Valuing Resource Access with Seminonparametric Techniques: An Application to Clear Lake

John R. Crooker

Working Paper 04-WP 352
January 2004

Center for Agricultural and Rural Development
Iowa State University
Ames, Iowa 50011-1070
www.card.iastate.edu

John Crooker is an assistant professor in the Department of Economics and Finance at Central Missouri State University.

For questions or comments about the contents of this paper, please contact John Crooker, Department of Economics and Finance, Central Missouri State University, Warrensburg, MO 64093; Ph: 660-543-8601; Fax: 660-543-8465; E-mail: crooker@cmsu1.cmsu.edu.

Abstract

The focus of this investigation is three-fold. First, seminonparametric techniques are considered in order to avoid the pitfall of imparting bias on value estimates. Second, the efficacy of seminonparametric techniques are explored in an empirical setting by valuing recreational access at Clear Lake, Iowa. Finally, a technique from Eastwood and Gallant (1991) is adapted and applied to the count data model.

Keywords: Clear Lake survey, nonmarket valuation, recreation valuation, seminon-parametric techniques, value estimates, welfare estimates.

VALUING RESOURCE ACCESS WITH SEMINONPARAMETRIC TECHNIQUES: AN APPLICATION TO CLEAR LAKE

Introduction and Motivation

In recreation valuation studies, accurate estimation of welfare measures is critically important. Oftentimes in the applied setting, welfare estimates are used to make policy decisions regarding the management of resources (and sometimes these management decisions are irreversible). The result in the literature suggests that the researcher must be careful in estimating welfare, as the potential for biased estimates is high (Bockstael, Hanemann, and Strand 1986; Graham-Tomasi, Adamowicz, and Fletcher 1988; Adamowicz, Fletcher, and Graham-Tomasi 1989; Kling 1989; Kling and Sexton 1990; Smith 1989).[1] Certainly, the presence of biased welfare estimates forces undesirable characteristics on decisions made based on these welfare estimates.

The focus of this investigation is threefold. First, we consider seminonparametric techniques in order to avoid the pitfall of imparting bias on value estimates. Second, we wish to explore the efficacy of the seminonparametric techniques in an empirical setting by valuing recreational access at Clear Lake, Iowa. Creel (1997) and Cooper (2000) found desirable characteristics of these techniques in Monte Carlo settings. However, actual parameterization of the seminonparametric models in purely empirical settings is lacking in the literature. One notable investigation in the money-demand literature is provided in Fisher, Fleissig, and Serletis 2001. Finally, determining the length of our sequence of periodic functions, we adapt a technique from Eastwood and Gallant (1991) and apply it to the count data model. Cooper (2000) investigated the Fourier Flexible Form (FFF) in the count data setting; however, his analysis did not consider adaptive techniques in selecting the seminonparametric truncation. These issues will be developed in what follows.

The remainder of this paper is divided into eight sections. The next section discusses why welfare estimation in the travel-cost model setting is problematic and explores approaches that researchers have used to address these concerns. Next, the Clear Lake application is discussed and the survey instrument used to gather data is described. Then an overview of count data and seminonaparametric modeling is presented. This is followed by a presentation of the results of these estimation techniques applied to the Clear Lake dataset. The pivotal statistics bootstrap approach for bounding welfare estimates in the applied setting is presented next. The final section presents remarks, conclusions, and extensions for future work.

Welfare Estimation in the Travel-Cost Model Setting

Typically, researchers assume a functional form for trip demand. Most often, this is assumed to be linear or semilog (McConnell 1985; Bockstael and Strand 1987). For example,

$$\begin{aligned} &\textit{Linear:} && y_i = a + b \cdot P_i + g \cdot M_i + e_i \text{; and} \\ &\textit{Semilog} && \ln(y_i) = a + b \cdot P_i + g \cdot M_i + e_i, \\ &\textit{Log-linear} && \ln(y_i) = a + b \ln(P_i) + g \ln(M_i) + e_i \end{aligned} \tag{1}$$

where y_i indicates individual *i*'s observed trips; P_i is the individual's price of attending the recreation site; M_i is individual *i*'s income; e_i captures individual *i*'s unsystematic stochastic deviation from the central tendency; and $\{a, b, g\}$ are unknown model parameters.[2] These functional forms imply consumer surplus (*c*) is as follows:

$$\begin{aligned} &\textit{Linear:} && c_i = -\frac{y_i^2}{2\beta}\text{; and} \\ &\textit{Semilog:} && c_i = -\frac{y_i}{\beta}, \\ &\textit{Log-linear:} && c_i = \frac{P_i y_i}{b+1}. \end{aligned} \tag{2}$$

Among the dilemmas of the traditional approaches is that these formulations involve nonlinear transformations of the parameter estimates. This is problematic, in general, as

$E\left[\frac{V}{W}\right] \neq \frac{E[V]}{E[W]}$, where *V* and *W* are stochastic variables. It follows that applying the plug-in principle to equation (2) will result in biased estimates in finite samples. Researchers have argued that Taylor series approximations may be appropriate for correcting the bias (Bockstael, Hanemann, and Strand 1986; Kealy and Bishop 1986). However, this approach also has its problems. Graham-Tomasi, Adamowicz, and Fletcher (1988) suggest that errors from the second-order approximations in a Taylor series are potentially large for both linear and semilog demand functions (Kling and Sexton [1990] detail this approach and discuss potential problems).

A bootstrap approach seems appealing to the applied researcher. Essentially, the distribution of the random variables in equation (2) is unknown. Even if we assume that the stochastic term e_i is normally distributed, analytically determining the distribution of the random variables in equation (2) appears formidable. A bootstrap procedure, however, is useful here because it does not require the analytic or asymptotic distribution to construct confidence intervals. Instead, we will rely on the plug-in principle by using the empirical distribution of the random variable in place of the unknown true distribution to generate our confidence intervals (using computer simulations). Provided the empirical distribution contains all relevant information from the population, the bootstrap technique will work well.[3] We will readdress the topic of a bootstrap procedure to construct bounds on welfare values in what follows. We will now turn our attention to the particular recreation demand application.

Clear Lake Application

Recently, the Iowa Department of Natural Resources became interested in improving water clarity at Clear Lake, which is located adjacent to the Iowa cities Clear Lake and Ventura in north-central Iowa. To gauge public attitudes concerning the recreational site, a survey was conducted. The survey effort was undertaken by researchers at Iowa State University (Azevedo et al. 2001) in 2000. While the survey effort consisted of a sample from the local population and visitors to the site, we use only the site visitors' database in estimating value.

From the visitors' dataset we eliminate observations from individuals who traveled more than five hours one way to reach Clear Lake.[4] This is done to exclude those whose primary reason for the trip may not have been to visit Clear Lake. If the trip was not simply to access Clear Lake, our valuation techniques would inappropriately attribute value to Clear Lake.

Site visitors were intercepted on-site from May to September of 2000. In total, 1,024 individuals agreed to participate in the survey effort. Each participant was mailed a survey in October of 2000. All participants were informed that if they returned a completed survey they would receive $5. Of the deliverable surveys, 662 were returned, for a response rate of 66 percent.

Count Data Models

An additional complication we face in this empirical setting is the fact that the data was collected on-site. This is problematic, as more frequent site visitors are more likely to be included (endogenous stratification). Also, site users likely have higher values for this resource; hence. valuation estimates based on this data cannot simply be extended to the greater population of Iowa.

A literature has emerged regarding correcting for the biases associated with on-site samples. Shaw (1988) modeled the count data by fitting a Poisson model. Englin and Shonkwiler (1995) fitted a negative binomial model, as it allows for overdispersion. That is, a characteristic of the Poisson model is that the mean equals the variance. In the recreation-demand setting (and many other settings it turns out), it is common to see many site visitors with few visits and a minority of site users with numerous visits. Hence, we empirically observe a low mean and high variance environment.

Gurmu and Trivedi (1994) take this further by noting that empirics demonstrate a fast decay rate. That is, the negative binomial model forces a relatively fat tail distribution. A fast decay model, on the other hand, better captures recreation data, as the tail collapses quickly away from the mean yet preserves the overdispersion feature.

Another issue explored in the literature is the nature of on-site samples to be endogenously stratified. That is, more frequent site visitors are more likely to be included in the survey effort. Hence, they are overly represented in the sample. Cameron and Trivedi

(1990) detail a negative binomial II model to account for this bias. Sarker and Surry (2003) suggest that the negative binomial II model is capable of fitting a fast decay process as well. For this reason, we will adapt the negative binomial II model in our valuation efforts of Clear Lake. The specification of the negative binomial II probability of site visits is

$$\Pr\left(y_i = k \middle| x_i\right) = \frac{\Gamma(q+k)}{\Gamma(k+1)\Gamma(q)} r^k (1-r)^q , \tag{3}$$

where $r = \frac{l}{l+q}$ (Greene 2000). The conditional mean of the distribution is l. The conditional variance is $\lambda\left(1+\lambda/\theta\right)$.

The parameter q allows for overdispersion and, under certain conditions, permits a fast-decay process. Traditional approaches in the travel demand literature are to model: $l_i = \exp(b' x_i)$. Cooper (2000), however, considers the application of the FFF to fitting the negative binomial II model. That is, more generally specify $l_i = \exp(g_J[x_i])$. Creel (1997) applied the FFF to the travel demand setting; however, his exploration of this technique did not consider the implications of on-site bias. We consider the seminonparametric fitting of this more general expression with on-site datasets in the following section.

Seminonparametric Demand Specification

The next issue we investigate in estimating consumer surplus in the applied setting is that the parametric form of demand is also unknown. The question then becomes, How does the researcher arrive at any particular specification?

In the literature, researchers typically estimate several different specifications and accept the model with the best goodness-of-fit measures. However, in many instances the difference in goodness-of-fit measures is extremely small, yet the estimators yield consumer surplus values that differ by large magnitudes. This is frustrating for the researcher as the assumption regarding the parametric form of demand is driving the estimates and may perhaps drive management decisions regarding the resource in question.

In this section, we will use an adaptive FFF to model trip demand by recreationists. The FFF is attractive as it is in a class of estimators that "…provide a second order approximation to an arbitrary twice differentiable function at any point" (Gallant 1981).

The FFF we consider is

$$g_{\infty}\left(v_i \,|\, q\right) = q_0 + \sum_{j=1}^{\infty}\left\{q_{j1}\cos\left(j \cdot v_i\right) + q_{j2}\sin\left(j \cdot v_i\right)\right\} + h_i \tag{4}$$

where q is a vector of model parameters; v_i are the scaled nonstochastic explanatory variables; and h_i is the error term. We use $g_{\infty}(\bullet)$ in place of *y* to signify that recreation trips are also scaled. The reason for the scaling is that cosine and sine are periodic functions. Thus, the range of the variables must be compressed into the $[0, 2p]$ interval.

In practice, we cannot estimate $g_{\infty}(\bullet)$, as it involves an infinite sum of variable transformations and would require infinitely many parameter estimates. So, a key question becomes where to truncate the infinite sum. Eastwood and Gallant (1991) demonstrate that an adaptive rule for determining the truncation point of the infinite sum dominates any deterministic choice in the applied setting. It is this adoptive rule that we consider. Essentially, we consider truncation points $J \in \left\{1, 2, ..., \frac{n}{2} - 1\right\}$. Eastwood and Gallant (1991) suggest selecting *J* such that it provides the maximum *F*-statistic in this range. The adoptive truncation model for recreation trip demand is then

$$g_J\left(v_i \,|\, q\right) = q_0 + \sum_{j=1}^{J}\left\{q_{j1}\cos\left(j \cdot v_i\right) + q_{j2}\sin\left(j \cdot v_i\right)\right\} + h_i. \tag{5}$$

To estimate consumer surplus, we use numerical methods to integrate the area under the estimated inverse demand curve.

In our present setting, we will be estimating our model parameters with maximum likelihood techniques. Hence, we will extend Eastwood and Gallant's (1991) approach to truncating the infinite periodic series by selecting J that maximizes the chi-square statistic.[5]

The natural question to ask is how well the seminonparametric specification allows us to estimate consumer surplus. In the next section, we present a bootstrap t-interval technique to bound our empirical welfare estimates.

Model Fitting

For exploratory purposes, we begin by fitting the linear and semilog demand specification as presented in equation (1). Notice that these results do not account for sample selection bias. Hence, we only report these results for comparison purposes. As the models we consider to account for sample selection bias and issues we encounter in performing the adaptive FFF techniques are considerably more complex, it will be interesting to discover how different the resulting valuation estimates are. The results are listed in the first three columns of Table 1.

TABLE 1. Travel-cost model results

	Linear	**Semilog**	**Adaptive FFF *J*=13, *A*=10**
R^2	0.14	0.22	0.53
Average individual's consumer surplus	$1,583.61	$223.50	$542.21
Intercept	*21.306** (5.538)	*2.812** (15.272)	–
Explicit price	*-0.089** (-5.304)	*-0.009** (-11.769)	–
Implicit price	-0.003 (-0.381)	-0.001 (-1.465)	–
Income	-0.000 (-0.280)	*0.000** (2.074)	–
Household size	-0.599 (-1.800)	*-0.0365** (-2.290)	–
Male	2.895 (1.438)	0.111 (1.154)	–
Age	1.583 (0.728)	-0.128 (-1.231)	–
Age^2	0.045 (0.678)	-0.002 (-0.647)	–

Before applying the adaptive FFF, we experimented with various multi-indices. The multi-indices we used allowed for interaction between the explicit and implicit prices, prices and income, income and age, and prices and age. The longest multi-index was 3. The complete list of multi-indices appears in Table 2. The adaptive FFF technique indicated the optimal truncation length of the sequence of periodic terms was 13 (i.e., J=13). Results of this regression appear in column 4 of Table 1. Valuation measures were determined by numerically integrating the area under the demand curve between the individual's choke price and the individual's price of attending Clear Lake for the semi-nonparametric model. For the parametric model, we used the appropriate consumer surplus estimate proscribed in equation (3).

Examining the results of these exploratory regressions reveals that the adaptive FFF considerably improves our measure of goodness of fit. Improving goodness of fit alone is not surprising, as the adaptive FFF includes many more right-hand side variables in the regression. However, the magnitude of improvement in *R*-square is noteworthy.[6] Before

TABLE 2. Multiple indices

k	Intercept	Explicit Price	Implicit Price	Income	Household Size	Male	Age	Age^2
k_1	0	1	0	1	0	0	0	0
k_2	0	1	0	-1	0	0	0	0
k_3	0	1	1	-1	0	0	0	0
k_4	0	1	1	1	0	0	0	0
k_5	0	0	0	1	0	0	1	0
k_6	0	0	0	1	0	0	-1	0
k_7	0	1	0	0	0	0	1	0
k_8	0	1	0	0	0	0	-1	0
k_9	0	1	1	0	0	0	1	0
k_{10}	0	1	1	0	0	0	-1	0
k_{11}	0	1	0	0	0	0	0	0

we can analyze the valuation estimates, we must consider corrections for sample selection bias. This is where we now direct our attention.

Essentially, to account for the on-site sample bias, we apply maximum likelihood techniques using the probability of site visits stated in equation (3) while simultaneously fitting the seminonparametric demand specification given in equation (5). For this application, we set J=1 and use the multi-index k_1, as presented in Table 2. Results from this seminonparametric version of the negative binomial II model appear in Table 3 with the linear version of the negative binomial II model.

The estimated parameter q in both models is quite small. This implies that overdispersion is indeed a problem, as the estimated variance is considerably larger than the estimated mean. This nature of overdispersion does suggest it may be appropriate

TABLE 3. On-site corrected-model results

	$l_i = \exp(b'x_i)$	$l_i = \exp(g_J[v_i])$
Mean log-likelihood	-3.68692	-3.6102263
$\hat{q}$	1/exp(29.2)	1/exp(32.2)
Average individual's consumer surplus	tba	$524
Intercept	-24.9561 (-0.838)	–
Explicit price	*-0.00586** (7.887)	–
Implicit price	-0.00122 (1.443)	–
Income	6.59E-07 (0.383)	–
Household size	*-0.03706** (2.676)	–
Male	0.146213 (1.588)	–
Age	*-0.0574** (4.578)	–
Age^2	*0.000647** (4.780)	–

to generalize the statement of the probability model, as discussed in Sarker and Surry 2003. In particular, adoption of a generalized Poisson model, which allows for a more flexible fast-decay process, may improve our model estimates. This adoption is outside the focus of the current paper.

For the FFF estimation we do perform, the estimated value for the average individual in our sample is presented in column 3 of Table 3. The estimated value is $524. The value from the adaptive FFF without correcting for on-site bias was $542.21. Hence, we do see some moderation in the average estimate, as we would expect the endogenously stratified sample to overmeasure value from more frequent site visitors. However, we have not performed any test to determine any statistically significant reduction. Next, we discuss methods for exploring the accuracy of these valuation estimates.[7] To do so, we suggest the aforementioned bootstrap techniques.

Bootstrap t-Intervals and Percentile Intervals for Welfare Estimates

In this section, we discuss how to implement the bootstrap t-interval technique for the travel cost model. To do so, we suppose we have observed a travel-cost model dataset with *n* observations. The procedure is as follows:

Step 1: Estimate the model parameters $\{a, b, g\}$ through maximum likelihood (ML). We call the ML estimates $\{\hat{a}, \hat{b}, \hat{g}\}$) and obtain the residuals. Also, estimate

$$\hat{c} = -\frac{y_i^2}{2\hat{b}}.$$

Step 2: Take *n* draws (independently and with replacement giving equal likelihood to each observation being selected on any draw) from the ordinary least squares (OLS) residuals.

Step 3: Form the simulated dataset according to $y_i^* = \hat{a} + \hat{b} \cdot P_i + \hat{g}_i \cdot M_i + e_i^*$, where e_i^* is the *i*th draw performed in Step 2.

Step 4: Using the simulated trip data in Step 3, estimate by means of OLS the model parameters (calling them $\{\hat{a}^*, \hat{b}^*, \hat{g}^*\}$) and obtain the bootstrap residuals. Also, estimate $\hat{c}^* = -\dfrac{y_i^{*2}}{2\hat{b}^*}$.

Step 5 (bootstrap within a bootstrap): Take *n* draws (independently and with replacement giving equal likelihood to each observation selected on any draw) from the bootstrap residuals formed in Step 4.

Step 6: Form the simulated dataset according to: $y_i^{**} = \hat{a}^* + \hat{b}^* \cdot P_i + \hat{g}^*{}_i \cdot M_i + e_i^{**}$, where e_i^{**} is the *i*th draw performed in Step 5.

Step 7: Using the simulated trip data in Step 6, estimate by means of OLS the model parameters (calling them $\{\hat{a}^{**}, \hat{b}^{**}, \hat{g}^{**}\}$). Also, estimate $\hat{c}^{**} = -\dfrac{y_i^{**2}}{2\hat{b}^{**}}$.

Step 8: Repeat Steps 5-7 many times, say, *BB* times, and calculate the standard error of the consumer surplus estimate as $\hat{s}^* = \left\{\dfrac{1}{n-1}\sum_{i=1}^{n}\left(\hat{c}^{**} - \hat{c}^{**}(\cdot)\right)^2\right\}^{\frac{1}{2}}$, where $\hat{c}^{**}(\cdot)$ is the mean $\hat{c}^{**}$ averaged over all *BB* trials.

Step 9: Calculate $t^* = \dfrac{\hat{c}^* - \hat{c}}{\hat{s}^*}$.

Step 10: Repeat Steps 2-9 many times, say, *B* times, storing the simulated distribution of t^*'s.

Step 11: Define $t^{(k)}$ such that $\dfrac{\#\{t^* \le t^{(k)}\}}{B} = k$. Construct the corresponding $(1-2\lambda)$ percent confidence interval for consumer surplus as $\{\hat{c} + t^{(l)}\hat{s}, \hat{c} + t^{(1-l)}\hat{s}\}$.

The bootstrap percentile intervals approach would mimic Steps 1-4 in the foregoing procedure. After Step 4, we do the following.

Step 5b: Repeat Steps 1-4 many times, say, *B* times, storing the simulated distributions of c^*'s.

Step 6b: Order the distribution of c^*'s from smallest to largest. The $(1–2\lambda)$ percent confidence interval is $\{c^*_{(l*B)}, c^*_{([1-l]*B)}\}$, where $c^*_{(i)}$ is the *i*th order statistic.

Among the advantages of the bootstrap t-interval procedure is that it has faster convergence properties than the bootstrap percentile interval technique (Efron and Tibshirani 1993; Shao and Tu 1995). Also, the t^*'s are asymptotically pivotal.[8] Disadvantages relative to the bootstrap percentile interval include that the t-interval approach is computationally more intensive (the t-interval approach requires $B \cdot BB$ bootstrap replications while the percentile interval approach requires only B replicates). Also, the percentile intervals are transformation-respecting while the t-intervals are not.

Bias-Corrected Bootstrap Intervals for Welfare Estimates

Another bootstrap procedure that may be relevant in bounding welfare estimates in the travel-cost model setting is the bias-corrected bootstrap intervals approach. As Li and Maddala (1999) stress, the notion of "bias correction" has been used differently in the bootstrap literature. Li and Maddala suggest that the bootstrap t-interval is easier to use and offers the same improvement as Efron's (1979) bias-corrected bootstraps (though this assertion has been challenged).

The "bias-corrected" method we will consider in this study is similar in nature to Killian's (1998) bootstrap-after-bootstrap procedure. The modification to the bootstrap procedure previously outlined is as follows:

Step 5c: Repeat Steps 1-4 many times, say, *B* times, storing the $\hat{c}^*$'s. Calculate the estimate of bias as $\Psi^* = \frac{1}{B}\sum_{b=1}^{B}\left(\hat{c}^*_b - \hat{c}\right)$.

Step 6c (bootstrap after bootstrap): Take *n* draws (independently and with replacement giving equal likelihood to each observation selected on any draw) from the OLS residuals formed in Step 1.

Step 7c: Form the simulated dataset according to $y_i^{**} = \hat{a} + \hat{b} \cdot P_i + \hat{g}_i \cdot M_i + e_i^{**}$, where e_i^{**} is the *i*th draw performed in Step 6c.

Step 8c: Using the simulated trip data in Step 7c, estimate by means of OLS the model parameters (calling them $\{\hat{a}^{**}, \hat{b}^{**}, \hat{g}^{**}\}$). Also, estimate the bias corrected welfare measure as $\hat{c}^{**} = -\frac{y_i^{**2}}{2\hat{b}^{**}} + \Psi^*$.

Step 9c: Repeat Steps 6c-8c many times, say, *BB* times, storing the $\hat{c}^{**}$'s.

Step 10c: Order the distribution of c^{**}'s from smallest to largest. The $(1\text{–}2\lambda)$ percent confidence interval is $\{c^{**}_{(l*BB)}, c^{**}_{([1-l]*BB)}\}$, where $c^{**}_{(i)}$ is the *i*th order statistic.

This bias-corrected approach will be slightly more computationally intensive than the percentile approach and considerably less intensive than the t-interval approach. The number of bootstrap replicates in the bias-corrected approach is $B + BB$.

Conclusions and Extensions

The goal of this research is to estimate accurate bounds on Clear Lake access by Iowans on the basis of a sample collected over the period 2000-2001. An examination of the literature pertaining to travel-cost models indicates that there are many issues of concern. First, the sample was collected on site. To assuage this issue, we adopt the negative binomial II model from the count data model literature. However, additional perils are encountered. Specifically, there do seem to be issues concerning the overdispersion of the empirical dataset. We suggest that further investigation into more general models allowing for fast decay is appropriate.

Next, we consider the implication that in the applied setting the researcher does not know the underlying demand specification that has generated the data. Also, past studies have indicated that the functional-form parameterization imparts value on the consumer surplus estimate (Kling 1989). To rid ourselves of this problematic assumption, we consider the FFF. As the Monte Carlo evidence of Creel (1997) and Cooper (2000) suggest, it appears that the FFF allows us to estimate consumer surplus accurately under varying true model specifications. Past work in recreation demand has not considered gains from modifying the adaptive rule in truncating the FFF. Thus, we present an innovation to the literature that allows for more flexible specification of trip demand.

As past studies have detailed the potential for straightforward estimates of consumer surplus to be biased, traditional confidence intervals are certainly not appropriate (and perhaps neither are traditional bootstrap techniques from the recreation demand literature).[9] Hence, we motivate the creation of a pivotal statistic bootstrap. The bootstrap technique alleviates the bias of parametric approaches.

This study explores the applicability of combining seminonparametric techniques with count data models in estimating resource value. As new issues emerge in estimating particular parameterizations of the model, it appears that refinements to the stochastic specification of the probability model are appropriate with this dataset. It does appear likely that these refinements will yield reasonable estimates with significantly less structural imposition from the researcher. In addition, future research could include performing pseudo–Monte Carlo experiments with an empirical dataset, similar to those of Cooper (2000) and Creel (1997), to assess the robustness of this technique.

Endnotes

1. Fundamentally, caution is appropriate because welfare estimates involve nonlinear transformations of the model parameters. For a discussion of this issue see Kling and Sexton 1990.

2. At this level of generality, we will not impose any distribution on the stochastic noise terms e_i. This is consistent with the nonparametric interpretation of the bootstrap techniques.

3. For a more comprehensive discussion of the properties of bootstrap techniques, see Efron and Tibshirani 1993, and Shao and Tu 1995.

4. This is consistent with Egan and Herriges (2003).

5. An investigation into the asymptotic properties of this approach is warranted.

6. Some have confused the notion of "overfitting" in seminonparametric settings to imply improvements in goodness-of-fit measures. However, the concept of overfitting in the seminonparametric literature actually means the fitting of equations that display nonlinearities not supported by the data. In general, if goodness-of-fit measures improve, then this suggests the data *is* supportive of the nonlinear estimations. Hence, high goodness of fit is not indicative of "overfitting." See, e.g., Jefferys et al. 2001.

7. This is a necessary step to statistically determine if the on-site sample corrections have an impact on our valuation estimates.

8. A statistic is "pivotal" if its distribution does not depend on the value of the population parameter. See Hartigan 1986 for a discussion on the importance of pivotal statistics.

9. Li and Maddala (1999) point out that several applications of the bootstrap in economic models fail to provide improvements over parametric techniques. They suggest that many of these findings occur because the bootstrap techniques considered do not explore techniques that offer the greatest improvements in efficiency. This is exactly what we do in the bootstrap methodologies explored here.

References

Adamowicz, W.L., J.J. Fletcher, and T. Graham-Tomasi 1989. "Functional Form and the Statistical Properties of Welfare Measures." *American Journal of Agricultural Economics* 71: 414-21.

Azevedo, C.D., J.A. Herriges, and C.L. Kling. 2001. "Valuing Preservation and Improvements of Water Quality in Clear Lake." CARD Staff Report 01-SR 94. Center for Agricultural and Rural Development, Iowa State University. March.

Bockstael, N.E., and I.E. Strand. 1987. "The Effect of Common Sources of Regression Error on Benefit Estimates." *Land Economics* 63: 11-20.

Bockstael, N.E., W.M. Hanemann, and I.E. Strand. 1986. "Measuring the Benefits of Water Quality Improvements Using Recreation Demand Models." Final Report to the Environmental Protection Agency, EPA Contract No. CR-811043-010-0.

Cameron, A.C., and P.K. Trivedi. 1990. "Regression-based Tests for Overdispersion in the Poisson Model." *Journal of Econometrics* 347-65.

Creel, M.D. 1997. "Welfare Estimation Using the Fourier Form: Simulation Evidence for the Recreation Demand Case." *Review of Economics and Statistics* 79: 88-94.

Cooper, J.C. 2000. "Nonparametric and Semi-Nonparametric Recreational Demand Analysis." *American Journal of Agricultural Economics* 82: 451-62.

Eastwood, B.J., and A.R. Gallant. 1991. "Adaptive Rules for Seminonparametric Estimators that Achieve Asymptotic Normality." *Econometric Theory* 7: 307-40.

Efron, B. 1979. "Bootstrap Methods: Another Look at the Jackknife." *Annals of Statistics* 7: 1-26.

Efron, B., and R.J. Tibshirani. 1993. *An Introduction to the Bootstrap*. New York: Chapman & Hall.

Egan, K., and J.A. Herriges. 2003. "Mixed Poisson Regression Models with Individual Panel Data from an On-Site Sample." Unpublished manuscript. Department of Economics, Iowa State University.

Englin, J., and J.S. Shonkwiler. 1995. "Estimating Social Welfare Using Count Data Models: An Application to Long-Run Recreation Demand under Conditions of Endogenous Stratification and Truncation." *Review of Economics and Statistics* 77: 104-12.

Fisher, D., A.R. Fleissig, and A. Serletis. 2001. "An Empirical Comparison of Flexible Demand System Functional Forms." *Journal of Applied Econometrics* 16: 59-80.

Gallant, A.R. 1981. "On the Bias in Flexible Functional Forms and an Essentially Unbiased Form: The Fourier Flexible Form." *Journal of Econometrics* 15: 211-45.

Graham-Tomasi, T., W.L. Adamowicz, and J.J. Fletcher. 1988. "On Approximations to Expected Consumer Surplus." Staff Paper No. P88-44, Department of Agricultural and Applied Economics, University of Minnesota.

Greene, W.H. 2000. *Econometric Analysis*, 4th ed. Upper Saddle River, NJ: Prentice Hall.

Gurmu, S., and P.K. Trivedi. 1994. "Recent Developments in Event Count Models: A Survey." Discussion Paper No. 261, Thomas Jefferson Center, Department of Economics, University of Virginia.

Hartigan, J.A. 1986. "Comment on the Paper by Efron and Tibshirani." *Statistical Science* 1: 75-76.

Jefferys, W.H., T.G. Barnes, R. Rodrigues, J.O. Berger, and P. Muller. 2001. "Model Selection for Cepheid Star Oscillations." In *International Society for Bayesian Analysis Proceedings.*

Kealy, M.J., and R.C. Bishop. 1986. "Theoretical and Empirical Specification Issues in Travel Cost Demand Studies." *American Journal of Agricultural Economics* 68: 660-67.

Killian, L. 1998. "Small-Sample Confidence Intervals for Impulse Response Functions." *Review of Economics and Statistics* 80: 218-30.

Kling, C.L. 1989. "The Importance of Functional Form in the Estimation of Welfare." *Western Journal of Agricultural Economics* 14: 161-68.

Kling, C.L., and R.J. Sexton. 1990. "Bootstrapping in Applied Welfare Analysis." *American Journal of Agricultural Economics* 72: 406-18.

Li, H., and G.S. Maddala. 1999. "Bootstrap Variance Estimation of Nonlinear Functions of Parameters: An Application to Long-Run Elasticities of Energy Demand." *Review of Economics and Statistics* 81: 728-33.

McConnell, K.E. 1985. "The Economics of Outdoor Recreation." In *Handbook of Natural Resource and Energy Economics*, vol. 1. Edited by A.V. Kneese and J.L. Sweeney. Amsterdam: North-Holland.

Sarker, R., and Y. Surry. 2003. "The Fast Decay Process in Outdoor Recreational Activities and the Use of Alternative Count Data Models." Forthcoming in *American Journal of Agricultural Economics*.

Shao, J., and D. Tu. 1995. *The Jackknife and Bootstrap*. Springer Series in Statistics, edited by I. Olkin. New York: Springer.

Shaw, D. 1988. "On-Site Samples' Regression: Problems of Non-Negative Integers, Truncation and Endogenous Stratification." *Journal of Econometrics* 37: 211-33.

Smith, V.K. 1989. "Nearly All Consumer Surplus Estimates Are Biased." Working paper, Department of Economics and Business, North Carolina State University.