DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

Welfare Economics for Tobit Models

Jeffrey A. Dubin and Louis L. Wilde
California Institute of Technology

SOCIAL SCIENCE WORKING PAPER 753

January 1991

Welfare Economics for Tobit Models

Jeffrey A. Dubin and Louis L. Wilde

Abstract

In this paper we demonstrate the correct calculation of consumer surplus in censored and truncated regression models, focusing on Tobit models. We review a variety of examples from the literature and isolate the nature of the bias associated with the incorrect calculation of consumer surplus in several of them.

Welfare Economics for Tobit Models

Jeffrey A. Dubin and Louis L. Wilde*

1 Introduction

There has been a recent recognition of the importance of censored and truncated regression models for the estimation of the demand for recreation.[1] The use of such models is now commonplace, due in part to the availability of statistical packages which have made methods of maximum likelihood estimation as straightforward as performing a linear regression. Quite naturally, what often follows the estimation of demand is a calculation of consumer surplus or the change in consumer surplus resulting from an increase or decrease in environmental quality.[2] Unfortunately these calculations are not always done correctly. When they are not done correctly, it is generally because they are based on the wrong demand curve. It is the purpose of this paper to demonstrate how to base calculations of consumer surplus in censored and truncated regression models on the right demand curve.

We focus on the Tobit model.[3] In the Tobit model the estimated demand curve represents two linked decisions. The first, the participation decision, is the choice by the individual to participate in an activity or to buy some positive quantity of a good. The second, the usage decision, determines the level of participation in an activity given that the individual participates or the amount of a good which is consumed given that the

*The authors gratefully acknowledge research support from the Environmental Quality Laboratory at the California Institute of Technology and from Putnam, Hayes, and Bartlett, Inc. Helpful comments made by Charlie Cicchetti and David Grether on an earlier draft are gratefully acknowledged.

[1]See for example Bockstael, Strand, McConnell, and Arsanjani (1990) on the application of sample selection methods to recreation demand estimation.

[2]See for example Burt and Brewer (1971) and Cicchetti, Fisher, and Smith (1976) for early examples of consumer surplus calculations undertaken within recreation demand systems.

[3]The Tobit model received its first articulation in the classic study by Tobin (1958). Tobin focused on the estimation of demand subject to the truncation of observed expenditures at zero. His original formulation was concerned with corner solutions in observed demand rather than the linkage between participation and usage decisions by the consumer.

individual consumes some amount. The correct calculation of consumer surplus must account for both of these components of demand.[4]

The Tobit model is a so-called single-index model since only one source of randomness affects both the participation and usage decisions.[5] A related class of multiple-index models that allow separate but dependent sources of randomness for the participation and usage decisions begins with the specification of the conditional indirect utility for a mutually exclusive set of discrete outcomes(Dubin and McFadden, 1984; Hanemann, 1984). The best known theoretical treatment of welfare economics for discrete choice models is due to Small and Rosen (1981).[6] Our theoretical analysis has some features in common with theirs, although, we focus on the calculation of consumer surplus for the individual rather than for population.

In Section 2 we develop the basic theory of consumer surplus when the underlying utility function is stochastic. In Section 3 we review several empirical examples to contrast the approaches employed by different authors in light of the theory as discussed in Section 2. In Section 4 we present the calculation of consumer surplus for Tobit models and in Section 5 we reanalyze some of the empirical examples to demonstrate the direction of bias their results have produced. We conclude in Section 6.

2 Review of Theory

The goal of this section is to demonstrate that for a wide class of models, an appropriate calculation of compensating variation from a change in price is given by the area below the expected unconditional demand curve between the initial price and the final price. We begin with the standard linear-in-parameters Tobit model:

$$Q_i = \begin{cases} \beta' X_i + U_i & \text{if} \quad \beta' X_i + U_i > 0 \\ 0 & \text{otherwise,} \end{cases}$$

where Q_i is observed demand, X_i is a vector of explanatory variables which includes price, and U_i is normally distributed with mean zero and variance σ^2. As we show in

[4]A recent example of an empirical analysis which calculates consumer surplus correctly is Rosenthal (1987), who considers models of the number of visits to U.S. Army Corps of Engineer reservoirs in Kansas and Missouri. While the focus of his paper is on the role of substitute prices in demand estimation, he correctly observes that consumer surplus is given by the area below the "estimated number of visits" curve, specified in his model as the product of conditional demand and the probability of nonzero participation.

[5]For direct extensions of the Tobit model see Cragg (1971), van de Ven and van Praag (1981), and Blundell and Meghir (1987). For a general critique of the Tobit model as applied to demand analysis see Deaton (1986, pp. 1807-1809).

[6]Extensive treatments of the welfare economics of discrete choice models are provided by Williams (1976) and McFadden (1981).

detail in Section 4, the proper calculation of consumer surplus for this model uses the integrand

$$E(Q_i) = \Phi(\beta' X_i/\sigma)[\beta' X_i + \sigma\phi(\beta' X_i/\sigma)] \tag{1}$$

in an expression such as $\int_p E(Q_i)dp$.[7] A common alternative that has been employed in the literature uses instead the integrand $(\beta' X_i)$ in an expression such as $\int_p(\beta' X_i)dp$; i.e., the deterministic component of the so-called notional demand $\beta' X_i + U_i$ forms the basis for the calculation of consumer surplus. The problem with the latter approach is that notional demands obey neither nonnegativity constraints nor the budget constraint.[8] To the contrary, for many consumers the optimal solution to the utility maximization problem yields a corner solution; i.e., $Q_i = 0$. But these demands are still valid revelations of preference. Thus a proper calculation of consumer surplus must recognize that as price is changed, many consumers for whom the corner initially was an optimal solution will continue to find it so, and the change in consumer surplus for these individuals as price is changed is zero.

Econometricians have made considerable progress in the analysis of demand systems subject to kinks, corners, rationing, and the like. For example, the analyses of Wales and Woodland (1983) and Lee and Pitt (1986, 1987) are suited to the specification and estimation of demand systems for samples in which a significant number of individuals choose zero consumption for one or more goods. While Tobit-type models are only a special case of the systems analyzed by these authors, it is useful to review a simple consumer optimization problem subject to nonnegativity constraints which generates a Tobit demand model.

Consider a consumer with utility function $U(X_n, X_1, \epsilon)$. The function U is assumed to be strictly quasi-concave, continuous, and strictly monotonic. It is defined on two commodities, a numeraire good X_n and another good X_1 which may be either zero or positive at the optimum, and a stochastic term ϵ. Utility is maximized subject to the budget constraint $X_n + P_1X_1 \leq y$, and two "nonnegativity" constraints, $X_n > 0$ and $X_1 \geq 0$. The "unconstrained" indirect utility function is defined by

$$V(P_1, y, \epsilon) = \max_{X_n, X_1} \{U(X_n, X_1, \epsilon) | \; X_n + P_1X_1 \leq y\}. \tag{2}$$

From Roy's identity we obtain the notional demand for good 1:

$$Q_1 = (\partial V/\partial P_1)/(\partial V/\partial y), \tag{3}$$

[7] Throughout this paper we use Φ to represent the standard normal distribution and ϕ its density.

[8] Indeed notional demands are frequently negative for individuals who have a low probability of participation.

which can be expressed as $Q_1 = f(P_1, y, \epsilon)$. Since the optimization in equation (2) does not embody the nonnegativity constraint $X_1 \geq 0$, it is possible for the notional demand Q_1 to be less than zero. Define the virtual price for good 1, $p_1(y, \epsilon)$, as that price for which $f(p_1, y, \epsilon) = 0$. Then as long as $P_1 \geq p_1(y, \epsilon)$, good 1 will not be consumed; i.e., observed demand is

$$X_1 = \begin{cases} f(P_1, y, \epsilon) & \text{if} \quad P_1 < p_1(y, \epsilon) \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Since the condition $P_1 < p_1(y, \epsilon)$ is stochastic, equation (4) is estimable by standard limited dependent variable techniques. If, for example, $f(P_1, y, \epsilon)$ is linear in P_1, y, and ϵ, and ϵ is normally distributed, then (4) has the form of a Tobit model.

In order to consider welfare measures in this system, it is somewhat easier to begin with the dual to the indirect utility function, the expenditure function. Provided U is strictly increasing in X_n and nondecreasing in X_1, the indirect utility function $V(P_1, y, \epsilon)$ can be inverted to find the expenditure function $e(P_1, u, \epsilon)$, where u is some specified level of utility. The derivative of e with respect to P_1 gives the compensated notional demand for good 1, Q_1^c, which can be expressed as $Q_1^c = g(P_1, u, \epsilon)$. As with the notional demand Q_1 defined in (3), Q_1^c can be negative. Thus we define the compensated virtual price for good 1, $p_1^c(u, \epsilon)$, as that price for which $g(p_1^c, u, \epsilon) = 0$. As long as $P_1 \geq p_1^c(u, \epsilon)$, good 1 will not be consumed in order to attain utility u; i.e.,

$$X_1^c = \begin{cases} g(P_1, u, \epsilon) & \text{if} \quad P_1 < p_1^c(u, \epsilon) \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

Another way to write (5) is

$$X_1^c = \delta_1^c(P_1, u, \epsilon) \cdot g(P_1, u, \epsilon) \tag{6}$$

where $\delta_1^c = 1$ if $P_1 < p_1^c(u, \epsilon)$ and zero otherwise. From Theorem 1 in Small and Rosen (1981, p.112), we obtain the compensating variation for a change in price from P_1^0 to P_1^1 as the integral below the unconditional compensated demand curve between P_1^0 and P_1^1; that is,

$$CV = e(P_1^1, u^0, \epsilon) - e(P_1^0, u^0, \epsilon) = \int_{P_1^0}^{P_1^1} \delta_1^c(P_1, u^0, \epsilon) \cdot g(P_1, u^0, \epsilon) dP_1 \tag{7}$$

where $u^0 = V(P_1^0, y, \epsilon)$.[9]

[9] We argue elsewhere that expected compensating variation, i.e. the expectation of equation (7), is

Since the compensating variation is stochastic from the point of view of the econometrician (but deterministic from the individuals vantage) we calculate the expected compensating variation as the integral of the expected unconditional demand, or

$$\begin{aligned} ECV = E[e(P_1^1, u^0, \epsilon) - e(P_1^0, u^0, \epsilon)] &= \int_{P_1^0}^{P_1^1} EX_1^c dP_1 \\ &= \int_{P_1^0}^{P_1^1} \pi_1^c \cdot E[X_1^c | \delta_1^c = 1] dP_1, \end{aligned} \tag{8}$$

where $\pi_1^c = E[\delta_1^c(P_1, u^0, \epsilon)]$.[10] For notational convenience, we write the expected compensating variation as $ECV = \int(\pi X)dp$. The approaches in the literature either have calculated $\pi \int X dp$ or $X \int \pi dp$ but only infrequently $\int(\pi X)dp$.[11] In some instances these formulae produce correct estimates of compensating variation (if either the participation probability π is independent of price p in which case $\pi \int X dp = \int(\pi X)dp$ or if the demand X can be assumed to be independent of price in which case $X \int \pi dp = \int(\pi X)dp$).[12]

The extension of this approach to the more general models considered by Wales and Woodland (1983) and Lee and Pitt (1986, 1987) is conceptually simple. First the expenditure function is defined either directly as the minimum expenditure required to achieve a given level of utility *subject to all constraints* or as the dual to the utility maximization problem *subject to all constraints.* Under sufficient regularity conditions, the envelope

the most appropriate welfare measure for Tobit-type models. It is, however, but one of several welfare measures one can define for stochastic demand models. For further discussion see Dubin and Wilde (1991).

[10]This represents a departure from the approach used by Small and Rosen (1981) who go from equation (7) to an equation similar to (8) via aggregation across consumers.

[11]For example, Bockstael, Hanemann, and Kling (1987) consider the joint decision of beach selection and number of trips using nested logit techniques; that is, the individual's choice problem is divided into two stages. In the first stage the individual chooses a beach from within a set of either all saltwater or all freshwater beaches. In the second stage the individual chooses between the fresh and saltwater beach types. Finally, the authors estimate a Tobit model for number of trips and the implicit participation for each beach. A measure of compensating variation is derived using a formula which equates the expected maximum utility from the discrete-choice model under hypothetical changes in water quality to an initial level of expected maximum utility. To obtain an annual or seasonal benefit, the calculated compensating variation is multiplied by "the predicted number of trips the individual takes" (Bockstael, Hanemann, and Kling, 1987, p.957). Even though the model as estimated is not consistent with a "common underlying utility maximization framework," (Bockstael, Hanemann, and Kling, 1987, p.954), the compensating variation calculation still can only be correct if one assumes that the number of trips remains constant as quality is changed. Since this is not supported in the authors own estimation (Bockstael, Hanemann, and Kling, 1987, p.957, Table 3), it is unclear how to interpret their results. More recently, Kling (1988), in a calculation of seasonal consumer surplus, multiplies the predicted number of trips times the consumer surplus per trip as estimated by a logit model. Each of these analyses relies on the incorrect formula $X \int \pi dp$. We discuss additional examples in Section 3 below.

[12]Small and Rosen (1981) make a similar point, noting that Feldstein and Friedman (1977) in their analysis of health insurance correctly use the formula $\pi \int X dp$ since in the Feldstein and Friedman model the probability of illness is assumed to be exogenous (Small and Rosen, 1981, footnote 8).

theorem provides the compensated unconditional demand (Yatchew, 1985). Compensating variation is then the integral of compensated unconditional demand. When stochastic, expected compensating variation is given by the integral of expected compensated unconditional demand, assuming the expectation and integration operations are interchangeable (Fubini's theorem).

In the next section we review some of the recreation demand literature which has considered the estimation of compensating variation from demand systems subject to either truncation or censoring.

3 Review of Literature

As noted in the introduction, many authors have begun to use censored and truncated regression models for the estimation of the demand for recreation. These estimated models are then used, in many cases, to calculate some measure of consumer surplus. As we argued in Section 2, this measure should be the expected compensating variation as given in equation (8). In this section we review some of the approaches to the calculation of welfare measures for Tobit models which have been used in the literature.[13]

Smith and Desvousges (1985) and Smith, Desvousges, and Fisher (1986) consider the demand for 33 water based recreation sites. Their estimation method correctly takes into account the "truncation in visits at low levels of use and of the censoring in the upper levels of use" (Smith and Desvousges, 1985, p. 372). Yet their calculation of consumer surplus still confounds the estimation method with the structural model. In their case, the censoring of observations at the high end (six or more trips were reported as a single class) is clearly an artifact of the data generating process and has nothing to do with the underlying preferences of the individual. On the other hand, the set of observations which record zero trips are revealing valid corner solutions for some individuals. Expected demand in this instance must take into account the probability of participation, an allowance which Smith and Desvousges do not make.

McConnell (1986) uses a Tobit model to estimate the demand for alternative beaches in the New Bedford area. McConnell bases his consumer surplus calculation on the area to the left of the deterministic component of the notional demand curve. McConnell uses as the quantity for his consumer surplus calculation the median number of trips for those individuals who use the given beach. His observed quantity is therefore very similar to a point on the conditional demand curve rather than a point on the expected unconditional demand curve. We discuss this substitution in detail in Section 5 below, and calculate explicitly the bias inherent in it.

Bockstael, Strand, and Hanemann (1987) develop an empirical model for annual pri-

[13] We follow the treatments as presented by the authors and purposely elide the issue of whether the empirical specifications represent Marshallian or Hicksian demand curves.

vate boat trips in Southern California in 1983. The authors derive demand under two types of labor market equilibria and fit essentially two truncated regression models. Their calculation of compensating variation is made conditional on the decision to take at least one boat trip—a probabilistic choice. Their calculation of compensating variation therefore does not make any adjustment for the *ex ante* decision to take a boat trip.

Bockstael, McConnell, and Strand (1989) estimate the demand for striped bass days in 1980. Their specification is:

$$Q_{ij} = \begin{cases} \beta_j' X_{ij} + U_{ij} & \text{if } \beta_j' X_{ij} + U_{ij} \geq 0 \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

where Q_{ij} is the ith individual's demand for striped bass days at site j, β_j is a vector of coefficients, and U_{ij} is normally distributed with mean zero and variance σ^2.

To calculate consumer surplus these authors use the formula

$$CS_{ij} = (Q_{ij})^2/(-2\beta_{j1}) \tag{10}$$

where β_{j1} is the coefficient on cost of access in the jth site demand function. This formula is simply the area of the triangle under the deterministic component of the notional demand curve up to the quantity Q_{ij}. The authors then use the predicted demand

$$\hat{Q}_{ij} = \Phi(\beta' X/\sigma)[\beta' X + \sigma\phi(\beta' X/\sigma)] \tag{11}$$

as the quantity measure in their consumer surplus formula in order to "adjust for the censored dependent variable." We argue below in Section 4 that $\hat{Q}_{ij}$ is an appropriate quantity on which to base the consumer surplus calculation since it represents a point on the expected unconditional demand curve. But we also argue that CS_{ij} as given in equation (10) miscalculates consumer surplus since it represents the integral of the expected notional demand curve and not that of the expected unconditional demand curve.[14]

Finally, in a more recent article, Bell and Leeworthy (1990) purposely eschew the truncated regression model (relying instead on OLS estimation methods) in their analysis of the demand for beach days. Their justification is that several studies have been inconclusive regarding the direction of bias in the calculation of consumer surplus when OLS and truncated maximum likelihood methods are compared. None of the studies ref-

[14]In their recent discussion of the application of sample selection methods to recreation demand estimation, Bockstael, Strand, McConnell, and Arsanjani (1990) make a similar miscalculation.

erenced by these authors, however, properly calculates consumer surplus when truncation is caused by corner solutions.

4 Consumer Surplus for Tobit Models

The demand for the good in question, denoted by Q, is assumed to be given by the Tobit model:

$$Q = \begin{cases} Q^* & \text{if} \quad Q^* \geq 0 \\ 0 & \text{otherwise,} \end{cases} \tag{12}$$

where $Q^* = \beta' X + U$ and U is normally distributed with mean zero and variance σ^2.

The unconditional expectation of Q^* is simply $\beta' X$—the deterministic component of notional demand curve. The conditional expectation of Q given $Q^* \geq 0$ is

$$E(Q|Q^* \geq 0) = \beta' X + \sigma\phi(\beta' X/\sigma)/\Phi(\beta' X/\sigma). \tag{13}$$

The unconditional expectation of Q is given by

$$\begin{aligned} E(Q) &= \text{Prob}(Q^* \geq 0)E(Q|Q^* \geq 0) \\ &= \Phi(\beta' X/\sigma)[\beta' X + \sigma\phi(\beta' X/\sigma)/\Phi(\beta' X/\sigma)]. \end{aligned} \tag{14}$$

Thus, if we let $Z = \beta' X/\sigma$, then we can write $E(Q) = \Phi(Z)[Z + \phi(Z)/\Phi(Z)]\sigma$.

In Figure 1, we illustrate three demand curves. The first, labeled $E(Q^*)$, is the expected notional demand, the second, labeled $E(Q)$, is the expected unconditional demand, and the third, labeled $E(Q|Q^* > 0)$, is the expected conditional demand. It is possible to show that the three demand curves generally lie relative to each other as drawn in Figure 1; i.e., with $E(Q|Q^* > 0)$ to the right of $E(Q)$, which is itself to the right of $E(Q^*)$.[15]

Of the three demand curves illustrated in Figure 1, as we argued in Section 2, the right one for the calculation of consumer surplus is $E(Q)$. Since $E(Q^*) \leq E(Q) \leq E(Q|Q^* \geq 0)$, the bias associated with using $E(Q^*)$ or $E(Q|Q^* \geq 0)$ instead of $E(Q)$ is easy to see for any given P; to wit, if $E(Q^*)$ is used consumer surplus will be underestimated

[15]This will always be the case since $E(Q) \leq E(Q|Q \geq 0) = E(Q|Q^* \geq 0)$ and since $Q^* \leq \max(Q^*, 0)$ implies $E(Q^*) \leq E[\max(Q^*, 0)] = E(Q)$.

and if $E(Q|Q^* \geq 0)$ is used consumer surplus will be overestimated. The tendency in the literature, however, is to start with some quantity, say Q_0, then to use one of the three demand curves to find an associated price, say P_0, and then to integrate one of the demand curves above P_0 to calculate consumer surplus. The demand curve which is integrated is not necessarily the demand curve which is used to calculate P_0. This makes the calculation of bias more difficult. We conclude this section with a derivation of the correct consumer surplus in a typical Tobit model. In Section 5 we use that derivation to calculate the bias for two specific examples from the literature.

Using (14) we define exact consumer surplus by

$$\begin{aligned} CS_E &= \int_{P_0}^{P_1} E(Q)dp \\ &= \int_{P_0}^{P_1} \sigma[\Phi(\beta' X/\sigma)(\beta' X/\sigma) + \phi(\beta' X/\sigma)]dp. \end{aligned} \tag{15}$$

Without loss of generality we rewrite $\beta' X$ in terms of a constant factor α_0 and its own price component $\alpha_1 P$ with $\beta' X = \alpha_0 + \alpha_1 P$. A change in variables with $Z = \beta' X/\sigma = (\alpha_0 + \alpha_1 P)/\sigma$ and $\sigma dZ/\alpha_1 = dp$ yields

$$CS_E = \frac{\sigma^2}{\alpha_1} \int_{Z_0}^{Z_1} [\Phi(Z)Z + \phi(Z)]dZ \tag{16}$$

where $Z_j = (\alpha_0 + \alpha_1 P_j)/\sigma$.

To further analyze the integral $\int_{Z_0}^{Z_1} \Phi(Z)ZdZ$ we will need the following result.

CLAIM

$$\int_{Z_0}^{Z_1} Z^2\phi(Z)dZ = [\Phi(Z_1) - \Phi(Z_0)] - [Z_1\phi(Z_1) - Z_0\phi(Z_0)]. \tag{17}$$

PROOF OF CLAIM

In general

$$\int_{-\infty}^{Z_j} Z^2\phi(Z)dZ = E[Z^2|Z \leq Z_j]\text{Prob}[Z \leq Z_j] \tag{18}$$

where $E[Z^2|Z \leq Z_j]$ is given by

$$E[Z^2|Z \leq Z_j] = \text{var}[Z^2|Z \leq Z_j] + E[Z|Z \leq Z_j]^2. \tag{19}$$

Let $M_j = E[Z|Z \leq Z_j]$. Then

$$\begin{aligned} \text{var}[Z^2|Z \leq Z_j] + E[Z|Z \leq Z_j]^2 &= [1 - M_j(M_j - Z_j)] + M_j^2 \\ &= 1 + M_j Z_j. \end{aligned} \tag{20}$$

Thus

$$\int_{-\infty}^{Z_j} Z^2 \phi(Z) dZ = [1 + M_j Z_j]\Phi(Z_j). \tag{21}$$

But

$$M_j = E[Z|Z \leq Z_j] = -\phi(Z_j)/\Phi(Z_j), \tag{22}$$

so that using (22), we can write (19) as

$$\int_{-\infty}^{Z_j} Z^2 \phi(Z) dZ = \Phi(Z_j) - \phi(Z_j) Z_j. \tag{23}$$

Hence,

$$\int_{Z_0}^{Z_1} Z^2 \phi(Z) dZ = [\Phi(Z_1) - \Phi(Z_0)] - [\phi(Z_1)Z_1 - \phi(Z_0)Z_0].$$

Q.E.D.

To calculate the exact consumer surplus, we integrate (17) by parts to obtain

$$\begin{aligned} CS_E &= \frac{\sigma^2}{\alpha_1}\left[Z^2\Phi(Z)/2|_{Z_0}^{Z_1} - \frac{1}{2}\int_{Z_0}^{Z_1} Z^2\phi(Z)dZ + \Phi(Z)|_{Z_0}^{Z_1} \right] \\ &= \frac{\sigma^2}{\alpha_1}\left[\frac{[Z_1^2\Phi(Z_1) - Z_0^2\Phi(Z_0)]}{2} + [\Phi(Z_1) - \Phi(Z_0)] \right. \end{aligned}$$

$$-\frac{1}{2}([\Phi(Z_1)-\Phi(Z_0)]-[Z_1\phi(Z_1)-Z_0\phi(Z_0)])\Big]$$

$$= \frac{-\sigma^2}{2\alpha_1}\Big[[Z_0^2\Phi(Z_0)-Z_1^2\Phi(Z_1)]+[\Phi(Z_0)-\Phi(Z_1)]+[Z_0\phi(Z_0)-Z_1\phi(Z_1)]\Big].$$

In the case where P_1 approaches infinity in the limit, Z_1 approaches $-\infty$, and standard limiting arguments produce the following formula for exact consumer surplus in the Tobit model.

$$CS_E = \frac{-\sigma^2}{2\alpha_1}[Z_0^2\Phi(Z_0)+\Phi(Z_0)+Z_0\phi(Z_0)].^{16} \tag{24}$$

5 Examples of Bias

In this section we calculate for two examples from the literature the nature of the bias associated with using the wrong demand curve to do welfare analysis in Tobit models. The first is McConnell (1986), the second is Bockstael, McConnell, and Strand (1989).

Example 1

As we noted in Section 3, McConnell (1986) uses a Tobit model to estimate the demand for alternative beaches in the New Bedford Area. McConnell's measure of consumer surplus (CS_M) for a given beach is based on a quantity Q_0 determined by the median number of trips calculated only from the group of households who believe that PCB's have contaminated the New Bedford Harbor and who plan to attend the particular beach. Thus Q_0 is approximately a conditional mean given attendance.[17] This quantity is consistent with the estimated Tobit model at the price level P_0 where expected conditional demand equals Q_0; i.e., at the price level P_0 such that $E(Q|Q^* \geq 0) = Q_0$. The price P_0 and quantity Q_0 are illustrated in Figure 1. But McConnell's consumer surplus calculation is based on the area ABC. This is the area under the deterministic component of the notional demand curve above price P_1, where P_1 is the price level such that $E(Q^*) = Q_0$. There are two basic errors in this calculation. First, the demand curve that should be integrated is $E(Q)$ rather than $E(Q^*)$ and second, the integral should be taken above the price level P_0 rather than above the artificially low price level P_1. The correct consumer surplus calculation is given by the area DEF. The question we next address is whether ABC under or over estimates DEF.[18]

[16]An alternative derivation of this basic result is given in the Appendix.

[17]We replace the concept of conditional median with that of conditional mean to simplify the presentation.

[18]These two errors tend to offset one another since the demand curve for $E(Q^*)$ is too low (which results in an underestimate of consumer surplus) but the price P_1 is also too low (which results in an

Since the formula for the exact consumer surplus in equation (25) is specified in terms the standardized variable $Z_0 = (\alpha_0 + \alpha_1 P_0)/\sigma$, we first find the value Z_0 consistent with McConnell's quantity level Q_0. This is given by the expression

$$\begin{aligned} Q_0 &= E(Q|Q^* \geq 0) \\ &= \sigma[Z_0 + \phi(Z_0)/\Phi(Z_0)]. \end{aligned}$$

Since McConnell's consumer surplus is the area of the triangle ABC, it is therefore given by

$$\begin{aligned} CS_M &= \frac{-Q_0^2}{2\alpha_1} \\ &= \frac{-\sigma^2}{2\alpha_1}[Z_0 + \phi(Z_0)/\Phi(Z_0)]^2. \end{aligned}$$

Thus $CS_M \geq CS_E$ whenever

$$Z_0^2 + (2Z_0\phi_0/\Phi_0) + (\phi_0^2/\Phi_0^2) \geq Z_0^2\Phi_0 + Z_0\phi_0 + \Phi_0. \tag{25}$$

Example 2

Bockstael, McConnell, and Strand (1989) also base their calculation of consumer surplus (CS_{BMS}) on the deterministic component of the notional demand curve. However, they begin with a quantity Q_0 which lies on the expected unconditional demand curve, $E(Q)$. A value of Z_0 consistent with this quantity satisfies $Q_0 = \Phi(Z_0)[Z_0 + \phi(Z_0)/\Phi(Z_0)]\sigma$. The area under the deterministic component of the notional demand curve above the price which supports Q_0 is then

$$\begin{aligned} CS_{BMS} &= \frac{-1}{2\alpha_1}Q_0^2 \\ &= \frac{-\sigma^2}{2\alpha_1}\Phi_0^2[Z_0^2 + 2Z_0\phi_0/\Phi_0 + \phi_0^2/\Phi_0^2]. \end{aligned}$$

Thus $CS_{BMS} \geq CS_E$ whenever

overestimate of consumer surplus).

$$\Phi_0^2 Z_0^2 + 2Z_0\phi_0\Phi_0 + \phi_0^2 \geq Z_0^2\Phi_0 + Z_0\phi_0 + \Phi_0. \tag{26}$$

Absent an examination of the underlying data, it is not possible to know whether the McConnell or Bockstael, McConnell, and Strand measures in fact produce economically significant biases. In the case of Example 1 it seems likely however that the McConnell measure of consumer surplus will overstate the exact measure. This follows since both the linear and quadratic terms in (26) are larger on the left-hand-side of the inequality than they are on the right-hand-side. But an exact comparison must rely on the values of Z_0.

In Figure 2 we graph the three measures of consumer surplus with respect to the standardized variable Z_0. Generally, CS_M is greater than CS_E when Z_0 is negative or only slightly positive. These are cases in which the predicted participation is low and demonstrate the importance of the participation component of demand for samples for which many individuals reveal optimal corner solutions. Comparing CS_{BMS} and CS_E we see that the Bockstael, McConnell, and Strand measure of consumer surplus is always too low relative to the exact measure.

6 Conclusion

In this paper we have derived a theoretically sound measure of consumer surplus for the Tobit demand specification. We have also shown that several studies which rely on the Tobit demand specification have inaccurately calculated consumer surplus because they have confounded aspects of the estimation method with the underlying structural model. Moreover, we do not view the choice of the demand curve to be used in the calculation of consumer surplus as a matter of taste (such as when one chooses between compensating and equivalent variation in deterministic welfare economics). The use of the right demand curve recognizes that both the participation and usage decisions made by the consumer are important parts of the consumer's choice process.

As analysts move to more complex demand and estimation methods, the calculation of the expected compensating variation is likely to become more complicated. In the case of the Tobit model we have shown that a simple closed-form solution exists that relies on standard functions which are either readily evaluated or tabled. In more general circumstances, numerical methods will be needed to make the calculations.

Deeper issues remain. For example, while the exact consumer surplus is deterministic from the individuals vantage, it still is known only up to its random distribution by the econometrician. As such it is possible that researchers may wish to consider the variation in exact consumer surplus as well as its expectation.

Appendix

In this Appendix we provide an alternative derivation of the exact consumer surplus for the Tobit model. Recall that the exact consumer surplus is given by the integral

$$CSE = \int_{P_0}^{\infty} E[\delta(\alpha_0 + \alpha_1 p + u)]dp. \qquad (A1)$$

In Section 4 we evaluated this expression by first performing the expectation operation and then integrating with respect to price. Here we proceed by interchanging the expectation and integration operations. Equation (A1) becomes:

$$CSE = E\int_{P_0}^{\infty} \delta(\alpha_0 + \alpha_1 p + u)dp = \Phi_0 E[\int_{P_0}^{\infty} \delta(\alpha_0 + \alpha_1 p + u)dp|\delta = 1]. \qquad (A2)$$

Define the price P^* such that $\alpha_0 + \alpha_1 P^* + u = 0$. For prices above P^* the optimal demand is zero. Thus we can rewrite (A2) as

$$CSE = \Phi_0 E\left[\int_{P_0}^{P^*} (\alpha_0 + \alpha_1 p + u)dp|\delta = 1\right].$$

We now make a change of variables and let $q = (\alpha_0 + \alpha_1 p + u)$. Then $dq/\alpha_1 = dp$ and

$$\begin{aligned} CS_E &= \tfrac{\Phi_0}{\alpha_1} E\left[\int_{Q_0}^{0} q dq|\delta = 1\right] = \tfrac{-\Phi_0}{2\alpha_1} E\left[Q_0^2|\delta = 1\right] \\ &= \tfrac{-\Phi_0}{2\alpha_1} E\left[(\alpha_0 + \alpha_1 P_0 + u)^2|\delta = 1\right] \end{aligned} \qquad (A3)$$

where $Q_0 = \alpha_0 + \alpha_1 P_0 + u$. Using equation (10.4.39) from Amemiya (1985, p.375) we have

$$E\left[(\alpha_0 + \alpha_1 P_0 + u)^2|\delta = 1\right] = \sigma^2\left[(Z_0^2 + (Z_0\phi_0/\Phi_0) + 1\right]$$

where again $Z_0 = (\alpha_0 + \alpha_1 P_0)/\sigma$. Thus,

$$CS_E = \frac{-\sigma^2}{2\alpha_1}(\Phi_0 Z_0^2 + \phi_0 Z_0 + \Phi_0). \qquad (A4)$$

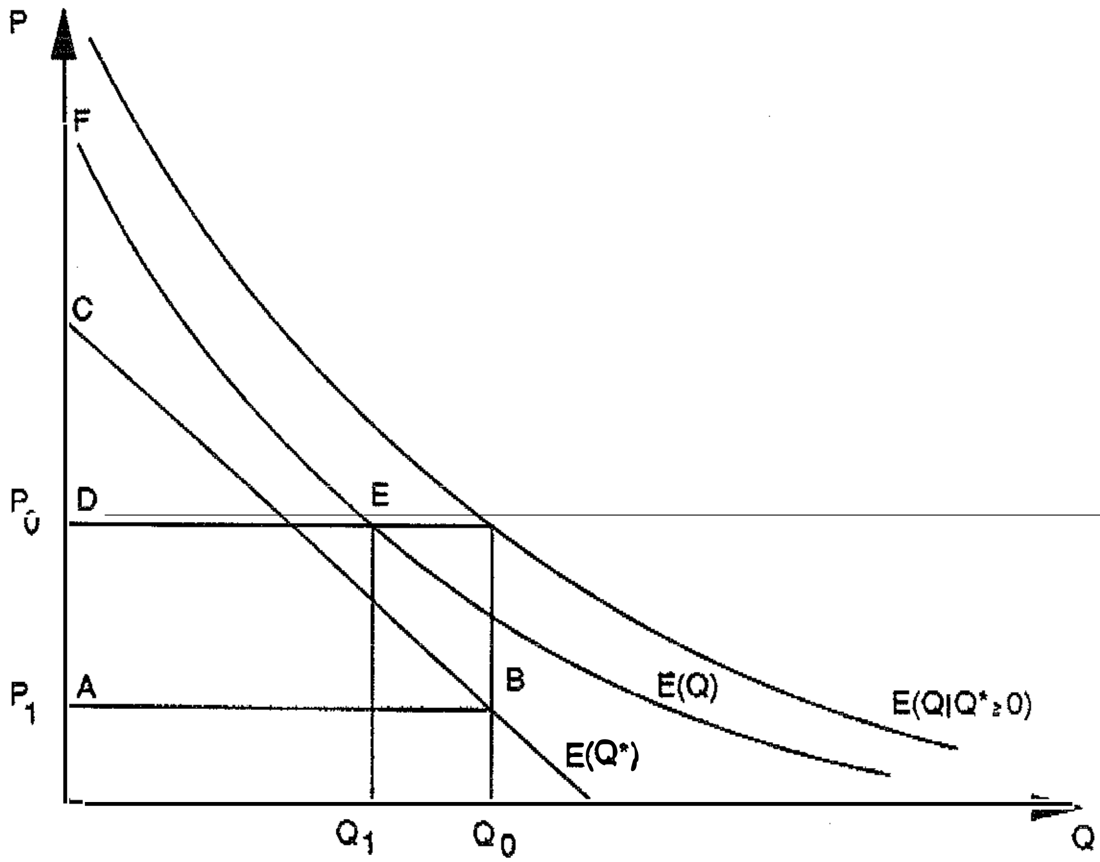

Figure 1

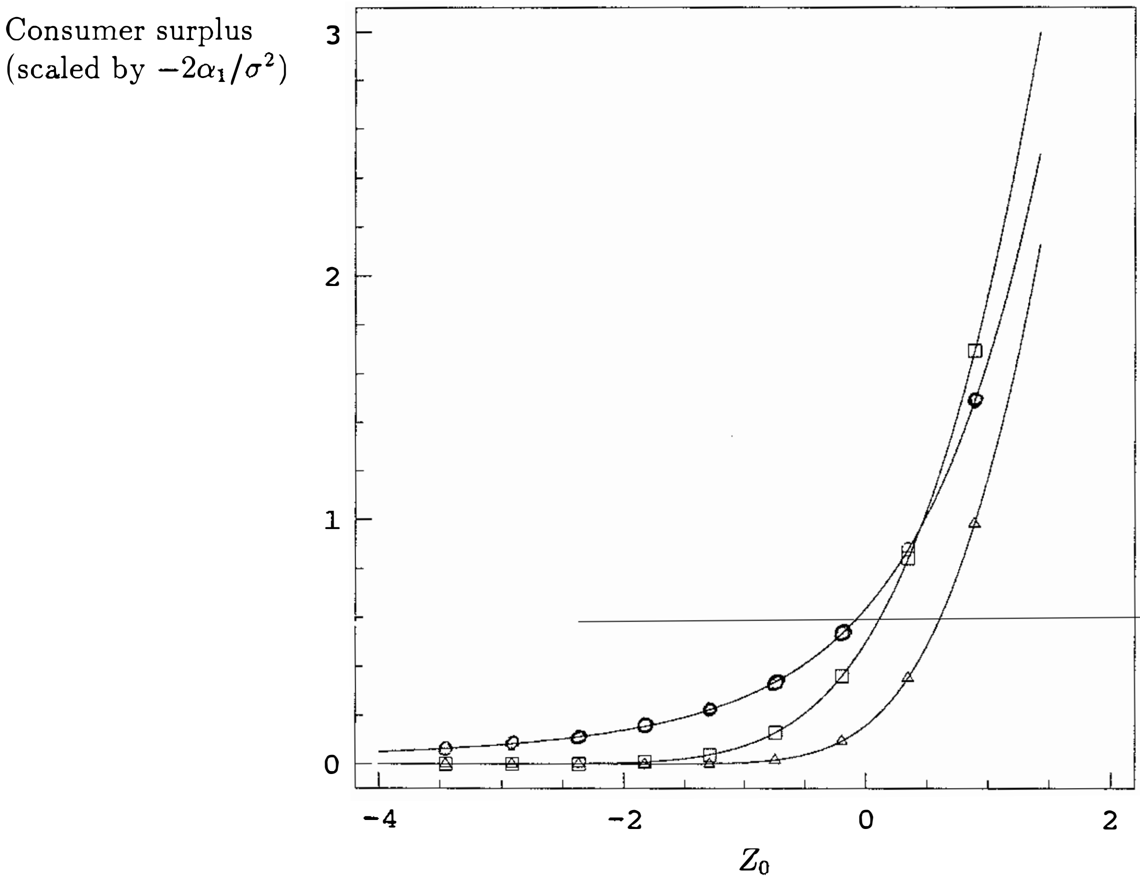

Figure 2

Legend

□ Exact consumer surplus
△ Bockstael, McConnell, and Strand consumer surplus
○ McConnell consumer surplus

References

Amemiya, T. *Advanced Econometrics*, Cambridge, Massachusetts: Harvard University Press, 1985.

Bell, F. W. and Leeworthy, V. R. "Recreational Demand by Tourists for Saltwater Beach Days." *Journal of Environmental Economics and Management* 18 (1990):189-205.

Bockstael, N. E.; Hanemann, W. M.; and Kling, C.L. "Estimating the Value of Water Quality Improvements in a Recreational Demand Framework." *Water Resources Research* 23 (1987):951-960.

Bockstael, N. E.; McConnell, K. E.; and Strand, I. E. "Measuring the Benefits of Improvements in Water Quality: The Chesapeake Bay." *Marine Resource Economics* 6 (1989):1-18.

Bockstael, N. E.; Strand, I. E.; and Hanemann, W. M. "Time and Recreational Demand." *American Journal of Agricultural Economics* 69 (1987):293-302.

Bockstael, N. E.; Strand, I. E.; McConnell, K. E.; and Arsanjani, F. "Sample Selection Bias in the Estimation of Recreation Demand Functions: An Application to Sportfishing." *Land Economics* 66 (1990):40-49.

Blundell, R. and Meghir, C. "Bivariate Alternatives to the Tobit Model." *Journal of Econometrics* 34 (1987):179-200.

Burt, O. and Brewer, D. "Estimation of Net Social Benefits from Outdoor Recreation." *Econometrica* 39 (1971):813-828.

Cicchetti, C.; Fisher, A.; and Smith, V. K. "An Econometric Evaluation of a Generalized Consumer Surplus Measure: The Mineral King Controversy." *Econometrica* 44 (1976):1259-1276.

Cragg, J. G. "Some Statistical Models for Limited Dependent Variables with Applications to the Demand for Durable Goods." *Econometrica* 39 (1971):829-44.

Deaton, A. "Demand Analysis." In *Handbook of Econometrics*, edited by Zvi Griliches and M. Intriligator. New York: Elsevier Science Publishers, 1986.

Dubin, J. A. and McFadden, D. L. "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption." *Econometrica* 52 (1984):345-362.

Dubin, J. A. and Wilde, L. L. "Stochastic Welfare Economics." Mimeograph. Division of the Humanities and Social Sciences, California Institute of Technology, January 1991.

Feldstein, M. S. and Friedman, B. "Tax Subsidies, the Rational Demand for Insurance and the Health Care Crisis." *Journal of Public Economics* 7 (1977):155-178.

Hanemann, W. M. "Discrete/Continuous Models of Consumer Demand." *Econometrica* 52 (1984):541-561.

Kling, C. L. "The Reliability of Estimates of Environmental Benefits from Recreation Demand Models." *Journal of Agricultural Economics* 70 (1988):892-901.

Lee, L. F. and Pitt, M. M. "Microeconometric Demand Systems with Binding Nonnegativity Constraints: The Dual Approach." *Econometrica* 54 (1986):1237-1242.

Lee, L. F. and Pitt, M. M. "Microeconometric Models of Rationing, Imperfect Markets, and Non-negativity Constraints." *Journal of Econometrics* 36 (1987):89-110.

McConnell, K. E. "The Damages to Recreational Activities from PCBs in New Bedford Harbor." Mimeograph. Department of Agricultural and Resource Economics, University of Maryland, December 1986.

McFadden, D. "Econometric Models of Probabilistic Choice." In *Structural Analysis of Discrete Data with Econometric Applications*, edited by C. F. Manski and D. McFadden. Cambridge, Massachusetts: The MIT Press, 1981.

Rosenthal, D. H. "The Necessity for Substitute Prices in Recreation Demand Analysis." *American Journal of Agricultural Economics* 69 (1987):828-837.

Small, K. A. and Rosen, H. S. "Applied Welfare Economics with Discrete Choice Models." *Econometrica* 49 (1981):105-130.

Smith, V. K. and Desvousges, W. H. "The Generalized Travel Cost Model and Water Quality Benefits: A Reconsideration." *Southern Economic Journal* 52 (1985):371-381.

Smith, V. K.; Desvousges, W. H.; and Fisher, A. "A Comparison of Direct and Indirect Methods for Estimating Environmental Benefits." *Journal of Agricultural Economics* 68 (1986):280-290.

Tobin, J. "Estimation of Relationships for Limited Dependent Variables." *Econometrica* 26 (1958):24-36.

van de Ven, W. P. M. M. and van Praag, B. M. S. "Risk Aversion and Deductibles in Private Health Insurance: Application of an Adjusted Tobit Model to Family Health Care Expenditures." In *Health, Economics, Heath Economics*, edited by J. van der Gaag and M. Perlman. Amsterdam, New York: North Holland, 1981, pp. 125-148.

Wales, T. J. and Woodland, A. D. "Estimation of Consumer Demand Systems with Binding Non-negativity Constraints." *Journal of Econometrics* 21 (1983):263-285.

Williams, H. C. W. L. "Travel Demand Models, Duality Relations and User Benefit Analysis." *Journal of Regional Science* 16 (1976):147-166.

Yatchew, A. J. "Applied Welfare Analysis with Discrete Choice Models: A Comment." *Economic Letters* 18 (1985):13-16.