Journal of Agricultural and Resource Economics 49(1):39–62 ISSN: 1068-5502 (Print); 2327-8285 (Online)

doi: 10.22004/ag.econ.337551

Distribution-Free Methods to Estimate Willingness-to-Pay Models Using Discrete Response Valuation Data

Samuel D. Zapata and Carlos E. Carpio

This study introduces distribution-free methods to estimate interval-censored willingness-to-pay (WTP) models. The approaches proposed encompass the recovery of WTP values using an iterated conditional expectation procedure and subsequent estimation of the mean WTP using parametric and nonparametric models. Methods allow us to estimate the effects of covariates on the mean WTP and the underlying probability distribution. We employ Monte Carlo simulations to compare the performance of the estimators proposed against standard parametric and nonparametric estimators. We illustrate the estimation techniques by assessing producers' WTP for services provided by an e-marketing website that helps connect farmers with local consumers.

Key words: additive models, double-bounded elicitation, kernel functions, iterated conditional expectation, nonparametric regression, Turnbull

Introduction

Contingent valuation (CV) is a survey-based method developed initially to elicit the value (i.e., willingness to pay, WTP) that people place on nonmarket resources such as environmental preservation (e.g., Carson et al., 1992; Hanemann, 1994; Bishop, 2018; Oviedo, Campos, and Caparrós, 2022). Although CV continues to be used to study issues related to agricultural and resource economics (e.g., demand for novel foods, technologies, and sustainable energy), new applications have also been found in other areas, such as health economics (Diener, O'Brien, and Gafni, 1998; Hudson and Hite, 2003; Sarasty et al., 2020; Pleeging et al., 2021).

Different elicitation formats can be used in CV (Carson and Hanemann, 2005). The double-bounded dichotomous choice (DBDC) approach is a popular elicitation option among CV practitioners because of its potential to reduce strategic bias (Hanemann, 1994; Boyle, 2003) and provide relatively more efficient estimates of central tendency (Hanemann, Loomis, and Kanninen, 1991). However, one drawback of the approach is that it generates interval-censored responses. Therefore, estimating measures of central tendency (e.g., mean WTP) as well as the covariates' marginal effects on mean WTP requires the use of specialized statistical techniques. Most empirical studies that use interval-censored responses from CV studies have used parametric methods, in which a distribution function for the WTP measure is specified, while other authors have advocated using distribution-free methods (e.g., Carson et al., 1992; Carson, Wilks, and Imber, 1994).

The nonparametric maximum likelihood (ML) estimation approach that Turnbull (1974, 1976) proposed is the standard method used to analyze interval-censored data, such as those collected using

Samuel D. Zapata (corresponding author, samuel.zapata@ag.tamu.edu) is an associate professor in the Department of Agricultural Economics at Texas A&M University. Carlos E. Carpio is a professor in the Department of Agricultural and Applied Economics at Texas Tech University.
We thank the anonymous reviewers for their insightful comments and suggestions.

Review coordinated by Christian A. Langpap.

DBDC. However, the Turnbull method has important limitations in CV applications. Although the Turnbull approach permits the recovery of the distribution of an interval-censored variable in the form of a discrete staircase probability distribution, it does not provide specific point estimates of central tendency, nor does it allow direct estimates of the effects that explanatory variables have on WTP estimates.

This study develops alternative distribution-free estimation approaches that can be used to analyze interval-censored WTP data. The models proposed involve iterated procedures to estimate the distribution of WTP and its corresponding conditional mean function. Compared to the Turnbull approach, the estimation approach proposed here provides a point estimate of the mean WTP and allows covariates' marginal effects on the mean WTP to be estimated. The approach also allows the estimation of the underlying WTP probability distribution function at any point, which is of especial interest in some CV applications using dichotomous choice elicitation methods (e.g., demand studies such as Lusk and Hudson, 2004; Lee et al., 2015; Sarasty et al., 2020). Moreover, the proposed methods can be used to evaluate the robustness of parametric procedures or as part of researchers' model-building efforts (e.g., to detect nonlinearities in the effects of covariates and the conditional distribution of the resulting errors).

We employ simulation techniques to compare the performance of the estimators proposed with the Turnbull approach and the conventional parametric linear model under different data-generation processes for the conditional mean and the error distribution. We also illustrate the use of the estimation techniques by studying producers' WTP for an e-commerce website. Although the paper focuses on modeling data obtained from the DBDC elicitation format, the methods can be applied to interval-censored WTP data obtained using other elicitation mechanisms, such as single-bounded dichotomous choice and payment card (Cameron, 1988; Cameron and Huppert, 1989).

Methods

WTP Function

The theoretical foundations of WTP functions are based on the consumer utility and producer profit maximization problems. For consumers, the WTP function is derived considering changes in the quality of goods or services consumed (Hanemann, 1991). In the case of producers, WTP is a function of the perceived economic benefits associated with alternative inputs used in the production process (Zapata and Carpio, 2014). For consumers and producers, WTP is defined as a function of several variables, including relevant prices and quality levels.

Throughout the paper, to simplify the mathematical notation, we use Y_i to refer to the WTP value of the ith individual (consumer or producer) and $\boldsymbol{X}_i$ for a vector of d explanatory variables. Moreover, for every $i = 1, \ldots, n$, Y_i is assumed to be related to $\boldsymbol{X}_i$ via the following model:

$$Y_i = g\left(\boldsymbol{X}_i\right) + \epsilon_i, \tag{1}$$

in which the ϵ_i are independent and identically distributed (*i.i.d.*) errors, with marginal density f_ε, 0 mean, and finite variance σ^2. It is also assumed that the ϵ_i are independent of the d-dimensional predictor vector $\boldsymbol{X}_i$. Further, $g(\boldsymbol{X}_i)$ is a function that represents the conditional mean function of Y_i given $\boldsymbol{X}_i$.

DBDC Elicitation and Estimation

Since Hanemann (1985) introduced the DBDC elicitation approach, it has been preferred over other elicitation methodologies, such as the single-bounded dichotomous choice formats, as it provides statistical efficiency gains (Hanemann and Kanninen, 1999).[1] The DBDC elicitation format asks

[1] However, DBDC could violate incentive compatibility (Johnston et al., 2017).

respondents two rounds of questions to bound their true WTP. First, every respondent i is presented with an initial bid (B_i) and asked whether she would be willing to pay that amount for the good or service in question. Based upon her initial response, the second question presents a follow-up bid, higher (B_i^u) or lower (B_i^l). Consequently, every Y_i (i.e., actual WTP) is observed to fall into one of four intervals: $(-\infty, B_i^l)$, $[B_i^l, B_i)$, $[B_i, B_i^u)$ and $[B_i^u, +\infty)$, $i = 1, \ldots, n$.

DBDC discrete responses are commonly analyzed using parametric ML estimation methods (e.g., Hanemann, Loomis, and Kanninen, 1991; Zapata et al., 2013). One of the primary advantages of the parametric approach is that it allows covariates to be included in the modeling process. Thus, the marginal effects of the explanatory variables, $\boldsymbol{X}_i$, on the conditional mean WTP function can be estimated. On the other hand, parametric ML methods rely on *a priori* assumptions about the underlying distribution function of respondents' WTP. Hence, if the distribution function is misspecified, parameter estimates and any subsequent estimated functions, including welfare estimates and marginal effects, may be inconsistent.

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

However, the Turnbull approach is not without shortcomings. First, the estimated CDF is defined only up to a discrete set of observed points given by the bid amounts used in the WTP questions (i.e., the estimated CDF is a staircase function). Second, the method does not provide a point estimate of the mean WTP value, only upper- and lower- bound estimates. Third, the assessment of systematic differences in WTP based on covariates is very involved (Haab and McConnell, 2003). Although conditional distribution functions could be estimated by limiting the Turnbull estimation to observations with similar covariate levels, implementing this stratified approach may not be feasible in practice, as the number of possible combinations of the covariate levels increases rapidly with each additional variable considered.

Distribution-free methods should overcome the Turnbull approach and traditional parametric methods' intrinsic limitations. However, few alternative distribution-free estimation procedures to analyze DBDC responses have been proposed (e.g., An, 2000; Burton, 2000; Watanabe, 2010). Although existing distribution-free methods allow covariates to be included in the analysis, they still rely on certain parametric assumptions to estimate either the WTP function's distribution or its conditional value. An (2000) and Burton (2000) opted to use semiparametric proportional hazard specifications, commonly employed in duration models, in which a component of the hazard function is defined as a functional parametric form. Watanabe (2010) assumed a parametric specification for the conditional mean WTP function (i.e., $g(\boldsymbol{X}_i)$ in equation 1), without making parametric assumptions about the error term's distribution. In this paper, we propose an alternative distribution-free estimation framework that does not impose parametric assumptions about the distribution of the WTP and can address a broad class of parametric and nonparametric techniques to estimate the conditional mean WTP function.

The methods proposed here rely on nonparametric kernel estimation techniques. None of the distribution-free estimation methods available currently to estimate DBDC data use kernel-based procedures. This may have been because traditional kernel estimation techniques require continuous observations of the dependent variable for calculating weighting functions, in contrast to the interval-censored observations obtained in DBDC CV studies. However, the iterative algorithms developed by Kang, Braun, and Stafford (2011) and Braun, Duchesne, and Stafford (2005) make it possible to adapt kernel estimation techniques to interval-censored data.

Iterative Estimation Approach

This study extends the current CV literature by introducing a flexible, distribution-free approach to analyze discrete WTP responses collected using the DBDC elicitation format. Two specific models are presented within the proposed estimation framework: the Semiparametric Iterated Linear Model (SPILM) and the Nonparametric Iterated Additive Model (NIAM). The SPILM assumes a parametric specification for the mean WTP function (semiparametric procedure), and in the NIAM, the mean WTP function is also estimated nonparametrically (nonparametric procedure). To the best of our knowledge, this is the first attempt to use fully nonparametric methods that allow covariates to be included in the analysis of DBDC data. On the other hand, the semiparametric method (SPILM) can be considered an alternative to the distribution-free models proposed by An (2000), Burton (2000), and Watanabe (2010).

The models proposed here do not impose arbitrary parametric assumption on the underlying distribution function of the errors, ϵ_is, as their marginal density function, f_ε, is estimated using a nonparametric iterated conditional expectation procedure (Braun, Duchesne, and Stafford, 2005). Further, the conditional mean WTP, $g(\boldsymbol{X})$, is estimated using linear regression techniques in the case of SPILM and nonparametric additive regression methods in the case of the NIAM.

The mathematical relation underlying the proposed procedure is given by

$$E[Y_i \mid Y_i \in I_i] = g(\boldsymbol{X}_i) + E[\epsilon_i \mid I_{\varepsilon_i}], \tag{2}$$

in which $E[Y_i \mid Y_i \in I_i]$ is the conditional expectation of Y_i given that $Y_i \in I_i$, 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, Braun, and Stafford, 2011). Note that equation (2) uses $E[Y_i \mid Y_i \in I_i]$ rather than Y_i because the Y_i are interval-censored and observed as $I_1, I_2, \ldots, I_n$. Estimating equation (2) involves eight major steps: Steps 1–3 provide starting values for the conditional mean function, $g(\boldsymbol{X})$ and distribution of the errors $f_\varepsilon(\varepsilon)$. Steps 4–8 estimate the conditional mean function and the distribution of the errors iteratively until convergence:

Estimate Starting Values to Initiate Iterations (denoted with 0 indices):

1. For all Y_i, compute the interval midpoints: $Y_i^o = \frac{L_i + R_i}{2}$.
2. Compute the initial conditional mean function estimates: $\hat{g}_0(\boldsymbol{X}_i)_\xi$, $\xi = SPILM, NIAM$, using $\boldsymbol{Y}^o = (Y_1^o, \ldots, Y_n^o)^T$. That is, $g(\boldsymbol{X}_i)$ is estimated using parametric regression in SPILM or nonparametric regression in the NIAM. See section below.
3. Set initial marginal density function, $\hat{f}_{\varepsilon;0}(\hat{\epsilon}_i^0)$, of the errors, $\hat{\epsilon}_i^0 = Y_i^o - \hat{g}_0(\boldsymbol{X}_i)_\xi$, as a uniform density function in the range $[\min(L_i - \hat{g}_0(\boldsymbol{X}_i)_\xi), \max(R_i - \hat{g}_0(\boldsymbol{X}_i)_\xi)]$.[2]

Conduct Iterative Steps (denoted with $j \geq 1$ indices)

4. Estimate the errors' marginal density, f_ε, using the iterated conditional expectation procedure developed by Braun, Duchesne, and Stafford (2005):

   a. Estimate the interval-censored errors as

$$I_{\varepsilon_i} = \left[L_i - \hat{g}_{j-1}(\boldsymbol{X}_i)_\xi, R_i - \hat{g}_{j-1}(\boldsymbol{X}_i)_\xi\right].$$

[2] Braun, Duchesne, and Stafford (2005) showed that the final estimate of f_ε does not depend upon the density function used in the initial iteration step.

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

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

in which $W_b(v) = b^{-1}W(v/b)$, $W(\cdot)$ is a kernel density function with scale parameter b and z is any real number. See the appendix for the method used to select b.

5. Compute the conditional expectation of the ϵ_i:

$$\hat{E}\left[\epsilon_i \mid I_{\varepsilon_i}\right] = \frac{\int_{I_{\varepsilon_i}} z\hat{f}_{\varepsilon;j}(z)d_z}{\int_{I_{\varepsilon_i}} \hat{f}_{\varepsilon;j}(z)d_z}.$$

6. Estimate the conditional expectation of Y_i,$\boldsymbol{Y}_{imp}$. At the jth iteration step, the ith element of $\boldsymbol{Y}_{imp}$ is given by $\hat{E}[Y_i \mid Y_i \in I_i] = \hat{g}_{j-1}(\boldsymbol{X}_i)_\xi + \hat{E}[\epsilon_i \mid I_{\varepsilon_i}]$.
7. Reestimate $\hat{g}_j(\boldsymbol{X})_\xi$ using the $\boldsymbol{Y}_{imp}$ estimated in the previous 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]

The sections below provide more details about the methods for estimating the conditional mean function (Steps 2 and 7) and the kernel functions needed for Step 4b, Step 2, and Step 7.

Estimating the Conditional Mean and Kernel Functions

In practice, the conditional mean WTP function, $g(\boldsymbol{X})$, can be defined and estimated using a broad range of functional forms and estimation approaches. In this study, we used two common and juxtaposed options: A fully parametric form estimated using linear regression and a flexible nonparametric specification estimated using additive regression.

Linear Regression

In the Semiparametric Iterated Linear Model (SPILM), the conditional mean function of Y is the ordinary multiple linear regression function:

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

and the estimates of the parameters β_0, β_1, …, β_d are obtained by least squares using $\boldsymbol{Y}_{imp}$ as the dependent variable. The SPILM estimate of the mean Y, $\hat{g}(\boldsymbol{X})_{SPILM}$, is calculated by averaging the estimate of equation (3), $\hat{g}(\boldsymbol{X}_i)_{SPILM}$, for all individuals:

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

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

[4] An absolute difference of less than 10^{-5} in successive objective function estimates (e.g., $|\hat{g}_j(\boldsymbol{X})_\xi - \hat{g}_{j-1}(\boldsymbol{X})_\xi|$, using the average conditional means across all observations as described in equations 4 and 9) was used to declare convergence on every iterative procedure employed in this study. Alternatively, instead of the average conditional mean, $\hat{g}(\boldsymbol{X})_\xi$ at each iteration can be evaluated at a given point of interest, $\boldsymbol{X} = \chi$.

Nonparametric Additive Regression

Although there are several options to estimate the $g(\boldsymbol{X})$ function nonparametrically, we used a nonparametric additive model rather than a multivariate kernel regression for the following reasons. First, additive models are affected less by the curse of dimensionality and multicollinearity. Second, their marginal effects are easier to interpret. Third, additive model estimates have a faster convergence rate than multivariate kernel estimates (Buja, Hastie, and Tibshirani, 1989; Cameron and Trivedi, 2005). Finally, the majority of WTP studies use an additive mean parametric function. The nonparametric additive model assumes that

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

in which the $\mu_k(\cdot)$ are standardized smooth functions so that $E[\mu_k(\cdot)] = 0$ for every k (Hastie and Tibshirani, 1986; Kauermann and Opsomer, 2004). As shown in Kauermann and Opsomer (2004), the $\mu_k(\cdot)$s can be estimated jointly. Consider the kth additive function estimator in particular:

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

in which $\widehat{\boldsymbol{\mu}}_k = \{\hat{\mu}_k(x_{1k}), \ldots, \hat{\mu}_k(x_{nk})\}^T, \widehat{\boldsymbol{\mu}}_{-k} = \sum_{r \neq k} \widehat{\boldsymbol{\mu}}_r$ 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 that the estimators are identifiable, $\boldsymbol{I}_n$ denotes an identity matrix, $\mathbf{1}$ is an n-vector of ones, and S_k is a $n \times n$ smoothing matrix, the ij element of which is given by

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

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

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

in which $\widehat{\boldsymbol{\mu}} = (\widehat{\boldsymbol{\mu}}_1^T, \ldots, \widehat{\boldsymbol{\mu}}_d^T)^T$, $\boldsymbol{S}^* = (\boldsymbol{S}_1^{*T}, \ldots, \boldsymbol{S}_d^{*T})^T$, and $M = \begin{pmatrix} \boldsymbol{I}_n & \boldsymbol{S}_1^* & \cdots & \boldsymbol{S}_1^* \\ \boldsymbol{S}_2^* & \boldsymbol{I}_n & \cdots & \boldsymbol{S}_2^* \\ \vdots & & \ddots & \vdots \\ \boldsymbol{S}_d^* & \boldsymbol{S}_d^* & \cdots & \boldsymbol{I}_n \end{pmatrix}$.

The Nonparametric Iterated Additive Model (NIAM) estimate of the mean Y, $\hat{g}(\boldsymbol{X})_{NIAM}$, also averages the individual $\hat{g}(\boldsymbol{X}_i)_{NIAM}$ in equation (5):

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

However, compared to the SPILM, in which the marginal effects are given by the coefficients $\hat{\beta}_1, \ldots, \hat{\beta}_d$, in the NIAM, the relations between covariates and the mean WTP are given by the smooth functions, $\mu_k(\cdot)$ (Buja, Hastie, and Tibshirani, 1989). Therefore, a covariate's marginal effect on the mean WTP is not constant. Consequently, the relations 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).

Kernel Functions

The computation of both the NIAM mean estimator, $\hat{g}(\boldsymbol{X}_i)_{NIAM}$, and the error density function estimator, $\hat{f}_\varepsilon(z)$, involves kernel functions (i.e., $K_k(\cdot)$ in equation 7 and $W_b(\cdot)$ in step 4b) of the iterative estimation approach. The kernel functions were selected based on asymptotic properties and their ability to model both continuous and categorical data. Three different kernel functions were used to estimate $\hat{g}(\boldsymbol{X}_i)_{NIAM}$. For continuous explanatory variables, we considered a second-order Epanechnikov kernel,[5] and for discrete variables with or without natural order we considered the kernel functions suggested by Racine and Li (2004). In the case of estimating the error density function, $\hat{f}_\varepsilon(z)$, the kernel function, $W_b(\cdot)$, was set to be equal to the second-order Epanechnikov kernel. The specific kernel functions used in this study are presented in the appendix.

Kernel functions depend upon the bandwidth or smoothing parameters, which are more important for the quality of the estimates than the kernel choice itself (Cameron and Trivedi, 2005). 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 the appendix). The bandwidth parameter, b, of the kernel density, $W_b(v)$, in the error density function was estimated using a modified version of the likelihood cross-validation method presented by Braun, Duchesne, and Stafford (2005) (see the appendix).

Estimating the Probability Distribution

The traditional focus of CV studies is the estimation of the mean WTP values; however, some applications also require the estimation of the probability distribution of the WTP values (e.g., Lusk and Hudson, 2004; Sarasty et al., 2020). The iterative processes of the SPILM and the NIAM can also be used to recover the conditional CDF of WTP at any point. The probability distribution of Y can be estimated as

$$\begin{aligned} \Pr(Y \leq y \mid \boldsymbol{X} = \chi) &= \Pr\{g(\chi) + \epsilon \leq y\} \\ &= \Pr\{\epsilon \leq y - g(\chi)\}; \end{aligned} \tag{10}$$

therefore, the following estimators for the CDF of Y are proposed:

$$\hat{F}_Y(y)_\xi = \int_{a_0}^{y} \hat{f}_\varepsilon\left(y - \hat{g}(\chi)_\xi\right) d_y, \tag{11}$$

in which $f_\varepsilon(\cdot)$ and $g(\chi)$ are replaced by their estimates and $\xi = SPILM, NIAM$.

Parametric and Nonparametric Maximum Likelihood Estimators

In this section, we describe sequentially the parametric and nonparametric maximum likelihood methods (i.e., Turnbull's nonparametric ML) traditionally used in CV studies to analyze DBDC interval-censored data. These models served as a benchmark for the proposed SPILM and NIAM. Since both Turnbull and parametric estimators are based on maximum likelihood, their derivation has many common elements.

Denoting the lower bound of the ith observed 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

$$P(L_i \leq Y_i < R_i) = F(R_i) - F(L_i) \;\; i = 1, \ldots, n, \tag{12}$$

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

in which $F(\cdot)$ is the CDF of Y. Because the number of different bids used in the DBDC questions is usually less than the number of observations in the sample, some of the intervals observed are the same across individuals, which results in $M \leq n$ unique observed intervals, J_m, $m = 1, \ldots, M$, with boundary values of $\mathcal{L}_m$ and $\mathcal{R}_m$. Therefore, the log-likelihood function for the interval-censored Y_is can be written as

$$(13) \qquad \begin{aligned} \ln L &= \sum_{i=1}^{n} \ln\left[F(R_i) - F(L_i)\right] \\ &= \sum_{m=1}^{M} n_m \ln\left[F(\mathcal{R}_m) - F(\mathcal{L}_m)\right], \end{aligned}$$

in which n_m, $m = 1 \ldots M$, is the number of observations in which both $L_i = \mathcal{L}_m$ and $R_i = \mathcal{R}_m$. Parametric models (PM) assume that Y_i follows a certain distribution (see, e.g., Zapata et al., 2013). For example, in the simulation exercises and empirical illustration that follow, it was assumed that Y_i was distributed normally, skewed, and log-logistically.

To specify the log-likelihood function of Turnbull's nonparametric ML procedure, each unique interval observed, J_m, $m = 1, \ldots, M$, needs to be expressed as a union of Q disjoint closed intervals of the form $A_q = [a_{q-1}, a_q)$, $q = 1, \ldots, Q$, referred to as innermost intervals.[6] For instance, the $\mathcal{J}_m$ interval can be represented as $\cup_{q=1}^{Q} d_{mq} A_q$, in which d_{mq} is a dummy variable that indicates whether the qth innermost interval, A_q, is used to express the mth unique interval, $\mathcal{J}_m$. Specifically,

$$(14) \qquad d_{mq} = \begin{cases} 1 \text{ if } \mathcal{L}_m \leq a_{q-1} \text{ and } \mathcal{R}_m \geq a_q, \\ 0 \text{ otherwise} \end{cases} \qquad m = 1, \ldots, M;\; q = 1, \ldots, Q.$$

Then, the log-likelihood function in equation (13) can be expressed with respect to the innermost intervals:

$$(15) \qquad \ln L = \sum_{m=1}^{M} n_m \ln \sum_{q=1}^{Q} d_{mq} \left[F\left(a_q\right) - F\left(a_{q-1}\right)\right].$$

The Turnbull procedure considers that each $F = F(a_q)$ in equation (15) is a parameter to estimate and imposes the restriction that $0 = F_0 \leq F_1 \ldots \leq F_Q = 1$. Estimation is then carried out using Turnbull's self-consistent algorithm (Turnbull, 1976; Gómez, Calle, and Oller, 2004; Day, 2007). As mentioned, a limitation of the traditional Turnbull approach is that it does not provide a point estimate of the mean WTP but only for upper and lower bounds of its value. Hence, the lower-bound estimate of WTP that Haab and McConnell (1997) proposed was adapted to obtain a nonparametric measure of central tendency and facilitate comparison across models. In this case, the Turnbull midpoint approximation of the expected value of Y is equal to $\widehat{E}(Y) = \sum_{q=1}^{Q} \frac{a_{q-1}+a_q}{2} (\widehat{F}_q - \widehat{F}_{q-1})$, in which the $\widehat{F}_q$s are the solution to the log-likelihood function in equation (15).[7]

Below, we compare the performance of the SPILM and the NIAM as estimators of the mean and marginal effects against the performance of the traditional parametric and nonparametric maximum likelihood estimators.

---

[6] Assuming that Y is nonnegative, the complete set of Q innermost intervals is $[a_0, a_1)$, $[a_1, a_2) \ldots [a_{Q-1}, a_Q)$, in which $0 = a_0 < a_1 < \cdots < a_Q$. In the case of DBDC data, the bid amounts used in the WTP questions give the boundaries of the innermost intervals (a_qs).

[7] Although this approximation is not used in practice, it allows us to include Turnbull in the performance evaluation process. Moreover, this approximation was found to work better than misspecified parametric models.

Data and Study Design

The relative performance of the Semiparametric Iterated Linear Model (SPILM), the Nonparametric Iterated Additive Model (NIAM), and the Turnbull model was evaluated using two procedures. First, their performance was assessed using Monte Carlo simulation procedures under different conditional mean and error distribution specifications. Second, the performance of the new methods was assessed using a dataset from an empirical study evaluating producers' WTP for the services provided by an electronic trade platform. All of the econometric models and methods described in the study were estimated using MATLAB's built-in matrix manipulation functions and optimization packages (MATLAB, 2022).

Monte Carlo Simulation

The proposed models were evaluated and compared with the Turnbull approach and the conventional parametric linear model under four different scenarios. The data-generation process in the simulated scenarios consisted of all the possible combinations of linear and quadratic conditional means with symmetric and asymmetric errors. Scenario 1 represented the commonly considered case in empirical work, in which valuation data were simulated using a linear conditional mean and symmetrically distributed (normal) errors. The remaining scenarios illustrated conditions that are less common (from the perspective of data analysts). A linear conditional mean with right-skewed errors was used in Scenario 2. In Scenarios 3 and 4, the conditional mean was specified as a quadratic function with symmetric and skewed errors, respectively.

In Scenario 1, we considered three sample sizes frequently observed in CV applications. This scenario also serves as the basis for describing the data-generation process used in all simulation cases: A total of 100 datasets (iterations) that contained n observations each, $\{Y_i, \boldsymbol{X}_i\}_{i=1}^{n}$, $n \in \{100,200,500\}$, were generated. The response variable, Y_i, was related to a set of both continuous and categorical predictor variables through the following multiple linear regression model:

$$Y_i = 40 + 3X_{1i} + 3X_{2i} + 3X_{3i}^{d1} - 2X_{3i}^{d2} + 2\epsilon_i, \tag{16}$$

in which the X_{1i} are *i.i.d.* observations from a 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 a mean of 0 and variance of 1.

The resulting Y_i from equation (16) can be seen as individuals' true, but unobserved, WTP value given a set of observable characteristics, $\boldsymbol{X}_i$. Continuous Y_i were censored using a set of predefined bids similar to those employed in CV applications using the DBDC elicitation format. Specifically, one of four potential initial bid amounts (i.e., 24, 36, 48, and 60) was assigned randomly to each Y_i. The initial bids are the 20th, 40th, 60th, and 80th percentiles, respectively, obtained from the resulting empirical distribution of a sample of 50 observations generated using the regression model in equation (16) with no error term.[8] The corresponding follow-up bid amounts were 18 (10th percentile), 24, 36, and 48 if the initial bid assigned to the ith observation was higher than the true Y_i. However, if the initial bid assigned to the ith observation was lower than Y_i, corresponding higher follow-up bids of 36, 48, 60, and 66 (90th percentile) were assigned. Based upon the sample distribution used to generate the bids, the lower bound was set to 0 for those Y_i less than the initial and lower follow-up bid and to 80 for all Y_i greater than the initial and higher follow-up bid.[9]

[8] The initial bids were chosen following the methods employed in Calia and Strazzera (1999).

[9] All models in Scenario 1 were reestimated using lower and higher distribution limits to evaluate the effect of the selection of outermost bounds on the mean WTP estimate. It was found that SPILM and NIAM mean estimates were less sensitive to the choice of the outermost bounds compared to the Turnbull estimator. Also, when the model was correctly specified, SPILM and PM responded similarly to the selection of the distribution limits. In practice, optimal bid structure is based on *a priori* knowledge of the distribution of the WTP values obtained through a pretest stage or sequential setting (Kanninen, 1993; Alberini, 1995; Boyle et al., 1998; Scarpa and Bateman, 2000; Haab and McConnell, 2003).

Scenario 2 considered the same conditional mean as Scenario 1, but its error term, ϵ_i, followed a skew-normal distribution with 0 mean,[10] unit variance, and skewness equal to 1. In Scenarios 3 and 4, the linear conditional mean used in the previous two scenarios was modified by replacing the term $3X_1$ in equation (16) by a quadratic form given by $-0.6X_1^2 + 30$. Under this specification, the total additive effect of X_1 on Y reaches its maximum at $X_1 = 0$ and its lowest observed values at the limits of X_1 (i.e., -10 and 10). The quadratic form was chosen to evaluate the performance of the estimators when nonlinearities are present in the relation between the mean WTP values and the explanatory variables. The function was parameterized to peak within the defined range of X_1 and to have marginal effects of similar magnitude across specifications. Regarding the distribution of the errors, ϵ_i was set to be normally distributed in Scenario 3 (equal to Scenario 1) and skew-normally distributed in Scenario 4 (equal to Scenario 2). Scenarios 2– 4 used a sample size of 500 in each of the 100 iterations.

Next, simulated censored data were employed to estimate the conditional mean of Y using SPILM, NIAM, and PM procedures, while the Turnbull approach was used to estimate the corresponding unconditional mean. Marginal effects were also estimated for the PM and the SPILM. In the case of the NIAM, the marginal effects are not unique and are given by the $\mu_k(\cdot)$ functions. Hence, the NIAM marginal effects were illustrated using a random iteration ($n = 500$) from each of the conditions considered in Scenarios 1–4.

Given their generality, the NIAM and Turnbull are appropriate for all four scenarios considered. However, the SPILM and the PM are not always appropriate, except in Scenario 1. In Scenario 2, the PM is incorrect because it assumes a normal distribution. In Scenario 3, the PM incorrectly assumed a linear conditional mean. The PM is also incorrect in Scenario 4 because of the normality and linear conditional mean assumptions. SPILM is incorrect in Scenarios 3 and 4 because it assumes a linear conditional mean. The less "ideal" modeling scenarios for the SPILM and the PM aim to evaluate their robustness to invalid assumptions.

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

$$RMSE(\hat{\theta}) = \sqrt{\frac{1}{100}\sum_{s=1}^{100}\left[\hat{\theta}^{(s)} - \theta^{(s)}\right]^2}; \tag{17}$$

bias,

$$bias(\hat{\theta}) = \frac{1}{100}\sum_{s=1}^{100}\left[\hat{\theta}^{(s)} - \theta^{(s)}\right]; \tag{18}$$

and standard error (SE),

$$SE(\hat{\theta}) = \sqrt{\frac{1}{100}\sum_{s=1}^{100}\left[\hat{\theta}^{(s)} - \bar{\hat{\theta}}\right]^2}, \tag{19}$$

in which $\hat{\theta}^{(s)}$ and $\theta^{(s)}$ are the estimated and true parameter functions of interest (i.e., mean or marginal effects) of the sth iteration and $\bar{\hat{\theta}} = \frac{1}{100}\sum_{s=1}^{100}\hat{\theta}^{(s)}$.

Empirical Application: Producers' WTP Study

We also evaluated the SPILM, NIAM, Turnbull, and PM estimators using an actual DBDC dataset obtained in an empirical study that used CV methods to assess the monetary value registered

[10] The density function of a skew-normal distribution with shape parameter λ is given by $\phi(x;\lambda) = 2\phi(x)\Phi(\lambda x)$, $-\infty < \lambda < \infty$, where $\phi(\cdot)$ and $\Phi(\cdot)$ are the standard normal density and distribution function, respectively (Azzalini, 1986).

producers placed on the services provided by MarketMaker, a free electronic trade platform that helps connect local consumers with farmers in their areas. By the time of the study, MarketMaker was available in 18 states. The data consisted of 227 interval-censored observations collected using a combination of electronic and mail surveys. The original study is described in Zapata et al. (2013). The data were analyzed using parametric techniques in which the log-logistic distribution was the probability distribution considered to fit the data best. The purpose of using this second dataset was to contrast the performance of the proposed distribution-free estimation framework versus well-accepted parametric methods in an actual application. Compared to the original study, a reduced set of the available explanatory variables was used to illustrate the estimation techniques proposed here. The covariates employed to estimate the WTP models were the type of user based on the intensity of use (*USER_TYPE*, passive = 0, active = 1), marketing contacts acquired because of participation in MarketMaker (*CONTACTS*), and the firm total annual sales ($thousands) (*SALES*). Proper kernel functions described in the appendix were used in the NIAM to model *SALES* as a continuous variable, and *USER_TYPE* and *CONTACTS* as ordered categorical variables. The outermost lower and upper bounds for the censored WTP variable were set at $0 and $300, respectively.

The mean WTP for the services received from MarketMaker was estimated using the SPILM, NIAM, Turnbull, and log-logistic PM methods. Further, marginal effects were estimated for SPILM and the log-logistic PM and covariate-mean functions were estimated for the NIAM. The standard errors of the estimated means and marginal effects were calculated using a bootstrapping procedure (Cameron and Trivedi, 2005, p. 362) with 100 replications. The point-wise standard error bands that Buja, Hastie, and Tibshirani (1989) recommended were used to measure dispersion of the estimated smooth functions in the NIAM. The standard error bands represent the fitted curve ± 2 times the estimated standard error. Each smooth function's standard error was estimated as the sample mean standard error of the 100 replications for each unique covariate value. The different bandwidths' parameters of the SPILM and NIAM estimators were calculated using the 227 observations in the original data and then fixed at these values in each replication of the bootstrapping procedure.[11] Finally, the underlying CDF of producers' WTP for MarketMaker was calculated using expression (11).

Results

Monte Carlo Simulation

Tables 1–4 present the RMSE, bias, and SE of the different mean and marginal effects estimators under the four scenarios considered. These results are discussed sequentially, starting with the results related to the mean estimators. Concerning the effect of sample size on the mean estimators' performance (Table 1), the RMSE and SE associated with the proposed SPILM and NIAM mean estimators decreased as the sample size increased, which is similar to the conventional parametric and nonparametric estimators. For all mean estimators, no clear pattern of changes in the bias is found related to sample sizes.

The simulation also showed that the conditional mean estimators of the SPILM and the NIAM outperformed the unconditional Turnbull mean estimator with respect to RMSE, bias, and SE in all four scenarios (Tables 1 and 2). Further, the SPILM mean estimator performed similarly to the benchmark linear normal PM when conditions for this latter estimator are ideal (correct conditional mean and distribution functions) in Scenario 1. Hence, when the conditional mean is properly specified, the SPILM mean estimator appears to be a more robust alternative to the PM that also does not sacrifice efficiency.

Scenarios 2–4 represent cases that are not ideal for the linear normal PM or SPILM because either the mean or the distribution is misspecified. The NIAM outperformed the SPILM and

[11] The bandwidth parameters, 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 to estimate the standard errors.

Table 1. Mean Estimators Comparison under Scenario 1 (linear conditional mean and symmetric error distribution)

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

Notes: Estimators are the semiparametric iterated linear model (SPILM), the nonparametric iterated additive model (NIAM), Turnbull's (1974; 1976) nonparametric maximum likelihood estimation approach, and the parametric model (PM).

Table 2. Mean Estimators Comparison under Scenarios 2–4 (S2–S4)

Scenario (conditional mean/ error distribution)	Estimator	RMSE	Bias	SE
S2. Linear/asymmetric	SPILM	0.210	−0.033	0.763
	NIAM	0.289	−0.039	0.788
	Turnbull	0.540	−0.250	0.930
	PM	0.208	0.035	0.766
S3. Nonlinear/symmetric	SPILM	0.511	−0.092	0.982
	NIAM	0.263	0.095	0.876
	Turnbull	0.528	0.282	0.957
	PM	5.932	3.878	4.532
S4. Nonlinear/asymmetric	SPILM	0.483	−0.139	0.971
	NIAM	0.275	0.067	0.881
	Turnbull	0.562	0.257	1.017
	PM	9.537	5.138	8.145

Notes: Estimators are the semiparametric iterated linear model (SPILM), the nonparametric iterated additive model (NIAM), Turnbull's (1974; 1976) nonparametric maximum likelihood estimation approach, and the parametric model (PM).

the linear PM under conditional mean misspecifications (Scenarios 3 and 4) but not when the misspecification is only related to the distribution (PM in Scenario 2). The PM appears to be more sensitive to conditional mean misspecifications: It performed poorly compared to all estimators in Scenarios 3 and 4.

Overall, if the conditional mean is known *a priori*, a semiparametric approach like the SPILM is a robust and efficient alternative to estimate the mean WTP. Otherwise, the NIAM represents a very robust option, followed by the Turnbull procedure; however, there are some trade-offs associated with the use of the NIAM, which is less efficient and more computationally intensive. Results also show that the PM procedures seem more sensitive to conditional mean than distribution misspecification.

Table 3. Comparison of Marginal Effect Estimators under Scenario 1 (linear conditional mean and symmetric error distribution)

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

Notes: Estimators are the semiparametric iterated linear model (SPILM) and the parametric model (PM).

With respect to the marginal effect estimators, the magnitude of their errors decreased with the sample size (Table 3). However, neither the SPILM nor the PM was superior for any of the sample sizes considered. Specifically, the marginal effects' RMSE and SE values obtained using the SPILM estimator were generally lower than those obtained using the PM. However, the biases in the SPILM's marginal effects were generally higher than their counterparts estimated using the PM, although the differences were minimal in both cases. Therefore, consistent with the mean estimators, the simulation results indicated that, relative to the PM, the increases in robustness when using the SPILM to estimate marginal effects do not result in significant efficiency losses, even when the PM is the correctly specified model (Scenario 1).

When the conditional mean for the PM was properly defined but the distribution of model errors was not (Scenario 2), the marginal effects obtained using the linear normal PM had slightly smaller RMSE, bias, and SE compared to those estimated with the SPILM (Table 4). The opposite results were found when the conditional mean was not properly specified, even in the scenario with symmetric errors (Scenario 3). When both the conditional mean and the error distribution are misspecified for the PM (Scenario 4), the SPILM clearly performs better, although its conditional mean is also misspecified. Hence, like its mean estimator, it seems that the marginal effects estimated using the SPILM are less affected by misspecifications.

The high RMSE and SE values in Scenarios 3 and 4 relative to those in Scenarios 1 and 2 reflect the fact that both the SPILM and the PM are unable to adequately capture the existing quadratic

Table 4. Comparison of Marginal Effect Estimators under Scenarios 2–4 (S2–S4)

Scenario (conditional mean/error distribution)	Estimator	Marginal Effect	RMSE	Bias	SE
S2. Linear/asymmetric	SPILM	X_1	0.046	0.008	0.045
		X_2	0.438	−0.027	0.437
		X_3^{d1}	0.558	−0.077	0.553
		X_3^{d2}	0.521	0.033	0.520
	PM	X_1	0.046	0.000	0.046
		X_2	0.431	−0.020	0.430
		X_3^{d1}	0.515	−0.029	0.514
		X_3^{d2}	0.474	0.013	0.473
S3. Nonlinear/symmetric	SPILM	X_1	6.928	−0.068	0.218
		X_2	2.589	−0.152	2.584
		X_3^{d1}	3.047	−0.053	3.047
		X_3^{d2}	2.802	0.163	2.797
	PM	X_1	6.978	0.018	0.850
		X_2	2.820	0.419	2.789
		X_3^{d1}	3.348	0.358	3.329
		X_3^{d2}	2.967	−0.250	2.956
S.4. Nonlinear/asymmetric	SPILM	X_1	6.928	−0.066	0.223
		X_2	2.389	−0.055	2.388
		X_3^{d1}	3.240	0.139	3.237
		X_3^{d2}	2.842	0.212	2.834
	PM	X_1	7.122	0.303	1.643
		X_2	2.589	0.442	2.551
		X_3^{d1}	3.289	0.232	3.280
		X_3^{d2}	2.930	−0.177	2.924

Notes: Estimators are the semiparametric iterated linear model (SPILM) and the parametric model (PM). The root mean square error (RMSE) and bias of the marginal effect of variable X_1 in the SPILM and PM under the nonlinear conditional mean scenarios were estimated as $\sqrt{\frac{1}{100}\sum_{s=1}^{100}\frac{1}{500}\sum_{i=1}^{500}[\hat{\theta}^{(si)}-\theta^{(si)}]^2}$ and $\frac{1}{100}\sum_{s=1}^{100}\frac{1}{500}\sum_{i=1}^{500}[\hat{\theta}^{(si)}-\theta^{(si)}]$, respectively, where $\hat{\theta}^{(si)}$ is the estimated coefficient associated with X_1 in the sth iteration and $\theta^{(si)}$ is its corresponding true marginal effect evaluated at X_{1i}.

Table 5. Summary of the Efficiency and Bias of the SPILM, NIAM, and Turnbull Relative to PM in the Simulation Exercise

		Scenario	
	Estimator	**PM Is Well Specified**	**PM Is Misspecified**
Mean	SPILM	Similar bias and efficiency	Lower bias, more efficient
	NIAM	Similar bias, less efficient	Lower bias (least biased), more efficient (most efficient)
	Turnbull	Biased, less efficient	Lower bias (most biased), more efficient (least efficient)
Marginal effects	SPILM	Similar bias and efficiency	Lower bias, more efficient
	NIAM	Not available	Not available
	Turnbull	Not available	Not available

Notes: Estimators are the parametric model (PM), the semiparametric iterated linear model (SPILM), the nonparametric iterated additive model (NIAM), and Turnbull's (1974; 1976) nonparametric maximum likelihood estimation approach. The terms "least" and "most" biased or efficient are used only to compare the proposed estimators used in the simulations.

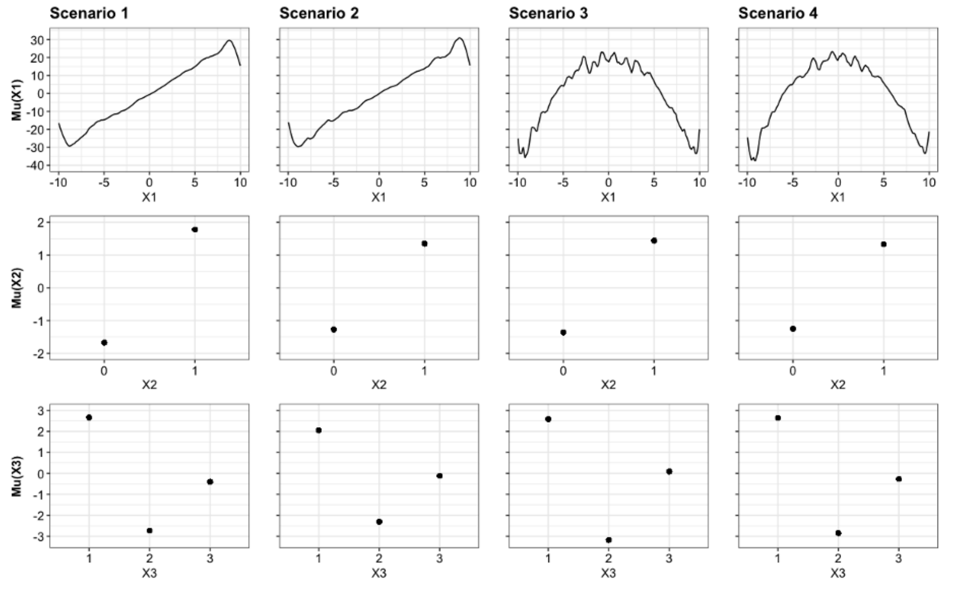

Figure 1. Estimated Nonparametric Iterated Additive Model (NIAM) Fitted Smooth Functions of a Random Sample of Each Scenario

relationship between X_1 and Y. It should be noted that the smaller magnitude of the bias associated with the marginal effect of the variable X_1 in Scenarios 3 and 4 was mainly caused by the fact that its true mean marginal effect (i.e., $E(-1.2X_1) = 0$) was designed to be equal to the linear marginal effects (i.e., $\frac{\partial Y}{\partial X_1} = 0$) estimated using the SPILM and the PM specifications. Table 5 summarizes the efficiency and bias of the proposed estimators (SPILM and NIAM) and Turnbull approach relative to the conventional PM.

Figure 1 illustrates the estimated fitted smooth functions, $\hat{\mu}_k(\cdot)$, obtained using the NIAM estimator in each scenario for a simulated random sample of 500 observations. As a reference, in Scenarios 1 and 2, the true effect of variable X_1 on Y implied by equation (17) is given by a straight line with a slope of 3. In Scenarios 3 and 4, the underlying relationship between X_1 on Y is given by a quadratic regression with a variable slope equal to $-1.2X_1$. In the case of the discrete variables, the true outcome differences between discrete levels are 3 for X_2 ($X_2 = 0$ serves as baseline) and 3 and -2 for X_3^{d1} and X_3^{d2} ($X_3 = 3$ serves as baseline), respectively. Figure 1 shows that the NIAM captured the true linear (Scenario 1 and 2) and quadratic (Scenario 3 and 4) relationships between X_1 and Y. Further, the estimated difference in Y between an observation in which $X_2 = 1$ and one in which $X_2 = 0$ was 3.45 units in Scenario 1, 2.62 units in Scenario 2, 2.80 units in Scenario 3, and 2.58 units in scenario 4. Similarly, compared to an observation in which $X_3 = 3$, the corresponding differences in Y when $X_3 = 1$ and $X_3 = 2$ were estimated to be 3.07 and -2.33 units, respectively, in Scenario 1, 2.17 and -2.19 units in Scenario 2, 2.50 and -3.27 units in Scenario 3, and 2.91 and -2.58 units in Scenario 4. Overall, the marginal effect plots presented in Figure 1 suggest that the NIAM adequately captured the underlying relations between the response variable Y and the predictors X_1, X_2, and X_3 in all the scenarios considered.

The same random samples employed to estimate the NIAM's smooth functions were also used to estimate the probability distribution of the error term in each scenario (Figure 2). Overall, and as discussed above, the PM provided the best fit to the sample datasets when the conditional mean was properly defined. However, the PM performed poorly under the model misspecification conditions presented in Scenarios 3 and 4. It was also observed that the SPILM and the NIAM were reasonable approximations of the true distribution function of the errors in Scenario 1. In Scenario 2, in which the conditional mean was specified correctly but the errors had an asymmetric distribution, some

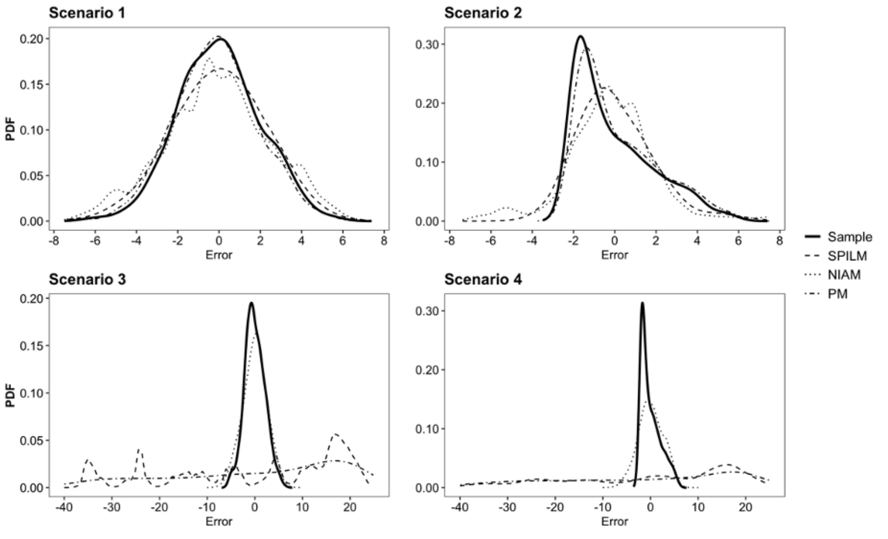

Figure 2. Estimated Distribution of the Error Term of a Random Sample of Each Scenario

Notes: Estimators are the semiparametric iterated linear model (SPILM), the nonparametric iterated additive model (NIAM), and the parametric model (PM).

Table 6. Mean Producers' WTP by Estimator, MarketMaker Valuation Data

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

Notes: Estimators are the semiparametric iterated linear model (SPILM), the nonparametric iterated additive model (NIAM), Turnbull's (1974; 1976) nonparametric maximum likelihood estimation approach, and the parametric model (PM).
[a]Turnbull's lower and upper bounds' mean estimates were 18.40 and 38.47, respectively. Turnbull's mean estimate shown in the table is the midpoint between the lower and upper bounds.

divergences were noted between the skewed sample errors and the more symmetric distributions estimated using the SPILM and the NIAM. Further, and like the PM, the SPILM was unable to capture the underlying distribution of the errors in Scenarios 3 and 4. On the other hand, the NIAM appeared to be a more robust estimator that closely followed the true distribution of the errors in the four scenarios considered. In short, these results suggest that the NIAM is a very robust estimator of the probability distribution of the errors.

Since researchers are unlike to know *a priori* the true conditional mean functions and the probability density function of error distributions (i.e., models will always be misspecified to a certain extent), the simulation results discussed above suggest that researchers should focus initial efforts on appropriately modeling the WTP mean function (e.g., exploring nonlinearities in the data) as this is more likely to affect mean and marginal effect estimates. In this sense, the NIAM estimator can be useful as part of researchers' modeling efforts. Efficiency gains can subsequently be obtained by using SPILM or PM methods.

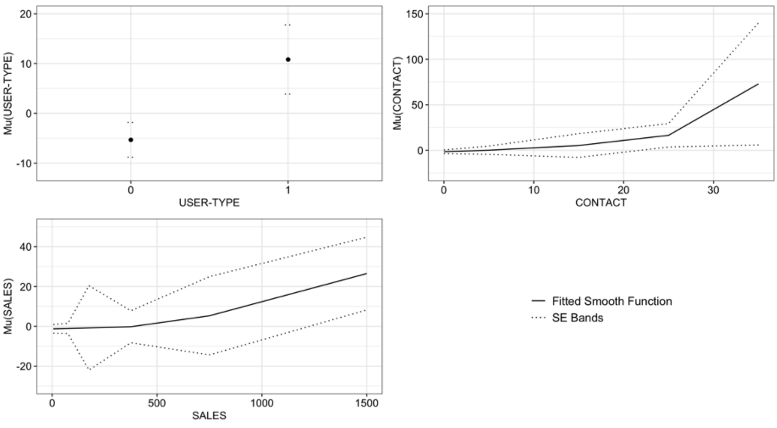

Figure 3. Nonparametric Iterated Additive Model (NIAM) Fitted Smooth Functions and Standard Error Bands, MarketMaker Valuation Data

Table 7. SPILM and Log-Logistic PM Marginal Effect Estimates Using the MarketMaker Valuation Data

	SPLIM		Log-Logistic PM	
Variable	**Marginal Effect**	**Standard Error**	**Marginal Effect**	**Standard Error**
USER_TYPE (active user =1, passive user = 0)	17.078**	9.493	31.363***	12.290
CONTACTS	1.584*	1.061	1.371*	0.889
SALES ($thousands)	0.026**	0.013	0.032***	0.014

Notes: Estimators are the semiparametric iterated linear model (SPILM) and the parametric model (PM). Single, double, and triple asterisks (*, **, ***) indicate significance at the 10%, 5%, and 1% level, respectively.

Empirical Application

The SPILM, NIAM, Turnbull, and log-logistic PM were also used to estimate producers' WTP for the services that MarketMaker provided. Table 6 reports the mean WTPs by estimator. The SPILM and the NIAM estimated that, on average, registered producers were willing to pay $36.82 and $36.58, respectively, annually for the services the e-marketing website provided. The SPILM and the NIAM mean estimates were higher than the Turnbull estimate (i.e., $28.44) and lower than the PM mean estimate (i.e., $41.20). Moreover, the Turnbull's mean interval estimate (i.e., [18.40, 38.47]) enclosed both the SPILM and the NIAM estimates. On the other hand, the log-logistic PM mean estimate fell outside Turnbull's mean interval estimate. However, all mean estimates had overlapping 95% confidence intervals.

In contrast to the Turnbull procedure, the SPILM and the NIAM allowed the effect of users' characteristics on their valuation of MarketMaker to be estimated. Table 7 presents the marginal effects of the different covariates employed in the SPILM, as well as those estimated using the log-logistic PM. The SPILM estimation results indicated that active users of MarketMaker are willing to pay $17.08 more per year than their passive counterparts. The SPILM also predicted that each additional marketing contact received because of MarketMaker would increase annual WTP by $1.58. Last, the SPILM results indicated that a $1,000 increase in total annual sales is expected to increase the annual WTP by $0.03. On the other hand, the log-logistic PM's estimated marginal effects suggested that active users are willing to pay $31.36 more per year, additional marketing contacts increased the annual WTP by $1.37, and additional sales also had the same moderate effect on WTP (i.e., $0.03/year). Although some differences were found between the SPILM and the log-

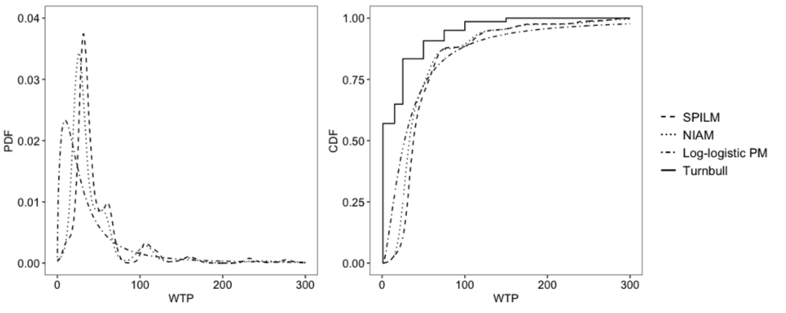

Figure 4. Distribution Function Estimates, MarketMaker Valuation Data

Notes: Estimators are the semiparametric iterated linear model (SPILM), the nonparametric iterated additive model (NIAM), Turnbull's (1974; 1976) nonparametric maximum likelihood estimation approach, and the parametric model (PM).

logistic PM (particularly with respect to user type), their marginal effects also had overlapping 95% confidence intervals.

In the case of the NIAM, Figure 3 presents the relations between each covariate and annual producers' WTP for MarketMaker's services. The NIAM results indicate that active users are willing to pay $16.13 more per year than passive users, which is closer to the marginal effect estimated using the SPILM than to the effect estimated using the log-logistic PM. The estimated smooth functions also suggested that producers' WTP is related positively to *CONTACTS* and *SALES*. Figure 3 highlights the NIAM's flexibility in identifying some nonlinearities in the relations between the dependent and explanatory variables.

Last, both the SPILM and the NIAM were used to recover the conditional underlying probability function of producers' WTP for the marketing services they receive, and their estimated distribution functions were compared with the counterpart distributions suggested by the log-logistic PM and Turnbull approach. For illustration purposes, the level of the explanatory variables was set to active *USER_TYPE*, 5 *CONTACTS*, and 75 *SALES* for the conditional models. Figure 4 displays the resulting probability density function (PDF) and CDF of the different models. Like the log-logistic PM, the PDFs of the SPILM and the NIAM suggest that the distribution of WTP values is skewed to the right, with most of the values concentrated in the lower end of the distribution; however, the PDF estimated using the log-logistic PM has a higher mass on the left side of the distribution and a lower level of peakedness. Further, the SPILM and the NIAM's estimated CDFs were very similar to each other and to the parametric log-logistic distribution but lower than that of the unconditional Turnbull. Compared to the Turnbull distribution, using the SPILM and the NIAM approach makes it possible to estimate a continuous CDF, which may be more appropriate for identifying specific valuation thresholds.

In summary, although the log-logistic PM estimator performs similarly to the NIAM and the SPILM estimators in this empirical application, there are some features of the WTP distribution that might not captured appropriately by this model (i.e., nonlinearities in the mean function and peakedness in the error distribution).

Summary and Conclusions

This paper introduced alternative distribution-free estimation methods that can be used to analyze interval-censored WTP data obtained using a variety of elicitation methods, including DBDC, the focus of this study. The estimators proposed involve iterated procedures that combine nonparametric kernel density estimation of the WTP function's errors with parametric or nonparametric estimation of its conditional mean function. Although estimating the mean WTP can be extended in principle to other functional forms and 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 commonly used parametric and nonparametric methods under different modeling scenarios. Overall, the simulation results showed that the proposed semiparametric (SPILM) and nonparametric (NIAM) estimators are valid alternatives to the parametric model and the Turnbull approach. Relative to the parametric model, the increased robustness when the SPILM is used to estimate the mean and marginal effects does not appear to result in significant efficiency losses. Moreover, when the linearity assumption is not valid, NIAM resulted in mean estimators with significantly less bias and more efficiency. Further, both the SPILM and the NIAM were shown to be significantly more efficient for the mean estimation than the Turnbull method in all the scenarios considered.

The estimation techniques proposed here were also shown to have three additional advantages over the nonparametric Turnbull approach: They (i) provide point estimates of the mean WTP, (ii) allow covariates' marginal effects on the mean WTP, and (iii) allow the estimation of the continuous underlying WTP probability distribution functions.

Further, an actual dataset was used to illustrate the proposed estimation techniques' properties in practice. Specifically, the models were employed to analyze producers' WTP for the services MarketMaker provided. The mean and marginal effects estimates obtained with the log-logistic PM, SPILM, and NIAM were not different from an economic and statistical point of view. However, the SPILM and the NIAM detected some features of the WTP mean function and the probability distribution of the errors that were not detected using the log-logistic PM. This added flexibility of the NIAM and the SPILM might prove helpful in other applications.

Although the proposed methods are presented as alternatives to procedures currently used for modeling CV data, some researchers might find them helpful to evaluate the robustness of their models or as part of their model-building efforts. For example, nonlinearities in the effects of covariates detected using NIAM can be used as a guide for specifying the conditional mean in the SPILM or PM procedures. Similarly, WTP distributional features observed using the NIAM or the SPILM can help with parametric distribution selection for PM procedures.

[First submitted March 2022; accepted for publication April 2023.]

References

Alberini, A. 1995. "Optimal Designs for Discrete Choice Contingent Valuation Surveys: Single-Bound, Double-Bound, and Bivariate Models." *Journal of Environmental Economics and Management* 28(3):287–306. doi: 10.1006/jeem.1995.1019.

An, M. Y. 2000. "A Semiparametric Distribution for Willingness to Pay and Statistical Inference with Dichotomous Choice Contingent Valuation Data." *American Journal of Agricultural Economics* 82(3):487–500. doi: 10.1111/0002-9092.00041.

Ayer, M., H. D. Brunk, G. M. Ewing, W. T. Reid, and E. Silverman. 1955. "An Empirical Distribution Function for Sampling with Incomplete Information." *Annals of Mathematical Statistics* 26(4):641–647. doi: 10.1214/aoms/1177728423.

Azzalini, A. 1986. "A Class of Distributions Which Includes the Normal Ones." *Statistica* 46(2): 199–208. doi: 10.6092/ISSN.1973-2201/711.

Bishop, R. C. 2018. "Warm Glow, Good Feelings, and Contingent Valuation." *Journal of Agricultural and Resource Economics* 43(3):307–320. doi: 10.22004/ag.econ.276497.

Boyle, K. J. 2003. "Contingent Valuation in Practice." In P. A. Champ, K. J. Boyle, and T. C. Brown, eds., *A Primer on Nonmarket Valuation,* No. 13 in The Economics of Non-Market Goods and Resources. Dordrecht, Netherlands: Springer, 111–169. doi: 10.1007/978-94-007-7104-8_4.

Boyle, K. J., H. F. MacDonald, H.-t. Cheng, and D. W. McCollum. 1998. "Bid Design and Yea Saying in Single-Bounded, Dichotomous-Choice Questions." *Land Economics* 74(1):49. doi: 10.2307/3147212.

Braun, J., T. Duchesne, and J. E. Stafford. 2005. "Local Likelihood Density Estimation for Interval-Censored Data." *Canadian Journal of Statistics* 33(1):39–60. doi: 10.1002/cjs.5540330104.

Buja, A., T. Hastie, and R. Tibshirani. 1989. "Linear Smoothers and Additive Models." *Annals of Statistics* 17(2):453–510. doi: 10.1214/aos/1176347115.

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. doi: 10.1111/1467-8454.00086.

Calia, P., and E. Strazzera. 1999. "Bias and Efficiency of Single vs Double Bound Models for Contingent Valuation Studies: a Monte Carlo Analysis." *Applied Economics* 32(10):1329–1336. doi: 10.1080/000368400404489.

Cameron, A. C., and P. K. Trivedi. 2005. *Microeconometrics: Methods and Applications*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9780511811241.

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. doi: 10.1016/0095-0696(88)90008-3.

Cameron, T. A., and D. D. Huppert. 1989. "OLS versus ML Estimation of Non-Market Resource Values with Payment Card Interval Data." *Journal of Environmental Economics and Management* 17(3):230–246. doi: 10.1016/0095-0696(89)90018-1.

Carson, R. T., and W. M. Hanemann. 2005. "Contingent Valuation." In *Handbook of Environmental Economics,* Vol. 2. Elsevier, 821–936. doi: 10.1016/S1574-0099(05)02017-6.

Carson, R. T., R. C. Mitchell, W. M. Hanemann, R. J. Kopp, S. Presser, and P. A. Ruud. 1992. *A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill*. A Report to the Attorney General of Alaska. A Report to the Attorney General of the State of Alaska. Available online at https://web.archive.org/web/20170810124216/http://www.evostc.state.ak.us/Universal/Documents/Publications/Economic/Econ_Passive.pdf.

Carson, R. T., L. Wilks, and D. Imber. 1994. "Valuing the Preservation of Australia's Kakadu Conservation Zone." *Oxford Economic Papers* 46(Supplement_1):727–749. doi: 10.1093/oep/46.Supplement_1.727.

Day, B. 2007. "Distribution-Free Estimation with Interval-Censored Contingent Valuation Data: Troubles with Turnbull?" *Environmental and Resource Economics* 37(4):777–795. doi: 10.1007/s10640-006-9061-8.

Diener, A., B. O'Brien, and A. Gafni. 1998. "Health Care Contingent Valuation Studies: A Review and Classification of the Literature." *Health Economics* 7(4):313–326. doi: 10.1002/(SICI)1099-1050(199806)7:4<313::AID-HEC350>3.0.CO;2-B.

Gómez, G., M. L. Calle, and R. Oller. 2004. "Frequentist and Bayesian Approaches for Interval-Censored Data." *Statistical Papers* 45(2):139–173. doi: 10.1007/BF02777221.

Haab, T. C., and K. E. McConnell. 1997. "Referendum Models and Negative Willingness to Pay: Alternative Solutions." *Journal of Environmental Economics and Management* 32(2):251–270. doi: 10.1006/jeem.1996.0968.

———. 2003. "The Distribution of Willingness to Pay." In *Valuing Environmental and Natural Resources,* Northampton, MA: Edward Elgar Publishing, 84–113. doi: 10.4337/9781840647044.00008.

Hanemann, W. M. 1985. "Some Issues in Continuous–and Discrete–Response Contingent Valuation Studies." *Northeastern Journal of Agricultural and Resource Economics* 14(1):5–13. doi: 10.1017/S0899367X00000702.

———. 1991. "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" *American Economic Review* 81(3):635–647.

———. 1994. "Valuing the Environment through Contingent Valuation." *Journal of Economic Perspectives* 8(4):19–43. doi: 10.1257/jep.8.4.19.

Hanemann, W. M., and B. Kanninen. 1999. "The Statistical Analysis of Discrete-Response CV Data." In I. J. Bateman and K. G. Willis, eds., *Valuing Environmental Preferences,* New York, NY: Oxford University Press, 302–441. doi: 10.1093/0199248915.003.0011.

Hanemann, W. M., J. Loomis, and B. Kanninen. 1991. "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation." *American Journal of Agricultural Economics* 73(4):1255–1263. doi: 10.2307/1242453.

Hastie, T., and R. Tibshirani. 1986. "Generalized Additive Models." *Statistical Science* 1(3): 297–310. doi: 10.1214/ss/1177013604.

Hudson, D., and D. Hite. 2003. "Producer Willingness to Pay for Precision Application Technology: Implications for Government and the Technology Industry." *Canadian Journal of Agricultural Economics* 51(1):39–53. doi: 10.1111/j.1744-7976.2003.tb00163.x.

Johnston, R. J., K. J. Boyle, W. Adamowicz, J. Bennett, R. Brouwer, T. A. Cameron, W. M. Hanemann, N. Hanley, M. Ryan, R. Scarpa, R. Tourangeau, and C. A. Vossler. 2017. "Contemporary Guidance for Stated Preference Studies." *Journal of the Association of Environmental and Resource Economists* 4(2):319–405. doi: 10.1086/691697.

Kang, X., W. J. Braun, and J. E. Stafford. 2011. "Local Regression When the Responses Are Interval-Censored." *Journal of Statistical Computation and Simulation* 81(10):1247–1279. doi: 10.1080/00949655.2010.482061.

Kanninen, B. J. 1993. "Optimal Experimental Design for Double-Bounded Dichotomous Choice Contingent Valuation." *Land Economics* 69(2):138–146. doi: 10.2307/3146514.

Kaplan, E. L., and P. Meier. 1958. "Nonparametric Estimation from Incomplete Observations." *Journal of the American Statistical Association* 53(282):457–481. doi: 10.1080/01621459.1958.10501452.

Kauermann, G., and J. D. Opsomer. 2004. "Generalized Cross-Validation for Bandwidth Selection of Backfitting Estimates in Generalized Additive Models." *Journal of Computational and Graphical Statistics* 13(1):66–89. doi: 10.1198/1061860043056.

Kristrom, B. 1990. "A Non-Parametric Approach to the Estimation of Welfare Measures in Discrete Response Valuation Studies." *Land Economics* 66(2):135. doi: 10.2307/3146363.

Lee, J.-S., V. Mogasale, J. K. Lim, M. Carabali, C. Sirivichayakul, D. D. Anh, K.-S. Lee, V. D. Thiem, K. Limkittikul, L. H. Tho, I. D. Velez, J. E. Osorio, P. Chanthavanich, L. J. Da Silva, and B. A. Maskery. 2015. "A Multi-Country Study of the Household Willingness-to-Pay for Dengue Vaccines: Household Surveys in Vietnam, Thailand, and Colombia." *PLOS Neglected Tropical Diseases* 9(6):e0003810. doi: 10.1371/journal.pntd.0003810.

Lusk, J. L., and D. Hudson. 2004. "Willingness-to-Pay Estimates and Their Relevance to Agribusiness Decision Making." *Review of Agricultural Economics* 26(2):152–169. doi: 10.1111/ j.1467-9353.2004.00168.x.

MATLAB. 2022. "MATLAB v. 9.12.2009381 (R2022a)." Natick, MA: MathWorks Inc.

Opsomer, J. D., and D. Ruppert. 1998. "A Fully Automated Bandwidth Selection Method for Fitting Additive Models." *Journal of the American Statistical Association* 93:605–619. doi: 10.1080/01621459.1998.10473714.

Oviedo, J. L., P. Campos, and A. Caparrós. 2022. "Contingent Valuation of Landowner Demand for Forest Amenities: Application in Andalusia, Spain." *European Review of Agricultural Economics* 49(3):615–643. doi: 10.1093/erae/jbab022.

Pleeging, E., J. Van Exel, M. J. Burger, and S. Stavropoulos. 2021. "Hope for the Future and Willingness to Pay for Sustainable Energy." *Ecological Economics* 181:106900. doi: 10.1016/ j.ecolecon.2020.106900.

Racine, J., and Q. Li. 2004. "Nonparametric Estimation of Regression Functions with Both Categorical and Continuous Data." *Journal of Econometrics* 119(1):99–130. doi: 10.1016/ S0304-4076(03)00157-X.

Sarasty, O., C. E. Carpio, D. Hudson, P. A. Guerrero-Ochoa, and I. Borja. 2020. "The Demand for a COVID-19 Vaccine in Ecuador." *Vaccine* 38(51):8090–8098. doi: 10.1016/j.vaccine.2020. 11.013.

Scarpa, R., and I. Bateman. 2000. "Efficiency Gains Afforded by Improved Bid Design versus Follow-Up Valuation Questions in Discrete-Choice CV Studies." *Land Economics* 76(2):299. doi: 10.2307/3147230.

Turnbull, B. W. 1974. "Nonparametric Estimation of a Survivorship Function with Doubly Censored Data." *Journal of the American Statistical Association* 69(345):169–173. doi: 10.1080/01621459.1974.10480146.

———. 1976. "The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data." *Journal of the Royal Statistical Society. Series B (Methodological)* 38(3): 290–295. doi: 10.1111/j.2517-6161.1976.tb01597.x.

Watanabe, M. 2010. "Nonparametric Estimation of Mean Willingness to Pay from Discrete Response Valuation Data." *American Journal of Agricultural Economics* 92(4):1114–1135. doi: 10.1093/ajae/aaq034.

Zapata, S. D., and C. E. Carpio. 2014. "The Theoretical Structure of Producer Willingness to Pay Estimates." *Agricultural Economics* 45(5):613–623. doi: 10.1111/agec.12110.

Zapata, S. D., C. E. Carpio, O. Isengildina-Massa, and R. D. Lamie. 2013. "The Economic Impact of Services Provided by an Electronic Trade Platform: The Case of MarketMaker." *Journal of Agricultural and Resource Economics* 38(3):359–378. doi: 10.22004/ag.econ.165933.

Appendix A. Kernel Functions and Bandwidth Selection

It is assumed that the d explanatory variable in $\boldsymbol{X}$ can be classified into d^c continuous variables, d^{od} ordered discrete variables, and d^{uod} unordered discrete variables, such that $d = d^c + d^{od} + d^{uod}$. In the NIAM, three different kernel functions were employed to model the relation between Y and each x_k independent variable. Specifically, the second-order Epanechnikov kernel function was used for continuous variables:

$$K_k^c\left(x_{ik}, x_{jk}, h_k^c\right) = \frac{3}{4h_k^c}\left\{1 - \left(\frac{x_{ik} - x_{jk}}{h_k^c}\right)^2\right\} \times \mathbf{1}_K\left(\left|\frac{x_{ik} - x_{jk}}{h_k^c}\right| < 1\right), \tag{A1}$$

in which $\mathbf{1}_K(\cdot)$ is an indicator function and $h_k^c > 0$. For discrete variables, the kernel functions Racine and Li (2004) proposed:

$$K_k^{od}\left(x_{ik}, x_{jk}, h_k^{od}\right) = h_k^{od|x_{ik}-x_{jk}|} \tag{A2}$$

and

$$K_k^{uod}\left(x_{ik}, x_{jk}, h_k^{uod}\right) = \begin{cases} 1 \text{ if } x_{ik} = x_{jk} \\ h_k^{uod} \text{ if } x_{ik} \neq x_{jk} \end{cases}, \tag{A3}$$

were considered to model explanatory variables with and without a natural order, respectively, where $0 \le h_k^{od} \le 1$ and $0 \le h_k^{uod} \le 1$.

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

$$\mathrm{GCV}(\boldsymbol{h}) = \frac{\left(\boldsymbol{Y}_{imp} - \widehat{\boldsymbol{g}}(\boldsymbol{X})_{NIAM}\right)^T\left(\boldsymbol{Y}_{imp} - \widehat{\boldsymbol{g}}(\boldsymbol{X})_{NIAM}\right)}{n\left\{1 - \Sigma_k \operatorname{tr}\left(\boldsymbol{S}_k^*\right)/n\right\}^2}, \tag{A4}$$

in which $\widehat{\boldsymbol{g}}(\boldsymbol{X})_{NIAM} = (\hat{g}(\boldsymbol{X}_i)_{NIAM}, \ldots, \hat{g}(\boldsymbol{X}_n)_{NIAM})^T$. Note that $\widehat{\boldsymbol{g}}(\boldsymbol{X})_{NIAM}$ and the $\boldsymbol{S}_k^*$s depend upon the bandwidth vector, $\boldsymbol{h}$, although this is suppressed in the notation.

The kernel function, $W_b(\cdot)$, used to estimate the error density function, $\hat{f}_\varepsilon(z)$, was also set to be equal to the second-order Epanechnikov kernel:

$$W\left(\frac{v}{b}\right) = \frac{3}{4}\left\{1 - \left(\frac{v}{b}\right)^2\right\} \times \mathbf{1}_W\left(\left|\frac{v}{b}\right| < 1\right), \tag{A5}$$

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

The bandwidth parameter, b, was selected by adapting the likelihood cross-validation (LCV) method developed by Braun, Duchesne, and Stafford (2005). In the original selection method, the overlapping intervals observed are redefined as a series of disjoint intervals and then specific intervals from the original data are omitted when estimating the CDF of each disjoint interval. Observed intervals are dropped based on their contribution to the presence of the selected disjoint interval. Rather than creating a series of disjoint intervals as in Braun , we proposed to evaluate the estimator of the error density, $\hat{f}_\varepsilon$, n times using the observed error intervals and leaving out of the estimation one error interval at a time. The original selection approach was modified because the error intervals in DBDC data show a high level of overlap, resulting in very small disjoint intervals, which makes it difficult or even impossible to observe error intervals in the original data that are not composed of the disjoint interval of interest.

The cross-validation method proposed is designed to prevent possible overfitting problems by maximizing the (leave-one-out) log-likelihood function:

$$\text{(A6)} \qquad \ln L(b) = \sum_{i=1}^{n} \ln \left[\int_{I_{\varepsilon_i}} \hat{f}_{\varepsilon}^{(-i)}(t) d_t \right],$$

with respect to b, in which $\int_{I_{\varepsilon_i}} \hat{f}_{\varepsilon}^{(-i)}(t) d_t$ is obtained by omitting the interval-censored error, I_{ε_i}, when estimating $\hat{f}_{\varepsilon}$. Dropping one error interval is achieved by removing the error interval of interest and all estimated error intervals that it encloses completely (in iterated step 4a). Note that the bandwidth, b, is suppressed in the notation as well, although $\hat{f}_{\varepsilon}^{(-i)}(t)$ depends on it.