Journal of Agricultural Economics, No. 42, December 1987. — 191 —

Estimation of Recreational Benefits from Individual Observations vs. Zonal Averages: Some Empirical Comparisons

Ching-Kai Hsiao*

(蕭 景 楷)

I. Introduction

The estimation of outdoor recreational benefits has been traditionally based upon average participation rates and travel costs for various distance zones (see, for example, Clawson, 1959; Knetsch, 1963; Brown, Singh, and Castle, 1964). According to Prais and Aitchison (1954), one reason to use grouped observations is to avoid the laborious numerical treatment of the individual observations, another reason is to keep the data confidential. Besides these two reasons given by Prais and Aitchison, one even more important reason to estimate the outdoor recreation demand from grouped observations is when it is difficult to obtain accurate measurements.

However, some researchers have suggested that substantial gains in efficiency in estimating outdoor recreational demand functions could be obtained by using individual observations instead of zone averages (Brown and Nawas, 1973; Gum and Martin, 1975). More recently, Brown et al. (1983) expressed concern about the use of unadjusted individual observations. They have argued that if individual observation are to be used, each observation on participation rates should be adjusted to a per capita basis. They stated that if the dependent variable is not adjusted to a per capita basis, then a biased estimate

* The author is an associate professor, Department of Agricultural Economics, National Chung Hsing University, Taichung, Taiwan, R.O.C.

of the travel cost coefficient will be obtained because the procedure would not properly account for the lower percentage of the people who come from more distant zones to participate in the outdoor recreational activity.

Nevertheless, several criticisms have been advanced concerning the preceding argument. One criticism was that there is no theoretical reason for expecting a decline in percentage participation from the more distant zones. Furthermore, no empirical analysis was presented to show a percentage participation decline from the more distant zones. Therefore, one objective of this study is to investigate possible reasons for expecting declining participation from more distant zones. A second objective is to estimate the impact of increasing distance upon the participation rates for ocean salmon sport fishing and to assess its effect upon consumer surplus estimates. A third objective is to show that estimates of demand based upon unadjusted individual observations can be used to properly compute consumer surplus if such estimates of demand are corrected by a separate distance and participation relationship. This correction procedure will be shown to be a reasonable alternative to the adjustment of the individual observations to a per capita basis or to the traditional zone average travel cost model.

II. Reasons for Declining Participation Rates

One possible reason to explain a declining percentage participation rates for the population of the more distant zones is as the following. If the individual demand functions in all distance zones were symmetrically distributed about some population mean demand function, then only those individuals with demand functions distributed above the population mean demand function would participate from those zones that were more than the average distance from the recreational site. Since the estimation of demand with unadjusted individual observations would not reflect the declining percentage of participants, an underestimate of the travel cost coefficient will be incurred.

To clarify and illustrate the preceding remarks, consider the following simple case. Suppose that the true individual demand functions for a given recreational activity is:

$$(1)\ q_i = 6 - 1.0\ TC_i + a_i$$

where a_i denotes a random "intensity" variable which represents the difference in intensity of preference among the various recreationists. If $E(a_i) = 0$, then the mean individual demand function would be $E(q_i) = 6 - 1.0\ TC_i$. (It is important to note that a_i is not an error term, but rather is a variable that denotes the difference in intensity or strength of demand for recreational activity among various individuals.) Suppose a_i is a discrete variable that takes certain values that are distributed symmetrically about zero with the following probabilities:

$$(2)\ E(a_i) = (1/16)\{1(-4)+4(-2)+6(0)+4(2)+1(4)\} = 0.$$

The variance of a_i would then be

$$(3)\ V(a_i) = E(a_i-0)^2 = (1/16)\{1(-4)^2+4(-2)^2+6(0)^2 +4(2)^2+1(4)^2\} = (1/16)\{64\} = 4$$

The distribution of a_i implied by (2) is the same as the distribution of the sum that could be obtained from flipping four unbiased coins where a tail would be assigned a value of −1.0 and a head a value of +1.0. Thus, the probability would be 1/16 of obtaining four tails equal to a sum of −4, and similarly for four heads giving a sum of +4. The probability of obtaining three tails and one head equal to a sum of -2 would be 4/16, with the same probability for three heads and one tail, giving a sum of +2. Finally, the probability of obtaining exactly two heads and two tails with zero sum is 6/16.

With the assumed true individual demand function of (1), then consider the simplest possible travel cost and distance zone data as shown in Table 1, generated from equations (1) and (2). For distance zone 1, the expected number of trips per recreationist is $E(q_i) = 6-1.0x1+0=5$, but some recreationists take more and some take less, depending upon their intensity of demand, that is their a_i value. For example, the first line of numbers in Table 1 corresponds to $a_i = -4$; hence, the number of visits per participant is equal to $q_1 = 6-1+(-4) = 1$. Note that the number of respondents is obtained based upon the binomial distribution. Since there is only one respondent for $a_i = -4$, the total number of visits would be 100 (i.e., the total number of visits are equal to q_i times number of respondents and then times the sample blow-up factor). For the second line of numbers in Table 1, corresponding to $a_i = -2$, $q_2 = 6-1+(-2) = 3$. The estimated total number of visits would be 3 times 4 times the expansion factor of 100 equals 1,200. The other numbers were generated in the same way.

It should be noted that zero trips would be assigned to those respondents who have negative or zero q_i. For example, the respondents in the first line of zone 2 with $a_i = -4$ would have $q_i = 6-4-4 = -2$, and zero trips would be indicated by such respondents. Thus, more respondents will take zero trips as the travel costs increase. (Actually, this is one of the driving forces behind the travel cost method.)

Fitting the data in Table 1 to the linear travel cost model based upon the unadjusted individual observation approach, the following OLS equation is obtained:

$$(4)\quad q_i = \underset{(14.42)}{5.5521} - \underset{(-7.50)}{0.5985}\, TC_{ij} \qquad n = 58 \qquad r^2 = 0.501$$

Table 1. Observations Generated for Three Distance Zones Where the True Individual Demand Functions Are Assumed to Be $q_i = 6-1.0TC_i + a_i$, $E(a_i) = (1/16)\{1(-4)+4(-2)+6(0)+4(2)+1(4)\}$

Main Distance Zone	Population	Intensity of Demand (a_i)	Average TC per visit	Total Visits per participant (q_i)	Number of Respondents	Total* # of visits	Zone Mean Visits per capita
		−4	$ 1	1	1	100	
		−2	1	3	4	1,200	
1	1,600	0	1	5	6	3,000	5.0
		2	1	7	4	2,800	
		4	1	9	1	900	
		−4	4	0	2	0	
		−2	4	0	8	0	
2	3,200	0	4	2	12	2,400	2.125
		2	4	4	8	3.200	
		4	4	6	2	1.200	
		−4	7	0	4	0	
		−2	7	0	16	0	
3	6,400	0	7	0	24	0	0.4375
		2	7	1	16	1,600	
		4	7	3	4	1,200	

*Assuming a random sampling of one percent from the general population and corresponding expansion factor of 100.

Computing the traditional consumer surplus from (4) for zone 1, an average consumer surplus per participating recreationist of about \$20.50 is obtained. Multiplying \$ 20.50 by the assumed 1,600 participating recreationists in zone 1 yields an estimated total consumer surplus for zone 1 yields an estimated total consumer surplus for zone 1 of about \$32,800. Following the same procedure for zone 2 gives \$8.33 x 2,200 = \$18,326, and \$1,55 x 2,000 = \$3,100 for zone 3. Thus, a total consumer surplus of about \$54,226 from the unadjusted individual observation approach is obtained.

Based upon the assumed true demand function, $q_i = 6 - 1.0TC_i + a_i$, the true individual consumer surplus can be computed. For the first line of numbers in Table 1, the true demand function is $q_1 = 2 - TC_i$, and it represents one actual participant. The consumer surplus is then equal to $(1)^2 / (2 \times 1) = 0.5$. (The formula for computing the consumer surplus from a linear travel cost demand function can be shown to be $CS_i = (q_i)^2/(2 \times \text{travel cost coefficient})$). Expanding this consumer surplus by multiplying the expansion factor of 100, a true consumer surplus of \$50 is obtained. Similarly, a total consumer surplus of \$1,800 for the second line is obtained by computing from $q_2 = 4 - TC_i$. Following the same procedure for the rest of the lines in Table 1, a total true consumer surplus of \$38,200 is obtained. Thus, the unadjusted individual observation approach overestimates the true consumer surplus by about 42 percent (54,226/38,200 = 1.42).

Of course, if distances and travel costs are such that the percentage participation rates for the more distant zones decline more rapidly than the case in Table 1, then an even greater overestimation of consumer surplus could result from the unadjusted individual observation approach.

It is interesting to compare the error in estimating consumer surplus from the unadjusted individual approach with that from the traditional zone average travel cost model.

Using the zone average visits per capita as the dependent variable, the following zone average travel cost estimate of the demand function is obtained:

$$(5)\ q_i = \underset{(10.38)}{5.5625} - \underset{(-6.66)}{0.7604}\, TC_i \qquad n = 3 \quad r^2 = 0.978.$$

Using (5), the traditional estimate of consumer surpluses for zone 1, zone 2, and zone 3, are \$24,256, \$13,376, and \$243, respectively. Thus, a total consumer surplus of about \$37,875 would be estimated by the zone average travel cost model, fairly close to the true consumer surplus of \$38,200. In other words, the error of estimation is only about one percent, much better than the 42 percent error in the unadjusted individual observation approach.

The results from Table 1 illustrates the fact that estimates of consumer surplus from the unadjusted individual observation approach can be substantially overestimated, when there is a significant decline in the percentage participation rates of the more distant zones. It should be noted that if one can assume that there is no variation in the intensity of demand among the participatns, i.e., $a_i = 0$ in (1), then the unadjusted observation approach will approximate the true demand function. However, this assumption may not be very realistic since one can always find great variation in the quantities taken by individuals with similar travel costs. Although the above example in Table 1 is enlightening, an empirical analysis by using actual data seems to be needed to obtain a better idea of the actual magnitude of bias that could be incurred by the unadjusted individual observation approach.

III. Effect of Distance Upon Participation Rate in Ocean Salmon Fishing

It is hypothesized that a larger percentage of the population would participate in ocean salmon sport fishing from nearby distance zones and a lower percentage of population would participate from more distant zones. Thus, the null hypothesis to be tested is that distance would have no effect upon the participation rate. Participation rate is defined as the proportion of the population who actually fished for ocean salmon during the survey year. This participation rate is then fitted by OLS regression as a function of measured round-trip distance from the zone of origin to the ocean port:

$$(6)\ \ln(y_{ij}) = -1.9818 - 0.003806\ DIST_{ij}$$
$$(-23.86)\quad(-12.60)$$
$$n = 211 \qquad r^2 = 0.432.$$

The participation rate is denoted by y_{ij} which is computed by multiplying the number of anglers in the ijth family observation by the sample blow up factor, then divided by its share of the population in its distance zone. $DIST_{ij}$ is the measured round-trip distance from the ith family observation's city of residence to the destination port j. In equation (6), the large negative value of $t = -12.60$ indicates that increasing distance had an extremely strong negative effect on the participation rate for ocean salmon fishing. It should be noted that there were no other independent variables that had a significant effect upon participation rate when $DIST_{ij}$ was included in the equation. For example, the income variable had an unexpected negative sign with $t = -1.09$, and the salmon-steelhead fishing equipment replacement value variable had a $t = 0.29$. No matter which explanatory variables were included in (6), only $DIST_{ij}$ remained significant with a stable coefficient and with absolute value of t always greater than 12. Thus, it is clear that measured distance is by far the most important factor affecting the participation rate.

One obvious way to cope with the problem of decreasing participation rates is to adjust the dependent variable by the sample blow up factor and shared population of each individual observation. (Brown et al., 1983). However, another way to solve this problem developed from correspondence between Professors William G. Brown and Kenneth E. McConnell. They concluded that the probability of whether or not to participate must also be considered if an unadjusted individual observation type of model is to be used. To scrutinize this idea, first denote the expected number of trips per capita, $TRPSCAP_{ij}$, as

$$(7)\ TRPSCAP_{ij} = (y_{ij} \times TRPS_{ij}) / NA_{ij}$$

were $y_{ij} = (NA_{ij} \times BLF_{ij}) / POP_{ij}$, denotes the number of anglers, NA_{ij} times the blow up factor BLF_{ij}, then divided by shared population POP_{ij}, i.e., the probability of participation as defined for equation (6); $TRPS_{ij}$ denotes the reported number of trips by the ijth respondent. Since both y_{ij} and $TRPS_{ij}$ are greatly affected by travel costs or distance, (7) can be further expressed as:

$$(8)\ TRPSCAP_{ij} = (y_{ij} \times TRPS_{ij}) / NA_{ij} = f(DIST_{ij}) \times g(DIST_{ij}).$$

Therefore, the computation of consumer surplus cannot validly be obtained by integrating $TRPS_{ij}$ only when y_{ij} is also a function of distance as shown in (8).

To see if a valid estimate of consumer surplus can be obtained based upon (7), the unadjusted individual data is fitted yielding the following linear equation:

$$(9)\ \ln(TRPS_{ij}) = \underset{(11.41)}{0.9288} - \underset{(-4.45)}{0.001318}\ DIST_{ij}$$

$$n = 211 \qquad r^2 = 0.087.$$

In (9), $DIST_{ij}$ denotes the two-way distance from the ith angler's city of residence to the jth port, while $TRPS_{ij}$ denotes the number of ocean salmon fishing trips reported by the ith respondent to the jth port. Thus, a valid estimate of consumer surplus can be obtained by integrating the product of equations (6) and (9), divided by the number of anglers:

$$(10)\quad TRPSCAP_{ij} = (y_{ij} \times TRPS_{ij}) / NA$$
$$= (0.34889e^{-0.005125}) / NA_{ij}.$$

Integrating equation (10) based upon the observed trips per capita, the Gum-Martin consumer surplus corresponding to each of the 211 individual observations was obtained. Multiplying each individual consumer surplus by its share of the population and summing all these values gave a total estimate of consumer surplus of about \$13.56 million. (An average reported travel cost of 27.5 cents per mile has been used here, based upon the following linearly homogeneous function:

$$(11)\quad RTC_{ij} = \underset{(30.63)}{0.2747}\, DIST_{ij} \qquad n = 211 \qquad r^2 = 0.817,$$

where RTC_{ij} is defined as revised travel costs.)

If the effect of distance upon the participation rates is ignored, i.e., if the y_{ij} term in (7) and (8) is deleted and the unadjusted individual observation approach is used to estimate consumer surplus, then a total Gum-Martin consumer surplus of \$52.71 million is obtained.

It is also interesting to compute and compare the estimates of consumer surplus from the traditional zone average travel cost model and the individual observations adjusted to a per capita approach. Fitting the average mumber of trips per capita per distance zone as a function of measured distance yields.

$$(12)\ \ln(TRPSCAP_{ij}) = -2.0280 - 0.005452\ DIST_{ij}$$
$$(-9.71) \qquad (-7.49)$$
$$n = 47 \qquad r^2 = 0.555.$$

Based upon (12), an average Gum-Martin consumer surplus of about $52 per trip and a total consumer surplus of about $13.16 million is obtained. (It should be noted that the cost per mile is about 28 cents for the zone average travel cost model.)

Consumer surplus can also be computed by using the adjusted individual observations approach. Fitting the individual trips per capita as a function of measured distance, the following equation is obtained,

$$(13)\ \ln(TRPSCAP_{ij}) = -2.2755 - 0.005161\ DIST_{ij}$$
$$(-20.01) \qquad (-12.48)$$
$$n = 211 \qquad r^2 = 0.427.$$

Based upon (13), a total Gum-Martin consumer surplus of $13.46 million is estimated, with a corresponding estimate of $53 per trip.

These four different estimates of consumer surplus are shown in Table 2. The last three estimates in Table 2 are very close to each other, which is an interesting result considering that different definitions and equations were used for those three estimates. By contrast, however, the unadjusted individual observation approach gave an estimate of about four times that of the other three Gum-Martin consumer surplus estimates. The comparison of various estimates of consumer surplus as shown in Table 2 thus has several implications. First, the estimate of consumer surplus based solely upon the unadjusted individual observations approach cannot be considered reliable. Second, if the unadjusted individual observation approach should be used, then the probability

Table 2. Estimated Consumer Surplus to Ocean Salmon Sport Anglers of Oregon, Based Upon Four Different Methods of Estimation

Method of Estimation	(1) Estimated Total Gum-Martin Consumer Surplus	(2) Estimated Total Traditional Consumer Surplus	(3) Estimatd Consumer Surplus Per Trip from(1)
	$ millions	$ millions	$
(UIO) Unadjusted Individual Observations Approach, Equation(9)	52.71	38.48	208
UIO Approach with Probability of Participation Rates Included, Equation(10)	13.56	13.52	54
Traditional Zone Average Travel Cost Model, Equation(12)	13.16	14.50	52
Individual Observations Adjusted to a Per Capita Basis, Equation(13)	13.46	12.35	53

of participation must be linked with it to estimate a valid consumer surplus. Third, using individual observations can lead to incorrect consumer surplus estimates unless they are adjusted to a per capita basis, just as for the zone average travel cost model. However, if there are equal percentages of participation from all distance zones, then individual observations would not need to be adjusted.

IV. Summary and Conclusions

Four different methods were used to estimate consumer surplus from ocean salmon sport fishing in Oregon. Two of the methods, individual observations adjusted to a per capita basis and the traditional zone average travel cost model, gave very similar estimates, as shown in Tables 2. A third method, the unadjusted individual observation (UIO) approach with probability of participation rates included, also gave quite similar results. However, the unadjusted individual observation method that has often been used in the past resulted in estimates of consumer surplus for ocean salmon sport fishing that were approximately four times higher than the estimates from the other three methods.

Although the "true" consumer surpluses are obviously unknown for ocean salmon sport fishing in Oregon, the estimates from the UIO method appear to be far too high when compared to the other three methods, especially considering the logical failings of the UIO approach noted in the earlier part of this paper. However, it cannot be implied that estimates based upon unadjusted individual observations would always be this much too high. In fact, the degree of overestimation of consumer surplus by the simple UIO approach is directly related to the impact of distance upon the per capita participation rate. The more sharply that the participation rate declines with increasing distance, the greater would be the overestimation expected from the UIO approach. Given the possibly large upward bias in consumer surplus estimates from earlier studies based upon the UIO approach, it would seem that these earlier UIO demand functions should be reestimated using the zone average travel cost model or the individual observations adjusted to a per capita basis.

With regard to the accuracy of estimation by the traditional zone average travel cost method versus the individual observations adjusted to a per capita basis versus the approach with probability of participation rates, small differences were observed. However, it needs to be kept in mind that if there were measurement error in the explanatory variable, such as for reported travel costs, then the traditional zone average travel cost model would be less biased than the models based upon individual observations because averaging the observations is one way to greatly reduce measurement error bias.

Aside from measurement error considerations, it is not clear from this study which of the last three methods of Tables 2 would ususally be the most accurate since all three are in relatively close agreement. An extensive set of Monte Carlo experiments would seem to be needed to determine which of the three methods would yield the most accurate estimates of consumer surplus under various specified conditions.

References

1. Bowes, M. D., and J. B. Loomis, "A Note on the Use of Travel Cost Models with Unequal Zonal Populations", Land Economics, 56(1980): 465-470.
2. Brown, W. G., C. Sorhus, B. Chou-Yang, and J. A. Richards, "Using Individual Observations to Estimate Recreation Demand Functions: A Caution", American Journal of Agricultural Economics, 65(1983): 154-157.
3. Brown, W. G., and F. Nawas, "Impact of Aggregation on the Estimation of Outdoor Recreation Demand Func-

tions", American Journal of Agricultural Economics, 55(1973):246-249.

4. Brown, W. G., A. Singh, and E. N. Castle, An Economic Evaluation of the Oregon Salmon and Steelhead Fishery, Oregon Agricultural Experiment Station Technical Bulletin 78, Corvallis, Sept. 1964.
5. Burt, O. R., and D. Brewer, "Estimation of Net Social Benefits from Outdoor Recreation", Econometrica, 39(1971): 813–827.
6. Clawson, M., Methods for Measuring the Demand for and Value of Outdoor Recreation. Reprint No. 10, Resources for the Future, Inc., Washington, D. C., 1959.
7. Gum, R. L., and W. E. Martin, "Problems and Solutions in Estimating the Demand for and Value of Rural Outdoor Recreation", American Journal of Agricultural Economics, 57(1975): 558-566.
8. Hsiao, C. K., An Evaluation of Alternative Estimates of Demand for and Benefits from Oregon Salmon Sport Fishing, Ph.D. Thesis, Oregon State University, Kerr Library, Corvallis, 1985.
9. Knetsch, J. L., "Outdoor Recreation Demand and Benefits", Land Economics, 39(1963): 387-396.
10. Prais, S. J., and J. Aitchison, "The Grouping of Observations in Regression Analysis", Review International Statistical Institute, 22(1954): 1-22.

Estimation of Recreational Benefits from Individual Observations vs. Zonal Averages: Some Empirical Comparisons

利用個別及分組平均資料推估娛樂效益：一些實證比較

Ching - Kai Hsiao
蕭 景 楷

摘 要

本文利用美國奧勒岡州鮭魚的娛樂性海釣活動資料，來實證旅行成本法中四種估算消費者剩餘的方法。結果顯示，其中三種方法所估算的值相當接近，只有未經調整的個別資料方法所估算的值，約為其他三種方法的四倍左右。這種高估的原因，主要是在模型中，沒有包括旅行距離對於參與活動率的影響。所以應用旅行成本法來估算效益時，如果要得到比較“可靠”的結果，除了採用傳統的區域平均資料，或是將個別資料調整為每人平均外，也可以在未經調整的個別資料方法中，考慮參與活動率的機率。