpluses were estimated for states with average cost coefficients that differed significantly from the pooled coefficient and significantly from zero (superscripted 'a' in Table 3).

The general demand model presented earlier could be rewritten as follows:

$$(3) \qquad \ln Q_{ih} = a_h + b_h\,(VC_{ih}) + \sum_{k=1}^{n-1} b_{kh}\,X_{ik} + u_{ih}$$

The only change from the model in equation (1) is that average cost per occasion for household 'i' is isolated from the other independent variables. Coefficient b_h is simply the regression coefficient for average cost for activity 'h'. Considering an added cost (c), such as an entrance fee, the equation for the estimated number of occasions demanded becomes:

$$(4) \qquad \ln \hat{q}_{ih} = a_h + b_h\,(VC_h + c) + \sum_{k=1}^{n-1} b_k\,X_{ik} + u_{ih}$$

where $\hat{q}_{ih}$ equals the number of occasions demanded by household 'i' for activity 'h' with added cost c. According to Gum and Martin (1975), an appropriate demand curve for activity 'h' for household 'i' can be obtained by subtracting equation (3) from equation (4). After simplification the following relationship is derived:

$$(5) \qquad \hat{q}_{ih} = e^{b_h c}\,Q_{ih}$$

Gum and Martin (1975) suggest that consumer surpluses be calculated with individual demand curves, as in equation (5), and then summed over all individuals to determine a regional consumer surplus value. To simplify this analysis, consumer surpluses were estimated using aggregate demand curves. An aggregate relationship for activity 'h' could be expressed as:

$$(6) \qquad \hat{q}_h = \Sigma\,\hat{q}_{ih} = e^{b_h c}\,Q_h$$

where $\hat{q}_h$ is the estimated number of occasions demanded in the geographical region under consideration for activity 'h' with added cost 'c', and Q_h is the actual number of occasions consumed in the region for activity 'h'. Integrating equation (6) over c from zero to a value at which no occasions are demanded, ($\hat{q}_h = 0$), yields an estimate of consumer surplus. Assuming b_h is negative (implying a downward sloping demand curve), no value of c will result in zero occasions demanded since $f(c) \rightarrow 0$ as $c \rightarrow \infty$. Letting t equal such a value of c, consumer surplus for activity h, (CS_h), can be estimated as:

$$(7) \qquad CS_h = \lim_{t\to\infty} \int_0^t e^{b_h c}\,Q_h\,dc$$

which converges to $-Q_h/b_h$.

Using equation (7), consumer surplus was estimated twice for each state and activity with a pooled coefficient that differed statistically from the state coefficient; the b_h appropriate for that state and the pooled b_h were used in these calculations. Results are presented in Table 4. In all cases, the pooled coefficient provided a much larger estimate of consumer surplus than the state coefficient. The difference was nearly forty-fold for Georgia's small game hunting activity and twenty-fold or more for some activities in other states. These differences resulted from the coefficients for these states being larger in absolute value, indicating greater cost sensitivity, than the corresponding pooled coefficients. The opposite result would have held for any state coefficient with a smaller absolute value that differed significantly from the pooled coefficient. Thus, substituting aggregate coefficients for a state coefficient will bias the state surplus estimate in the same direction as the difference in the absolute value of these coefficients. In other words,

TABLE 4. Consumer Surplus Estimates Utilizing Significantly Different State and Pooled Cost Coefficients by Activity and State[a]

State	Coefficient	
	State	Pooled
	(dollars)	
Warm Water Fishing		
Georgia	339,535	1,294,169
Louisiana	422,695	1,456,076
West Virginia	75,921	486,212
Small Game Hunting		
Georgia	107,010	4,132,512
Louisiana	658,067	9,118,644
Tennessee	211,773	5,909,091
West Virginia	364,724	8,408,321
Big Game Hunting		
Georgia	85,466	1,253,425
Kentucky	27,557	494,521
Louisiana	329,323	2,587,671
Maryland	59,971	917,671
Mississippi	226,395	2,552,055
Virginia	123,615	2,186,302
Wildlife Enjoyment		
West Virginia	37,636	506,678

[a]All figures represent sample surpluses only; State estimates would require adjustment by an appropriate population expansion factor.

using a pooled coefficient results in a positive surplus bias if the larger population as a whole is less cost sensitive and a negative bias if the larger population is more cost sensitive.

To observe the impact of using coefficients from disaggregated samples to estimate regional surplus, consumer surplus was calculated for all states with both the state and pooled coefficients. These state estimates were then summed for each activity and are presented in Table 5. Results indicate that consumer surplus is higher for warm water fishing and wildlife enjoyment when estimated with the appropriate state coefficients, but lower for the other two activities. Fur-

TABLE 5. Consumer Surplus Estimates for the Southeastern United States, Utilizing State and Pooled Cost Coefficients by Activity[a]

Activity	Coefficient	
	State	Pooled
	(dollars)	
Warm Water Fishing	23,173,244	10,003,742
Small Game Hunting	48,787,445	56,867,488
Big Game Hunting	9,113,668	16,734,247
Wildlife Enjoyment	22,936,165	5,830,027

[a]All figures represent sample estimates.

thermore, the relative difference between the two regional estimates for each activity is less than the differences for the state estimates presented in Table 4. This smaller difference is not surprising since many of the state coefficients used in the regional estimates did not differ significantly from the pooled coefficients.

Since all pooled coefficients were significantly different from zero, the resulting consumer surplus estimates for the Southeastern region as a whole, presented in Table 5, should be more reliable than the sum of the estimates made with the individual state coefficients. In contrast, the state surplus estimates made from those state coefficients which statistically differed from the corresponding pooled coefficient and from zero, presented in Table 4, should be more reliable than estimates made from the pooled coefficients.

Conclusions

This study demonstrates that reducing region (population) size, or in an opposite sense, pooling sample data, can significantly affect the parameters of population-specific recreation demand models. Such differences can result in quite different consumer surplus estimates, which are often used as indicators of recreation values. The estimates of state consumer surplus with average cost coefficients from a state model differed considerably from estimates made with coefficients from a pooled model. These results indicated that use of recreational demand parameters in analysis for a smaller population than the sample on which the parameters were estimated can lead to severe biases. The analysis of regional surpluses was less conclusive. While differences in estimates from state and pooled coefficients were again obtained, the relative differences were less pronounced. However, the lack of statistical significance for many of the coefficients in the state models suggested that the pooled model is more appropriate for regional analysis. More research on the effects of different population definitions on demand parameters and consumer surplus estimates appears warranted, particularly for geographical areas other than the Southeast.

These results also point toward a much more significant conclusion about the overall validity and usefulness of population-specific models. Since reduced region (population) size can significantly alter final results, the same logic would imply that average cost coefficients could be significantly different when determined at particular recreation sites within a specified region. These coefficients could differ both among sites and from the region as a whole. Consumer surpluses for a particular site could therefore be biased if estimated with a population-specific rather than a site-specific model. Depending on their use, these results imply that site-specific models, which can be considered a reduction in region (population) size, may be less biased and therefore more appropriate than population-specific models.

These findings do not necessarily suggest that population-specific studies have no merit. In many situations, particularly those concerned with aggregate questions, recreation sites may not be well enough defined to implement site-specific studies. Even on a local level, surveying participants at all sites may be impossible; for example, surveying all recreation sites in a particular state, county, or other geographical region, may be methodologically difficult and prohibitively costly. In such cases, a population specific approach appears to be the only appropriate alternative. However, the researcher should be aware that the consumer surplus values from such studies may be biased. Such bias may be reduced if the analysis is disaggregated as much as possible. This paper has presented a method for determining whether such a bias exists within the sample population definition under consideration.

A fruitful area for future research appears to be explaining the differences in recreation demand functions found in this study. Some hypotheses on the sources of these differences are suggested by the theory of consumer behavior. The price and availability of substitute recreation activities, which has received limited attention in wildlife recreation studies, may be one reason. Another standard hypothesis is that systematic differences in preferences between populations may exist. Finally, the quality of recreational activity likely differs across populations due to differences in physical environment, access to recreational areas, and congestion of use. Methodology used by Gum and Martin (1977), Oliveira and Rausser, and Wetzstein and Green may be useful in testing these hypotheses. As factors causing regional differences in parameters are identified and accounted for in pooled recreation demand functions, the bias arising from applying such demand functions to smaller populations will likely be reduced.

References

Brown, William G., *et. al. An Economic Evaluation of the Oregon Salmon and Steelhead Sport Fishery.* Ore. Agr. Exp. St. Tech. Bul. 78, September, 1964.

Cicchetti, C. S., *et. al.* "Recreation Benefit Estimation and Forecasting: Implications of the Identification Problem." *Water Res. Res.* 8(1972): 840-850.

Clawson, Marion. *Methods of Measuring the Demand for and Value of Outdoor Recreation.* Resources for the Future, Inc., Washington, D.C., No. 10, February 1959.

Clawson, Marion, and J. L. Knetsch. *Economics of Outdoor Recreation.* Johns Hopkins Press, Baltimore, Md., 1966.

Edwards, J. A., *et. al. The Demand for Non-Unique Outdoor Recreational Services Methodological Issues.* Ore. Agr. Exp. St. Tech. Bul. 133, May, 1976.

Environmental Research Group. *A Guide to the Data Collecting and Processing System. Economic Evaluation of Wildlife and Wildlife Oriented Resources in the Southeastern U.S.* Georgia State University, Atlanta, 1972.

Gum, Russell L., and William E. Martin. "Problems and Solutions in Estimating the Demand for and Value of Rural Outdoor Recreation." *Amer. J. Agr. Econ.* 57(1975): 558-566.

Gum, Russell L., and William E. Martin. "Structure of Demand for Outdoor Recreation," *Land Econ.* 53(1977): 43-55.

Kalter, R. J., and L. E. Gosse. *Outdoor Recreation in New York State: Projections of Demand, Economic Value and Pricing Effects.* Cornell University, Ithaca, N.Y., Special Series, 1969.

Oliveira, Ronald A., and Gordon C. Rausser. "Daily Fluctuations in Campground Use: An Econometric Analysis." *Amer. J. Agr. Econ.* 59(1977): 283-293.

Outdoor Recreation Resources Review Commission. *Outdoor Recreation for America.* U.S. Government Printing Office, Washington, D.C., 1962.

Sawyer, Thomas G., and Robert N. Shulstad. *Economic Feasibility of Developing Additional Public Outdoor Recreation Areas at Beaver Lake, Arkansas.* Ark. Exp. St. Bul. 813, Nov., 1976.

U.S. Bureau of Outdoor Recreation. *Outdoor Recreation, Appendix "A" An Economic Analysis,* U.S. Government Printing Office, Washington, D.C., Dec. 1973.

Wetzstein, Michael E., and Richard D. Green. "Use of Principal Components Attractiveness Indexes in Recreation Demand Functions." *West. J. Agr. Econ.* 3(1978): 11-21.