FUNCTIONAL FORM AND THE STATISTICAL PROPERTIES OF WELFARE MEASURES

Wiktor L. Adamowicz, Jerald J. Fletcher
and
Theodore Graham-Tomasi

Working Paper #87-28

November 6, 1987

Respectively, the authors are:

Assistant Professor, Department of Rural Economy, University of Alberta and Research Assistant, Department of Agricultural and Applied Economics, University of Minnesota;
Assistant Professor, Department of Agricultural Economics, Purdue University; and
Assistant Professor, Department of Agricultural and Applied Economics and Department of Forest Resources, University of Minnesota, and Visiting Assistant Professor, School of Natural Resources, University of Michigan.

Authors are listed alphabetically; senior authorship is not assigned.

The authors would like to thank Jim Binkley, Jay Coggins, Marie Livingston, and Ted McConnell for helpful discussions.

FUNCTIONAL FORM AND THE STATISTICAL PROPERTIES OF WELFARE MEASURES

Introduction

A large literature exists on the problem of specifying economic relationships for purposes of statistical estimation and testing. Regarding demand relationships, economic theory provides some restrictions on admissible specifications, but still allows considerable leeway concerning variables to be included and the appropriate choice of functional form. In the absence of specific guidance in these respects, a variety of statistical criteria for model selection have been suggested, but the problem essentially remains unresolved.

In an influential article, Ziemer, Musser, and Hill note the importance of functional form specification for the magnitude of welfare measures. They report results for demand for a recreation site based on the travel cost model [see McConnell (1985) for an overview of this model] which exhibit nearly a four-fold difference between consumer's surplus based on a linear demand curve and surplus computed from a semi-log demand. Ziemer, et al., go on to show that for their data, the semi-log demand statistically out-performs (in terms of t and F statistics) the then popular linear form. Moreover, they estimated a restricted Box-Cox transform (discussed below) and found the estimated Box-Cox parameter to be relatively close to zero, which also indicates a semi-log form. A number of subsequent investigators in the area of recreation demand modeling have followed the lead of Ziemer, et al., and employed the semi-log form.

Smith and Desvousges rationalize the use of the semi-log on the basis of the work by Ziemer, Musser, and Hill and that of Vaughan, Russell, and Hazilla (Smith and Desvousges, page 254). Several studies use the semi-log and/or linear form without consideration of any other forms (e.g. Kealy and Bishop; Wilman and Pauls; Bockstael and Strand; Smith, Desvousges, and Fisher). McConnell summarizes the past efforts in functional form choice by stating that, "...a researcher who was forced to choose from the literature...would find that the bulk of the evidence supports a semi-log form" (page 701). These conclusions may lead researchers to employ the semi-log form without full consideration of its effects on the welfare measure.

The research of Ziemer, et al., emphasizes that in many circumstances estimation of the demand curve is only an intermediate step in the overall research process. Often, consumer surplus computed from the estimated demand curve is used to make resource allocation decisions. While considerable attention has been devoted to obtaining good estimates of demand, what one really wants is a good estimate of the welfare measure. However, the former does not necessarily lead to the latter; the estimate of consumer's surplus is a random variable and alternative functional form specifications affect not only its mean, but also its variability.

That estimates of consumer's surplus are random variables is, of course, well-recognized. For example, Bockstael and Strand discuss the impact that alternative interpretations of the error term in the demand equation have on methods of computing expected consumer's surplus. In one interpretation, the error term is due to individual-specific excluded variables and the analyst should compute consumer's surplus for each observation and then take the expected value. In another interpretation, sources of error are not individual-specific and it is appropriate to base surplus estimates on the expected demand, i.e., the estimated demand. In general, these two will differ due to nonlinearity in the surplus function and Jensen's Inequality. As well, the issue of stochastic welfare measures has received attention in the context of discrete-choice, random utility models by Hanemann (1982a).

While these analyses have addressed some of the stochastic properties of welfare estimates, to our knowledge the influence of model specification on properties other than the mean or some other measure of central tendency has not been assessed. In this paper we simply note that different functional forms imply different transformations from demand parameters to welfare measures and that these transformations map instability in parameter estimates into instability of welfare estimates in different ways.

We show below that this insight can alter assessments of fits of alternative specifications to the sample data. The key results are driven by the fact that the coefficient on the price variable appears in the denominator of the consumer's surplus equation. In some instances (linear and semi-log) this parameter appears alone or multiplied by a constant. Hence, if the parameter is not significantly different from zero, it often will be the case (in a repeated sampling sense) that near-zero coefficients will be realized and the welfare measure will exhibit marked instability. However, for a double-log form, the denominator is one plus this coefficient. Thus, if the coefficient is far away (in terms of numbers of standard deviations) from minus one, the consumer's surplus estimate remains relatively stable. These effects may be quite large. Below we present Monte Carlo estimates for recreation data which show that, although the semi-log form is superior to the double-log form in terms of overall fit (as judged by t and F statistics), the coefficient of variation ($\sigma/|\mu|$) of consumer's surplus for the semi-log form is 24 times that for the double-log form.

Similar results can hold for decisions regarding whether to include a variable. Exclusion of an important variable is known to bias parameter estimates, but if the excluded and included variables have an appropriate correlation structure, the ratio of the estimated price coefficient to its standard error might rise by excluding a variable. Thus, excluding a variable could greatly decrease the variance of the welfare measure, a fact that may at first glance appear counter intuitive.

Why should one care about the variability of the welfare measure and not just its mean? Presumably, in many instances the welfare measure is to be used in making a resource allocation decision in a benefit-cost framework. The decision-maker can then be viewed as a statistician who is testing the hypothesis that the true population welfare measure exceeds the cost of the project or policy under consideration. Exactly how this statistical decision problem should be formulated raises complex questions regarding the treatment of uncertainty in benefit cost analysis which are beyond the scope of this paper. But in many formulations of this problem, there will be curvature in the associated statistical loss function and the variance of the estimates will matter.

This discussion reveals that a trade-off may exist between a bias in consumer's surplus estimates from choosing an incorrect specification and the variance of these estimates induced by the welfare transformation. This would suggest that a minimum mean square error criterion is appropriate. Unfortunately, economic theory cannot be used to deduce a true form and therefore bias cannot be assessed. We offer our analysis to provoke consideration of this issue in assessing alternative demand specifications. No exact guidance concerning the best solution to this problem is offered here.

The remainder of this paper is organized as follows. In the next section the consumer surplus functions for linear, semi-log, double-log, linear-log, and restricted Box-cox forms are presented and discussed in terms of the implication for the variance of surplus and statistical assessment of "best fits." In addition, the issue of variable inclusion is analyzed briefly. Then, Monte Carlo estimates of the importance of functional form choice on welfare variance are presented. The data used concerns the demand for hunting of big game in Canada. A final section presents the conclusions reached.

Welfare Measures and Demand Functional Form

The most commonly used functional forms for demand functions are the linear and the semi-log. As discussed above, these forms have been advocated both for how they fit the data (Ziemer, Musser, and Hill) and for their theoretical properties (see Bockstael, Hanemann, and Strand). In this section we consider the properties of the consumer's surplus functions from these two functional forms as well as the double-log and the linear-log forms.

The linear, semi-log, double-log and linear-log functional forms for a simple demand equation are:

(1) Linear $Q = \alpha_1 + \beta_1 P$

(2) Semi-log $\ln Q = \alpha_2 + \beta_2 P$

(3) Double-log $\ln Q = \alpha_3 + \beta_3 \ln P$

(4) Linear-log $Q = \alpha_4 + \beta_4 \ln P$

The derivation of the consumer's surplus functions for each of these forms requires an assumption regarding the source of error in the equation. Bockstael and Strand show that the estimated consumer's surplus function depends on whether the predicted quantity for an individual or their actual quantity is used. We present the consumer's surplus functions based on the assumption of "omitted variables", using the actual value of quantity in the surplus function. An alternative is to use the predicted quantity consumed which Bockstael and Strand argue is appropriate if the equation error is due to errors in measurement.[1] As is common in applied work, the demand functions which are asymptotic to the price axis are evaluated at the maximum and the average price.

The consumer's surplus functions for each of these forms are evaluated using:

(1-a) Linear $CS1 = Q^2/(-\beta_1 * 2)$

(2-a) Semi-log $CS2 = Q/(-\beta_2)$

(3-a) Double-log $CS3 = e^{\alpha_3} * P^{(1+\beta_3)}/(1+\beta_3)$ {eval at MAX(P),AVG(P)}

(4-a) Linear-log $CS4 = \alpha_4 P + \beta_4 * P * (\ln P - 1)$ {eval at MAX(P),AVG(P)}

---

1. For each sample of consumers we use the actual quantity of visits to estimate the consumer's surplus. This hypothesis does not preclude repeated sampling experiments to investigate the distribution of the welfare measure. This may not be the case for the "errors in measurement" hypothesis which uses predicted quantities. That is, in the absence of measurement error, the latter approach implies the absence of an inference problem; all individuals are identical and sampling variability is not an issue. This seems extreme. In the absence of a measurement model to identify the different sources of error, we believe that the omitted variable approach is more reasonable.

where Q is the actual quantity. These consumer's surplus functions are functions of the estimated slope and intercept of the demand functions; thus, they themselves are random variables. The distributions of these random variables are unknown. The linear form, for example, is a constant over a random variable. The expected value of this new random variable can be approximated but we know little of the other parameters of its distribution.

In typical demand studies the estimated parameters are of primary concern and hence the statistical fit has been used to discriminate between models. However, in cases where the consumer's surplus is of interest, both the expected value and the variance of the surplus measure are of importance. A significant t-statistic on a price parameter in a demand equation may not ensure low variance in the consumer's surplus. In particular, when discriminating between functional forms the most common tactic has been to examine t-values and measures of overall fit of the demand parameters. Some studies have used Box-Cox estimation to choose functional form (Ziemer, Musser, and Hill). This method of choosing functional form may be inappropriate when higher moments of consumer's surplus variables are of interest.

In both the linear and semi-log functional forms the β parameter cannot become zero, since as β approaches zero the estimate of consumer's surplus becomes infinite. This indicates the potential for a large variance of the consumer's surplus when the demand parameter β has a low t-statistic. For the double-log function the demand parameter must not approach -1. An appropriate test for the double-log model is to test the null hypothesis of β_3 not equal to -1 as well as β_3 not equal to zero. For the linear-log model such considerations do not appear to arise.

In addition to the requirements on the β parameters, one should consider the effect of other included variables on the price variable if consumer's surplus is the final interest of the estimation. Specification bias will occur by excluding important variables, but including variables that lower the significance levels of the price parameters can inflate the variance of the consumer's surplus. For example, if the researcher is interested in a confident measure of surplus, the addition of socioeconomic factors to a demand function must be approached carefully.

An alternative to estimating several functional forms is the Box-Cox estimation procedure applied in many tests of functional form. The restricted Box-Cox form[2] is:

$$(5) \qquad \frac{Q^{\lambda} - 1}{\lambda} = \alpha + \beta P$$

The consumer's surplus for this general form (again using actual quantity) is:

$$(5\text{-a}) \qquad CS5 = -\frac{1}{(1 + \lambda)\beta} * Q^{(1 + \lambda)}$$

For the case of $\lambda=0$ this becomes the semi-log form and for $\lambda=1$ this is the linear form. Once again the consumer's surplus is highly sensitive to the value of β. If β approaches zero, the consumer's surplus measure is unbounded. Thus for any transformation within this simple Box-Cox form, caution must be taken regarding the choice of the demand form.

Bockstael and Strand have shown that the expected value of the linear and semi-log consumer's surpluses can be approximated using the

---

2. This is distinguished from the Generalized Box-Cox form in which both the dependent and independent variables have Box-Cox parameters attached to them.

form for the expectation of the ratio of two random variables (derived by taking the expectation of a second order Taylor expansion of xy^{-1} around $(\bar{x},\bar{y})$

$$(6)\quad E(x/y) \approx E(x)/E(y) - cov(x,y)/E(y)^2 + E(x)\ var(y)/E(y)^3.$$

This approximation can be used to derive the following approximation to the variance of (x/y).[3]

$$(7)\quad Var(\frac{x}{y}) = E(\frac{x^2}{y^2}) - [E(\frac{x}{y})]^2$$

$$\approx \frac{E(x^2)}{E(y^2)} - \frac{cov(x,y)^2}{E(y^2)^2} + \frac{E(x^2)var(y^2)}{E(y^2)^3} - \frac{E(x)^2}{E(y)^2} - \frac{cov(x,y)}{E(y)^4}$$

$$- \frac{E(x)^2 var(y)^2}{E(y)^6} + 2\frac{E(x)}{E(y)^3}\left[cov(x,y) + \frac{var(y)\ cov(x,y)}{E(y)^2} - \frac{E(x)\ var(y)}{E(y)}\right]$$

In a model with measurement error, the numerators for the consumer's surplus estimates in (1a) to (4a) are random variables. In the interpretation used here, we may assume x to be constant and terms involving covariances drop out of (7), but terms involving the mean and variance of the square of the coefficient β on the price variable remain. Therefore we shall proceed to perform Monte Carlo analysis on these distributions for a given data set to observe the distributions of the welfare measures.

Empirical Analysis

In order to illustrate the effects of functional form on the statistical properties of welfare measures, we estimate several functional forms of demand functions for recreation. These functions correspond to the travel cost demand model popular in the recreational

---

3. We thank Kenneth McConnell for pointing out this possibility to us.

demand literature. The data, collected by mail survey, are the number of visits in a season and travel costs to a bighorn or Rocky Mountain sheep hunting site in Alberta, Canada.[4] The travel costs are expressed in 1981 dollars. The travel cost model is a simple one estimated with visits the measure of quantity and travel cost the measure of price. The consumer's surplus is the value of the site for recreational use.

The results of estimations of the linear, semi-log, double-log, and linear-log models are reported in table 1. Inspection of the results indicates that the linear and semi-log models perform well in terms of t-statistics on the price variable and F-tests. The double-log and linear-log models do not perform as well under these criteria. Typically, a researcher would choose the semi-log model in this case as it has the highest F-value and the highest t-statistic on the travel cost variable.

The point estimates of consumer's surplus are provided in table 2. As found in other studies, the (expected) consumer's surplus measure is quite sensitive to the choice of functional form. We also estimated the Box-Cox form described above and provide figure 1 which shows the consumer's surplus measure as a function of the Box-Cox parameter. As the Box-Cox parameter increases the measure of consumer's surplus declines[5]. We are interested in describing the variance around selected points on the curve drawn in figure 1, as well as the variance of consumer's surplus for other forms not nested within the Box-Cox framework.

---

4. Further details on the data are available from the authors upon request.
5. Implicit in this figure are the changing parameters in the demand funciton.

In order to estimate the variance of the consumer's surplus measures, for each model we generate a new series of dependent variables using the non-random design matrix and a randomly generated error from a normal distribution which has mean zero and variance equal to the variance of the error of the regression. This new dependent variable is then used to determine a new set of demand coefficients; the new demand parameters are in turn used to calculate new estimates of the consumer's surplus for each functional form. This procedure is replicated 5000 times. The result is a distribution of welfare measures for each functional form[6].

The results of the Monte Carlo analysis are presented in table 3. The mean, standard deviation, minimum, maximum and coefficient of variation are presented for each functional form. The size of the standard deviation of the welfare measure for the linear and semi-log model is most apparent. The large variances occur because the travel cost parameter is only significantly different from zero at a 1 percent level. Therefore, many replications of the model result in travel cost parameters near zero. The semi-log and linear-log models do not suffer from this inflation.

In the case of the double-log and linear-log models, the means of the Monte Carlo analyses and the point estimates of the means provided by equations (3-a) and (4-a) are not very different. For the linear and semi-log models, the means of the Monte Carlo analyses are somewhat larger than the point estimates. This is due to the fact that the Monte Carlo results provide a direct estimate of, rather than an approximation

6. The data are provided in the Appendix. This procedure can be replicated except for the random number generation procedure chosen. The algorithm and any further information desired on the Monte Carlo procedures used are available from the authors upon request.

for, the expected consumer's surplus. While equations (1-a) to (4-a) provide the point estimates usually used in welfare analysis, the expected value of consumer's surplus can be approximated for the linear and semi-log functional forms using relation (6) above. However, this approximation may not be appropriate for cases in which the price parameter approaches zero.[7] The Monte Carlo analysis will choose some values of the price coefficient which are very close to zero and this will have a severe effect on the estimate of the mean and variance of the consumer's surplus. For this reason, the expected value of consumer's surplus calculated in the Monte Carlo analysis for the linear and semi-log models is somewhat higher than the value provided by the approximation. The approximated expected values for the semi-log and linear models are \$2971.14 and \$1601.12, respectively. Comparing these to the mean values of consumer's surplus estimates reported in table 3 shows that the difference between using the approximation and the Monte Carlo analysis is as large as the difference in consumer's surplus between two different functional forms.

The analysis of the distribution of the welfare measures has shown that even though the parameters of the demand equation may appear preferable, the distribution of the consumer's surplus measure may not

---

7. The approximation presented in (6) is a second order Taylor series approximation of the ratio of random variables evaluated around the expected values. Extension of this Taylor series to the third order yields additional terms which involve the variance of y, the covariance of y and y^2, and the expected value of x in the numerator with $E(y)^3$ and $E(y)^4$ in the denominator. Larger variances increase the size of this term. Division by $E(y)^3$ and $E(y)^4$ when $E(y)$ is small may also lead to large values for this term. The significance of the higher order terms is illustrated by the fact that the first order term for the semi-log model is \$2,593 and the second order term adds \$378; the Monte Carlo mean is \$3,282 indicating terms higher than 2nd order add \$311 to the expected value. For the double-log model, the Monte Carlo estimate and the point estimate differ by only \$10 or 0.6%.

be appealing. The investigator must realize that the curve drawn in figure 1 also has a distribution around each point and that altering the choice of functional form will affect the distribution of the estimate of the welfare measure. For the data analyzed above, the demand equations for the double-log and linear-log appear more appealing on the grounds of confidence in the welfare measure. The parameter of the double-log is strongly significantly different from unity and thus the inflation of the welfare measure does not occur as it does in the linear and semi-log models[8].

Discussion

Most analyses of different functional forms for demand equations have labeled the form that "best fits" the data as the "true" form and the welfare measure computed from it as the "true" measure. This is not strictly true, since the true forms are unknown. The use of statistical demand analyses results in a probability distribution for the welfare measure. In many situations, the variance of this distribution, as well as its mean, will matter to the analyst. In this paper we point out that different function forms imply different dispersion of this distribution in addition to the different means noted previously in the literature (Ziemer, Musser, and Hill). Our results indicate that this effect may be substantial. Of course these specific Monte Carlo results may not generalize to all data sets.

One method of choosing functional form is to use the Generalized Box-Cox. Unfortunately, integrating the generalized Box-Cox form to

---

8. We have also estimated these models using Bockstael and Strand's measurement error form with predicted quantity (number of trips taken) in the welfare measure. The results obtained do not differ qualitatively from the results presented here.

find the consumer's surplus measure requires integration by reduction which implies that the Box-Cox parameters must be known before integration. Thus, the general form of the consumer's surplus for the Generalized Box-Cox cannot be defined. Once the welfare measure is determined for each specific combination of Box-Cox parameters, however, the variance can be estimated using the procedures outlined above -- a very cumbersome recommendation. A simple approach is to estimate a few functional forms and follow some rules to determine how sensitive the welfare measure is, namely, the price coefficient must be different from zero for the linear or semi-log and must be different from one for the double-log. A more complete analysis must address the question of how different these should be.

This analysis does not preclude the results of Hanemann (1982b) and others who have investigated the restrictions that demand parameters must satisfy in order to be consistent with an underlying utility function. Our analysis only addresses the question of functional form and the statistical properties of the welfare measure within the class of integrable demands. However, not all forms are consistent with utility theory and some investigators may be willing to sacrifice theoretical consistency for good fits and small variance. Also, some forms, such as the double-log form, imply that the resource is "essential" (Bockstael, Hanemann and Strand). These theoretical considerations may influence the choice of functional form. The results from our experiments imply that if one chooses a functional form on the basis of the variance of consumer's surplus, for our data the double-log appears to be the appropriate choice among the functional forms considered.

Our results imply that there is a tradeoff between the utility-theoretic models and models which provide appealing statistical results. While we do not provide any final recommendations on this problem, we have noted that the problem exists, may be important, and is worthy of further attention via both basic and applied research.

Table 1. Regression Results of Alternate Functional Forms

Dependent variable: VISITS
Independent variable: TRAVEL COST

Model	Constant*	Slope*	R^2	Adj R^2	F	Standard Error
Linear	2.869 (0.1809)	-.0025 (.00103)	.043	.036	5.84	1.658
Semi-log	0.864 (0.0677)	-.0010 (.00038)	.050	.042	6.79	0.621
Double-log	1.126 (0.2380)	-.2381 (.05573)	.019	.011	2.52	0.631
Linear-log	3.652 (0.6326)	-.2516 (.14810)	.022	.014	2.89	1.677

* Standard errors in parentheses

Table 2. Point Estimates of Consumer's Surplus

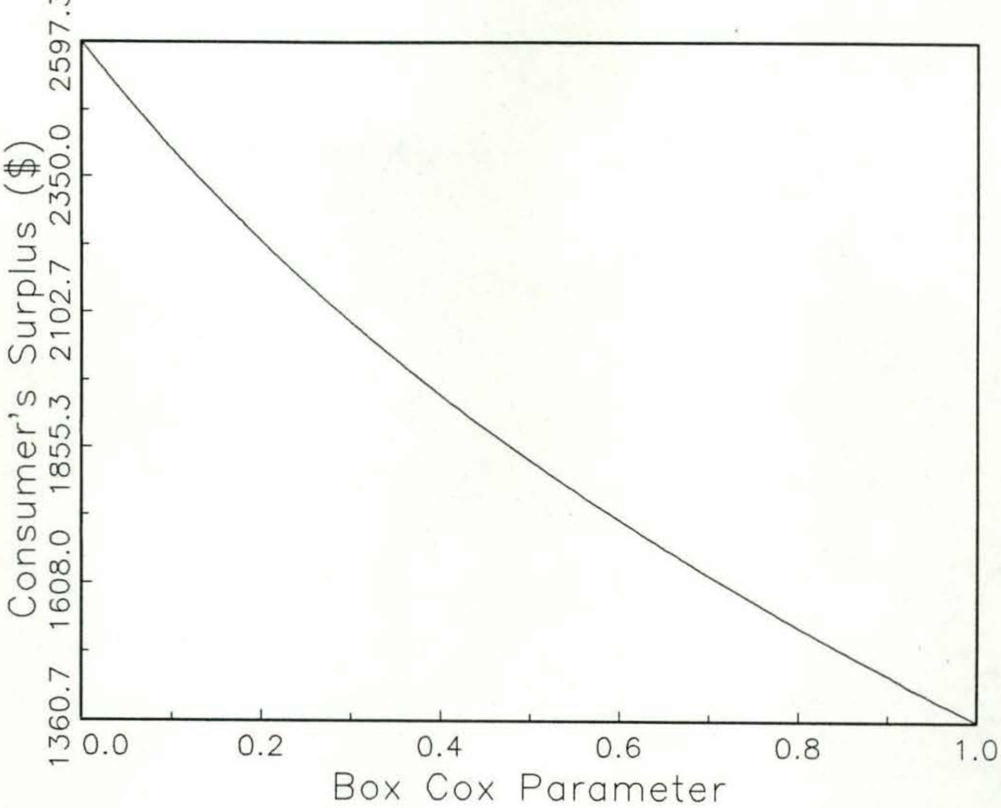
Functional Form	Consumer's Surplus
Linear Model	$1367
Semi-log Model	$2593
Double-log Model	$1598
Linear-log Model	$1876

Table 3. Measures of Welfare*

Model	Mean	Standard Deviation	Minimum	Maximum	Coefficient of Variation
Linear	1,974	12,090	-184,766	624,327	6.12
Semi-log	3,282	10,379	-320,804	365,844	3.16
Double-log	1,608	203	1,021	2,540	0.13
Linear-log	1,873	300	963	2,990	0.16

* From a Monte Carlo experiment designed to calculate estimates of consumer's surplus from 5,000 sets of randomly generated observations.

FIGURE 1: Consumer's Surplus



References

Bockstael, Nancy E., W. Michael Hanemann, and Ivar E.Strand. "Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models." Draft Report EPA No. CR-811043-01-0, Department of Agricultural and Resource Economics, University of Maryland. College Park, Maryland, 1984.

Bockstael, Nancy E. and Ivar E. Strand. "The Effect of Common Sources of Regression Error on Benefit Estimates." Land Econ., 63(1987):11-20.

Hanemann, W. Michael. "Applied Welfare Analysis with Qualitative Response Models." Giannini Foundation of Agricultural Economics, University of California-Berkeley, WP #241, 1982a.

Hanemann, W. Michael. "Some Further Results on Exact Consumer's Surplus." Giannini Foundation of Agricultural Economics, University of California-Berkeley, WP #190, 1982b.

Kealy, Mary Jo and Richard C. Bishop. "Theoretical and Empirical Specifications in Travel Cost Demand Studies." Amer. J. of Agr. Econ., 69(1986):660-667.

McConnell, Kenneth E. "The Economics of Outdoor Recreation." In A.V. Kneese and J.L. Sweeney eds., Handbook of Natural Resource and Energy Economics, Elsevier Science Publishers, 1985.

Smith, V. Kerry and William H. Desvousges. Measuring Water Quality Benefits. Kluwer-Nijhoff Publishing, Boston, 1986.

Smith, V. Kerry, William H. Desvousges, and Ann Fisher. "A Comparison of Direct and Indirect Methods for Estimating Environmental Benefits." Amer. J. of Agr. Econ., 68(1986):280-290.

Vaughan, William J., Clifford S. Russell, and M. Hazilla. "A Note on the Use of Travel Cost Models with Unequal Zonal Population." Land Econ., 58(1982):400-407.

Wilman, Elizabeth A. and R.J. Pauls. "Sensitivity of Consumers' Surplus Estimates to Variation in the Parameters of the Travel Cost Model." Cand. J. of Agr. Econ., 35(1987):197-212.

Ziemer, Rod F., Wesley N. Musser, and R. Carter Hill. "Recreation Demand Equations: Functional Form and Consumer Surplus." Amer J. of Agri. Econ. 62(1980):136-141.

Appendix - Data on Visits and Trip Costs

VISITS	TRAVEL COST	VISITS	TRAVEL COST	VISITS	TRAVEL COST	VISITS	TRAVEL COST
6.00	66.67	1.00	133.33	3.00	37.50	2.00	44.12
7.00	40.00	7.00	25.00	6.00	23.33	2.00	141.67
3.00	100.00	1.00	37.50	2.00	37.50	4.00	75.00
2.00	250.00	6.00	30.70	4.00	46.15	1.00	27.27
1.00	71.43	1.00	11.94	1.00	51.72	2.00	90.00
6.00	91.67	1.00	25.00	2.00	83.33	3.00	232.56
1.00	61.97	2.00	50.00	1.00	350.00	5.00	52.17
7.00	21.43	1.00	93.75	2.00	76.92	2.00	12.68
1.00	71.94	2.00	70.00	1.00	50.00	2.00	120.00
5.00	157.45	7.00	71.43	1.00	108.02	1.00	291.97
2.00	28.26	1.00	123.76	2.00	30.00	3.00	18.01
2.00	123.08	1.00	462.96	6.00	140.00	2.00	150.00
1.00	71.43	2.00	72.29	2.00	42.55	2.00	20.00
4.00	23.08	3.00	78.36	1.00	62.23	6.00	66.67
4.00	347.22	2.00	113.33	2.00	180.49	2.00	100.52
3.00	119.57	3.00	58.33	4.00	62.50	4.00	60.87
1.00	50.00	5.00	13.20	1.00	764.71	2.00	55.93
1.00	95.24	2.00	560.75	1.00	26.67	2.00	40.12
2.00	321.10	4.00	49.85	3.00	25.40	2.00	27.48
2.00	20.66	3.00	52.17	5.00	59.49	1.00	6.86
1.00	1000.00	2.00	50.00	3.00	66.67	5.00	34.87
4.00	100.00	1.00	121.21	3.00	100.00	2.00	24.19
4.00	49.87	2.00	40.00	4.00	52.82	2.00	50.00
4.00	145.45	1.00	50.00	1.00	350.88	1.00	30.61
3.00	81.70	2.00	65.43	3.00	18.50	2.00	4.03
2.00	250.00	5.00	32.00	3.00	104.13	2.00	23.16
5.00	9.00	1.00	63.06	1.00	200.00	5.00	30.00
1.00	79.11	1.00	121.62	1.00	600.00	7.00	85.71
2.00	48.95	1.00	40.85	4.00	125.00	1.00	22.73
3.00	198.11	5.00	40.95	3.00	259.26	1.00	37.97
1.00	35.71	1.00	196.51	1.00	27.27	4.00	17.21
1.00	264.08	3.00	116.67	1.00	30.38	2.00	180.00
4.00	85.71	1.00	45.00	1.00	2.94	4.00	30.41

Descriptive Statistics (Number of observations = 132)

Variable	Mean	Standard Deviation	Minimum	Maximum
VISITS	2.61	1.69	1.00	7.00
TRAVEL COST	106.05	140.91	2.94	1000.00