Distribution-free Methods for Estimation of Willingness to Pay Models Using Discrete Response Valuation Data

Samuel D. Zapata
Former Research Assistant, School of Agricultural, Forest, and Environmental Sciences, Clemson University, Samuel Zapata, Clemson, SC 29634-0313, samueldzapata@gmail.com

Carlos E. Carpio
Associate Professor, Department of Agricultural and Applied Economics, Texas Tech University, Lubbock, TX79409-2132, carlos.carpio@ttu.edu

Selected Paper prepared for presentation at the Agricultural & Applied Economics Association's 2014 AAEA Annual Meeting, Minneapolis, Minnesota, July 27-29, 2014

Distribution-free Methods for Estimation of Willingness to Pay Models Using Discrete Response Valuation Data

Abstract

The Turnbull method is the standard approach used in contingent valuation studies to estimate willingness to pay (WTP) models using discrete responses without making assumptions about the distribution of the data. However, this approach has several limitations. The purpose of this study is to develop alternative distribution-free methods for the estimation of WTP models using nonparametric conditional imputation and local regression procedures. The proposed approaches encompass the recovery of the individuals' WTP values using an iterated conditional expectation procedure and subsequent estimation of the mean WTP using linear and nonparametric additive models. In contrast to the Turnbull approach, the proposed estimation methods allow the inclusion of covariates in the modeling of WTP estimates, as well as the complete recovery of its underlying probability distribution. Monte Carlo simulations are employed to compare the performance of the proposed estimators with that of the Turnbull estimator. We also illustrate the use of the proposed estimation techniques using a real data set.

Key words: Additive models, double-bounded elicitation, kernel functions, iterated conditional expectation, non-parametric regression, Turnbull method.

1. Introduction

Contingent valuation (CV) is a survey-based method initially developed to elicit the value (i.e., willingness to pay, WTP) that people place on non-market resources such as environmental preservation (e.g., Carson et *al*., 1992; Hanemann, 1994). New applications of CV are found in other areas such as health economics and agribusiness (Diener et *al*., 1998; Hudson and Hite, 2003).

The standard elicitation format used by CV practitioners is the double-bounded dichotomous choice (DBDC) approach. One drawback of the DBDC approach, as well as of other “closed-ended” elicitation formats, is that it generates interval-censored responses; hence, the estimation of measures of central tendency (e.g., mean WTP) as well as the marginal effects of covariates on the mean WTP requires the use of specialized statistical techniques. The majority of empirical studies using interval-censored responses from CV studies have utilized parametric methods, in which a distribution function for the WTP measure is specified; however, some authors have advocated the use of distribution-free methods (e.g., Carson et *al*., 1992; Carson et *al*., 1994).

With regard to distribution-free methods used to analyze CV interval-censored data, most of the literature is based on the nonparametric maximum likelihood (ML) estimation approach proposed by Turnbull (1974, 1976). However, the Turnbull approach has several important limitations. The purpose of this study is to develop alternative distribution-free estimation approaches that can be used to analyze interval-censored WTP data. The proposed estimators involve iterated procedures that combine nonparametric kernel density estimation of the errors of the WTP function with parametric linear or nonparametric kernel regression of its conditional mean function. In contrast to the Turnbull approach, the proposed estimation approach provides a point estimate of the mean WTP, allows the estimation of the marginal effects of covariates on the mean WTP, as well as the estimation of the underlying WTP probability distribution function at any point. Simulation techniques are employed to compare the performance of the proposed estimators with that of the Turnbull approach and the true parametric model. We also illustrate the use of the propose estimation techniques using a real data set. Although the paper focuses on modeling data obtained from the DBDC elicitation method, the procedures proposed in this article can be applied with WTP data obtained using other dichotomous choice elicitation mechanisms.

2. Literature Review

2.1. WTP Theoretical Framework

The theoretical foundations of WTP functions are based on the consumer utility and producer profit maximization problems. The WTP functions in both contexts are derived considering changes in the quality level of goods or services consumed or inputs used in the production process, for consumers and producers, respectively. For both consumers and producers, WTP can be shown to be a function of several variables including relevant prices and quality levels (Hanemann, 1991; Zapata and Carpio, 2014). To simplify mathematical notation, for the reminder of the paper we will use Y_i for the WTP value of the ith individual (consumer or producer) and $\boldsymbol{X}_i$ for the vector of arguments. Moreover, we will assume that Y_i is related to a set of explanatory variables $\boldsymbol{X}_i$ via the following model

(1) $$Y_i = g(\boldsymbol{X}_i) + \epsilon_i \qquad i = 1, \dots, n,$$

where the ϵ_i's are independent and identically distributed (i.i.d.) errors, with marginal density f_ε, zero mean and finite variance σ^2. It is also assumed that the ϵ_i's are independent of the d-dimensional predictor vector $\boldsymbol{X}_i$. Furthermore, $g(\boldsymbol{X}_i)$ is a function that represents the conditional mean function of Y_i given $\boldsymbol{X}_i$.

2.2. DBDC Approach and Estimation

Since its introduction by Hanemann (1985), the DBDC elicitation approach has gradually replaced other elicitation methodologies such as the open-ended and single-bounded dichotomous choice formats (Hanemann and Kanninen, 1999). DBDC responses have been mainly analyzed using parametric ML estimation methods (e.g., Hanemann et *al*., 1991; Zapata et *al*., 2013). One of the main advantages of the parametric ML estimation is that this estimation technique allows the inclusion of covariates in the modeling process, thus marginal effects are usually easy to estimate. On the other hand, the parametric ML method relies on *a priori* assumptions about the underlying distribution function of respondents' WTP. Hence, if the distribution function is misspecified, parameter estimates and any function of them, including welfare estimates and marginal effects, might be inconsistent.

An alternative to parametric ML estimation is the use of distribution-free methods which do not place any parametric assumptions on the distribution of the error ϵ_i. Distribution-free estimation procedures initially used in CV studies were adapted from the survival analyses models proposed by Ayer et *al*. (1955)**,** Kaplan and Meier (1958) and Turnbull (1974, 1976) (e.g., Kristom, 1990; Carson et *al*., 1992). In the case of DBDC responses, the preferred distribution-free estimation method used by practitioners has been the nonparametric ML estimator proposed by Turnbull (1976) (e.g., Carson et *al*., 1992; Carson et *al*., 1994). Unlike the parametric ML that seeks particular values of the distribution parameters, the Turnbull method directly estimates the underlying cumulative density function of respondents' WTP.

The Turnbull approach is not without shortcomings. First, the estimated cumulative density function is only defined up to a discrete set of observed points given by the bid amounts used in the WTP questions (i.e., the estimated CDF function is a step function). Second, the Turnbull approach does not allow the inclusion of covariates in the modeling of the mean WTP function. Furthermore, the Turnbull approach does not provide a point estimate of the mean WTP, but only upper and lower bound estimates.

More recently, researchers have proposed two types of alternative distribution-free estimation procedures to analyze DBDC responses. The first type includes distribution-free methods that assume a parametric specification for the conditional mean WTP function (i.e., $g(\boldsymbol{X}_i)$ in equation (1)) (Watanabe, 2010). The second type of procedures use semiparametric proportional hazard specifications commonly employed in duration models (e.g., An, 1996; Burton, 2000).

In this study we propose two distribution-free methods using kernel based procedures: one that assumes a parametric specifications for the mean WTP function (semiparametric procedure), and another where the mean WTP function is also estimated nonparametrically (nonparametric procedure). Hence, to the best of our knowledge this is the first study that uses fully nonparametric methods that allow the inclusion of covariates for the analysis of DBDC data. The semiparametric method can be considered an alternative to the distribution-free models proposed by Watanabe (2010), An (1996) and Burton (2000). None of the distribution-free estimation methods currently available for the estimation of DBDC data use kernel based procedures. A possible limitation to the lack of adoption of kernel based procedures is the fact that the weighting functions employed by these approaches usually require continuous observations of the dependent variable contrary to the interval-censored observations obtained in DBDC CV studies. However, recently developed algorithms make possible the adaptation of these techniques to interval-censored data (e.g., Kang et *al*., 2011; Braun et *al*., 2005).

3. Methodology

We propose to estimate the WTP function described in (1) using two novel distribution-free estimation techniques: the Semiparametric Iterated Linear Model (*SPILM*) and the Nonparametric Iterated Additive Model (*NIAM*). These models do not impose any arbitrary parametric assumption on the underlying distribution function of the errors $\epsilon_i's$ since its marginal density function (f_ε) is estimated using the nonparametric iterated conditional expectation procedure proposed by Braun et *al*. (2005). The function $g(\boldsymbol{X})$ is estimated using linear regression techniques in the case of SPILM and nonparametric additive regression methods in the case of *NIAM*.

The mathematical relation underlying the proposed procedure is given by:

(2) $$E[Y_i|Y_i \in I_i] = g(\boldsymbol{X}_i) + E\left[\epsilon_i\middle|I_{\varepsilon_i}\right],$$

where $E[Y_i|Y_i \in I_i]$ is the conditional expectation of Y_i given $Y_i \in I_i$, I_i is the observed interval of Y_i with boundary values L_i and R_i (i.e., $I_i = [L_i, R_i]$), and $I_{\varepsilon_i} = [L_i - g(\boldsymbol{X}_i), R_i - g(\boldsymbol{X}_i)]$ (Kang et *al*., 2011)[1]. It is important to note that equation (2) uses $E[Y_i|Y_i \in I_i]$ instead of Y_i since the Y_i's are interval-censored, i.e., observed as I_1, I_2,…,I_n. If the true value of the Y_i's were observed, *SPILM* and *NIAM* reduce to the standard linear regression and nonparametric additive estimators, respectively.

The proposed procedures involve four major steps which are iterated until convergence: 1) Start with an estimate of $E[Y_i|Y_i \in I_i]$ ($\hat{E}[Y_i|Y_i \in I_i]$); 2) Use the estimates of $E[Y_i|Y_i \in I_i]$ instead of the unobserved Y_i's to estimate $g(\boldsymbol{X}_i)$ using regression procedures (parametric regression in *SPILM* or nonparametric regression in *NIAM*); 3) Use the estimates of $g(\boldsymbol{X}_i)$ to obtain an estimate $E\left[\epsilon_i \middle| I_{\varepsilon_i}\right]$ using nonparametric kernel density estimation procedures; and 4) Use estimates of $g(\boldsymbol{X}_i)$ and $E\left[\epsilon_i \middle| I_{\varepsilon_i}\right]$ obtained in step 2 and 3, respectively, to obtain a new estimate of $E[Y_i|Y_i \in I_i]$.

In the sections below, we describe in detail all the steps of the proposed nonparametric iterative estimation procedure. For comparison purposes, we also briefly describe the Turnbull's nonparametric ML estimator and the standard parametric approach. The Monte Carlo experiments used to evaluate the proposed procedures are explained at the end of the section.

---

[1] The determination of L_i and R_i in the context of DBDC data is explained in detailed in Section 3.3.

3.1. Iterated Conditional Expectation Procedure

The algorithm employed to estimate the conditional expected value of the Y_i's, $\boldsymbol{Y}_{\mathbf{imp}} = \left(\hat{E}[Y_1|Y_1 \in I_1], \ldots, \hat{E}[Y_n|Y_n \in I_n]\right)^t$, and subsequently the *SPILM* mean estimator ($\hat{g}(\boldsymbol{X})_{SPILM}$) and *NIAM* mean estimator ($\hat{g}(\boldsymbol{X})_{NIAM}$), includes eight major steps:

Estimate starting values to initiate iterations (denoted with zero indices):

Step 1) For all Y_i's compute the interval midpoints: $Y_i^o = \frac{L_i+R_i}{2}$.

Step 2) Compute the initial mean function estimate: $\hat{g}_0(\boldsymbol{X})_\xi$, ξ = *SPILM*, *NIAM*, using $\boldsymbol{Y^o} = (Y_1^o, \ldots, Y_n^o)^t$.

Step 3) The error marginal density at the initial step $\hat{f}_{\varepsilon;0}(\omega)$ is taken as a uniform density[2] on the range $\left[min\left(L_i - \hat{g}_0(\boldsymbol{X}_i)_\xi\right), max\left(R_i - \hat{g}_0(\boldsymbol{X}_i)_\xi\right)\right]$.

Conduct iterative steps (denoted with j indices):

Step 4) Estimate the marginal density of the errors f_ε using the iterated conditional expectation procedure developed by Braun et *al*. (2005):

a) Estimate the interval-censored errors as $I_{\varepsilon_i} = \left[L_i - \hat{g}_{j-1}(\boldsymbol{X}_i)_\xi, R_i - \hat{g}_{j-1}(\boldsymbol{X}_i)_\xi\right]$.

b) Compute the error marginal density function using the fixed point estimator[3]:

$$\hat{f}_{\varepsilon;j}(z) = \frac{1}{n}\sum_{i=1}^{n} \frac{\int_{I_{\varepsilon_i}} W_b(z-\omega)\hat{f}_{\varepsilon;j-1}(\omega)d_\omega}{\int_{I_{\varepsilon_i}} \hat{f}_{\varepsilon;j-1}(\omega)d_\omega},$$

---

[2] Braun et *al*. (2005) show that the final estimate of f_ε does not depend on the density function used on the initial iteration step.

[3] Computation of all integrals was carried out using the trapezoid rule.

where $W_b(v) = b^{-1}W(v/b)$, $W(\cdot)$ is a kernel density function with scale parameter b, and z is any real number.

Step 5) Compute the conditional expectation of the ϵ_i's: $\hat{E}\left[\epsilon_i \middle| I_{\varepsilon_i}\right] = \frac{\int_{I_{\varepsilon_i}} z\hat{f}_\varepsilon(z)d_z}{\int_{I_{\varepsilon_i}} \hat{f}_\varepsilon(z)d_z}$.

Step 6) Estimate the conditional expectation of the Y_i's, $\boldsymbol{Y}_{\mathbf{imp}}$. At the j[th] iteration step the i[th] element of $\boldsymbol{Y}_{\mathbf{imp}}$ is given by: $\hat{E}[Y_i|Y_i \in I_i] = \hat{g}_{j-1}(\boldsymbol{X}_i)_\xi + \hat{E}\left[\epsilon_i \middle| I_{\varepsilon_i}\right]$.

Step 7) Compute $\hat{g}_j(\boldsymbol{X})_\xi$ using the estimate $\boldsymbol{Y}_{\mathbf{imp}}$ from the previous step using regression procedures (parametric regression in *SPILM* or nonparametric regression in *NIAM:* see section below)

Step 8) Set $\hat{g}_{j-1}(\boldsymbol{X})_\xi = \hat{g}_j(\boldsymbol{X})_\xi$ and return to step (4) or stop if the convergence criterion is satisfied[4].

3.2. Conditional Mean Function Estimation

3.2.1. *Linear regression*

In the *SPILM* the conditional mean function of Y given $\boldsymbol{X}$, $g(\boldsymbol{X})$, is estimated using the standard linear regression model

(3) $$g(\boldsymbol{X}_i) = \beta_0 + \sum_{k=1}^{d} \beta_k x_{ik} ,$$

where the estimates of the parameters $\beta_0, \beta_1, \ldots, \beta_d$ are obtained by least squares using $\boldsymbol{Y}_{\mathbf{imp}}$ as the dependent variable. The *SPILM* mean estimator $\hat{g}(\boldsymbol{X})_{SPILM}$ is calculated by averaging the estimate of (3), $\hat{g}(\boldsymbol{X}_i)_{SPILM}$, for all individuals

(4) $$\hat{g}(\boldsymbol{X})_{SPILM} = n^{-1} \sum_{i=1}^{n} \hat{g}(\boldsymbol{X}_i)_{SPILM}.$$

[4] An absolute difference of less than 10^{-5} in successive objective function estimates (e.g., $\left|\hat{g}_j(\boldsymbol{X})_\xi - \hat{g}_{j-1}(\boldsymbol{X})_\xi\right|$) was used to declare convergence on every iteration procedure employed in this study.

3.2.2. Nonparametric Additive Regression

There are several options for the nonparametric estimation of the $g(\boldsymbol{X})$ function. In this study, we use a nonparametric additive model instead of a multivariate kernel regression for several reasons. First, additive models are less affected by the curse of dimensionality and multicollinearity. Second, their marginal effects are easier to interpret. Third, additive model estimates possess a faster convergence rate than multivariate kernel estimates (Buja et *al*., 1989; Cameron and Trivedi, 2005, p. 319). Finally, the majority of WTP studies use an additive mean parametric function. The additive model assumes that

(5) $$g(\boldsymbol{X}_i) = \mu_0 + \sum_{k=1}^{d} \mu_k(x_{ik}) \,,$$

where the $\mu_k(\cdot)$'s are standardized smooth functions so that $E[\mu_k(\cdot)] = 0$ for every k. These functions are estimated one at a time using a backfitting algorithm as suggested by Hastie and Tibshirani (1986) and Kauermann and Opsomer (2004).

As shown in Kauermann and Opsomer (2004), the $\mu_k(\cdot)$'s can be jointly estimated. First, consider the kth additive function estimator

(6) $$\widehat{\boldsymbol{\mu}}_k = \boldsymbol{S}_k^{*}\left\{\left(\boldsymbol{Y}_{\mathbf{imp}} - \hat{\mu}_0\right) - \widehat{\boldsymbol{\mu}}_{-k}\right\}$$

where $\widehat{\boldsymbol{\mu}}_k = \{\hat{\mu}_k(x_{1k}), \ldots, \hat{\mu}_k(x_{nk})\}^t$, $\widehat{\boldsymbol{\mu}}_{-k} = \sum_{r \neq k} \widehat{\boldsymbol{\mu}}_r$ is an estimator of the sum of the remaining $d-1$ additive functions, $\hat{\mu}_0 = n^{-1}\sum_{i=1}^{n} Y_i$ and $\boldsymbol{S}_k^{*} = (\boldsymbol{I}_n - \mathbf{1}\,\mathbf{1}^t/n)\boldsymbol{S}_k$ is a centered smooth matrix to ensure identifiability of the estimators, $\boldsymbol{I}_n$ denotes an identity matrix, and $\boldsymbol{S}_k$is a $n \times n$ smoothing matrix whose ij element is given by

(7) $$\boldsymbol{S}_{k,ij} = K_k\left(x_{ik}, x_{jk}, h_k\right) / \sum_{j=1}^{n} K_k\left(x_{ik}, x_{jk}, h_k\right),$$

where $K_k(\cdot)$ is a kernel density function with scale parameter h_k (i.e., a bandwidth). Joint estimation of the additive functions $\widehat{\boldsymbol{\mu}}_1, \dots, \widehat{\boldsymbol{\mu}}_d$ entails finding the solution to the normal equations

(8) $$M\widehat{\boldsymbol{\mu}} = \boldsymbol{S}^*\left(\boldsymbol{Y}_{\mathbf{imp}} - \hat{\mu}_0\right),$$

with $\widehat{\boldsymbol{\mu}} = (\widehat{\boldsymbol{\mu}}_1^t, \dots, \widehat{\boldsymbol{\mu}}_d^t)^t$, $\boldsymbol{S}^* = \left(\boldsymbol{S}_1^{*\,t}, \dots, \boldsymbol{S}_d^{*\,t}\right)^t$ and

$$M = \begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{S}_1^* & \cdots & \boldsymbol{S}_1^* \\ \boldsymbol{S}_2^* & \boldsymbol{I}_n & \cdots & \boldsymbol{S}_2^* \\ \vdots & & \ddots & \vdots \\ \boldsymbol{S}_d^* & \boldsymbol{S}_d^* & \cdots & \boldsymbol{I}_n \end{pmatrix}.$$

As the *SPILM* estimator, the *NIAM* mean estimator $\hat{g}(\boldsymbol{X})_{NIAM}$ also averages the estimated of expression (5), $\hat{g}(\boldsymbol{X}_i)_{NIAM}$, for all individuals

(9) $$\hat{g}(\boldsymbol{X})_{NIAM} = n^{-1}\sum_{i=1}^{n}\hat{g}(\boldsymbol{X}_i)_{NIAM}.$$

Whereas in *SPILM* the marginal effects are given by the coefficients $\hat{\beta}_1, \dots, \hat{\beta}_d$, in *NIAM* the relationships between covariates and mean WTP are given by the smooth functions $\mu_k(\cdot)$'s (Buja et *al*.,1989). Therefore, the marginal effect of a covariate on the mean WTP changes from point to point. Consequently, the relationships between explanatory variables and smooth functions in additive models are presented in the form of plots (e.g., Opsomer and Ruppert, 1998; Kauermann and Opsomer, 2004).

3.2.3. Kernel functions and bandwidth selection

The computation of both the *NIAM* mean estimator $\hat{g}(\boldsymbol{X}_i)_{NIAM}$ and the error density function estimator $\hat{f}_\varepsilon(z)$ involve kernel functions: $K_k(\cdot)$'s in equation (7) and $W_b(\cdot)$ in step 4b). The kernel functions were selected based on asymptotic properties and on their ability to model both continuous and categorical data. Specifically, three different kernel functions were used to estimate $\hat{g}(\boldsymbol{X}_i)_{NIAM}$: for continuous explanatory variables we consider a 2th-order Epanechnikov kernel[5], and for discrete variables with or without natural order we consider the kernel functions proposed by Racine and Li (2004). In the case of the estimation of the error density function $\hat{f}_{\varepsilon}(z)$, the kernel function $W_b(\cdot)$ is set to be equal to the 2th-order Epanechnikov kernel (see Appendix A).

All kernel functions considered in this study depend on the bandwidth or smoothing parameters which are more crucial for the quality of the estimates than the kernel choice itself (Cameron and Trivedi, 2005, p. 303). The bandwidth parameters for the kernels used to estimate $\hat{g}(\boldsymbol{X}_i)_{NIAM}$ were selected by the generalized cross-validation procedure described in Kauermann and Opsomer (2004) (see Appendix A). The bandwidth parameter b needed for the estimation of the error density function (step 4b) was estimated using a modified version of the likelihood cross-validation method proposed by Braun et *al*. (2005) (see Appendix A).

3.3. Parametric and Nonparametric Maximum Likelihood Estimators

In this section we describe the parametric and nonparametric maximum likelihood methods traditionally used in CV studies which will be subsequently compared with the proposed distribution estimation procedures. In the DBDC elicitation format every respondent *i* is presented with an initial bid B_i and asked if he is willing to pay that amount. If the respondent answers “yes” to the first bid, a second WTP question is asked using a higher bid amount B_i^u. If the respondent answers “no” to the first bid, the second

[5] The 2th-order Epanechnikov kernel function is referred as the “optimal kernel” because it possesses the minimum mean integrated squared error (MISE) among available kernel functions (Cameron and Trivedi, 2005, p. 303).

WTP question uses a lower bid B_i^l. Consequently, every Y_i (i.e., WTP) is observed to fall into one of the four intervals: $(-\infty, B_i^l)$, $[B_i^l, B_i)$, $[B_i, B_i^u)$ and $[B_i^u, +\infty)$, $i = 1, \ldots, n$.

Denoting the lower bound of the observed i^th interval (I_i) as L_i and the upper bound as R_i, the probability that Y_i is in the I_i interval is given by

(10) $$P(L_i \leq Y_i < R_i) = \mathrm{F}(R_i) - \mathrm{F}(L_i) \quad i = 1, \ldots, n,$$

where $\mathrm{F}(\cdot)$ is the cumulative density function (CDF) of Y. Since the number of different bids used in the DBDC questions is usually less than the number of observations in the sample, some of the observed intervals are the same across individuals; resulting in $M \leq n$ unique observed intervals $\mathcal{I}_m$, m=1, …, M, with boundary values of $\mathcal{L}_m$ and $\mathcal{R}_m$. Therefore, the log-likelihood function for the interval-censored Y_i's can be written as

(11) $$\ln L = \sum_{i=1}^{n} ln[\mathrm{F}(R_i) - \mathrm{F}(L_i)]$$
$$= \sum_{m=1}^{M} n_m ln[\mathrm{F}(\mathcal{R}_m) - \mathrm{F}(\mathcal{L}_m)],$$

where n_m, $m = 1, \ldots M$, is the number of observations for whom both $L_i = \mathcal{L}_m$ and $R_i = \mathcal{R}_m$. Parametric models (*PM*) assume that Y_i follow a certain distribution (see e.g., Zapata, 2012).

To specify the log-likelihood function of Turnbull's nonparametric ML procedure each unique observed interval $\mathcal{I}_m$, m=1,…, M, needs to be expressed as an union of Q disjoint closed intervals of the form $[a_{q-1}, a_q)$, $q = 1, \ldots, Q$, called innermost intervals[6], such that $\mathcal{I}_m = \bigcup_{q=1}^{Q} d_{mq} [a_{q-1}, a_q)$, where d_{mq} is a dummy variable that indicates

[6] Assuming that Y is non-negative, the complete set of Q innermost intervals is $[a_0, a_1), [a_1, a_2) \ldots [a_{Q-1}, a_Q)$, where $0 = a_0 < a_1 < \cdots < a_Q$. In the case of DBDC data, the boundaries of the innermost intervals (a_q's) are given by the bid amounts used in the WTP questions.

whether the qth innermost interval (A_q) is used to express the mth unique interval. Specifically,

(12) $$d_{mq} = \begin{cases} 1 & if\ \mathcal{L}_m \leq a_{q-1}\ and\ \mathcal{R}_m \geq a_q \\ 0 & otherwise \end{cases}, m = 1, \dots, M; q = 1, \dots, Q.$$

The log likelihood function in (11) can then be expressed in terms of the innermost intervals

(13) $$\ln L = \sum_{m=1}^{M} n_m ln \sum_{q=1}^{Q} d_{mq}[\mathrm{F}(a_q) - \mathrm{F}(a_{q-1})].$$

The Turnbull procedure considers each $F_q = F(a_q)$ in (13) as a parameter to be estimated and imposes the restriction that $0 = F_0 \leq F_1 \dots \leq F_Q = 1$. Estimation is then carried out using Turnbull's self-consistent algorithm (Day 2007; Gomez et *al*., 2004; Turnbull 1976). The mean value of Y can thus be written as $E(Y) = \int_0^{a_Q} Y\, d_{F(Y)} = \sum_{q=1}^{Q} \int_{A_q} Y d_{F(Y)}$ (Haab and McConnell, 1997).

As mentioned earlier, the Turnbull approach does not provide a point estimate of the mean WTP, but only upper and lower bounds of its value. Therefore, to facilitate comparison across models, we used the Turnbull midpoint approximation of the expected value of Y ($\hat{g}_T$): $\hat{g}_T = \hat{E}(Y) = \sum_{q=1}^{Q} \frac{a_{q-1}+a_q}{2} (\hat{F}_q - \hat{F}_{q-1})$, where the $\hat{F}_q$'s are the solution to the log likelihood function in (13).

3.4.Probability Distribution Estimation

The iteration process used in the *SPILM* and *NIAM* approaches can also be used to recover the CDF and probability density function (PDF) of WTP at any point. Estimation of the probability distribution of Y is possible since

(14) $$f_Y(y) = f_\varepsilon(y - g(\boldsymbol{X})),$$

where f_Y is the PDF of Y.

Equation (14) suggests the following estimators for the PDF and CDF of Y:

(15) $$\hat{f}_Y(y)_\xi = \hat{f}_\varepsilon\big(y - \hat{g}(\boldsymbol{X})_\xi\big)$$

and

(16) $$\hat{F}_Y(y)_\xi = \int_0^y \hat{f}_\varepsilon\big(y - \hat{g}(\boldsymbol{X})_\xi\big)d_y,$$

respectively, where $f_\varepsilon(\cdot)$ and $g(\boldsymbol{X})$ in (14) are replaced by estimates and ξ = *SPILM*, *NIAM*.

3.5. Data and Study Design

The relative performance of the *SPILM*, *NIAM* and Turnbull estimation procedures was evaluated using Monte Carlo simulation procedures. Estimated mean values and marginal effects were compared to those obtained from the true underlying parametric model. The three models were also employed to estimate producers' WTP for the services provided by an Electronic Trade Platform in the data set described and analyzed in Zapata et *al.* (2013).

3.5.1. Monte Carlo Simulation

A total of 100 data sets (simulations) containing n observations each, $\{Y_i, \boldsymbol{X}_i\}_{i=1}^n$, $n \in \{100, 200, 500\}$, were generating using the following regression model containing both continuous and categorical predictor variables

(17) $$Y_i = 40 + 3X_{1i} + 3X_{2i} + 3X_{3i}^{d1} - 2X_{3i}^{d2} + 2\epsilon_i,$$

where the X_{1i}'s are i.i.d. observations from an Uniform distribution in the range [-10, 10], $X_{2i} \in \{0,1\}$ with $Pr(X_{2i} = 0) = Pr(X_{2i} = 1) = 0.5$, $X_{3i}^{d_j} \in \{0,1\}$, $j = 1, 2$, indicate the occurrence of the jth category of X_{3i}, $X_{3i} \in \{1,2,3\}$ with $Pr(X_{3i} = \iota) = 1/3$ for $\iota = 1,2,3$, and ϵ_i is an i.i.d. observation from a Normal distribution with mean zero and variance equal to one.

The resulting Y_i's from (17) can be seen as the individuals' true valuation (e.g., individuals' WTP value) given a set of observable characteristics, $X_{ji}'s$. The data generating process considered in this study mimics the one employed in CV using a DBDC elicitation format. Four initial bid amounts were randomly assigned to each observation in the generated data: \$24, \$36, \$48 and \$60. The initial bids, respectively, are the 20th, 40th, 60th and 80th percentiles of an empirical distribution in a 50 observation sample simulated with the regression model in (17) with no error term[7]. The corresponding follow-up bid amounts were \$18 (10th percentile), \$24, \$36 and \$48 if the initial bid assigned to the observation was higher than the true WTP value. On the other hand, if the initial bid assigned to the observation was lower than the true WTP value, corresponding higher follow-up bids of \$36, \$48, \$60 and \$66 (90th percentile) were assigned. Based on the sample distribution used to generate the bids, the lower bound for those observations answering "no/no" was set to \$0 and the upper bound for those answering "yes/yes" was set to \$80 in the *SPILM*, *NIAM* and Turnbull approaches.[8]

Using the DBDC WTP data generated from (17) we estimate both the mean WTP using *SPILM*, *NIAM*, the Turnbull procedure and the true normal parametric model (*PM*) ($\hat{g}(\boldsymbol{X})_{SPILM}$, $\hat{g}(\boldsymbol{X})_{NIAM}$, $\hat{g}_T$ and $\hat{g}(\boldsymbol{X})_{PM}$), and the marginal effects from the true *PM* and *SPILM*.

[7] The initial bids were chosen following the methods employed in Calia and Strazzera (2012).
[8] Results were not sensitive to the choice of the maximum value.

The performance of all four mean estimators and marginal effect estimators from the *SPILM* and *PM* were analyzed using the squared-root of the mean squared error (RMSE),

(18) $$RMSE(\hat{\theta}) = \sqrt{\frac{1}{100}\sum_{s=1}^{100}\left[\hat{\theta}^{(s)} - \theta^{(s)}\right]^2},$$

bias

(19) $$bias(\hat{\theta}) = \frac{1}{100}\sum_{s=1}^{100}\left[\hat{\theta}^{(s)} - \theta^{(s)}\right]$$

and standard error (SE)

(20) $$SE(\hat{\theta}) = \sqrt{\frac{1}{100}\sum_{s=1}^{100}\left[\hat{\theta}^{(s)} - \bar{\hat{\theta}}\right]^2},$$

where $\hat{\theta}^{(s)}$ and $\theta^{(s)}$ are the estimated and true parameter function of interest (e.g., mean or marginal effect) of the s^{th} data set, and $\bar{\hat{\theta}} = \frac{1}{100}\sum_{s=1}^{100}\hat{\theta}^{(s)}$. Since in the case of *NIAM* there are not unique marginal effects estimates, we only estimated the functions ($\mu_k(\cdot)$'s) for one randomly generated datasets of each sample size ($n =$ 100, 200, 500); hence, we did not calculate RMSE, bias and SE for the estimated marginal effects using NIAM.

3.5.2. Empirical application: producers' WTP study

SPILM, NIAM, Turnbull and PM estimators were also evaluated using a real DBDC data set. The data was described and analyzed in Zapata et *al.* (2013) using parametric techniques, where the WTP measure was found to follow a log-logistic distribution. The main objective of the study was to estimate the monetary value that registered producers placed on the services provided by an Electronic Trade Platform (i.e., MarketMaker). A reduced set of available explanatory variables was used as an illustration of the attributes of the proposed estimation techniques. Covariates employed in the estimation of the WTP models are type of user based on intensity of use (USER_TYPE), marketing contacts gained due to participation in MarketMaker (CONTACTS), and firm total annual sales (SALES). Kernel selection for *NIAM* was based on the fact that SALES is a continuous variable and USER_TYPE and CONTACTS are ordered categorical variables.

The mean WTP was estimated for *SPILM*, *NIAM*, Turnbull and log-logistic PM[9] methods. Marginal effects were estimated for SPILM and the log-logistic PM, and covariate-mean relationships were estimated for *NIAM*. The standard errors of the estimated means and marginal effects in *SPILM* and log-logistic *PM* were calculated using the bootstrapping procedure outlined by Cameron and Trivedi (2005, p.362) using a total of 100 replications. The point wise standard error bands suggested by Buja et *al*. (1989) were used as a measure of dispersion of the estimated smooth functions in *NIAM*. The standard error bands represent the fitted curve ± 2 estimated standard error. The standard error of each smooth function was estimated as the mean standard error across the 100 replications at each unique covariate value. Finally, the underlying PDF and CDF of the producers' WTP for MarketMaker were calculated using expressions (15) and (16), respectively.

The different bandwidths parameters of the *SPILM* and *NIAM* estimators were calculated using the 227 observations in the original data, then the smooth parameter (*b*)

[9] The log-logistic*PM* assumes that producers' WTP follows a log-logistic distribution rather than a normal distribution,

of the error density function in iteration step 4b (Section 3.1) was fixed at these values in each replication of the bootstrapping procedure.[10]

4. Results

4.1. Monte Carlo Simulation

The RMSE, bias and standard error of the different mean estimators are presented in Table 1. Simulation results show that the conditional mean estimators of *SPILM* and *NIAM* dominate the unconditional Turnbull mean estimator in terms of RMSE, bias and SE. Furthermore, the *SPILM* mean estimator performed as well as the benchmark correctly specified parametric model even for the small sample size (100 observations). Hence, the *SPILM* mean estimator seems to provide a more robust alternative to *PM* without sacrificing efficiency. Moreover, and as expected, the robustness gains to misspecification of the mean and distribution function when using the *NIAM* and Turnbull approaches result in efficiency losses.

The RMSE, bias and standard error of the marginal effects estimated using *NIAM* and the *PM* are shown in Table 2. None of the models is clearly superior for any of the sample sizes considered. The RMSE and SE values of the marginal effects estimated using *SPILM* were generally lower than those of the *PM*. On the other hand, the biases of the marginal effects on *SPILM* were generally higher than their counterparts estimated using the parametric model. However, in both cases the differences are very small. Therefore, and consistent with the mean estimators comparison, the simulation results

[10] The bandwidth parameter b in the *SPILM* and *NIAM* were estimated to be equal to 5.30 and 7.01, respectively. Fixing the bandwidth at predetermined values reduces the time needed for estimation of the standard errors.

indicate that relative to *PM*, the gains in robustness when using *SPILM* do not result in significant efficiency losses.

The fitted smooth functions $\mu_k(\cdot)'$s using *NIAM* are displayed on Figure1. The true effect of X_1 on Y implied by equation (17) is given by a straight line with slope of 3. In the case of the discrete variables, the true "marginal effects" are 3 for X_2, and 3 and -2 for X_3^{d1} and X_3^{d2}, respectively. Overall, these plots of the marginal effects estimated using *NIAM* suggest that the true effects implied by (17) seem to be adequately captured by the estimated relationships betweenX_1, X_2 and X_3. For illustration purposes consider the random sample of size 500, where the estimated difference in Y between an observation with $X_2 = 1$ and one with $X_2 = 0$ is estimated to be 2.29 units (compared to a difference of 3 units in the true model). Similarly, $Y's$ value of an observation with $X_3 = 3$ is estimated to be 2.68 units less and 2.31 units more than an observation with $X_3 = 1$ and $X_3 = 2$, respectively (compared to the corresponding true differences of 3 and 2 units).

The same data sets employed to estimate *NIAM*'s $\mu_k(\cdot)'$s functions were also utilized to estimate the CDF and PDF of Y using the PM, the Turnbull approach and the two proposed distribution free approaches (Figure 2). The CDF and PDF functions estimated using *SPILM* and *NIAM* are very similar. Furthermore, both approaches seem to provide reasonable approximations of the true underlying distribution functions. The marked difference between *SPILM*, *NIAM* and the true CDF estimates and those from the Turnbull approach are attributed to the fact that the formers are conditional estimates while the Turnbull CDF is estimated without considering the effect of covariates.

4.2. MarketMaker Data WTP Results

The *SPILM*, *NIAM*, Turnbull and log-logistic PM mean estimates are reported in Table 3. The *SPILM* estimates that registered producers, on average, are willing to pay $36.82 annually for the services provided by MarketMaker, and the *NIAM* estimates that, on average, producers are willing to pay $36.58 for such services. The *SPILM* and *NIAM* mean estimates were higher than the Turnbull estimate and lower than the parametric estimate. Moreover, both *SPILM and NIAM* estimates are within Turnbull's mean interval estimate. On the other hand, the PM mean estimate of $41.20 lies outside Turnbull's mean interval estimate [18.40; 38.47]. However, all mean estimates have overlapping 95% confidence intervals.

In contrast to the Turnbull procedure, the *SPILM* and *NIAM* approaches allows the estimation of the effect of producers' characteristics on their valuation of MarketMaker. Table 4 presents the marginal effects of the different covariates employed in the *SPILM*, as well as those estimated using the parametric model. *SPILM* estimation results indicate that active users of MarketMaker are willing to pay $17.08 more per year than their passive counterparts. The *SPILM* also predict that each additional marketing contact received due to participation with MarketMaker increases the annual WTP by $1.58. Lastly, *SPILM* results indicate that a $1,000 increase in total annual sales is expected to increase the annual WTP by only $0.03. Regarding the marginal effects estimated using the PM, two of the estimated effects are very similar to those estimated using *SPILM* and one of the effects is almost double (the effect of total sales). However, as in the case of the mean estimates, the marginal effects estimated using both models have overlapping 95% confidence intervals.

In the case of *NIAM*, the relationships between each covariate – USER_TYPE, CONTACTS and SALES – and annual producers' WTP for the serviced provided by MarketMaker are presented in Figure 3. In term of USER_TYPE, *NIAM* estimation results indicate that active users are willing to pay $16.13 more per year than passive users. *NIAM* results also indicate that producers' WTP is positively related to CONTACT and SALES. Figure 3 also highlights the flexibility of *NIAM* to identify nonlinearities in the relationship between the dependent and explanatory variables.

Finally, as an illustration, both *SPILM* and *NIAM* approaches were used to recover the conditional underlying probability function of producers' WTP for the services provided by MarketMaker. The PDF and CDF estimates of producers' WTP for the different models are displayed in Figure 4.

5. Summary and Conclusions

The purpose of this study was to develop alternative distribution-free estimation approaches that can be used to analyze interval-censored WTP data obtained using the DBDC elicitation method. The proposed estimators involve iterated procedures that combine nonparametric kernel density estimation of the errors of the WTP function with parametric or nonparametric estimation of its conditional mean function. Although estimation of the mean WTP can be extended in principle to other modeling techniques, this study focused on parametric linear and nonparametric additive models.

Monte Carlo simulation techniques were employed to compare the performance of the proposed estimators with those of the true parametric model and the Turnbull approach (the standard distribution-free approach used to analyze WTP data obtained with the DBDC elicitation method). A real data set was also used to illustrate the usefulness of the proposed estimation techniques in practice.

Overall, the simulation results show that the proposed semiparametric (*SPILM*) and nonparametric (*NIAM*) estimators are valid alternatives to the Turnbull approach. Relative to the correctly specified parametric model, the robustness gains of using *SPILM* to estimate the mean and marginal effects do not seem to result in significant efficiency losses. *SPILM* was also shown to be significatively more efficient than the Turnbull method but it requires that the mean function is correctly specified. The relatively mild assumption used in *NIAM*, that the mean distribution function is of the additive form, results in significant efficiency gains and bias reduction relative to the Turnbull approach which does not require the specification of the mean function. The proposed estimation techniques were also shown to have three additional advantages relative to the Turnbull approach: 1) they provide point estimates of the mean WTP; 2) allow the estimation of the marginal effects of covariates on the mean WTP; and 3) allow the estimation of the underlying WTP probability distribution functions at any point. Finally, results of the empirical analysis demonstrate the advantages of the proposed methods.

Table 1. Mean Estimators Comparison using Monte Carlo Simulation

N	**Estimator**	**RMSE**	**Bias**	**SE**
100	SPILM	0.456	0.003	1.752
	NIAM	0.686	0.022	1.774
	Turnbull	1.205	-0.096	2.145
	PM	0.456	0.008	1.775
200	SPILM	0.327	-0.004	1.323
	NIAM	0.406	0.010	1.310
	Turnbull	0.772	-0.213	1.548
	PM	0.323	0.002	1.328
500	SPILM	0.206	0.018	0.830
	NIAM	0.273	0.026	0.840
	Turnbull	0.518	-0.135	0.963
	PM	0.198	0.013	0.830

Table 2. Marginal Effect Estimators Comparison using Monte Carlo Simulation

N	**Estimator**	**Marginal Effect**	**RMSE**	**Bias**	**SE**
100	SPILM	X_1	0.124	0.038	0.119
		X_2	1.096	0.098	1.097
		X_3^{d1}	1.186	-0.162	1.181
		X_3^{d2}	1.111	-0.096	1.112
	PM	X_1	0.132	0.025	0.130
		X_2	1.122	0.105	1.123
		X_3^{d1}	1.199	-0.140	1.197
		X_3^{d2}	1.111	-0.108	1.111
200	SPILM	X_1	0.075	0.017	0.073
		X_2	0.807	0.077	0.807
		X_3^{d1}	0.858	0.042	0.862
		X_3^{d2}	0.852	-0.080	0.853
	PM	X_1	0.077	0.007	0.077
		X_2	0.815	0.068	0.816
		X_3^{d1}	0.870	0.034	0.874
		X_3^{d2}	0.822	-0.068	0.823
500	SPILM	X_1	0.039	0.003	0.039
		X_2	0.482	0.031	0.483
		X_3^{d1}	0.552	0.076	0.550
		X_3^{d2}	0.499	-0.020	0.501
	PM	X_1	0.038	-0.002	0.039
		X_2	0.478	0.053	0.477
		X_3^{d1}	0.534	0.093	0.529
		X_3^{d2}	0.494	-0.029	0.496

Table 3. Mean Producers' WTP by Estimator, MarketMaker Valuation Data.

Estimator	Mean Estimate	SE
SPILM	36.815	3.675
NIAM	36.584	3.849
Turnbull	28.435 [a]	3.166
Log-logistic PM	41.197	6.772

[a]Turnbull's lower and upper bounds mean estimates were 18.40 and 38.47, respectively. Turnbull's mean estimate shown in the Table was calculated as $\hat{E}(Y) = \sum_{q=1}^{Q} \frac{a_{q-1}+a_q}{2} (\hat{F}_q - \hat{F}_{q-1})$.

Table 4. *SPILM* and log-logistic *PM* Marginal Effect Estimates using the MarketMaker Valuation Data.

Variable	***SPLIM***			**Log-logistic *PM***		
	Marginal Effect		**SE**	**Marginal Effect**		**SE**
USER_TYPE (Active user =1, Passive user=0)	17.078	**[a]	9.493	31.363	***	12.290
CONTACTS	1.584	*	1.061	1.371	*	0.889
SALES ($1,000)	0.026	**	0.013	0.032	***	0.014

[a] Significance levels of 0.01, 0.05 and 0.10 are indicated by ***, ** and * respectively.

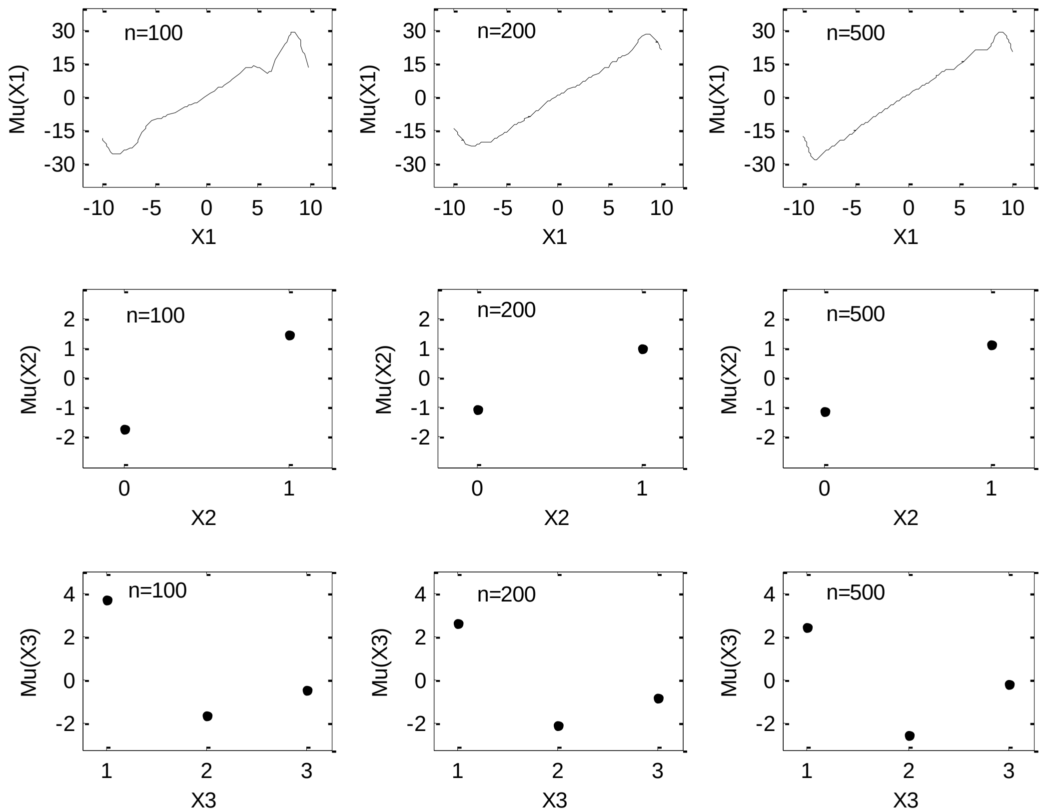

Figure 1. *NIAM* Fitted Smooth Functions Using Three Random Monte Carlo Finite Samples.

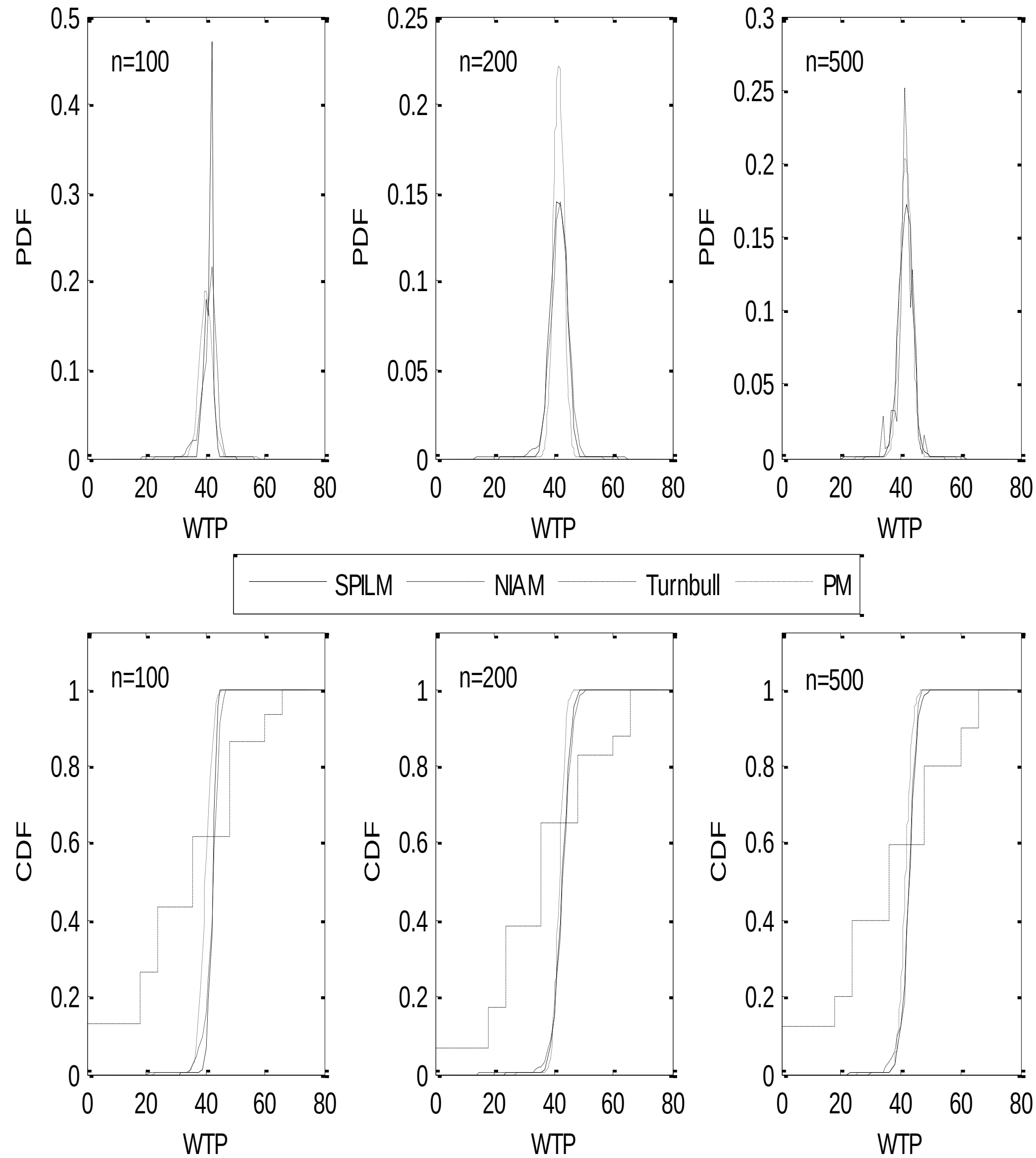

Figure 2. Distribution Function Estimates Using Three Random Monte Carlo Finite Samples.

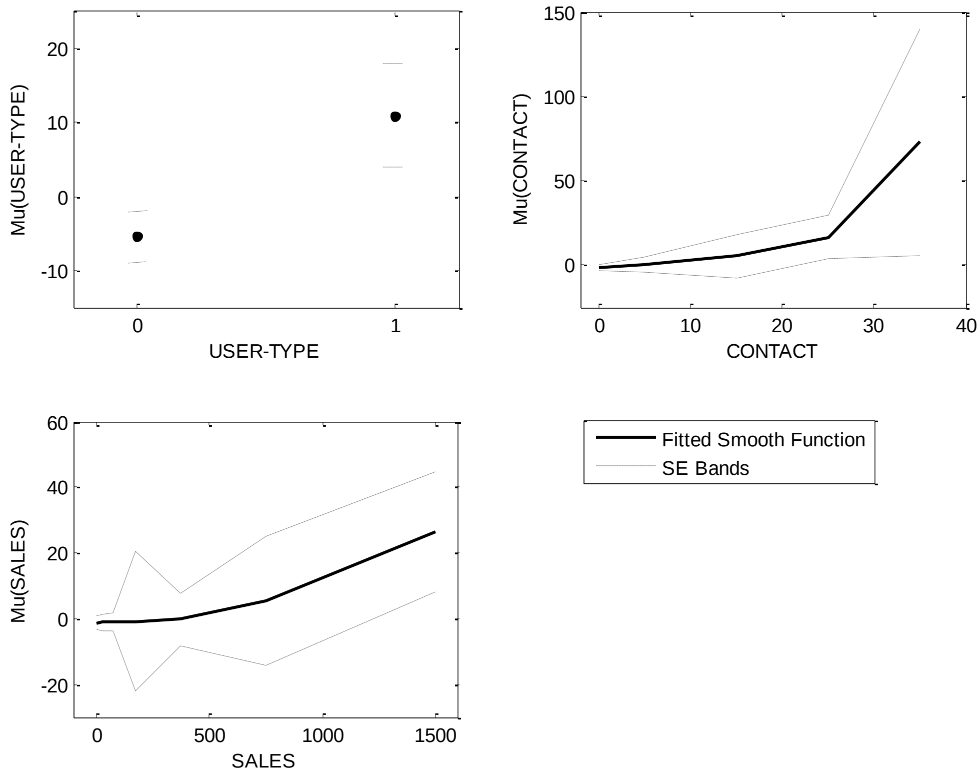

Figure 3. *NIAM* Fitted Smooth Functions and Standard Error Bands, MarketMaker Valuation Data.

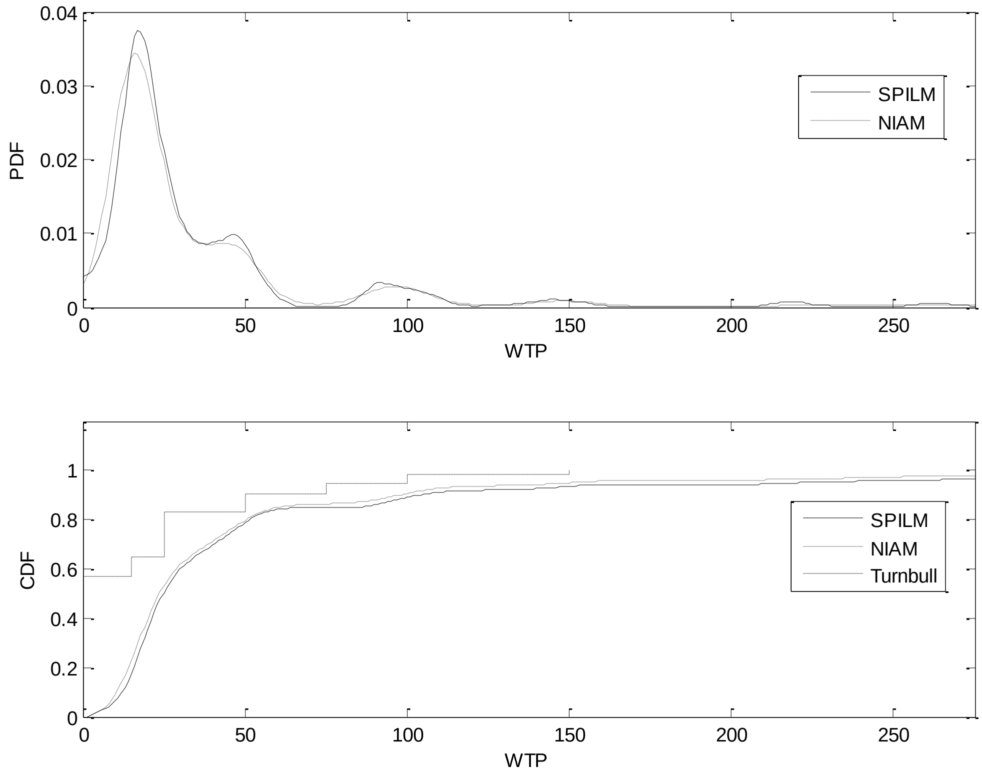

Figure 4. Distribution Function Estimates, MarketMaker Valuation Data.

6. References

An, M. 1996. Semiparametric estimation of willingness to pay distribution. Department of Economics Working Paper, Duke University.

Ayer, M., Brunk, H.D., Ewing, G.M., Reid, W.T., Silverman, E. 1955. An empirical distribution function for sampling with incomplete information. *The Annals of Mathematical Statistics* 26(4): 641-647.

Banfi, S., Farsi, M., Filippini, M., Jakob, M. 2008. Willingness to pay for energy-saving measures in residential buildings. *Energy Economics* 30: 503-506.

Boyle, K.J. 2003. Contingent valuation in practice. In *A primer on nonmarket valuation*. Edited by Champ, P.A., Boyle, K.J., Brown, T.C. Kluwer Academic Publishers, Dordrecht.

Braun, J., Duchesne, T., Stafford, J.E. 2005. Local likelihood density estimation for interval censored data. *The Canadian Journal of Statistics* 33(1):39-60.

Breffle, W.S., Morey, E.R., Lodder, T.S. 1998. Using contingent valuation to estimate a neighbourhood's willingness to pay to preserve undeveloped urban land. *Urban Studies* 35(4): 715–727.

Buja, A., Hastie, T., Tibshirai, R. 1989. Linear smoothers and additive models. *The Annals of Statistics* 17(2): 453-510.

Burton, M. 2000. A semi-parametric estimator of willingness-to-pay applied to dichotomous choice contingent valuation data. *Australian Economic Papers* 39(2): 200-214.

Calia, P., Strazzera, E. 2000. Bias and efficiency of single versus double bound models for contingent valuation studies: a Monte Carlo analysis. *Applied Economics* 32(10): 1329-1336.

Carson, R.T., Mitchell, R.C., Hanemann, W.M., Kopp, R.J., Presser, S., Ruud, P.A. 1992. A contingent valuation study of lost passive use values resulting from the Exxon Valdez oil spill. A report to the attorney general of the state of Alaska. Available at : http://www.evostc.state.ak.us/Universal/Documents/Publications/Economic/Econ_Passive.pdf. [Retrived April 2012].

Carson, R.T., Wilks, L., Imber, D. 1994. Valuing the preservation of Autralia's Kakadu conservation zone. *Oxford Economic Papers* 46: 727-749.

Cameron, A.C., Trivedi, P.K. 2005. Microeconomics: methods and applications. Cambridge University Press, Cambridge.

Cameron, T.A. 1988. A new paradigm for valuing non-market goods using referendum data: Maximum likelihood estimation by censored logistic regression. *Journal of Environmental Economics and Management* 15(3): 355-379.

Day, B. 2007. Distribution-free estimation with interval-censored contingent valuation data: troubles with Turnbull? *Environmental and Resource Economics* 37(4): 777-95.

Diener, A., O'Brien, B., Gafni, A. 1998. Health care contingent valuation studies: a review and classification of the literature. *Health Economics*, 7: 313–326.

Escanciano, J.C., Jacho-Chávez, D.T. 2012. Root-n uniformly consistent density estimation in nonparametric regression models. *Journal of Econometrics* 167(2): 305-316.

Gomez, G., Calle, L.M.and Oller, R. 2004. Frequentist and bayesian approaches for interval-censored data. *Statistical Papers* 45(2): 139-73.

Haab, T. C. and McConnell, K.E. 1997. Referendum models and negative willingness to pay: alternative solutions. *Journal of Environmental Economics and Management* 32(2): 251-70.

Hanemann, W.M. 1985. Some issues in continuous- and discrete-response contingent valuation studies. *Northeastern Journal of Agricultural Economics* 14: 5-13.

Hanemann, M., Loomis, J., Kannimen, B. 1991. Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation. *American Journal of Agricultural Economics* 73(4): 1255-1263.

Hanemann, W. M. 1991. Willingness to pay and willingness to accept: How much can they differ? *American Economic Review* 81: 635–47.

Hanemann, W.M. 1994. Valuing the environment through contingent valuation. *The Journal of Economic Perspectives* 8(4): 19-43.

Hanemann,W. M., Kanninen, B. 1999. The statistical analysis of discrete-response cv data', in *Valuing environmental preferences: theory and practice of the contingent valuation method in the usa, ec, and developing countries*. Edited by I. J. Bateman and K. G. Willis. New York: Oxford University Press.

Hastie, T., Tibshirani, R. 1986. Generalized additive models. *Statistical Science* 1(3):287-318.

Hudson, D., Hite, H. 2003. Producer willingness to pay for precision application technology: implications for government and the technology industry. *Canadian Journal of Agricultural Economics*, 51, 39–53.

Kang, X., Braun, W.J., Stafford, J.E. 2011. Local regression when the responses are interval-censored. *Journal of Statistical Computation and Simulation* 81(10): 1247-1279.

Kaplan, E.L., Meier, P. 1958. Nonparametric estimation from incomplete observations. *Journal of the American Statistical Association* 53(282): 457-481.

Kauermann, G., Opsomer, J.D., 2004. Generalized cross-validation for bandwidth selection of backfitting estimates in generalized additive models. *Journal of Computational and Graphical Statistics* 13(1): 66-89.

Kenkel, P.L., Norris, P.E. 1995. Agricultural Producers' Willingness to Pay for Real-Time Mesoscale Weather Information. *Journal of Agricultural and Resource Economics* 20(2), 356-372.

Kristom, B. 1990. A non-parametric approach to the estimation of welafare measures in discrete response valuation studies. *Land Economics* 66(2): 135-139.

Krupnick, A., Alberini, A., Cropper, M., Simon, N., O'Brien, B., Goeree, R., Heintzelman, M. 2002. Age, health and the willingness to pay for mortality risk reductions: a contingent valuation study of Ontario residents, *Journal of Risk and Uncertainty*, 24(2):161–186.

Lipscomb, C. 2011. Using contingent valuation to measure property value impacts. *Journal of Property Investment and Finance* 29: 448-459.

Patrick, F.G. 1988. Mallee wheat farmers' demand for crop and rainfall insurance. *Australian Journal of Agricultural Economics*, 32(1):37–49.

Racine, J., Qi, L. 2004. Nonparametric estimation of regression functions with both categorical and continuous data. *Journal of Econometrics* 119: 99 – 130.

Turnbull, B.W. 1974. Nonparametric estimation of a survivorship function with doubly censored data. *Journal of the American Statistical Association* 69(345): 169-173.

Turnbull, B.W. 1976. The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data. *Journal of the Royal Statistical Society* Series B 38 (3): 290-95.

Watanabe, M. 2010. Nonparametric estimation of mean willingness to pay from discrete response valuation data. *American Journal of Agricultural Economics* 92(4): 1114-1135.

Yu, Q., Li, L., Wong, G. Y. C. 1998. Asymptotic variance of the GMLE of a survival function with interval-censored data. *Sankhyā: The Indian Journal of Statistics, Series A* 60(2): 184-197.

Zapata, S.D. 2012. The economic impact of the services provided by an electronic trade platform: the case of MarketMaker. In *Three essays on contingent valuation*. Ph.D. dissertation. Clemson University.

Zapata, S.D., Benavides, H.M., Carpio, C.E., Willis, D.B. 2012. The economic value of basin protection to improve the quality and reliability of potable water supply: the case of Loja, Ecuador. *Water Policy* 14, 1-13.

Zapata, S.D. and C.E. Carpio. The Theoretical Structure of Producer Willingness to Pay Estimates. *Agricultural Economics* (*Forthcoming*).

Zapata, S.D., C.E. Carpio, O. Isengildina-Massa and R.D. Lamie. The Economic Impact of the Services Provided by an Electronic Trade Platform: The Case of MarketMaker. *Journal of Agricultural and Resource Economics (Forthcoming)*.

7. Appendix

Appendix A. Kernel Functions and Bandwidth Selection

Three different kernel functions were used to compute the *NIAM* mean estimator $\hat{g}(\boldsymbol{X}_i)_{NIAM}$ depending on the type of variable under consideration. The 2th-order Epanechnikov kernel function for the kth continuous variable $K_k^c(\cdot)$ is given by

(A.1) $$K_k^c\left(x_{ik}, x_{jk}, h_k^c\right) = \frac{3}{4h_k^c}\left\{1 - \left(\frac{x_{ik}-x_{jk}}{h_k^c}\right)^2\right\} \times \mathbf{1}_K\left(\left|\frac{x_{ik}-x_{jk}}{h_k^c}\right| < 1\right),$$

where $\mathbf{1}_K(\cdot)$ is an indicator function and $h_k^c > 0$. For the kth unordered discrete variable the kernel function $K_k^{uod}(\cdot)$ is given by (Racine and Li, 2004)

(A.2) $$K_k^{uod}\left(x_{ik}, x_{jk}, h_k^{uod}\right) = \begin{cases} 1 & if\ x_{ik} = x_{jk} \\ h_k^{uod} & if\ x_{ik} \neq x_{jk} \end{cases},$$

where $0 \leq h_k^{uod} \leq 1$. Finally, the kernel function for the kth ordered discrete variable $K_k^{od}(\cdot)$ is given by (Racine and Li, 2004)

(A.3) $$K_k^{od}\left(x_{ik}, x_{jk}, h_k^{od}\right) = {h_k^{od}}^{|x_{ik}-x_{jk}|},$$

where $0 \leq h_k^{od} \leq 1$.

The bandwidth parameters h_k^c , h_k^{uod} and h_k^{od} were selected by the generalized cross-validation (*GCV*) procedure described in Kauermann and Opsomer (2004). The objective of the this procedure is to find the vector $\boldsymbol{h} = (h_1^c, \dots, h_{d^c}^c, h_1^{uod}, \dots, h_{d^{uod}}^{uod}, h_1^{od}, \dots, h_{d^{od}}^{od})$ that minimizes the adjusted mean squared error

(A.4) $$GCV(\boldsymbol{h}) = \frac{\left(\boldsymbol{Y}_{\mathbf{imp}} - \hat{\boldsymbol{g}}(\boldsymbol{X})_{NIAM}\right)^t\left(\boldsymbol{Y}_{\mathbf{imp}} - \hat{\boldsymbol{g}}(\boldsymbol{X})_{NIAM}\right)}{n\left\{1 - \sum_k \mathrm{tr}\left(\boldsymbol{S}_k^*\right)/n\right\}^2},$$

where $\hat{\boldsymbol{g}}(\boldsymbol{X})_{NIAM} = (\hat{g}(\boldsymbol{X}_i)_{NIAM}, \dots, \hat{g}(\boldsymbol{X}_n)_{NIAM})^t$, and for illustration purposes it is assumed that there are d^c continuous variables, d^{uod} unordered categorical variables and d^{od} ordered categorical variables such that $d^c + d^{uod} + d^{od} = d$. Note that $\widehat{\boldsymbol{g}}(\boldsymbol{X})_{NIAM}$ and the $\boldsymbol{S}_k^*$'s depend on bandwidth vector $\boldsymbol{h}$, even though this is suppressed in the notation.

The kernel function $W_b(\cdot)$ needed for estimation of the error density function $\hat{f}_\varepsilon(z)$ is set to be equal to the 2nd-order Epanechnikov kernel

(A.5) $$W\left(\frac{v}{b}\right) = \frac{3}{4}\left\{1 - \left(\frac{v}{b}\right)^2\right\} \times \mathbf{1}_W\left(\left|\frac{v}{b}\right| < 1\right),$$

where $\mathbf{1}_W(\cdot)$ is an indicator function.

The bandwidth parameter b was selected by adapting the the likelihood cross-validation (LCV) method developed by Braun et *al*. (2005) who proposed to redefine the observed intervals in terms of a series of disjoint intervals and then drop specific intervals form the original data based on their contribution to the presence of the created disjoint intervals. Instead of creating a series of disjoint intervals as in Braun et *al*. (2005), we propose to evaluate the estimator of the error density, $\hat{f}_\varepsilon$, *n* times using the observed error intervals and leaving out one error interval from the estimation at a time. Braun et *al*. (2005) This original approach was modified because the error intervals in DBDC data present a high level of overlapping, resulting in very small disjoint intervals which makes difficult or even impossible to observe error intervals in the original data that are not composed by the disjoint interval of interest. Specifically, the cross-validation method proposed aims to prevent possible overfitting problems by maximizing the (leave-one-out) log likelihood function given by

(A.6) $$\ln L(b) = \sum_{i=1}^{n} ln\left[\int_{I_{\varepsilon_i}} \hat{f}_\varepsilon^{(-i)}(t) d_t\right],$$

with respect to b, where $\int_{I_{\varepsilon_i}} \hat{f}_\varepsilon^{(-i)}(t) d_t$ is obtained by dropping the interval-censored error I_{ε_i} when estimating $\hat{f}_\varepsilon$. Dropping an error interval is achieved by removing that particular error interval in addition to all estimated error intervals on iteration step (0.a) that are completely enclosed by the error interval of interest. Once again, the bandwidth b is suppressed in the notation, even though $\hat{f}_{\varepsilon}^{(-i)}(t)$ depends on it.