THE SENSITIVITY OF
CONSUMER SURPLUS ESTIMATION
TO FUNCTIONAL FORM SPECIFICATION

Duangkamon Chotikapanich and William E. Griffiths

No. 94 - December 1996

ISSN 0 157 0188
ISBN 1 86389 399 7
THE SENSITIVITY OF CONSUMER SURPLUS ESTIMATION
TO FUNCTIONAL FORM SPECIFICATION

Duangkamon Chotikapanich
Department of Economics
Curtin University of Technology

William E. Griffiths
Department of Econometrics
University of New England
The Sensitivity of Consumer Surplus Estimation
to Functional Form Specification

Abstract

The travel cost method utilized by Beal (1995) for estimating consumer surplus is re-examined. An alternative more algebraically consistent estimation strategy is suggested. We demonstrate that estimation of consumer surplus is inherently unreliable when only a small number of observations are available, and when the consumer surplus estimate depends critically on extrapolation of the demand curve beyond the sample observations.
1. INTRODUCTION

In a recent paper in the *Review of Marketing and Agricultural Economics*, Beal (1995) assesses the value of Carnarvon Gorge National Park for recreational use by estimating the consumer surplus associated with selected demand equations. She was kind enough to provide us with her data so that we could use her model as one of our examples in a paper we are writing on using Bayesian techniques for estimating areas in economics. One of the things that struck us when we looked more closely at her paper was how sensitive consumer surplus estimates can be to functional form specification. In addition, there were some methodological issues which we thought could benefit from further discussion. These issues are (i) provision of standard errors for consumer surplus estimates, (ii) choice of functional form, and (iii) prediction in log-log models.

To provide what we hope will be regarded as useful insights on these questions, we have organized this paper as follows. In the next section we summarize the Beal methodology for estimating consumer surplus associated with the use of the National Park. We point out that the 2-step, 2-equation strategy used by Beal is internally inconsistent and we estimate consumer surplus using an alternative more consistent strategy. We also indicate how a standard error for the consumer surplus estimate can be calculated. The large discrepancy between the consumer surplus estimates from the two strategies made us wonder how sensitive such estimates are to the functional form specification. So, in the section that follows, we estimate the consumer surplus implied by each of the functional forms tested by Beal. We also compute standard errors of the estimates and give the maximized log-likelihood values for each of the functions. These values are likely to be a better guide to functional form selection than the $R^2$’s used by Beal. In a final section, we comment on prediction in log-log models; our comments help explain why Beal’s equation seemed to give a biased prediction of total visits for a zero price.
2. AN ALTERNATIVE CONSUMER SURPLUS ESTIMATE

2.1 Beal's Estimation Technique

Beal estimates two demand functions, a camping demand function and a demand function for day visits. We restrict our discussion to the camping demand curve. The points we make are equally valid for the demand for day visits.

Potential visitors to the park are divided into 12 geographical zones. Associated with travel to the park from each zone is a travel and time cost variable \((TC_i)\) that takes the place of price in a conventional demand equation. A demand equation for camp use is estimated by relating the \(i\)-th zone's average annual visitation rate per thousand population \((V_i)\) to travel cost, say,

\[
V_i = f(TC_i) \quad i = 1,2,\ldots,12
\]

(2.1)

The relationship between aggregate demand \(Q\) and the travel costs for each zone is given by

\[
Q = \sum_{i=1}^{12} pop_i V_i = \sum_{i=1}^{12} pop_i f(TC_i)
\]

(2.2)

where \(pop_i\) is the population of the \(i\)-th zone. Because it represents just one point on an aggregate demand curve, equation (2.2) by itself cannot be used to estimate consumer surplus. It describes the existing demand that relates to the existing range of zonal travel costs. The "travel cost method" overcomes this problem by assuming that consumers of recreation react to changes in a hypothetical entry price \((P)\) in the same way as they would to changes in travel costs. If such is the case, augmenting the travel cost variable with an entry price variable yields the function

\[
Q = \sum_{i=1}^{12} pop_i f(TC_i + P)
\]

(2.3)

which can be used to trace out a complete demand function where aggregate demand \(Q\) is related to price \(P\). The relevant consumer surplus measure in Beal's study is the area under this aggregate demand function.

The first point we make in this paper is that Beal's specification for the aggregate demand function in (2.3) is algebraically inconsistent with her choice of the functional form for \(f(TC_i)\). In addition, specifying (2.3) in an algebraically consistent
way has a considerable effect on the estimate of consumer surplus. More explicitly, after some preliminary testing, Beal chooses the log-log equation

$$\ln(\hat{V}_i) = \hat{\beta}_0 + \hat{\beta}_1 \ln(TC_i)$$

(2.4)

for the visitation rate demand curve. The estimates obtained were $\hat{\beta}_0 = 9.8426$ and $\hat{\beta}_1 = -1.8743$. Equation (2.1) becomes, therefore,

$$\hat{V}_i = \exp\left\{\hat{\beta}_0 + \hat{\beta}_1 \ln(TC_i)\right\} = e^{\hat{\beta}_0} (TC_i)^{\hat{\beta}_1}$$

(2.5)

and the corresponding aggregate demand curve, with the entry price variable $P$ included, becomes

$$Q = e^{\hat{\beta}_0} \sum_{i=1}^{12} pop_i (TC_i + P)^{\hat{\beta}_1}$$

(2.6)

If equation (2.5) is the legitimate function to use for $V_i$, then, presumably, the relevant consumer surplus measure is the area under the aggregate demand function in equation (2.6). However, this was not the function utilized by Beal. Instead, she used equation (2.6) to predict values for $Q$ for 7 selected values of $P$. An 8th value was the observed value for $Q$ when $P=0$. These values of $Q$ and $P$ were used to estimate a number of alternative aggregate demand functions from which the linear function

$$\hat{Q} = 15764.7 - 252.09 P$$

(2.7)

was chosen. Based on this estimated demand curve, the consumer surplus, which is the total area under the curve, was estimated as 492,931.

2.2 An Alternative Consumer Surplus Estimate

From section 2.1 it can be seen that Beal chose to approximate the aggregate exponential demand function (2.6) with the linear function (2.7) and to use the latter to calculate the consumer surplus. An alternative way which is internally more consistent is to calculate the consumer surplus directly from the total area under the exponential demand function (2.6). Figure 1 shows the two curves drawn from (2.6) and (2.7). It can be seen that the linear function is only a reasonable approximation for prices up to $20 and that, by using the linear function to approximate the exponential one, we largely underestimate the total area under the curve.
To calculate the area under the exponential curve (2.6) we need to be aware that this exponential curve is asymptotic to the price axis. That is, as \( P \to \infty, Q \to 0 \).

The area under the curve is given by

\[
\hat{CS} = e^{\hat{\beta}_0} \lim_{P^* \to \infty} \int_{0}^{P^*} \left( \sum_{i} (pop_i)(TC_i + P) \right) dP
\]

\[
= \left( \frac{e^{\hat{\beta}_0}}{\hat{\beta}_1 - 1} \right) \lim_{P^* \to \infty} \left[ \sum_{i} (pop_i) \left\{ (TC_i + P^*)^{\hat{\beta}_1+1} - TC_i^{\hat{\beta}_1+1} \right\} \right]
\]

Since \( \hat{\beta}_1 \) is less than 1, \( \lim_{P^* \to \infty} (TC_i + P^*)^{\hat{\beta}_1+1} \) is zero and the \( \hat{CS} \) becomes

\[
\hat{CS} = \frac{e^{\hat{\beta}_0}}{\hat{\beta}_1 + 1} \sum_{i} (pop_i) (TC_i)
\]

The consumer surplus calculated from this definition is 2,954,560. By using the linear function to approximate the exponential function the consumer surplus has been underestimated by as much as 6 times.

2.3 Standard Error for the Consumer Surplus

When estimating any quantity it is important to have some idea of the reliability of that estimate. One indicator of reliability is provided by the standard error and a consequent interval estimate. An approximate standard error for the consumer surplus estimate can be obtained from the squared root of the asymptotic variance which is given by (Judge et al 1988, p. 542)

\[
\text{var} (\hat{CS}) = \left( \frac{\partial (CS)}{\partial \beta_0} \right)^2 \text{var} (\hat{\beta}_0) + \left( \frac{\partial (CS)}{\partial \beta_1} \right)^2 \text{var} (\hat{\beta}_1) + 2 \left( \frac{\partial (CS)}{\partial \beta_0} \right) \left( \frac{\partial (CS)}{\partial \beta_1} \right) \text{cov} (\hat{\beta}_0, \hat{\beta}_1)
\]

(2.9)

For the consumer surplus defined in (2.8), the derivatives are
\[
\frac{\partial (CS)}{\partial \beta_0} = - \frac{e^\beta_0}{\beta_1 + 1} \sum_i (\text{pop}_i)(TC_i)^{\beta_1+1}
\]

\[
\frac{\partial (CS)}{\partial \beta_1} = \frac{e^\beta_0}{(\beta_1 + 1)^2} \sum_i (\text{pop}_i)(TC_i)^{\beta_1+1} - \frac{e^\beta_0}{(\beta_1 + 1)} \sum_i (\text{pop}_i)(TC_i)^{\beta_1+1} \ln(TC_i)
\]

Using \( \text{vâr}(\hat{\beta}_0) = 1.1005, \text{vâr}(\hat{\beta}_1) = 0.0377, \text{cov}(\hat{\beta}_0, \hat{\beta}_1) = -0.2016 \) and evaluating the derivatives at \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \), the standard error calculated for the consumer surplus (2.8) is 792,214. This value, along with the estimated consumer surplus gives a 95% confidence interval as \( (1,189,507, 4,719,613) \). This very wide interval shows there is considerable uncertainty associated with our consumer surplus estimate.

3. SENSITIVITY OF CONSUMER SURPLUS ESTIMATION

The enormous difference between the estimate of consumer surplus from Beal’s linear aggregate demand function and the estimate implied by her visitation rate function raises questions about the sensitivity of consumer surplus estimation to functional form specification. In this section, we look at the consumer surplus implied by the other five functional forms of the demand function that were considered by Beal. For each of the functional forms, we provide the point estimate, the standard error and the 95% interval estimate of the consumer surplus. A comparison of these results reveals the critical nature of functional form selection. Also, for a goodness-of-fit comparison, we give values of the maximized log-likelihoods under an assumption of normally distributed errors. Since the six functional forms do not have the same dependent variables, the maximized log-likelihood provides a better basis for comparison than the \( R^2 \)'s used by Beal.

3.1 Consumer Surplus Estimates and their Standard Errors

The procedure for obtaining the consumer surplus estimated in the previous section can be summarized as follows. If the zonal demand is the following function

\[ V_i = f(TC_i) \] (3.1)
then the aggregate demand for the site is

$$ Q = \sum_i \left( pop_i \right) f \left( T C_i + P \right) \quad (3.2) $$

And consumer surplus is defined as

$$ CS = \int_0^{P^*} \sum_i \left( pop_i \right) f \left( T C_i + P \right) dP \quad (3.3) $$

where $P^*$ is that price for which $Q = 0$ in equation (3.2). If the demand function in equation (3.2) is asymptotic to the price axis, then it is necessary to take limit of equation (3.3) as $P^* \to \infty$. The expression for the estimated variance of $CS$ from which we can obtain standard errors is given in equation (2.9). Specific expressions for $V, Q$ and $CS$ for each of the functional forms considered by Beal are summarized in the Appendix, together with the derivatives of $CS$ with respect to $\beta_0$ and $\beta_1$. These derivatives are used in the calculation of the standard error of the estimated consumer surplus.

Table 1 reports the values of the point estimates, their standard errors and corresponding 95% interval estimates of consumer surplus obtained from the different functional forms. The interval estimates are derived assuming, in each case, that $\hat{CS}$ has a limiting distribution that is normal. It should be noted that the demand curve for equation 4 is asymptotic to the price axis, and, as price approaches infinity, the curve does not approach the axis quickly enough for the area to be finite. The most striking thing about the entries in Table 1 is that the estimated $CS$'s obtained from the different functional forms vary considerably, from 442,324 to 4,105,566. Also, from the large standard errors and the wide 95% interval estimates, we see that all consumer surplus estimates are very unreliable, even if the underlying functional form is known. This is particularly the case for equations 1, 2 and 5 where the 95% interval estimates contain meaningless negative ranges. In the sampling theory approach to inference (as distinct from Bayesian inference), there does not appear to be any simple way of incorporate prior information to ensure that interval estimates include only a positive range. Also, it appears that a small degree of uncertainty or unreliability in the estimation of the parameters $\beta_0$ and $\beta_1$ can lead to a large degree of uncertainty in the estimation of $CS$. Very few of Beal's estimates for $\beta_0$ and $\beta_1$ are not significantly different from zero, yet three out of five of the $CS$ estimates derived from them have standard errors
that are bigger than the estimates themselves. It is clear that any valuation of the Carnarvon Gorge National Park on the basis of consumer surplus estimation is particularly tenuous.

TABLE 1

<table>
<thead>
<tr>
<th>Functional Form</th>
<th>$\hat{CS}$ (se($\hat{CS}$))</th>
<th>95% interval estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>equation 1</td>
<td>4,105,566 (5,933,944)</td>
<td>(-7,524,965) - (15,736,100)</td>
</tr>
<tr>
<td>$V = \beta_0 + \beta_1 TC$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>equation 2</td>
<td>1,002,818 (1,742,918)</td>
<td>(-2,413,301) - (4,418,937)</td>
</tr>
<tr>
<td>$V = \beta_0 + \beta_1 \ln TC$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>equation 3</td>
<td>2,875,357 (983,023)</td>
<td>(948,632) - (4,802,082)</td>
</tr>
<tr>
<td>$\ln V = \beta_0 + \beta_1 TC$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>equation 4</td>
<td>$\infty$</td>
<td>$\infty$</td>
</tr>
<tr>
<td>$1 / V = \beta_0 + \beta_1 TC$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>equation 5</td>
<td>442,324 (1,442,770)</td>
<td>(-2,385,506) - (3,270,153)</td>
</tr>
<tr>
<td>$V = \beta_0 + \beta_1 (1 / TC)$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>equation 6</td>
<td>2,954,560 (792,214)</td>
<td>(1,401,821) - (4,507,299)</td>
</tr>
<tr>
<td>$\ln V = \beta_0 + \beta_1 \ln TC$</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.2 Comparing Functional Forms

As mentioned earlier, the use of $R^2$ as a criterion for model selection is questionable for models with different dependent variables. An alternative, which makes comparisons in terms of comparable units of measurement, is the value of the maximised log-likelihood function. We are not suggesting that choosing the model that has the largest maximized log-likelihood is a foolproof model selection criterion. Such a procedure is subject to sampling error and does not consider whether the data “significantly” favour one model over another (see, for example, Box and Cox, 1964
and Fisher and McAleer, 1981). However, as a descriptive goodness-of-fit measure, it is preferable to $R^2$. To give an expression for the maximized log-likelihood consider the equation:

$$g_1(V) = \beta_0 + \beta_1 g_2(TC) + e$$

where it is assumed that $e$ is an independent normal random variable with mean zero and variance $\sigma^2$, and $g_1(V)$ and $g_2(TC)$ are functions of $V$ and $TC$, respectively. For example, for equation 2, $g_1(V) = \ln V$ and $g_2(TC) = TC$. After substituting maximum likelihood estimators for $\beta_0$, $\beta_1$ and $\sigma^2$ into the log-likelihood function, this function becomes

$$L = -\frac{n}{2} \left[ \ln(2\pi) + 1 + \ln(n) \right] - \frac{n}{2} \ln(SSE) + \sum \ln \left( \left| \frac{\partial g_1(V)}{\partial V} \right| \right)$$

where $n = \text{total number of observations}$, $SSE = \text{sum of squared errors}$, and $\left| \frac{\partial g_1(V)}{\partial V} \right|$ is the Jacobian of the transformation from $g_1(V)$ to $V$. It is the presence of this last term that makes a ranking of models on the basis of $L$ possibly different from a ranking of models on the basis of $R^2$.

Table 2 reports the values of the log-likelihood function, along with the $R^2$s from Beal’s paper. Based on these values of the maximized log-likelihood functions, equation 6 is the best equation, a choice that is consistent with that of Beal. However, the ranking of the other models has changed.

<table>
<thead>
<tr>
<th>Functional Form</th>
<th>$R^2$</th>
<th>$L$</th>
</tr>
</thead>
<tbody>
<tr>
<td>eqn 1: $V = \beta_0 + \beta_1 TC$</td>
<td>0.23</td>
<td>-34.1261</td>
</tr>
<tr>
<td>eqn 2: $V = \beta_0 + \beta_1 \ln TC$</td>
<td>0.56</td>
<td>-30.7955</td>
</tr>
<tr>
<td>eqn 3: $\ln V = \beta_0 + \beta_1 TC$</td>
<td>0.62</td>
<td>-14.1678</td>
</tr>
<tr>
<td>eqn 4: $1/V = \beta_0 + \beta_1 TC$</td>
<td>0.86</td>
<td>-13.5402</td>
</tr>
<tr>
<td>eqn 5: $V = \beta_0 + \beta_1 (1/TC)$</td>
<td>0.83</td>
<td>-25.0763</td>
</tr>
<tr>
<td>eqn 6: $\ln V = \beta_0 + \beta_1 \ln TC$</td>
<td>0.90</td>
<td>-5.9777</td>
</tr>
</tbody>
</table>
4. PREDICTION IN LOG-LOG MODELS

One last point of interest is the characteristics of prediction in log-log models and their implications for consumer surplus estimation. As mentioned in Section 2, Beal uses a 2-stage procedure where, to begin the second stage, $\hat{Q}_j$ is obtained using the predictive model

$$\hat{Q}_j = \sum_i \text{pop}_i \hat{V}_i$$

with $\hat{V}_i$ based on equation 6 (the log-log model), and given by the expression

$$\hat{V}_i = \exp\{\hat{\beta}_0 + \hat{\beta}_1 \ln(TC_i + P_i)\}$$

(4.1)

She notes that the result obtained from this predictor for $P_j = 0$ is 14,843. This value largely underestimates the true value of 17,000. A possible reason for this outcome is that the predictor $\hat{V}_i$ is not an unbiased predictor. Consider the zonal demand function

$$\ln(V_i) = \beta_0 + \beta_1 \ln(TC_i) + e_i$$

where we assume $e_i \sim N(0, \sigma^2)$ and hence

$$\ln(V_i) \sim N(\beta_0 + \beta_1 \ln(TC_i), \sigma^2)$$

$$V_i \sim \log N(\beta_0 + \beta_1 \ln(TC_i), \sigma^2)$$

From the properties of lognormal distributions (Atchison and Brown, 1966), the mean of $V_i$ is

$$E(V_i) = \exp\{\beta_0 + \beta_1 \ln(TC_i) + \frac{1}{2} \sigma^2\}$$

Thus, an unbiased predictor for $V_i$ is

$$\tilde{V}_i = \exp\{\beta_0 + \beta_1 \ln(TC_i + P_j) + \frac{1}{2} \sigma^2\}$$

(4.2)

When $\beta_0$, $\beta_1$ and $\sigma^2$ are replaced by their estimators, $\tilde{V}_i$ no longer retains its unbiasedness property, but it is still likely to be a better predictor than $\hat{V}_i$. Table 3 shows the new values of $\tilde{Q}_j$ obtained using the feasible version of $\tilde{V}_i$, alongside those obtained by Beal. Note that the actual number of visits demanded $Q_j$ at zero additional entry fee ($P_j = 0$) is 17,000. The new prediction for $Q_j$ is very accurate
with a value of $\tilde{Q}_j = 17,003$. If a new linear function is estimated on the basis of those new values of $\tilde{Q}_j$, the estimated consumer surplus is $697,771$. This value emphasizes once more how sensitive consumer surplus estimation can be to choice of methodology.

### Table 3

*Alternative Predictions from the Log-Log Model*

<table>
<thead>
<tr>
<th>P</th>
<th>$\tilde{Q}$</th>
<th>Beal's $\hat{Q}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>17,003</td>
<td>17,000*</td>
</tr>
<tr>
<td>5</td>
<td>15,735</td>
<td>13,736</td>
</tr>
<tr>
<td>8</td>
<td>15,068</td>
<td>13,156</td>
</tr>
<tr>
<td>10</td>
<td>14,656</td>
<td>12,769</td>
</tr>
<tr>
<td>12</td>
<td>14,267</td>
<td>12,456</td>
</tr>
<tr>
<td>15</td>
<td>13,723</td>
<td>11,981</td>
</tr>
<tr>
<td>18</td>
<td>13,219</td>
<td>11,540</td>
</tr>
<tr>
<td>20</td>
<td>12,904</td>
<td>11,265</td>
</tr>
</tbody>
</table>

* Observed value

**CS** 697,771 492,931

### 4.1 Stochastic Considerations in Consumer Surplus Estimation

The possible use of properties of the log-normal distribution to predict quantity for Beal's second stage of estimation raises questions about whether such properties should be used for the more algebraically consistent procedure that we proposed. The underlying issue here is how to treat the stochastic error term $e$ in the visitation rate demand function $\ln V_i = \beta_0 + \beta_1 \ln TC_i + e_i$. One approach is to ignore it. This is the approach taken by Beal and by us so far. Another approach is to carry it through to the aggregate demand function and consider the estimation of consumer surplus on the basis of the average or expected demand function. In this case the function becomes

$$\hat{Q} = \sum_i pop_i \exp\left\{\beta_0 + \beta_1 \ln(TC_i + P) + e_i\right\}$$

(4.3)

And, if we assume $e_i \sim N(0, \sigma^2)$, then the expected aggregate demand function is

$$E(\hat{Q}) = \sum_i pop_i \exp\left\{\beta_0 + \beta_1 \ln(TC_i + P) + \frac{1}{2} \sigma^2\right\}$$

(4.4)
Consumer surplus is obtained by integrating $E(Q)$ in (4.4) between 0 and $P^*$ and taking the limit of the integral as $P^* \to \infty$.

A third approach is to first obtain consumer surplus by integrating $Q$ in (4.3) between 0 and $P^*$ and taking the limit as $P^* \to \infty$. The resulting $CS$ depends on the error term $e_i$. Taking the expected value of this result gives an expected or average $CS$. The two approaches are conceptually different. One measures $CS$ as the area under the "average" demand curve. The other recognizes that different errors lead to different demand curves that lead to different consumer surpluses, and then takes the average $CS$. Despite this conceptual difference, for the model we are considering, both approaches lead to the same result, namely

$$\hat{CS} = -\frac{e}{\beta_1 + 1} \sum_i pop_i T C_{i}^{\beta_1 + 1}$$

To derive the standard error of a $CS$ estimate based on this expression would involve partial derivatives of $CS$ with respect to $\beta_0$, $\beta_1$ and $\sigma^2$ as well as the variance of $\hat{\sigma}^2$.

5. CONCLUSIONS

It is evident from the results in this paper that estimating consumer surplus can be tricky. The results can be very sensitive to demand function specification and to the chosen estimation methodology. The purpose of this paper was to expose this sensitivity as well as to suggest ways in which the methodology could be improved by computing standard errors and developing consumer surplus estimates which are internally consistent with visitation rate equations. There are two general observations that we would also like to make.

One of likely reasons for the sensitivity of Beal's results to model specification, and the impreciseness of the consumer surplus estimates as reflected in the width of interval estimates, is the small number of observations. With only 12 observations it is difficult to estimate most things accurately. In Beal's case, however, the 12 observations were obtained via an aggregation of individual responses. The effect of such an aggregation may have been a loss of information. A possible direction for the future that might yield better, more reliable estimates is through the development and estimation of models that utilize single observations on individuals rather than
observations that are computed by aggregating information. Also, it would be good to get consumer surplus estimates and reliability measures that are not dependent on functional form selection. We are now to investigate a Bayesian approach that may yield some success along this lines.

Another fundamental difficulty associated with estimating consumer surplus from models of the type considered here is the need to extrapolate the demand function beyond the region of the observed sample. The problem is well depicted in Figure 1. While any number of functions may fit the data reasonably well in the region of the observed data points, it is the behaviour of the function beyond these data points that is critical for the estimation of consumer surplus. Since the data contain no information on this behaviour, the most critical part of the estimation procedure is fraught with uncertainty. While this uncertainty is reflected by the sensitivity of the consumer surplus estimates to functional form specification, it is not reflected in the width of an interval estimate for a given functional form. The fact that these interval estimates are very wide means that estimation of consumer surplus is subject to a great deal of uncertainty even if we have perfect knowledge of the correct functional form. However, as mentioned in the previous paragraph, it may be possible to correct this problem with a larger number of observations.
APPENDIX

The expressions for $V$, $Q$, $P^*$ and $CS$ for each of the functional forms considered by Beal are summarized below, together with the derivatives of $CS$ with respect to $\beta_0$ and $\beta_1$. These derivatives are used in the calculation of the standard error of the $CS$ in equation (3.3). In some cases (equations 2 and 5) where it is not possible to derive an analytical expression for $P^*$, its value was obtained numerically. For these cases, the derivatives of $CS$ with respect to $\beta_0$ and $\beta_1$ depend on the derivatives of $P^*$ with respect to $\beta_0$ and $\beta_1$. These derivatives of $P^*$ were obtained using the differentiation of implicit functions.

Equation 1

\[ \dot{V}_i = \hat{\beta}_0 + \hat{\beta}_1 TC_i \]

\[ \dot{Q} = \sum_i (\text{pop}_i) \left( \hat{\beta}_0 + \hat{\beta}_1 (TC_i + P) \right) \]

\[ P^* = \frac{-\hat{\beta}_0}{\hat{\beta}_1} - \frac{\sum_i \text{pop}_i TC_i}{\sum_i \text{pop}_i} \]

\[ \dot{CS} = \hat{\beta}_0 P^* \sum_i \text{pop}_i + \hat{\beta}_1 P^* \sum_i \text{pop}_i TC_i + \frac{1}{2} \hat{\beta}_1 (P^*)^2 \sum_i \text{pop}_i \]

\[ \frac{\partial \dot{CS}}{\partial \hat{\beta}_0} = \frac{-\hat{\beta}_0 \sum_i \text{pop}_i}{\hat{\beta}_1} - \sum_i \text{pop}_i TC_i \]

\[ \frac{\partial \dot{CS}}{\partial \hat{\beta}_1} = \frac{\hat{\beta}_0^2}{2\beta_1^2} \sum_i \text{pop}_i - \frac{\left( \sum_i \text{pop}_i TC_i \right)^2}{2\sum_i \text{pop}_i} \]

Equation 2

\[ \dot{V}_i = \hat{\beta}_0 + \hat{\beta}_1 \ln TC_i \]

\[ \dot{Q} = \sum_i (\text{pop}_i) \left[ \hat{\beta}_0 + \hat{\beta}_1 \ln (TC_i + P) \right] \]
\( P^* = 90.8365 \) is obtained numerically as the solution to

\[
h(\hat{\beta}_0, \hat{\beta}_1, P^*) = \sum_i (\text{pop}_i) [\hat{\beta}_0 + \hat{\beta}_1 \ln(\text{TC}_i + P^*)] = 0
\]

\[
\hat{C}S = \hat{\beta}_0 P^* \sum_i \text{pop}_i + \hat{\beta}_1 \sum_i \text{pop}_i \left\{ (\text{TC}_i + P^*) \ln(\text{TC}_i + P^*) - P^* - \text{TC}_i \ln(\text{TC}_i) \right\}
\]

\[
\frac{\partial \hat{C}S}{\partial \hat{\beta}_0} = P^* \sum_i \text{pop}_i + \hat{\beta}_0 \frac{\partial P^*}{\partial \hat{\beta}_0} \sum_i \text{pop}_i + \hat{\beta}_1 \sum_i \text{pop}_i \ln(\text{TC}_i + P^*) \frac{\partial P^*}{\partial \hat{\beta}_0}
\]

\[
\frac{\partial \hat{C}S}{\partial \hat{\beta}_1} = \sum_i \text{pop}_i \left\{ (\text{TC}_i + P^*) \ln(\text{TC}_i + P^*) - P^* - \text{TC}_i \ln(\text{TC}_i) \right\} + \hat{\beta}_0 \frac{\partial P^*}{\partial \hat{\beta}_1} \sum_i \text{pop}_i
\]

where

\[
\frac{\partial P^*}{\partial \hat{\beta}_0} = -\frac{\partial h/\partial \hat{\beta}_0}{\partial h/\partial P^*} = \frac{-\sum_i \text{pop}_i}{\hat{\beta}_0 \sum_i \text{pop}_i \frac{1}{\text{TC}_i + P^*}}
\]

\[
\frac{\partial P^*}{\partial \hat{\beta}_1} = -\frac{\partial h/\partial \hat{\beta}_1}{\partial h/\partial P^*} = \frac{-\sum_i \text{pop}_i \ln(\text{TC}_i + P^*)}{\hat{\beta}_1 \sum_i \text{pop}_i \frac{1}{\text{TC}_i + P^*}}
\]

Equation 3

\[\ln \hat{V}_i = \hat{\beta}_0 + \hat{\beta}_1 \text{TC}_i\]

\[\hat{Q} = e^{\hat{\beta}_0} \sum_i (\text{pop}_i) e^{\hat{\beta}_1 (\text{TC}_i + P)}\]

As \( \hat{Q} \rightarrow 0, P \rightarrow \infty \). Integrating and taking limits yields

\[
\hat{C}S = -\frac{e^{\hat{\beta}_0}}{\hat{\beta}_1} \sum_i (\text{pop}_i) e^{\hat{\beta}_1 \text{TC}_i}
\]
\[ \frac{\partial \hat{CS}}{\partial \hat{\beta}_0} = -\frac{\hat{\beta}_0}{\beta_1} \sum_i pop_i e^{\hat{\beta}_1 TC_i} \]

\[ \frac{\partial \hat{CS}}{\partial \hat{\beta}_1} = \frac{\hat{\beta}_0}{\beta_1} \sum_i pop_i e^{\hat{\beta}_1 TC_i} - \frac{\hat{\beta}_0}{\beta_1} \sum_i pop_i TC_i e^{\hat{\beta}_1 TC_i} \]

Equation 4

\[ \frac{1}{\hat{\nu}} = \hat{\beta}_0 + \hat{\beta}_1 TC_i \]

\[ \hat{Q} = \sum_i \frac{1}{pop_i} \frac{1}{\hat{\beta}_0 + \hat{\beta}_1 (TC_i + P)} \]

As \( P \to \infty \), \( Q \to 0 \). We find that

\[ \hat{CS} = \lim_{P^* \to \infty} \sum_i \left[ \frac{1}{\hat{\beta}_1} \left( pop_i \ln \{ \hat{\beta}_0 + \hat{\beta}_1 TC_i + \hat{\beta}_1 P^* \} - pop_i \ln \{ \hat{\beta}_0 + \hat{\beta}_1 TC_i \} \right) \right] \]

As \( P^* \to \infty \), \( \ln \{ \hat{\beta}_0 + \hat{\beta}_1 TC_i + \hat{\beta}_1 P^* \} \to \infty \), and \( \hat{CS} \to \infty \).

Equation 5

\[ \hat{\nu}_i = \hat{\beta}_0 + \hat{\beta}_1 \frac{1}{TC_i} \]

\[ \hat{Q} = \sum_i (pop_i) \left( \hat{\beta}_0 + \hat{\beta}_1 \frac{1}{(TC_i + P)} \right) \]

\( P^* = 67.3087 \) is obtained numerically as the solution to

\[ h(\hat{\beta}_0, \hat{\beta}_1, P^*) = \sum_i (pop_i) \left[ \hat{\beta}_0 + \hat{\beta}_1 \left( \frac{1}{(TC_i + P^*)} \right) \right] = 0 \]

\[ \hat{CS} = \hat{\beta}_0 P^* \sum_i pop_i + \hat{\beta}_1 \sum_i pop_i \left[ \ln (TC_i + P^*) - \ln (TC_i) \right] \]
\[
\frac{\partial \hat{C}}{\partial \beta_0} = P^* \sum_i pop_i + \hat{\beta}_0 \frac{\partial P^*}{\partial \beta_0} \sum_i pop_i + \hat{\beta}_1 \sum_i \frac{1}{(TC_i + P^*)} \frac{\partial P^*}{\partial \beta_0}
\]

\[
\frac{\partial \hat{C}}{\partial \beta_1} = \sum_i pop_i \left[ \ln(TC_i + P^*) - \ln TC_i \right] + \hat{\beta}_1 \sum_i \frac{1}{(TC_i + P^*)} \frac{\partial P^*}{\partial \beta_1}
\]

where

\[
\frac{\partial P^*}{\partial \beta_0} = -\frac{\partial h}{\partial \hat{\beta}_0} = \frac{\sum_i pop_i}{\hat{\beta}_1 \sum_i \left( \frac{1}{(TC_i + P^*)^2} \right)}
\]

\[
\frac{\partial P^*}{\partial \beta_1} = -\frac{\partial h}{\partial \hat{\beta}_1} = \frac{\sum_i \left( \frac{pop_i}{(TC_i + P^*)} \right)}{\hat{\beta}_1 \sum_i \left( \frac{1}{(TC_i + P^*)^2} \right)}
\]

REFERENCES


WORKING PAPERS IN ECONOMETRICS
AND APPLIED STATISTICS


A Note on A Bayesian Estimator in an Autocorrelated Error Model. William Griffiths and Dan Dao, No. 3 - April 1979.


Bayesian Econometrics and How to Get Rid of Those Wrong Signs. William E. Griffiths, No. 31 - November 1987.


Bayesian Methodology for Imposing Inequality Constraints on a Linear Expenditure System with Demographic Factors. William E. Griffiths and Duangkamon Chotikapanich, No. 93 - December 1996.