The Sensitivity of Travel Cost Estimates of Recreation Demand to the Functional Form and Definition of Origin Zones

Ronald J. Sutherland

The travel-cost method of estimating a recreation demand function requires specifying the functional form of the first-stage demand curve and defining the width of the concentric origin zones. A Monte Carlo approach is used to determine the sensitivity of demand and valuation estimates to alternative choices about these two issues. Demand and valuation estimates are shown to be sensitive to the definition of the origin zone and to the use of a semilog versus a double log first-stage demand curve. The proper choice or origin zones is unclear, but a semilog form is more appropriate than a double log form.

Since Clawson's paper in 1959, the travel cost method (TCM) of obtaining a recreation demand curve has been used frequently to estimate demand and value of a recreation site. Despite widespread acceptance and the official sanction of the Water Resources Council, the TCM has been subject to numerous criticisms, for example, the time bias and the identification problem. The implication of these criticisms is that the TCM is not sufficiently rigorous and comprehensive to produce reliable demand and variation estimates.

The focus of this study is on two specification choices required by the TCM which may influence estimates of the demand curve and consumers' surplus. The issues investigated here are (1) the functional form of the first-stage demand curve; and (2) the width of the concentric origin zones. The objective is to determine the sensitivity of travel cost demand and valuation estimates to various assumptions concerning these points. The method of analysis is to apply the TCM to several sites under various assumptions and to contrast the results.

Most empirical demand curves in the economics literature are specified in double log form, perhaps because the coefficients may be interpreted as elasticities.[1] In the recreation literature the semilog specification is most prevalent, although linear functions have been used.[2] The relative merit of the semilog and double log specification of the first-stage demand curve and the sensitivity of the valuation estimates to the choice of these two functional forms are considered in this study.

In the TCM, visit rates from various origins are regressed against corresponding travel costs. Since the pioneering work of

Ronald J. Sutherland is currently with the Los Alamos National Laboratory. The research for this paper was conducted while the author was with the Environmental Protection Agency, Corvallis Environmental Research Laboratory.

[1] In their literature surveys on the demand for money, Laidler and Goldfeld present empirical evidence in favor of a double log specification.

[2] Linear demand curves were used by Burt and Brewer and by Cicchetti, Fisher and Smith because this specification is required by some properties of their models.

Clawson and Knetsch, origins have been defined by a series of concentric rings around the recreation site. For instance, if recreationists travel a maximum of 200 miles and rings are defined every 20 miles, then there are 10 origin zones and 10 observations for the experience demand schedule. Similarly, if a ring is defined every 10 miles, there will be 20 travel zones and 20 observations for estimating the visit rate schedule. As an alternative to a system of concentric rings, each population centroid may be construed as a separate origin, and the number of observations is therefore determined by the number of such centroids. The sensitivity of the demand and valuation estimates to the definition of the origin zone is also examined in this study.

An Overview of a Regional Travel Cost Model

The simulation methodology used in this study requires the capability to estimate a travel cost demand curve for a large number of recreation sites. In two recent papers (1981, 1982), I present a regional recreation demand and benefits model, which is designed to estimate demand and valuation for each of 179 centroids. The model uses the TCM to analyze camping, fishing, boating and swimming in the Pacific Northwest. The essence of the approach is that household recreation surveys are used to estimate visitor days emanating from each origin zone, and a gravity model is used to estimate visits by origin to each recreation centroid in the region. The estimates of visit rates by origin and the corresponding travel costs are used to construct travel cost demand and valuation estimates.

Only a brief overview of the model is presented here. A gravity model is specified in the form

$$T_{ij} = P_i \frac{A_j F_{ij}}{\sum_j A_j F_{ij}}, \quad (1)$$

where

T_{ij} = number of activity days produced at origin i and attracted to destination j

P_i = number of activity days produced at i

A_j = number of activity days attracted to the jth recreation centroid

F_{ij} = a calibration term reflecting spatial impedance for interchange ij.

The gravity model, as specified in (1), is a distribution model; that is, it distributes a given number of trips according to relative attractiveness and relative effect of spatial impedance of each origin and destination centroid. As employed in the regional recreation model, the gravity model estimates the number of visits to each of 179 recreation destinations from each of 144 origins in the system. Total visitor days weighted by corresponding populations is the visit rate, which is the critical input in the TCM.

Travel costs are estimated as the round trip cost per person per mile times the minimum driving distance.[3] Population centroids define trip origin points and at least one population centroid is defined for each county. Recreation centroids are the center of recreation activity in the county, and, where appropriate, multiple recreation centroids are defined for each county.[4] Each origin and destination centroid was identified and a highway network was constructed showing various possible routes from the origins to the destinations. The minimum distance routes were estimated via a computer program, and these distances form an impedance matrix. This

[3]Travel costs are defined here as pecuniary costs and no allowance is made for the time value of travel. Using data from a recent household recreation survey in the Northwest, I show — in a paper in progress — that recreationists do not consider time spent for recreation travel as a cost.

[4]There are 119 counties in the three states of Washington, Oregon and Idaho, but 179 recreation centroids and 132 internal population centroids and 12 external zones. A list of the population and recreation centroids is given in Sutherland (1981).

matrix is used to estimate travel costs and to calibrate the gravity model.

Although the regional model considers four activities, this study will consider only boating. Estimates of the number of boating trips by origin were obtained from three household recreation surveys taken in 1976 by the state parks authorities in Washington, Oregon and Idaho. These surveys also include information on travel distances by activity. Distance data, along with the corresponding number of trips, were used to estimate a relative trip length frequency distribution for boating. The F_{ij} values in (1) are obtained by substituting the impedance (travel distance) values into the trip length frequency distribution. The F_{ij} values can be interpreted as the probability that a boater will travel the distance from origin i to destination j.

The boating attractiveness of each recreation centroid is hypothesized to be a multiplicative function of the number of boat ramps and the accessibility of the centroid. Boat ramps are a proxy for a combination of factors — for example, water quality, acres of water — which determine the attractiveness of a boating site. Accessibility (RA_j), or total potential demand, is estimated as the sum of boating visitor days from each population centroid in the region weighted by the probability that a boater will drive to the particular centroid. That is, the accessibility of the jth centroid is

$$\text{(2)} \qquad RA_j = \sum_{i}^{132} F_{ij}P_i ,$$

where 132 internal origins have been defined for Washington, Oregon and Idaho. Recreation accessibility varies positively with the nearness of population centers and with the number of boating days produced by these origins.

U.S. Forest Service data on visitor days and boat ramps (BR_j) by ranger district were used with accessibility estimates of (2) to estimate an attractions model.[5] The regression results are

$$\text{(3)} \qquad A_j = 0.001\, \underset{(4.33)}{BR_j^{0.60}}\; \underset{(4.79)}{RA_j^{1.74}} ,$$

$$R^2 = 0.52, \qquad n = 37$$

where the t values (in parentheses) indicate that each coefficient is highly significant. Boat ramp and accessibility data for each recreation centroid in the region were substituted into (3) to estimate the relative attractions.

Travel cost demand schedules are estimated by first regressing visit rates by origin (V_i) against travel cost (C_i) per person per visitor day (5.52 cents per mile).[6] In general form, the experience or first-stage demand equation is

$$\text{(4)} \qquad \hat{V}_i = f(C_i),$$

where the hat means predicted and i refers to the origin zone. Equation (4) is estimated for each recreation centroid in the region with ordinary least squares and is used to analyze the issues of functional form and definition of origin zone. Multiplying (4) by the population of origin i and summing gives

$$\text{(5)} \qquad \sum_i T_i = \sum_i N_i\hat{V}_i ,$$

which estimates total visits as a function of the estimated visit rate of each origin zone

[5]These data are unpublished and are part of the Recreation Information Management (RIM) system of the U.S. Forest Service. The data were obtained from the regional office in Portland and reflect ranger districts in Washington and Oregon.

[6]According to the 1979 edition of *Principles and Standards,* of the Water Resources Council, the cost per vehicle mile for a standard vehicle was 8.4 cents in 1976. Using the U.S. Department of Commerce gas and oil price deflator, this figure was adjusted to 9.66 cents per mile in 1979. Next, 9.66 cents was doubled to adjust for round trip costs and divided by the average number of persons per vehicle (3.5) to obtain 5.52 cents.

and its corresponding population. Total quantity demanded can also be estimated exogenously by site attendance data, which in this study is the sum of the appropriate column vector in the T_{ij} matrix produced by the gravity model. Although the regression estimate of (4) estimates visit rates with an average error of zero, total visits are not necessarily estimated correctly, that is, $N_iV_i \neq N_i\hat{V}_i$.

The effect of various hypothetical prices on total quantity demanded is estimated via

$$(6) \qquad T_i \neq N_i f(C_i + \Delta P).$$

The quantities obtained from (6) and the corresponding price increments are used to estimate a recreation site demand curve. Instead of using regression analysis to estimate a sign demand curve, a fourth degree polynomial is fit to every five consecutive price-quantity observations and Bode's Rule is used to measure the area under the polynomial.[7] The integral under this polynomial estimates consumers' surplus.

The site visit rate data obtained here via the gravity model would not correspond identically to data obtained from attendance surveys. However, the gravity model is calibrated so that the trip length frequency distribution formed from the trip interchange matrix (T_{ij}) corresponds closely to the distribution estimated from household survey data. The travel cost estimates presented here should, on the average, be representative of those obtained using site attendance data or household survey data. For purposes of a simulation analysis, the visit rate and travel cost data may be considered exact.

Sensitivity of Travel Cost Estimates to Various Assumptions

Since the 179 recreation centroids and 4 activities included in the regional model are more than sufficient for this analysis, I arbitrarily consider the demand for boating at 20 Washington recreation centroids, numbered 17.0 to 26.0 (column 1 in the accompanying tables). These centroids include those in King County, which contains Seattle and is heavily populated, as well as sparsely populated counties east of the Cascade Mountains. By including both urban and rural counties in the sample, the travel cost estimates reflect a diversity of realistic conditions. The rationale for sampling a relatively large number of centroids (20) is that certain adverse consequences may be observed only occasionally, and a large sample increases the likelihood of such a result. Also, results based on a single site may reflect a special case, inconsistent with results obtained over a wide range of experience.

The sensitivity of travel cost estimates to each of the two issues being considered depends upon the assumption made on the other issue. The interdependence of these issues precludes analyzing them individually. The functional form of the first stage demand curve is considered first by focusing on the semilog form and the double log form. Travel cost estimates will then be presented using various size origin zones. The results are shown to be sensitive to the definition of origin zone, and this sensitivity in turn depends on the functional form.

Functional Form of the First-Stage Demand Curve

The proper form of a recreation demand curve has been studied by Ziemer *et al.* and by Smith. The studies are similar in that only one site was considered and a statistical analysis, namely a Box-Cox transformation, was used to statistically estimate the most appropriate functional form. Smith rejected the linear form because it provided a poorer fit of the data than the double log and semilog form. However, Smith also concluded that even though the latter two forms fit the data and provided reasonable results, each form must be considered inappropriate. Ziemer *et*

[7]Bode's Rule is given in Davis and Rabinowitz (p.30), in Abramowitz and Stegun (p. 886), and discussed in Sutherland (1981).

al. used the Box-Cox transformation procedure and concluded that a semilog form is appropriate and a linear form is inappropriate, and further that consumers' surplus estimates are highly sensitive to the choice of functional form.

In considering the various functional forms, double log and semilog (logarithm of the dependent variable) are candidates, but the linear form need not be considered. Ziemer *et al.* and Smith provide evidence against the linear form, and scatter plots of several sites indicate a distinct curvilinear relationship. The evidence against the appropriateness of the linear form is persuasive, and in this study we consider the double log and semilog functional form.

Analyzing these two forms determines if the results are sensitive to the choice of functional form, and if so, which of the two forms seems most appropriate. Four criteria are suggested which may be useful in identifying the most appropriate form. The coefficients of determination are a relevant but not decisive indicator, particularly if estimated over several sites. Secondly, estimates of consumers' surplus per trip should be somewhat stable across sites and should be similar to those reported elsewhere in the literature. Thirdly, the first-stage demand curve should estimate closely known quantity demanded at a zero price when $P = 0$ is used in equation (6). Finally, goodness of fit and consumers' surplus estimates should be insensitive to other computational decisions, particularly if the decisions are made arbitrarily. These properties are not asserted to be rigorous statistical criteria that will necessarily determine the unambiguous superiority of one functional form. Since previous studies have not been able to resolve this issue on statistical or theoretical grounds, it is appropriate to employ a Monte Carlo analysis, where a demand curve for several sites is estimated with each functional form and the results are compared.

First-stage demand curves for boating, equation (4), are estimated for 20 centroids using both double log and semilog forms, where the logarithm is taken of the dependent variable. These estimates are based on population centroids as origin zones and quantity demanded estimated exogeneously. The results are presented in Table 1. The coefficients of determination, columns (4) and (7), indicate that each form fits the data reasonably well, but the semilog model has more explanatory power in 19 of the 20 cases. The semilog surplus per day estimates are more stable than the corresponding double log estimates. Dwyer, Kelley and Bowes review several empirical studies of recreation behavior, but only a few of these studies deal specifically with boating. If we presume that other water-based activities have a value comparable to boating or that boating is typical of outdoor recreation in general, we may conjecture on the basis of Dwyer *et al.* that value per day estimates below \$1 or above \$10 are outside the range of most existing studies. The double log estimate of surplus per day of \$68.95 for centroid 24.3 is clearly untenable, and the double log surplus per day estimates of \$10.85 and \$11.05 appear suspiciously high.

A few of the surplus per day estimates, such as those for centroids 18.0 are highly sensitive to this choice. This result exposes the inadequacy of analyzing the issue of functional form by considering only one site. The results in Table 1 do not establish that either form is correct, but the consistently lower explanatory power of the double log form and the wide variation in surplus per day estimates cast some doubt about the appropriateness of this form.

Size of Origin Zone

When the travel cost method was presented by Clawson and by Clawson and Knetsch, origins were aggregated into zones defined by a series of concentric circles. To my knowledge there has been no serious analysis of the appropriate size of these origin zones nor of the sensitivity of the results to

TABLE 1. Annual Valuation Estimates for Boating in Selected Washington Centroids Using a Semilog and Double Log Functional Form.

Recreation Centroid Number (1)	County (2)	Recreation Centroid (3)	Semilog Results* R^2 (4)	Semilog Results* Consumers' Surplus (in $000) (5)	Semilog Results* Surplus per Day $ (6)	Double Log Results R^2 (7)	Double Log Results Consumers' Surplus (in $000) (8)	Double Log Results Surplus per Day $ (9)
17.0	King	Snoqualm	0.851	$2,163	$6.22	0.645	$2,757	$7.94
17.1	King	Lake Sammamish	0.863	5,239	5.48	0.605	5,577	5.83
17.2	King	Lake Washington	0.873	5,596	5.12	0.666	11,564	10.58
18.0	Kitsap	Horseshoe Lake	0.885	283	4.50	0.724	334	5.31
19.0	Kittitas	Wawapum State Park	0.740	141	4.28	0.660	142	4.78
19.1	Kittitas	Lake Kachees	0.829	536	6.15	0.686	620	7.11
20.0	Klickitat	Horseshoe Lake	0.727	548	5.20	0.582	526	5.00
21.0	Lewis	Ike Kinswa State Park	0.888	1,008	5.83	0.702	886	5.13
22.0	Lincoln	Grand Coulee Dam	0.766	15	2.57	0.744	18	3.18
22.1	Lincoln	Fort Spokane	0.751	27	2.56	0.729	58	5.57
22.2	Lincoln	Sprague Lake	0.673	34	2.32	0.699	69	4.79
23.0	Mason	Lake Cushman	0.884	228	5.44	0.695	220	5.24
23.1	Mason	Belfair	0.884	231	5.08	0.691	331	7.26
23.2	Mason	Dash Point State Park	0.874	1,253	5.78	0.636	1,975	9.12
24.0	Okanogan	Pearrygin Lake	0.841	19	2.67	0.781	35	4.97
24.1	Okanogan	Conconolly State Park	0.810	31	2.55	0.740	136	11.05
24.2	Okanogan	Alta Lake State Park	0.805	39	2.72	0.747	60	4.11
24.3	Okanogan	Osoyoos Lake State Park	0.812	67	2.87	0.747	1,606	68.95
25.0	Pacific	Fort Stevens State Park	0.874	205	4.89	0.732	168	4.00
26.0	Pend Oreille	Skookum Lakes	0.785	52	2.42	0.724	109	5.17
		column mean	0.820	$886	$4.24	0.696	$1,360	$9.23

*In the semilog form the logarithm is taken of the dependent variable.

various size zones.[8] The above results use each population centroid in the region as a potential origin zone. Evidence on the sensitivity of travel cost demand and valuation to the definition of the origin zone is obtained by comparing the above results to those obtained using 10-mile and 20-mile origin zones. Consider two systems of concentric circles, one at 10-mile and one at 20-mile intervals from the recreation centroid. Origin zones are now defined as the area between each ring; visit rates, as total trips from each zone per 1,000 population of the zone. The travel cost from each zone is the weighted average travel cost of all centroids within the zone where the weights are the number of trips per centroid.

Travel cost valuation estimates using a semilog form and 10- and 20-mile origin zones are presented in Table 2. Comparing the results using a 10-mile zone with those of a 20-mile origin shows similar estimates for several sites but quite dissimilar estimates for others. The most important result in Table 2 is the instability of consumers' surplus per trip estimates using both 10- and 20-mile origin zones. Columns (4) and (7) indicate that surplus per trip estimates range from just over $1 to above $20. An estimation procedure which occasionally gives unstable results cannot be relied upon when analyzing a single site.

The estimates in Table 2 are comparable to the semilog results in Table 1, with the difference being that population centroids are used as origins in the analysis reported in Table 1. A comparison of the results on these two tables indicates that aggregating population centroids into concentric zones increases consumers' surplus by an average of over $1 per trip. Furthermore, consumer's surplus estimates on Table 1 appear uncorrelated with those on Table 2. Estimates of total surplus for centroids 17.1 and 17.2 are over $1 million lower when population centroids are aggregated into zones. However, the aggregation process increases the surplus estimates per trip for centroids 22.0 and 22.2 by over 300 percent. The surplus per trip estimates for these two recreation centroids exceed $20, and the coefficients of determination cast doubt on the reliability of these estimates. The results for these two centroids may be regarded as outliers and therefore dismissed, but it is significant that aggregating population centroids into zones produces outliers while use of population centroids as origins did not.

The conclusion that travel cost valuation estimates are sensitive to the definition of the origin zone is significant, and raises the question of which definition is most appropriate. The average of the coefficients of determination favor the use of population centroids as origin zones, but the differences in R^2 values between models do not provide sufficient evidence to resolve this issue. The two extreme estimates (centroid 22.0 and 22.2) obtained from the 10- and 20-mile origin zone equations raise a question about aggregating, but are not compelling evidence against it. A third potential indicator of the proper model is the ability of the statistical estimate of equation (6) to estimate known quantity demanded at a zero price.

Table 3 depicts the assumed known quantities demanded, column (2), and endogenous estimates of this variable using 10- and 20-mile origin zones and recreation centroids as origin zones. This table also depicts quantity estimates using semilog and double log functional forms. Regardless of the choice of the origin zone, the semilog form predicts total quantity demanded more accurately than the double log form. This result is further evidence in favor of the semilog form over the double log form. An additional result is that aggregating population centroids into either 10- or 20-mile zones substantially improves the ability of the model to predict total use at a zero price. Although aggregat-

[8]Brown and Newas have argued that observations should be based on individuals, rather than aggregations of people. Since they use site attendance data, visit rates reflect the frequency of participation of recreators and of the participation rate of the entire population, which may also decline with distance from the site.

TABLE 2. Semilog Valuation Estimates Using 10-Mile and 20-Mile Origin Zones.

	10 Mile Origin Zones			20 Mile Origin Zones		
Recreation Centroid Number (1)	R^2 (2)	Consumers' Surplus (in $000) (3)	Surplus Per Trip (4)	R^2 (5)	Consumers' Surplus (in $000) (6)	Surplus Per Trip (7)
17.0	0.675	$1,394	$4.01	0.787	$1,415	$4.07
17.1	0.764	3,463	3.62	0.749	3,606	3.77
17.2	0.768	3,775	3.45	0.868	2,830	2.59
18.0	0.935	111	1.77	0.977	91	1.44
19.0	0.631	82	2.47	0.668	81	2.45
19.1	0.547	444	5.10	0.470	587	6.73
20.0	0.619	394	3.75	0.751	337	3.21
21.0	0.824	767	4.44	0.827	882	5.11
22.0	0.258	117	20.52	0.216	217	38.12
22.1	0.620	60	5.70	0.716	44	4.21
22.2	0.170	361	24.87	0.474	313	21.58
23.0	0.882	118	2.82	0.948	114	2.73
23.1	0.895	120	2.62	0.928	109	2.39
23.2	0.815	890	4.11	0.913	868	4.01
24.0	0.916	9	1.36	0.919	7	0.97
24.1	0.903	53	4.31	0.903	44	3.56
24.2	0.835	29	2.02	0.877	23	1.61
24.2	0.830	139	5.97	0.806	106	4.56
25.0	0.840	98	2.35	0.915	75	1.79
26.0	0.699	102	4.80	0.698	80	3.76
Mean	0.720	$626	$5.50	0.771	$591	$5.93

*These estimates are based on a $1 price increment in equation (6), and quantity demand estimated exogeneously.

ing populations into zones improves the predictive ability of the model in this sense, the quantity estimates for several centroids still contain substantial errors.

The result — that aggregating population centroids into concentric zones does not improve the R^2 values as we may expect, but does improve the estimates of total quantity demanded — is easily explained. Visit rates diminish with distance from the site, but the number of population centroids increases with distance from the site. When population centroids are used as origins, there are many observations of low visit rates which are close to the regression line. The very few origin zones which have high visit rates and account for most of the total visits have relatively little influence on the regression line. The visit rates of the close origin zones are often estimated with large residuals. Aggregation results in a large number of good fitting observations being combined into a few observations and hence reduces their influence on R^2.

Aggregation decreases the total number of observations and thereby increases the relative weight of the close origins in determining the regression line. The error in estimating these visit rates thereby decreases; with it, the error in estimating total visits. The "solution" to the visit estimation problem is not obtained by aggregating because aggregating from a 10-mile origin zone to a 20-mile origin zone actually decreases the reliability of predicting total visits, see Table 3, columns (4) and (5). Indeed, total visits could

TABLE 3. Estimates of Quantity Demanded by Centroid Using Semilog and Double Log Forms and Various Definitions of Origin Zones (in Thousands of Visitor Days).

		Semilog Results			Double Log Results		
Recreation Centroid Number (1)	Exogeneous Quantity Demanded (2)	Recreation Centroids (3)	Ten Mile Origin Zone (4)	Twenty Mile Origin Zone (5)	Recreation Centroids (6)	Ten Mile Origin Zone (7)	Twenty Mile Origin Zone (8)
17.0	347	534	352	313	3,751	492	540
17.1	956	1,270	912	829	17,657	1,357	1,381
17.2	1,092	1,354	1,009	833	111,807	1,656	1,333
18.0	63	70	47	34	572	56	44
19.0	33	36	32	27	122	45	59
19.1	87	134	97	87	475	134	155
20.0	105	135	114	98	882	261	240
21.0	113	244	148	145	800	243	337
22.0	6	4	18	24	15	34	51
22.1	10	8	18	14	130	39	48
22.2	15	9	43	40	102	101	113
23.0	42	57	37	31	389	48	51
23.1	45	59	39	33	1,279	55	50
23.2	217	303	206	185	10,549	345	325
24.0	7	5	7	4	52	6	4
24.1	12	9	26	21	1,223	20	18
24.2	14	10	16	13	84	15	12
24.3	23	19	61	43	7,081	68	43
25.0	42	49	30	24	91	37	34
26.0	21	17	32	26	1,015	106	116
Mean	166	216	171	141	8,234	256	248

NOTE: The quantity estimates in columns (3) through (8) are obtained by letting $\Delta P = 0$ in the appropriate least squares estimate of equation (6).

be predicted exactly if populations were of constant size across origins.[9]

The composite influence of a double log specification and 10- and 20-mile origin zones on valuation estimates is depicted in Table 4. The coefficients of determination, columns (2) and (5), are lower for a double log model than for a semilog model (Table 2) when population centroids are aggregated into 10- or 20-mile zones. Furthermore, most of the surplus per day estimates are higher than one could reasonably expect, and they are very unstable across sites. Overall, the use of a double log form together with 10- or 20-mile concentric origin zones results in very untenable valuation estimates. This result also follows when we consider the double log estimates of total use at a zero price. As seen in Table 3, aggregating population centroids into 10- or 20-mile origin zones improved the predictability of the model in terms of total use. However the double log model predicts total use with a larger error than a semilog model, regardless of the choice of origin zone. Overall, the double log estimates are much more sensitive to the definition of the origin zone than are semilog estimates.

[9]The estimated residuals in predicting visit rates necessarily sum to zero, i.e.,

$$\Sigma(V_i - \hat{V}_i) = \Sigma(T_i/N_i - \hat{T}_i/N_i) = 0.$$

If the population of each origin is identical, $N\Sigma(T_i - \hat{T}_i) = 0$, visits ($T_i$) are also predicted exactly. Bowes and Loomis (1980) show that specifying populations in square root form results in a first-stage demand curve which necessarily predicts total visits identically.

TABLE 4. Double Log Valuation Estimates Using 10-Mile and 20-Mile Origin Zones.

	10 Mile Origin Zones			20 Mile Origin Zones		
Recreation Centroid Number (1)	R^2 (2)	Consumers' Surplus (in $000) (3)	Surplus Per Trip (4)	R^2 (5)	Consumers' Surplus (in $000) (6)	Surplus Per Trip (7)
17.0	0.556	$8,810	$25.36	0.596	$18,546	$53.39
17.1	0.599	34,694	36.29	0.588	38,203	39.96
17.2	0.608	36,984	33.84	0.673	19,616	17.95
18.0	0.803	405	6.44	0.804	770	12.25
19.0	0.493	500	15.10	0.454	1,631	49.38
19.1	0.536	2,312	26.54	0.434	5,044	57.90
20.0	0.535	8,386	79.72	0.555	7,732	73.51
21.0	0.723	7,171	41.52	0.668	15,291	88.54
22.0	0.366	1,050	184.24	0.283	2,114	371.10
22.1	0.491	906	87.75	0.469	1,658	158.72
22.2	0.175	4,326	298.29	0.399	5,345	368.54
23.0	0.775	902	21.52	0.878	1,445	34.46
23.1	0.844	926	20.31	0.903	1,150	25.25
23.2	0.736	9,984	46.10	0.855	9,324	43.06
24.0	0.797	15	2.04	0.776	15	2.15
24.1	0.813	180	14.69	0.792	227	18.49
24.2	0.782	62	4.26	0.822	60	4.15
24.3	0.711	238	10.23	0.670	230	9.92
25.0	0.764	505	12.05	0.772	561	13.39
26.0	0.461	3,242	153.28	0.459	4,083	193.05
Mean	0.627	$6,080	$55.93	0.643	$6,653	$81.76

Conclusions and Implications

This study presents travel cost demand and valuation estimates for boating in 20 recreation centroids in Washington. The objective of the analysis is to determine the sensitivity of the results to the functional form of the first-stage demand curve and to the definition of the origin zone.

The preference of most recreation analysts for a semilog specification of the first-stage demand function over a double log is confirmed by the results of this study. In terms of goodness of fit, stability of results across sites, accuracy of predicting quantity demanded at a zero price, and *a priori* reasonableness of results, this specification is clearly superior to the double log.

Some recreation analysts, such as Common, have used a double log specification with satisfactory results. However, Common and others have tried alternative specifications for only one site. For some centroids, consumers' surplus estimates are insensitive to the specification, but this result is a special case which may be observed in a sample of size one. A particularly serious problem with the double log specification is that on occasion it produces unrealistic results. The source of this problem is unclear and cannot be determined from the regression estimates of the experience demand schedule. The cause of these occasional untenable results may, to a lesser extent, affect the apparently tenable results; hence, these estimates should also be considered suspect.

This analysis of boating at 20 recreation centroids reflects a small sample of the 179 centroids and four recreation activities considered in my regional model. Recreation experience demand curves were estimated for each centroid and for each activity using both a double log and semilog specification.

The results are similar to those reported here. Some estimates of consumers' surplus are sensitive to the choice of functional form; some, insensitive. About five percent of the results using a double log specification were unreasonable.

The most disconcerting result of this study is that valuation estimates are sensitive to the definition of the origin zone. When each population centroid is construed as a separate zone, the explanatory power of the model is higher on the average than when centroids are aggregated into 10- or 20-mile zones. Furthermore, aggregating centroids results in a substantial loss of degrees of freedom, which *ceteris paribus* is undesirable and in this case causes the results to become unstable. However, aggregating population centroids into origin zones improves the accuracy by which total use is predicted at a zero price.

Most travel cost studies have been based on an aggregation of population centroids into concentric zones. The choice of a 10-mile versus a 20-mile system of concentric circles affects the results, but there is a greater disparity between using zones and using population centroids as origins. A consequence of using each centroid as an origin is that a large proportion of the centroids account for a small proportion of the trips. In rough numbers, about 96 percent of the centroids account for only 10 to 15 percent of the trips. The experience demand curve is therefore influenced disproportionately by centroids which account for very few trips. There is some justification for using each population centroid as an origin zone and for aggregating centroids into concentric zones. The best choice is unclear. Since travel cost valuation estimates are sensitive to the definition of the origin zone, this is an important topic for future work.

References

Abramowitz, Milton and Irene Stegun, *Handbook of Mathematical Functions,* Dover Publications Inc., 1965.

Brown, William B. and Farid W. Newas, "Impact of Aggregation on Estimation of Outdoor Recreation Demand Functions." *American Journal of Agricultural Economics,* 55(1973):246-49.

Bowes, Michael D. and John B. Loomis, "A Note on the Use of Travel Cost Models with Unequal Zonal Populations." *Land Economics,* 56(1980):465-70.

Burt, O. D. and D. Brewer, "Evaluation of Net Social Benefits from Outdoor Recreation." *Econometrica,* 39(1971):813-27.

Cicchetti, C. J., A. C. Fisher and V. K. Smith, "An Econometric Evaluation of a Generalized Consumer Surplus Measure: The Mineral King Controversy," *Econometrica* 44(1976):1259-76.

Clawson, Marion, "Methods of Measuring the Demand for and Value of Outdoor Recreation," Resources for the Future, Inc., Reprint #10, Washington, D.C., 1959.

Clawson, Marion, and Jack Knetsch, *Economics of Outdoor Recreation,* Baltimore: The Johns Hopkins University Press, Baltimore 1966.

Common, M. S., "A Note on the Use of the Clawson Method for the Evaluation of Recreation Site Benefits." *Regional Studies* 7(1973):401-406.

Davis, Phillip J. and Phillip Rabinowitz, *Numerical Integration,* Blaisdell Publishing Co., 1967.

Dwyer, John F., John R. Kelley and Michael D. Bowes, *Improved Procedures for Valuation of the Contribution of Recreation to National Economic Development,* University of Illinois, Water Resources Center, September 1977.

Goldfeld, Stephen M. "The Demand for Money Revisited." *Brookings Papers on Economic Activity* 3(1973):577-646.

Laidler, David E. W., *The Demand for Money: Theories and Evidence* New York: Dun-Donnelly, Second Edition, 1977.

Smith, V. Kerry, "Travel Cost Demand Models for Wilderness Recreation: A Problem for Non-Nested Hypotheses," *Land Economics,* 51(1975):103-11.

Sutherland, Ronald J., *A Regional Recreation Benefits Model,* U.S. Enviromental Protection Agency, Corvallis Environmental Research Laboratory, 1981.

Sutherland, Ronald J., "A Regional Approach to Estimating Recreation Benefits of Improved Water Quality." *Journal of Economics and Environmental Management,* forthcoming.

Water Resources Council, "Procedures for Evaluation of National Economic Development (NED) Benefits and Costs in Water Resources Planning (Level C)," *Federal Register*, (December, 1979):72,950-72,965.

Ziemer, Rod, F., Wesly N. Musser and R. Carter Hill. "Recreation Demand Equations: Functional Form and Consumer Surplus." *American Journal of Agricultural Economics*, 62(1980):136-41.